独立行政法人国立病院機構
水戸医療センター

研 究 業 績 集

令和6年度

巻頭言

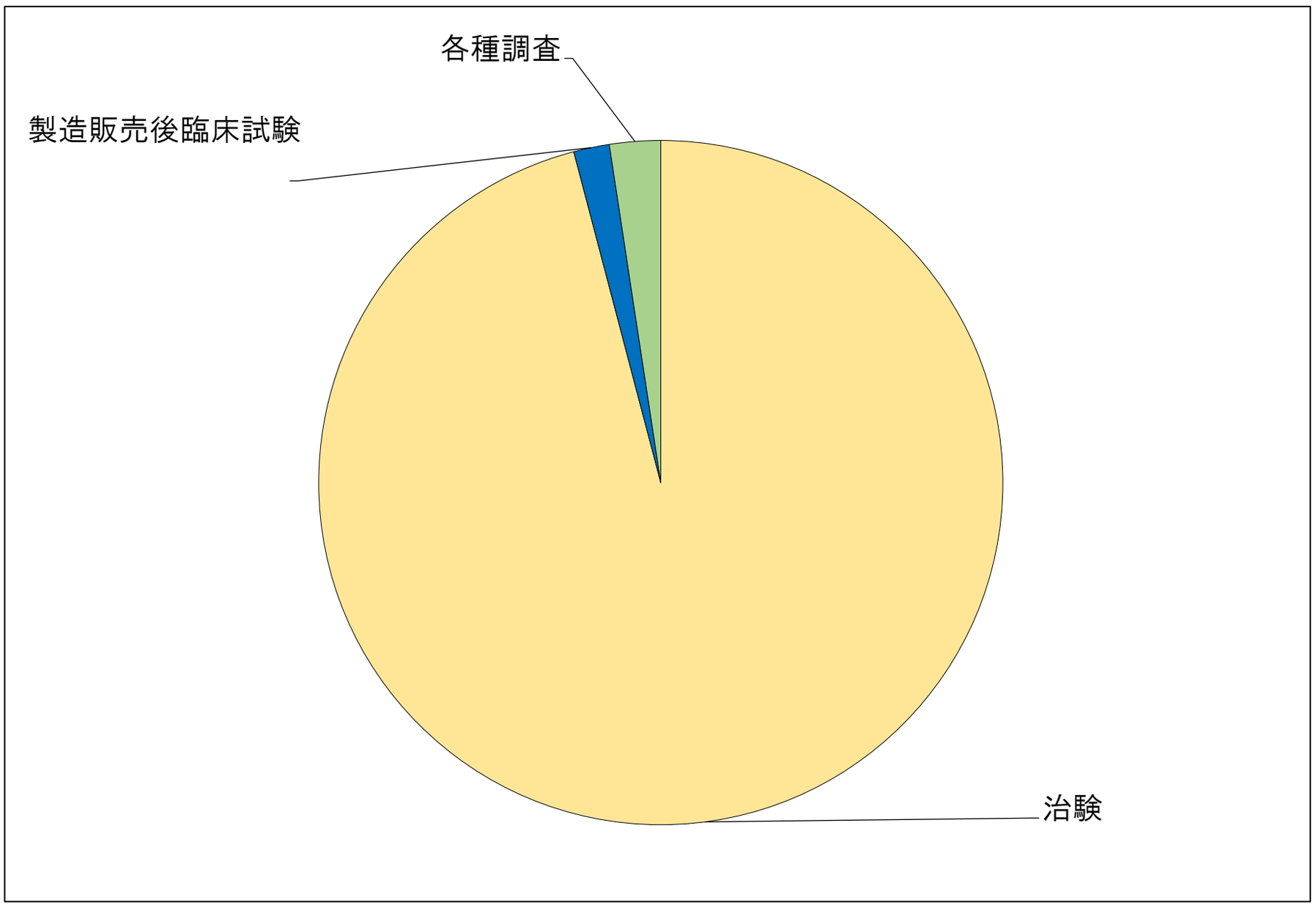
病院長　米野琢哉

水戸医療センターの基本方針として、「臨床研究を積極的に推進します」を掲げています。働き方改革で研究に対する時間を捻出するのが難しくなっておりますが、日常診療を黙々とこなし、臨床研究にも真摯に取り組んでいただきました。職員の皆様のご努力に敬意を払いたいと思います。

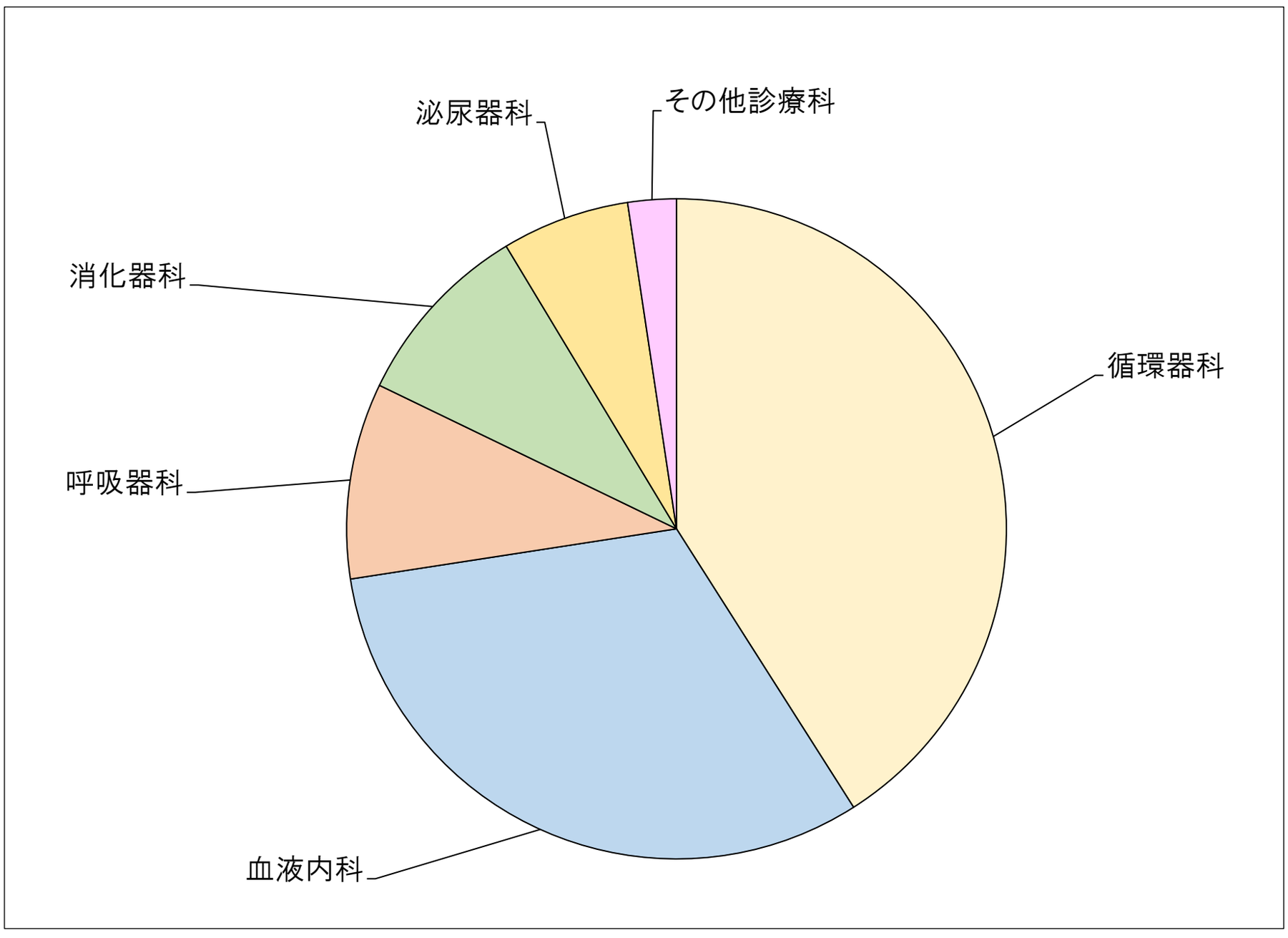
　臨床研究の推進は、診療の質向上のためにも必須です。様々な職種の方々が研究に取り組むことによって、診療の活性化、当院のブランディングにもつながると期待しております。病院としても、臨床研究部を中心に資金供給・研究に必要な文書作成のアドバイス等、研究実施のサポートを継続していきます。今後も是非臨床研究にチャレンジしてみてください。

2024年度　　受託研究実績金額　10,076万円

契約種類別グラフ

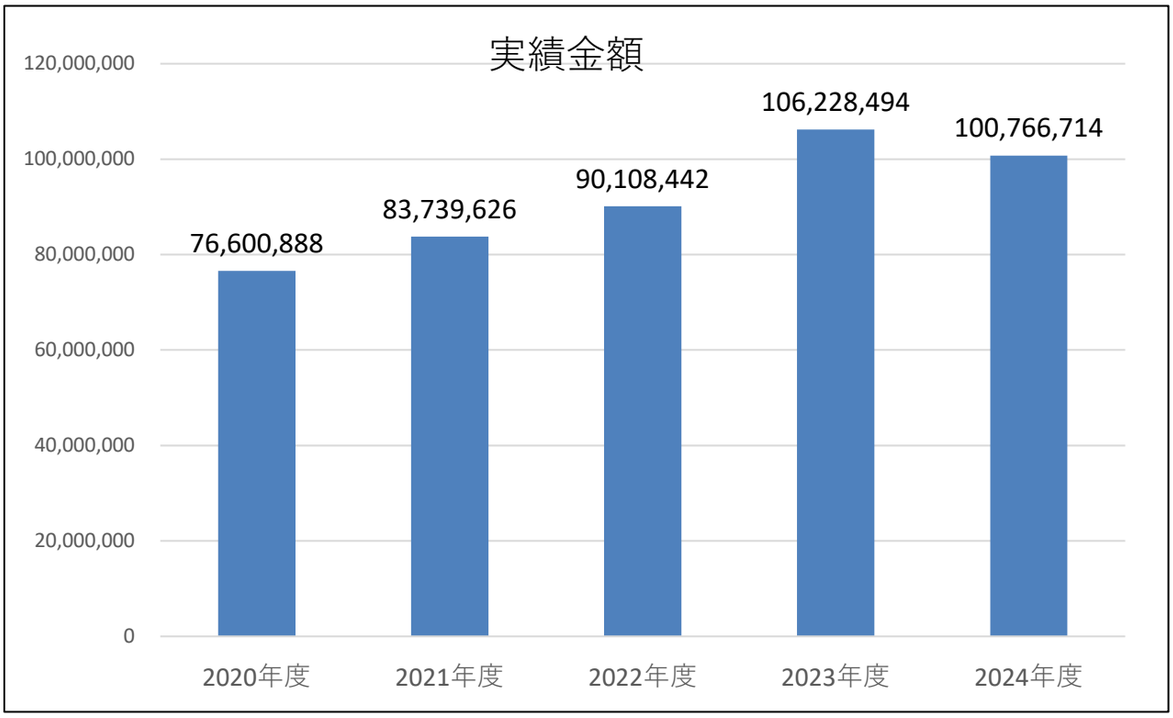


診療科別グラフ

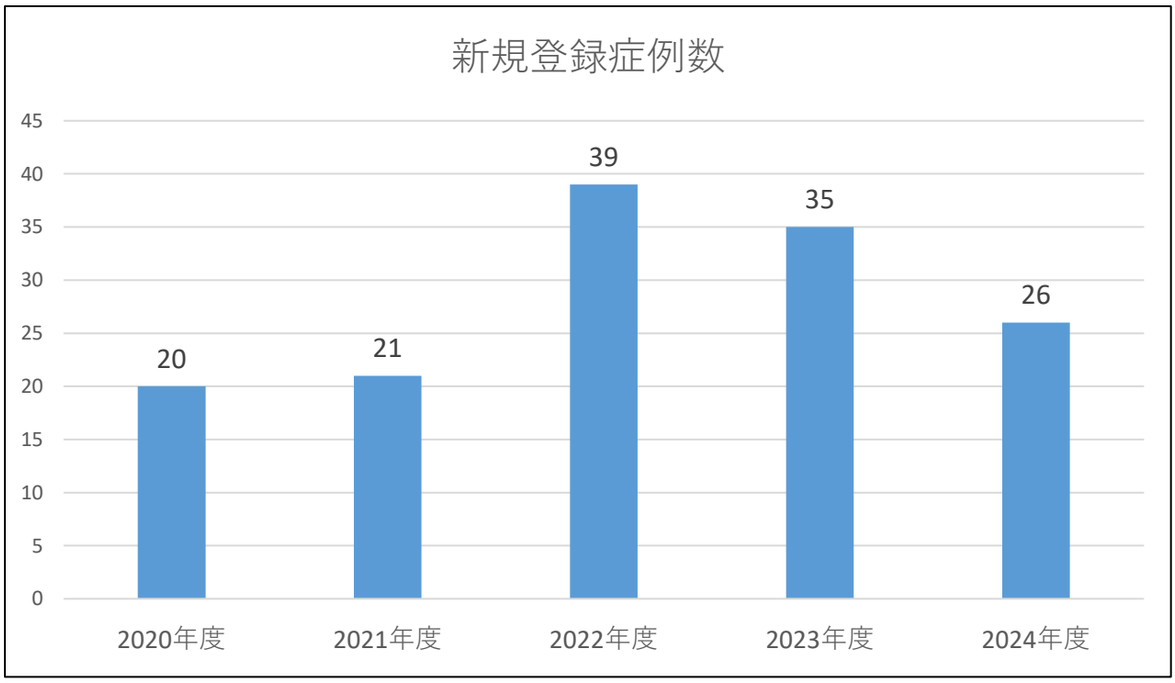


受託研究実績報告

①　受託研究実績金額（治験・製造販売後臨床試験・製造販売後調査）



②　治験・製造販売後臨床試験　新規登録症例数



NHOネットワーク共同研究　新規症例登録数

研究責任者	研究課題名(採択番号)	研究代表者(施設名	文書同意 有・無	当該施設 新規症例 登録数
福本　英樹	DOAC服用患者における抜歯の安全性の確立に関する研究:ガイドライン確立のための多施設共同前向き研究 （採択課題：R3-NHO(他研)-01）	吉川 博政 （九州医療センター）	有	5
田代 裕一	後期パーキンソン病の予後に関する多施設共同前向き研究 （採択課題：R5-NHO(神経)-01）	饗場　郁子 （東名古屋病院）	有	2

競争的研究費

項目	研究課題名	研究者名	研究事業者名 （依頼者名）	主任 分担	研究費 受領日	研究費 単位：円
科学研究費助成事業 （学術研究助成基金助成金）	包括的外傷長期予後データベースを用いたテーラーメイド型社会復帰支援システムの確立（22K10476）	堤　悠介	東海大学 （土谷飛鳥）	分担	2024/8/9	65,000
厚生労働科学研究費	HAMならびに類縁疾患の患者レジストリによる診療連携体制および相談機能の強化と診療ガイドラインの改訂（ 22FC0201 ）	湯沢賢治	聖マリアンナ医科大学（山野嘉久）	分担	2024/8/16	200,000
科学研究費助成事業 （学術研究助成基金助成金）	文献レジストリ構築とリアルワールドデータによる膠原病予後因子の網羅的負荷推計（22K10423）	堤　悠介	昭和大学 （辻本　康）	分担	2024/9/30	65,000
科学研究費助成事業 （学術研究助成基金助成金）	子宮移植に関する医事刑法上の諸問題の総合的検討（24K21398）	湯沢賢治	北海道大学 （城下裕二）	分担	2024/10/10	195,000
科学研究費助成事業 （学術研究助成基金助成金）	病態に強固な関連がある敗血症新規ｻﾌﾞｸﾗｽ分類の開発（23K27696）	堤　悠介	東北大学 （工藤大介）	分担	2024/11/25	260,000

英文論文

No.	タイトル	著者	ポイント
1	Magnetic resonance imaging predicts outcomes of conservative treatment in patients with lateral epicondylitis	小川 健	4.400
2	Pathophysiology of sex difference in refractoriness in lateral epicondylitis: Biomechanical study of wrist torque	小川 健	5.300
3	Is a Novel Fluoroscopic Intraoperative Reference System Superior to Conventional Management for Distal Radius Fracture Reduction? A Propensity-matched Comparative Study	小川 健	7.400
4	Development and evaluation of a rapid one-step high sensitivity real-time quantitative PCR system for minor BCR-ABL (e1a2) test in Philadelphia-positive acute lymphoblastic leukemia (Ph plus ALL)	吉田 近思	5.200
5	Female and preserved platelet count subgroups of myelodysplastic syndrome patients benefit from standard-dose azacitidine	吉田 近思	4.900
6	The Effect of Axial Traction MRI on the Articular Cartilage Visibility in Thumb Carpometacarpal Arthritis	小川 健	4.300
7	Efficacy and Safety of Lumbar Drainage before Endovascular Treatment for Ruptured Intracranial Aneurysms	加藤 徳之	3.500
8	Effect of Preoperative Oral Antibiotics and Mechanical Bowel Preparations on the Intestinal Flora of Patients Undergoing Laparoscopic Colorectal Cancer Surgery: A Single-Center Prospective Pilot Study	伊瀬谷 和輝	4.300
9	Prognostic Impact of Preoperative Assessment of Muscle Mass and Strength in Surgically Resected Lung Cancer	栗原 秀輔	11.400
10	Exploring the relationship between plasma substance P and glottal incompetence in the elderly	瀬成田 雅光	6.600
11	Proton Pump Inhibitors and Cyclin-Dependent Kinase 4/6 Inhibitors in Patients With Breast Cancer	小坂 真吉	7.200
12	Optimal Limb Position for the Stress Ultrasound Evaluation of Elbow Valgus Laxity in Baseball Players	小川 健	5.500
13	Excess mortality in COVID-19-affected solid organ transplant recipients across the pandemic	湯沢 賢治	11.200
14	Patient Age and EGFR-positive Non-small Cell Lung Cancer: A Multicenter Retrospective Study	遠藤 健夫	4.700

No.	タイトル	著者	ポイント
15	The impact of continuity correction methods in Cochrane reviews with single-zero trials with rare events: A meta-epidemiological study	堤　悠介	9.100
16	Association between a history of major osteoporotic fractures and subsequent hip fracture: a systematic review and meta-analysis	堤　悠介	2.400
17	Real Clinical Practice of Combined Atezolizumab Plus Chemotherapy in Patients With Small Cell Lung Cancer	遠藤　健夫	4.700
18	Investigation of age and smoking in NSCLC patients with uncommon EGFR mutations	遠藤　健夫	3.700
19	Serum C-reactive protein and procalcitonin levels in patients with pneumonia and anastomotic leakage in the postoperative period after esophagectomy	福富　俊明	4.300
20	Mechanisms of resistance and correlation between pre-treatment co-alterations and p-prognosis to osimertinib in chemo-naïve advanced non-small cell lung cancer	箭内　英俊	7.400
21	Asciminib: the next-generation bullet for first-line treatment of chronic myeloid leukemia	吉田　近思	13.800
22	High plasma BNP concentration associates with clinical outcome after mechanical thrombectomy: Post hoc analysis of SKIP	加藤　徳之	4.800
23	Delayed peak antibody titers after the second dose of SARS-CoV-2 vaccine in solid organ transplant recipients: Prospective cohort study	湯沢　賢治	6.500
24	Baseline genetic abnormalities and effectiveness of osimertinib treatment in patients with chemotherapy-naïve EGFR-mutated NSCLC based on performance status	箭内　英俊	5.800
25	Relationship between Physical Characteristics and Morphological Features of the Articular Radius Surface: A Retrospective Single-Center Study	小川　健	6.300
26	Association between helicopter medical services for pediatric trauma patients and mortality: Systematic review and meta-analysis	堤　悠介	2.100
27	Machine learning approaches to evaluate heterogeneous treatment effects in randomized controlled trials: a scoping review	堤　悠介	3.600
28	Patient and Healthcare Professional Satisfaction, Acceptability, and Preference Experiences With Mirikizumab Administration for Ulcerative Colitis: An International Survey	伊藤　有香	5.600
29	Clinical Outcomes of Endovascular Coil Embolization for Ruptured Middle Cerebral Artery Aneurysms	加藤　徳之	3.500

No.	タイトル	著者	ポイント
30	Prognostic Factors for Patients with Small-Cell Lung Cancer Treated with Chemoimmunotherapy: A Retrospective Multicenter Study	羽鳥　貴士	14.800
31	Outcomes of pregnancy in patients with chronic myeloid leukemia in the era of tyrosine kinase inhibitors	吉田　近思	8.100
32	Influence of Diabetes Mellitus on Neurological Recovery in Older Patients With Cervical Spinal Cord Injury Without Bone Injury: A Retrospective Multicenter Study	江藤　文彦	6.000
33	A Phase 1/2 study of teclistamab, a humanized BCMA x CD3 bispecific Ab in Japanese patients with relapsed/refractory MM	米野　琢哉	4.800
34	A practice-oriented genome-profiling study for acute myeloid leukemia using the novel HANDLE system: HM-screen-JAPAN02	吉田　近思	4.800
35	Radiolunate Fusion After Pyogenic Arthritis Caused by Pinning for Scapholunate Dissociation	小川　健	1.500
36	Outcomes of Combined Atezolizumab Plus Chemotherapy in Non-small Cell Lung Cancer Patients in Clinical Practice	沼田　岳士	1.500
37	Endoscopic assessment of the J pouch in ulcerative colitis: A narrative review	小野田　翼	1.750
38	Cardiac Multiple Micro-Scars: An Autopsy Study	小泉　智三	1.500
39	Anaphylaxis-Associated Cerebral Infarction: A Case Report.	渡邊　達也	1.500

論文

No.	論文名	著者	ポイント
1	疫学研究から得られるもの-疫学とリアルワールドデータ	堤　育代	1.500
2	腎移植臨床登録集計報告(2024) 2023年実施症例の集計報告と追跡調査結果	湯沢　賢治	1.000
3	抗体関連型拒絶反応を発現した腎移植患者に対するIVIGの有効性及び安全性	湯沢　賢治	1.000
4	レシピエント温阻血時間を用いた移植後早期腎機能予測モデル	湯沢　賢治	1.000
5	欧州における現況からわが国の心停止後ドナーからの移植を考える	湯沢　賢治	1.000
6	急性下腿コンパートメント症候群に対する多数小切開を用いた筋膜切開の有用性についての検討	森田 純一郎	1.500
7	手指化膿性腱鞘炎の治療成績―持続局所抗菌薬還流療法（iSAP）有効性の検討―	小川　健	1.000
8	上腕骨外側上顆炎の再々手術で除神経が奏功した一例	小川　健	1.500
9	Gadolinium 造影dynamic MRI によるキーンベック病術後の月状骨髄内血流動態の評価	小川　健	1.500
10	【総説】変形性肘関節症の診断と関節形成術	小川　健	1.500
11	診断学と臨床検査	江藤　文彦	1.000
12	痙性尖足に対する鏡視下腓腹筋膜切離術の経験	大山　和生	1.500
13	両側上部尿路癌摘出後に透析を導入した3例	高橋　佳子	1.500

国際学会

学会名	演題名	演者名	発表年月日
26th Annual John Goldman Conference on CHRONIC MYELOID LEUKEMIA: Biology and Therapy	Asciminib for Chronic Myeloid Leukemia in Real-World Clinical Practice: Interim Analysis of a Single-Center Observational Study	吉田　近思	2024/9/28
第86回日本血液学会学術集会	Asciminib for chronic myeloid leukemia in clinical practice: a retrospective observational study	小田　卓弥	2024/10/13
Korean Society of Radiology 2024,Seoul,Korea	Ecessity of remote 3D imaging creation in Emergency Radiology –An optimal way to enhance patient satisfaction for quality of emergency cace delivery by improving availability.	田中　善啓	2024/10/2

国内学会

学会名	演題名	演者名	発表年月日
第43回茨城造血器疾患研究会	移植10年後に頸部局所再発した急性リンパ性白血病	原　浩ノ輔	2024/5/25
第701回日本内科学会関東地方会	移植10年後に頸部局所再発した急性リンパ芽球性白血病	原　浩ノ輔	2024/12/14
第64回日本呼吸器学会学術講演会	当院におけるCiliated muconodular papillaly tumor(CMPT)の2症例	羽鳥　貴士	2024/4/8
第64回日本呼吸器学会学術講演会	急激な転機をたどり剖検により診断しえた自験例PTTMの2症例と、本症の臨床的特徴に関する考察	山岸　哲也	2024/4/8
第259回日本呼吸器学会関東地方会	肺底動脈大動脈起始症の1例	藤田　弘輝	2024/5/11
第260回日本呼吸器学会関東地方会	PembrolizumabによるirAEギラン・バレー症候群(GBS)の肺癌術後再発の一例	菊池　柊汰	2024/7/6
第262回日本呼吸器学会関東地方会	鉄剤誤嚥による気管支上皮障害をきたした1例	藤田　弘輝	2024/11/30
第263回日本呼吸器学会関東地方会	腸重積をきたした肺腺癌小腸転移の一例	山崎　健斗	2025/2/8
第702回日本内科学会関東地方会	検診異常で発見された胸部脾症の1例	渡邊　崚	2025/2/8
第226回茨城県内科学会	肺底動脈大動脈起始症の1例	藤田　弘輝	2024/6/22
Lung Cancer Expert Symposium	進展型小細胞肺癌の治療戦略～細胞障害性抗癌剤とPD-L1阻害剤併用療法の予後因子を明らかにするための多施設共同研究の結果を踏まえて～	羽鳥　貴士	2024/4/9
第2回肺がん診療セミナー	Ⅳ期NSCLCに対するICIの治療戦略を考える	羽鳥　貴士	2024/6/5
AZ Immuno-Oncology Online Seminar	進展型小細胞肺癌治療の進展～CASPIAN試験と実臨床経験を踏まえて～	沼田　岳士	2024/6/17
SANOFI WEB 喘息治療 UPDATE	当院におけるデュピルマブの使用状況	太田　恭子	2024/7/9
第258回水戸チェストカンファレンス	免疫チェックポイント阻害薬(ICI)に関連した呼吸器病変あれこれ	遠藤　健夫	2024/7/18
Lung Cancer Seminar	ILD合併肺癌における化学療法	沼田　岳士	2024/7/23
Ibaraki NHO Symposium Vol.2	進行・再発非小細胞肺癌における患者背景毎の治療方針を考える	山岸　哲也	2024/7/30
AstraZeneca Immuno-Oncology Web Symposium	ES-SCLC 1st line treatment -イミフィンジの新たなリアルワールドデータ-	沼田　岳士	2024/8/21
Respiratory Conference in 関東	喘息診療の現状と課題	遠藤　健夫	2024/10/29
茨城肺癌治療カンファレンス	Ⅳ期NSCLC治療に抗CTLA-4抗体は必要か？ 抗CTLA-4抗体が適した症例とは？	山岸　哲也	2024/11/19
第9回桜の郷チェストカンファレンス	呼吸器領域における興味深い症例の検討	山岸　哲也	2024/11/27

学会名	演題名	演者名	発表年月日
第9回桜の郷チェストカンファレンス	ガイドライン改定を踏まえた市中肺炎の診断と治療	遠藤　健夫	2024/11/27
第5回 Lung Cancer Expert Symposium in Mito	高齢非小細胞肺癌における課題、個別化治療を考える	沼田　岳士	2024/11/28
第60回日本胆道学会学術集会	総胆管結石を契機に発見された十二指腸乳頭部mixed neuroendocrine-non-neuroendocrine neoplasm (MiNEN)の一切除例	岸本　希実	2024/10/10
第703回日本内科学会関東地方会	内視鏡生検および悪性リンパ腫の治療期間を経て消失し、8年後に再確認された早期胃癌の一切除例	岸本　希実	2025/3/8
第695回　内科学会関東地方会	視力正常で自覚症状はないものの視野検査異常を呈した巨細胞性動脈炎の1例	米野 友一朗	2024/5/11
第16回日本Acute care surgey学会	Rb直腸癌穿孔による膿瘍形成に対して緊急で腹会陰式直腸切断術を施行し、良好な経過をたどった1例	成田　保和	2024/9/27
第61回日本腹部救急医学会	重症膵炎による腹部コンパートメント症候群(ACS)に対してOpen Abdominal　Management(OAM)を施行した1例	成田　保和	2025/3/21
第78回日本食道学会学術集会	有鉤義歯誤飲による胸部上部食道の食道異物に対して胸腔鏡下摘出術を施行した1例	福富　俊明	2024/7/5
第77回日本胸部外科学会学術集会	右総腸骨動脈瘤に対するステントグラフト内挿術後の食道胃接合部癌に対して食道切除胃全摘有茎空腸再建術を施行した1例	福富　俊明	2024/11/3
第41回日本呼吸器外科学会学術集会	術前CXIと非小細胞肺がん切除例の予後との関係	栗原　秀輔	2024/6/1
第41回日本呼吸器外科学会学術集会	肺切除検体におけるPD-L1発現と臨床像の比較検討	中村　亮太	2024/6/1
第77回胸部外科学会定期学術集会	サルコペニアを有する肺癌手術症例における性差についての検討	中村　亮太	2024/11/3
第254回茨城外科学会	鼠径ヘルニア偽還納の1例	大曽根 龍汰	2024/10/20
第79回日本消化器外科学会総会	当院における閉鎖孔ヘルニアの治療成績	米山　智	2024/7/17
第37回日本内視鏡外科学会総会	腹腔鏡下鼠径ヘルニア修復術にける助手の立ち位置について	米山　智	2024/12/7
第32回日本乳癌学会学術総会	当院の乳腺外科における医師業務のタスクシェアと、診療連携の現状・課題	森　千子	2024/7/11
第83回日本脳神経外科学術集会	Gentle combination Solitaire X &REACT71	佐藤　允之	2024/10/16
第83回日本脳神経外科学術集会	急性期脳梗塞の閉塞遠血管のMRI造影3D画像による評価	佐藤　允之	2024/10/18
第111回茨城県脳神経外科集談会	橈骨動脈穿刺における合併症の要因と予防対策の検討	丸山　沙彩	2024/10/26
STROKE2025	IC-PC 動脈瘤治療を考える -課題と当院の戦略-	佐藤　允之	2025/3/6
STROKE2025	急性期血栓回収療法におけるリアルタイム3Dナビゲーションによる治療支援	佐藤　允之	2025/3/8
第50回日本骨折治療学会	急性下腿コンパートメント症候群に対する多数小切開を用いた筋膜切開の有用性についての検討	森田 純一郎	2024/6/28

学会名	演題名	演者名	発表年月日
橈骨遠位端骨折を語る会2024	掌側locking plateと背側distraction plateを併用した小経験	森田 純一郎	2025/2/15
第133回茨城県整形外科集談会	胸腰椎損傷（AO分類typeB3）における 椎体前方開大の術後変化	江藤　文彦	2024/5/11
第33回日本脊椎インストゥルメンテーション学会	胸腰椎過伸展損傷に対する腹臥位手術後の矯正損失	江藤　文彦	2024/9/20
第67回日本手外科学会学術集会．2024年4月，奈良	MRIによるキーンベック病術後の月状骨髄内血流動態の評価	小川　健	2024/4/25
日本スポーツ整形外科学会2024	JSOA-USA Travelling Fellow 2023 「手肘肩外科医の立場から」	小川　健	2024/9/11
第39回東日本手外科研究会	キーンベック病の月状骨血流動態は病期により異なるのか？―ガドリニウム造影ダイナミックMRIでの検討―	小川　健	2025/2/22
第37回日本肘関節学会学術集会	小児上腕骨顆上骨折に対し成人手関節用創外固定器を用いた治療成績	小川　健	2025/3/14
第37回日本臨床整形外科学会学術集会	舟状月状骨間解離に対する経皮的pinning抜去後に月状骨骨髄炎から化膿性手関節炎に至った一例	小川　健	2024/7/14
第134回茨城県整形外科集談会	載距突起スクリューの至適刺入位置とプレートスクリューホールとの位置関係	大山　和生	2024/10/6
第49回日本足の外科学会学術集会	載距突起スクリューの至適刺入位置とプレートスクリューホールとの位置関係	大山　和生	2024/11/8
第131回日本泌尿器科学会茨城地方会	転移性尿管腫瘍に対して尿管部分切除を施行し腎機能を温存した1例	高橋　祥太	2025/2/1
第130回日本泌尿器科学会茨城地方会	右側腹部痛を契機に発見された右副腎骨髄脂肪腫の1例	高橋　祥太	2024/10/20
第129回日本泌尿器科学会茨城地方会	precaval right renal arteryの1例	高橋　祥太	2024/6/15
第28回秋田腎不全研究会	当院における透析腎癌6例の検討	高橋　佳子	2024/12/8
第37回日本老年泌尿器科学会	75歳以上高齢者に対するカボザンチニブの治療経験	飯沼　昌宏	2024/5/17
第89回日本泌尿器科学会東部総会	ロボット支援腹腔鏡下膀胱全摘術術後に広範囲な播種をきたしたinflammatory myofibroblastic tumorの1例	飯沼　昌宏	2024/10/4
第38回日本泌尿器内視鏡・ロボティクス学会総会	当院におけるロボット支援腹腔鏡下腎部分切除術(RAPN)の治療成績についての検討	飯沼　昌宏	2024/11/28
第40回日本脳神経血管内治療学会	橈骨動脈穿刺による脳血管造影を受ける患者の苦痛軽減に向けた看護支援	和田 由樹子	2024/11/21
第27回茨城県総合リハビリテーションケア学会学術集会	リハビリテーション科スタッフのNST参加がもたらした効果	永山　愛子	2025/2/16
第14回日本臨床腫瘍薬学会学術大会2025	トリプルネガティブ早期乳がんに対する ペムブロリズマブ併用レジメンの有害事象状況調査	小島　卓也	2025/3/15
第38回日本外傷学会総会・学術集会	外傷に対する3DCTを用いた骨折の初期評価法	杉原　理菜	2024/4/25
第38回日本外傷学会総会・学術集会	肋間動脈損傷の緊急IVRにおける仮想透視画像の有用性	田中　善啓	2024/4/25

学会名	演題名	演者名	発表年月日
第49回日本超音波検査学会学術集会	バルサルバ洞動脈瘤に限局的解離が生じた一症例	渡邉　隼	2024/7/21
第3回全国国立病院機構臨床工学技士関東信越グループ学術大会	業務実績とデータから今後の臨床工学技士業務内容の検討	山内　和樹	2025/3/16
第65回日本神経学会学術大会	フィッシャー症候群とギラン・バレー症候群 overlap症例の急性期からの運動療法の経験	井口　朋重	2024/6/1
第79回国立病院総合医学会	FS/GBS overlap症例の神経症状の急性期からの運動療法の経験	井口　朋重	2024/10/19
第80回国立病院総合医学会	排尿ケアチーム療法士としてのリハビリテーション科への取り組み～排尿ケアチーム介入で排尿自立まで改善した1症例～	淺野 花耶子	2024/10/18
第24回CRCと臨床試験のあり方を考える会議2024 in Sapporo	CRCと薬剤部の実施体制強化 ～薬剤師向け活用ツールを作ってみました！～	吉野 有美子	2024/9/15
第78回国立病院総合医学会	院内危機管理体制及び災害時情報共有体制構築の検証	野崎　基亜	2024/10/19
第78回国立病院総合医学会	茨城県におけるIMATの運用体制構築について	石上　耕司	2024/10/19
第78回国立病院総合医学会	サステナブルな医療機関であるために	米野　琢哉	2024/10/19
第78回国立病院総合医学会	看護師が実施するNPPVマスク装着時のベルト固定や留意点の現状を明らかにする	阿部　沙雪	2024/10/18
第78回国立病院総合医学会	手術本位の側臥位における陰圧式固定具（マジックベッド）使用時の体圧分散について～アクションマットとソフトナースとの比較～	柏　綾友美	2024/10/19
第78回国立病院総合医学会	骨シンチグラフトのコントラストに関する検討	天野　祥吾	2024/10/19
第78回国立病院総合医学会	初回化学療法実施後にIRを発現した患者の特徴～HP+DTX療法を実施した乳癌患者～	中山　舞	2024/10/18

令和6年度院内臨床研究課題

研究代表者	課題名	配分額
呼吸器科 沼田　岳士	80歳以上の進行肺癌患者における化学療法治療前骨格筋量と治療への忍容性や予後との関連性についての検討	460,000
救急科 堤　悠介	機械学習を用いた外相患者に対する全身CTの効果に関する異質性の検証	640,000
外科 中村　亮太	肺癌手術検体におけるPD-L1発現と臨床像の比較検討	550,000
外科 米山　智	ロボット支援下大腸・胃手術の導入・手技習得・教育に対する臓器モデルを用いたシミュレーションの有用性	550,000
外科 田部田　厚史	上部消化管穿孔で手術を施行された患者における術中の腹腔内洗浄量と術後の腹腔内腫瘍の発生の有無の関係を探索する過去起点コホート研究	450,000
脳神経外科 佐藤　允之	カテーテル治療後止血部の観察可能な固定テープの開発	660,000
脳神経外科 丸山　沙彩	橈骨動脈穿刺における合併症の要因と予防対策の検討	750,000
整形外科 小川　健	整形外科医のメンタルヘルスについて ーシステマティックレビューによる解析ー	300,000
整形外科 江藤　文彦	脊髄手術における外視鏡の導入 ー術野の共有がもたらす利点を検証するー	460,000
整形外科 大山　和生	載距突起スクリューの至適挿入位置とプレートスクリューホールとの位置関係に関する臨床研究	700,000
整形外科 平林　匠	大腿骨転子下骨折における骨形態、骨密度に関する後ろ向き観察研究	640,000
整形外科 薬師寺　亮	手根管症候群と心アミロイドーシスの関連についての検討	460,000
看護部 木村　梨奈	当院における終末期化学療法の現状ー医療用麻薬使用との比較ー	360,000
放射線科 田中　善啓	STAT画像所見報告における遠隔画像処理システムを用いた診療放射線技師の教育体制構築	600,000
放射線科 杉原　理菜	心筋シンチグラフィ検査における呼吸アーチファクトの低減方法	200,000
栄養管理室 鴨志田　純子	大学生のプレコンセプションケアへの意識調査と体組成・栄養摂取状況との関連について	210,000
事務部 野崎　基亜	災害時におけるSNSを用いた院内情報共有ポータルサイトの構築、利用範囲・有用性等の検証及び利用コストの最小化	360,000

水戸医療センター

令和6年度　代表的論文

Journal of Orthopaedic Science 29 (2024) 795–801

Contents lists available at ScienceDirect

Journal of Orthopaedic Science

journal homepage: http://www.elsevier.com/locate/jos

Original Article

Magnetic resonance imaging predicts outcomes of conservative treatment in patients with lateral epicondylitis☆

Kazuhiro Ikeda [a, b], Takeshi Ogawa [c], Akira Ikumi [b], Yuichi Yoshii [d, *], Sho Kohyama [a], Reimi Ikeda [e], Masashi Yamazaki [b]

[a] Department of Orthopedic Surgery, Kikkoman General Hospital, Noda-city, Chiba-Pref., Japan
[b] Department of Orthopedic Surgery, Faculty of Medicine, University of Tsukuba, Tsukuba-city, Ibaraki-Pref., Japan
[c] Department of Orthopedic Surgery, Mito Medical Center, Ibaraki-town, Japan
[d] Department of Orthopedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ami-town, Ibaraki-Pref., Japan
[e] Department of Orthopedic Surgery, Moriya Daiichi General Hospital, Moriya-city, Ibaraki-Pref., Japan

ARTICLE INFO

Article history:
Received 27 December 2022
Received in revised form
11 March 2023
Accepted 17 March 2023
Available online 4 April 2023

Keywords:
Lateral epicondylitis
Tennis elbow
Magnetic resonance imaging
Conservative treatment
Surgery

ABSTRACT

Background: The clinical validity of positive magnetic resonance imaging findings in lateral epicondylitis is controversial. We hypothesized that magnetic resonance imaging could predict the outcome of conservative treatment. This study determined the relationship between magnetic resonance imaging-defined disease severity and treatment outcomes in patients with lateral epicondylitis.
Methods: This retrospective single-cohort study included 43 conservatively managed and 50 surgically treated patients with lateral epicondylitis. The magnetic resonance imaging scores and clinical outcomes were examined six months post-treatment, and the former was compared between patients with good and poor treatment outcomes. We developed operating characteristic curves of magnetic resonance imaging scores for treatment outcomes, and divided patients into magnetic resonance imaging-mild and severe groups according to the obtained cut-off value of the scores. We compared the outcomes of conservative treatment with that of surgery for each magnetic resonance imaging severity.
Results: Twenty-nine (67.4%) conservatively treated patients had good outcomes, while 14 (32.6%) had poor outcomes. The magnetic resonance imaging score was higher in patients with poor outcomes; the cut-off value was 6. Forty-three (86.0%) surgically treated patients had good outcomes, while 7 (14.0%) had poor outcomes. There was no significant difference in magnetic resonance imaging scores between patients with good and poor surgical outcomes. In the magnetic resonance imaging-mild group (score $\leq$ 5), the outcome showed no significant difference between the conservative and surgical treatment groups. In the magnetic resonance imaging-severe group (score$\geq$6), the outcome of conservative treatment was significantly worse than that of surgical treatment.
Conclusions: The magnetic resonance imaging score was associated with conservative treatment outcomes. A treatment strategy that includes surgery should be considered for patients with severe magnetic resonance imaging findings; this is not recommended for those with mild magnetic resonance imaging findings. Magnetic resonance imaging is helpful in determining the best treatment strategies for patients with lateral epicondylitis.
Level of evidence: III, Retrospective cohort study.

1. Introduction

Lateral epicondylitis (LE) is the tendinopathy of the common extensor tendon (CET) of the forearm [1], and the elbow's lateral collateral ligament (LCL) is susceptible to injury [2–4]. Repetitive eccentric contractions of the CET induce micro rupture and degeneration at the enthesis [5], and a pathologically disrupted healing response potentiates LE [6–8]. Diagnosis of LE is based on

Abbreviations: LE, lateral epicondylitis; CET, common extensor tendon; LCL, lateral collateral ligament.

☆ Institutional review board of Mito Clinical Education and Training Center, University of Tsukuba Hospital, Mito Kyodo General Hospital approved this study (approval no: 16-25, date: September 7, 2016).

* Corresponding author. Tokyo Medical University Ibaraki Medical Center, Department of Orthopaedic Surgery, 3-20-1 Chuo, Ami, Inashiki, Ibaraki 300-0395, Japan.
E-mail address: yyoshii@tokyo-med.ac.jp (Y. Yoshii).

https://doi.org/10.1016/j.jos.2023.03.014
0949-2658/

physical examinations, such as the observation of tenderness on the lateral epicondyle, Thomsen test, and Maudsley's test [9–12]. Imaging examinations are performed supplementally for patients with refractory LE who have not responded to conservative treatment [3,9,13]. Magnetic resonance imaging (MRI) detects degeneration and rupture of the CET/LCL complex as high-signal changes. Since MRI has excellent sensitivity for LE (93%), negative findings are useful in ruling out LE [14].

However, the clinical validity of positive MRI findings for LE remains controversial. Since MRI captures asymptomatic structural abnormalities and degeneration, its specificity is low [14–16]. The relationship between MRI findings and the clinical severity of LE differs between studies [2,3,17–19]. Although several studies have reported various positive MRI findings and quantitative evaluation methods for LE [2,3,11,16,20,21], none have shown the clinical validity of MRI. Notably, evaluation of positive MRI findings cannot be used to formulate a treatment strategy; instead, surgical indications are based on the symptoms' clinical severity and treatment progress [10], regardless of the MRI findings. Currently, there is no evidence to support changing LE treatment strategy based on positive MRI findings.

As progressive LE results in an irreversible histological degeneration of the tendon attachment, some studies suggest surgery in such cases [6–8]. Although histological severity can only be determined by postoperative pathology, MRI captures pathological changes, allowing us to predict the severity. Therefore, we hypothesized that MRI could predict the outcomes of patients undergoing conservative treatment for LE.

This study aimed to determine the association of MRI severity with treatment outcomes in patients with LE. A previous study [14] presented an MRI scoring system (LE-MRI scoring), which accurately reflects imaging severity, and demonstrated its high association with clinical severity. The present study investigated the association between LE-MRI scoring and treatment outcomes.

2. Materials and methods

2.1. Study design

This was a retrospective single-cohort study. The study protocol conformed to the principles of the 1964 Declaration of Helsinki. Written informed consent was obtained from all patients enrolled in the study. Our institutional review board approved this study.

2.2. Study participants

The inclusion criterion was patients with LE who underwent high-resolution MRI and were treated at our hospital. The exclusion criteria were a history of elbow trauma, elbow osteoarthritis (Kellgren-Lawrence classification 2 or higher) [22], osteochondritis dissecans of the humeral capitellum, or rheumatoid arthritis. To align the criteria for treatment outcome, we excluded patients who discontinued treatment in under six months before the symptoms resolved to the Nirschl Phase Rating Scale 3 or lower [23].

Fig. 1 shows the flow diagram of the study. Our medical database identified 366 patients diagnosed with LE between January 2013 and December 2020. We excluded 258 patients without MRI, seven with inappropriate MRI, and one with rheumatoid arthritis. Of the remaining 100 patients, 38 underwent surgery before six months of conservative treatment at our institution; treatment was terminated in 15 patients without improvement of symptoms within six months of post-conservative treatment. Additionally, we excluded four cases in which the description of specific symptoms was insufficient to determine treatment outcomes. Finally, this study included 43 elbows in 40 patients who underwent conservative treatment (18 males and 22 females; median (25–75 percentile) age, 50 (48–56)).

This study included 50 elbows of 43 patients who underwent surgery (23 males and 20 females; median (25–75 percentile) age, 49 (45–54)). Of these 50 cases, 38 were referred to our hospital as they did not respond to conservative treatment at a different center and therefore underwent surgery. The remaining 12 cases did not respond to conservative treatment for six months in our center and had to undergo surgery. Finally, the present study included 81 affected elbows of 72 patients with LE (38 males and 34 females; median (25–75 percentile) age, 49 (45–56)).

2.3. Treatment strategy for patients with LE

Two upper extremity orthopedic surgeons with 23 and 28 years of experience made all clinical decisions for patients with LE. LE diagnosis was based on physical examination findings, positive Thomsen or Maudsley's test, and tenderness at the lateral epicondyle [9–11]. MRI was indicated for persistent pain of the Nirschl Phase Rating Scale 3 or higher [23]. The median (25–75 percentile) duration between LE onset and MRI was 7.8 (4.5–14.5) months.

The median (25–75 percentile) duration between LE onset and the starting date of conservative treatment was 3.3 (1.8–6.0) months. Twenty-seven (62.8%) patients were initially diagnosed with LE at our hospital, and the other 16 (37.2%) were referred by previous physicians. Orthotic therapy with the tennis elbow brace was prescribed for all patients (n = 43, 100%). The patients wore this brace for the entirety of the symptomatic period; patients who frequently visited the hospital received occupational therapy (n = 20, 46.5%). Occupational therapy consisted mainly of stretching and massage of the forearm extensor muscles and supinator. Moreover, training to improve coordination from the shoulder to the upper limb girdle and instruction in daily activities were included. Steroid injections (8 mg triamcinolone) were administered to patients who did not respond to occupational therapy or orthotics once extra-articularly and once intra-articularly (n = 23, 67.4%); we prescribed oral analgesics (n = 27, 62.8%) and topical medications (n = 29, 65.4%) for patients who experienced severe pains. Although the details of treatment from the previous physicians for the 16 referred patients were unavailable, 14 (87.5%) patients received local steroid injections, with a median (25–75 percentile) number of injections of 1 (1–2).

Indications for surgery were resistance to conservative treatment for at least six months and severely limited activity levels due to pain (corresponding to the Nirschl Phase Rating Scale 5 or higher) [23]. We performed intra-articular steroid injections preoperatively to differentiate the intra-articular lesion involvement for patients with severe tenderness of the humeroradial joint and positive fringe impingement test. We indicated arthroscopic surgery for patients diagnosed with intra-articular lesion involvement.

We primarily performed the modified Nirschl's procedure from April 2013 to March 2019 in 28 (56%) patients, resected the degenerated CET/LCL complex under direct vision, and sutured the remaining healthy tendon to the lateral epicondyle using anchors [24]. From April 2019 to March 2020, we primarily performed arthroscopic LE release in 16 (32%) patients and resected the anterolateral articular capsule lining the CET/LCL complex and synovial folds [25]. In six (12%) patients with a massive CET/LCL complex tear, as observed on MRI, we anchored the remaining healthy tendon to the lateral epicondyle without LCL reconstruction [4]. The median (25–75 percentile) duration between LE onset to surgery date was 12.8 (6.8–19.7) months.

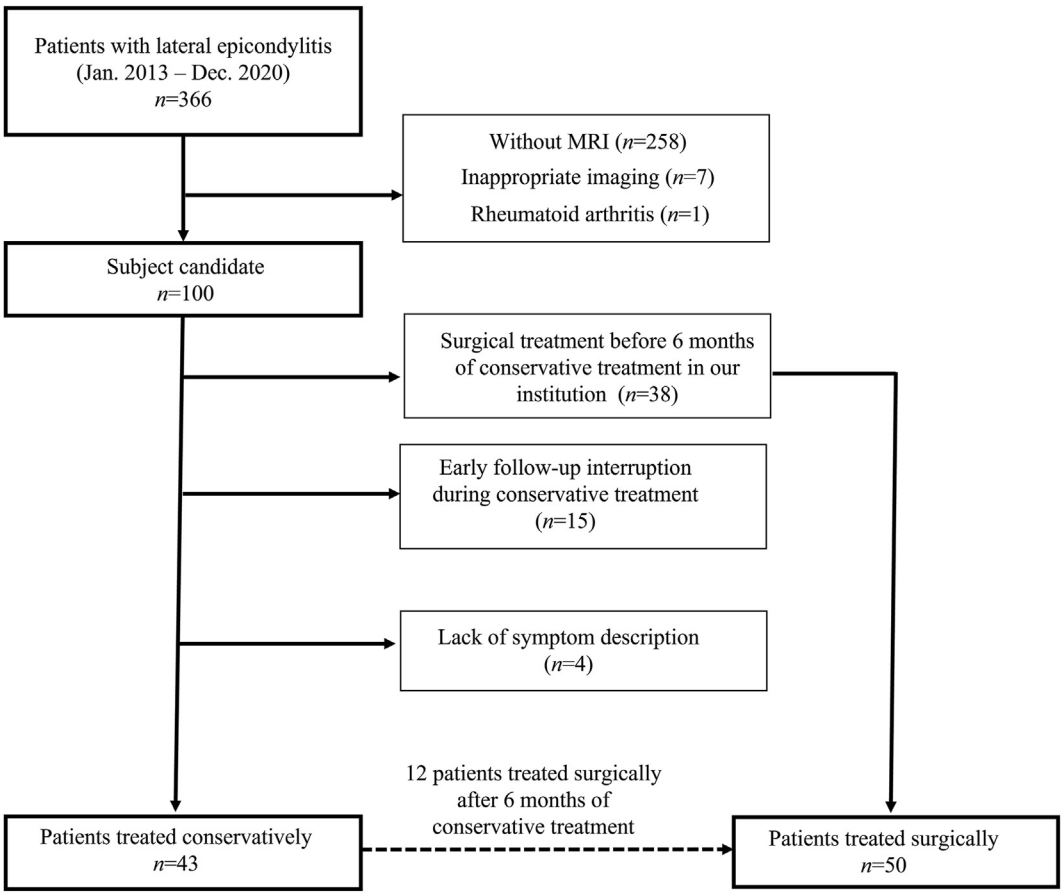


Fig. 1. Flow diagram of enrollment of the study participants. MRI, magnetic resonance imaging.

2.4. Determination of treatment efficacy

We defined the date of LE diagnosis as day zero of conservative treatment at our hospital. Similarly, the surgery date was defined as day zero of treatment for surgical patients. Treatment outcomes were determined at six months post-treatment by investigation of the medical record entries from five to seven months post-treatment. In patients whose LE resolved early, we referred to the date of treatment termination for outcome. Subsequently, the median value of the treatment outcome determination period and its 25–75 percentile was 5.6 (4.2–6.7) months and 6.0 (5.6–6.7) in conservatively and surgically treated patients, respectively. Treatment outcomes were defined as "good" or "poor." A clear statement of no pain or no activity limitation (corresponding to the Nirschl Phase Rating Scale 3 or lower) [23] was considered to indicate "good" treatment outcomes. A clear statement of activity limitation due to pain (corresponding to the Nirschl Phase Rating Scale 4 or higher) [23] was considered to indicate "poor" treatment outcomes. Additionally, the quick disabilities of the arm, shoulder, and hand scores were used to determine treatment efficacy in 27 patients. Following previous reports, a disability/symptom score of <15 was considered good, and a score ≥15 was considered poor [26,27]. Two orthopedic surgeons, Examiner 1 (with 10 years experience) and Examiner 2 (with 9 years experience), independently determined the treatment outcome. The examiners repeated the analysis, with the second taking place one month after the initial analysis. We adopted the treatment outcome determined by Examiner 1 for further analysis.

2.5. MRI protocol and scoring

We used a clinical 3.0 T imager (Magnetom Symphony, SIEMENS, Munchen, Germany) with a small-diameter surface coil (Loop Flex Coil, SIEMENS) above the lateral epicondyle of the humerus. Patients' elbows were placed at the center of the MRI scanner, extended beside the trunk with the forearm supinated. We obtained a coronal section of the lateral elbow under the following sequences: T2*-weighted images (T2*WI) using gradient echo to evaluate synovial folds, proton density-weighted images (PDWI) using high-speed spin echo to recognize the morphology of the CET/LCL complex attachment, and T2 fat-saturated weighted images (T2FSWI) to evaluate LE severity. MRI parameters are shown in the supplemental data (Table S1).

We evaluated the LE-MRI scoring for quantitative severity evaluation with MRI findings [14]. The region of interest for LE-MRI scoring was defined as the area between the level of the articular surface of the radial head and CET- and LCL attachment. To reflect the imaging severity of the three-dimensional structure of the CET/LCL complex [28], LE-MRI scoring was used to evaluate CET and LCL individually, and the scores were subsequently added. The LE-MRI scoring methodology is described in Table 1 and Fig. 2. LE-MRI scoring reflects both the intensity and extent of signal changes in the CET/LCL complex, and is highly associated with clinical severity [14].

Two examiners, an orthopedic surgeon with a 10-year experience (Examiner 1) and a hand surgeon with a 23-year experience (Examiner 3), independently assessed the images. The examiners

Table 1
LE-MRI scoring.[a].

Scoring	Condition	Definition
0	Normal	Dark, linear low-signal structure without changes in signal intensity
1	Mild degeneration	Thickening or mild signal change below the signal intensity of the muscle.
2	Severe, localized degeneration	High signal change above the signal intensity of the muscle, localized below 50% of the evaluation range.[b]
3 (a)	Severe, extensive degeneration	High signal change above the signal intensity of the muscle, beyond 50% of the evaluation range.[b]
3 (b)	Partial tear	High signal change equivalent to joint fluid, within 75% of the tendon or ligament width.
4	Extensive tear	High signal change equivalent to joint fluid, more than 75% of the tendon or ligament's width.

LE, lateral epicondylitis; MRI, magnetic resonance imaging; CET, common extensor tendon; LCL, lateral collateral ligament.
[a] The CET and LCL were individually evaluated on a scale of 0–4, and the individual scores were subsequently added to obtain a total score of 0–8.
[b] The region of interest for MRI scoring was from the level of the articular surface of the radial head to the CET/LCL attachment site of the lateral epicondyle.

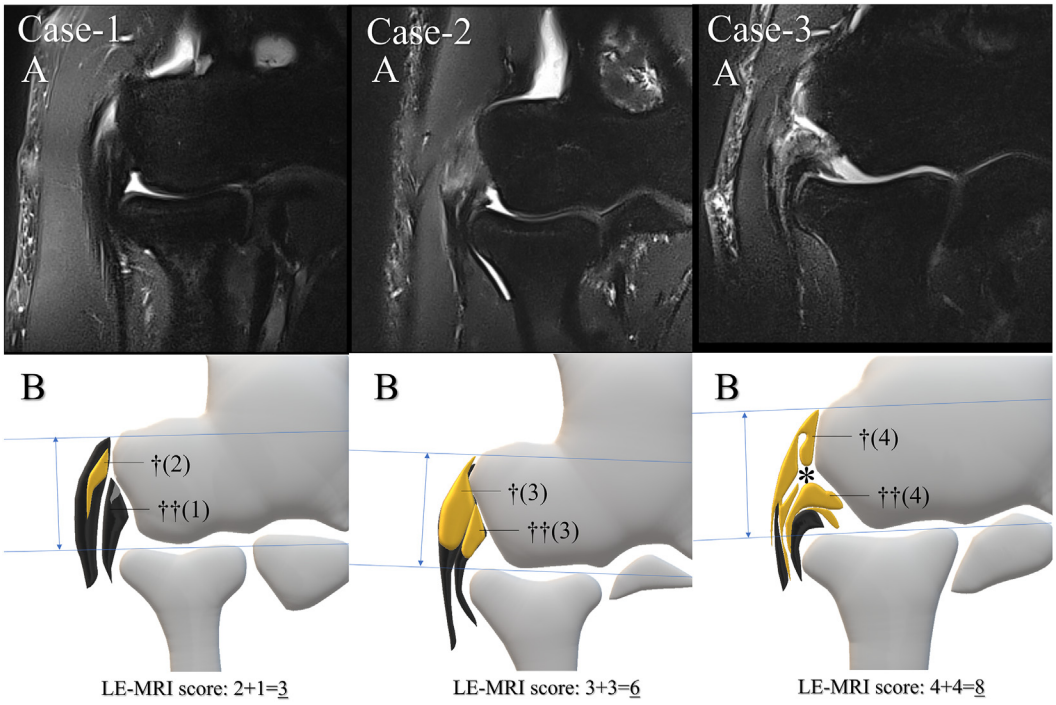


Fig. 2. Examples of magnetic resonance imaging severities. A, coronal images of T2-fat saturated weighted images; B, corresponding schema; †, common extensor tendon; ††, lateral collateral ligament; (number), score; *, tear; yellow area, severe degeneration; arrow, the region of interest for scoring. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

repeated the image analysis twice, with the second taking place one month after the initial analysis. During MRI evaluation, a third person, blinded to any clinical data, randomized the MR images. We adopted the MRI scores measured by Examiner 1 for further analyses.

2.6. Statistical analysis

We performed the Shapiro-Wilk test as a normality test for each evaluated item; none followed a normal distribution. All collected parameters, including clinical characteristics and measured values, were measured. We used Fisher's exact test for categorical parameters that included five or fewer study items and the chi-squared test for those with more items. We used the Mann-Whitney *U* test for continuous parameters between the two groups.

We developed a receiver operating characteristic (ROC) curve consisting of LE-MRI score and treatment outcomes, with the point closest to the upper left corner of the curve as the cut-off value. We categorized patients into MRI-mild and MRI-severe groups according to the obtained cut-off value; patients with LE-MRI scores less than the cut-off value were classified as MRI-mild patients, while those with LE-MRI scores at or above the cut-off value were classified as MRI-severe patients. We compared treatment outcomes between conservative treatment and surgery based on MRI severity.

We evaluated the intrarater and interrater reliability of treatment outcomes and MRI scores using Fleiss' kappa analysis. The kappa coefficients were interpreted as follows: 0.41–0.60, fair; 0.61–0.80, good; and 0.81–1.00, excellent agreement. *P*-values <0.05 were considered statistically significant.

All statistical analyses were performed using Bell Curve for Excel version 3.20 (SSRI Co., Tokyo, Japan).

3. Results

3.1. Conservative treatment outcomes and LE-MRI scores

Of the 43 patients treated conservatively, 29 (67.4%) and 14 (32.6%) had good and poor outcomes, respectively. The demographic and clinical data of each group are shown in the supplemental data (Table S2). The median (25–75 percentile) values of the LE-MRI scores were 4 (3–5) and 6 (6–6) in patients with good

and poor outcomes, respectively. LE-MRI scores in the poor outcome patients were higher than in the good outcome patients with conservative treatment ($p < 0.001$).

In the ROC curve of the LE-MRI score for predicting conservative treatment outcomes, the area under the curve (AUC) and its 95% confidence interval was 0.81 (0.68–0.93) ($p < 0.001$), as shown in Fig. 3. The cut-off value was 6, and its positive and negative likelihood ratio was 3.80 and 0.27, respectively.

3.2. Surgical treatment outcomes and LE-MRI scores

Of 50 surgically-treated patients, 43 (86.0%) had good outcomes, and seven (14.0%) had poor outcomes. The demographic and clinical data of each group are shown in the supplemental data (Table S3).The median (25–75 percentile) values of the MRI scores were 6 (5–6) and 6 (5–6) in patients with good and poor treatment outcomes, respectively. The LE-MRI scores were not significantly different between patients with good and poor outcomes ($p = 0.35$).

In the ROC curve of the LE-MRI score for predicting surgical outcomes, the AUC and its 95% confidence interval was 0.60 (0.40–0.81); the ROC curve for predicting surgical treatment outcomes was not significant ($p = 0.32$).

3.3. Comparison of conservative and surgical treatment outcomes for each LE-MRI severity group

Based on the cut-off value of the ROC curve for conservative treatment, we defined MRI-mild patients as those with LE-MRI scores ≤5 and MRI-severe patients as those with LE-MRI scores ≥6. The demographic and clinical characteristics of MRI-mild and severe patients are shown in Tables 2 and 3.

Fig. 4 shows the treatment outcomes for each LE-MRI severity. There were no significant difference in treatment outcomes between conservative and surgical treatments in the MRI-mild group ($p = 0.46$). Of the 26 MRI-mild patients who underwent conservative treatment, 23 had good and 3 had poor outcomes; Of the 15 patients with surgery, 12 had good, and 3 had poor treatment outcomes.

In MRI-severe patients, the outcomes of conservative treatment were significantly lower than that of surgical treatment ($p < 0.001$). Of the 17 MRI-severe patients with conservative treatment, 6 had good, and 11 had poor outcomes. Of the 35 patients with surgery, 31 had good, and 4 had poor outcomes.

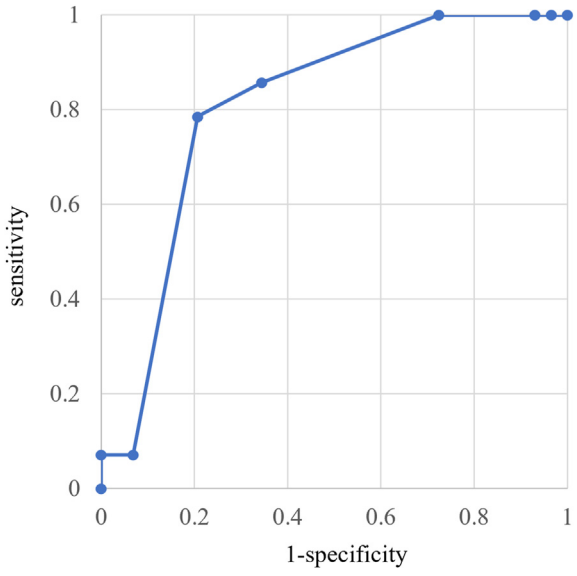


Fig. 3. The receiver operating characteristic curve of the LE-MRI score to predict conservative treatment outcomes.LE, lateral epicondylitis; MRI, magnetic resonance imaging.

3.4. Reliability of treatment outcome determination and the LE-MRI scoring system

Values for the kappa statistic showed excellent intra- and inter-observer agreement in terms of treatment outcome determination. The kappa values and their 95% confidence intervals were 0.94 (0.85–1.02; p < 0.001) between Examiner 1-1 and 2-1; 0.94 (0.85–1.02; p < 0.001) between Examiner 1-1 and 1-2; and 0.91 (0.80–1.01; p < 0.001) between Examiner 2-1 and 2-2.

The reliability of LE-MRI scoring system in the intra- and inter-observer agreement was also excellent. The kappa values and their 95% confidence intervals were 0.81 (0.77–0.85; p < 0.001) between Examiner 1-1 and 3-1; 0.84 (0.79–0.89; p < 0.001) between Examiner 1-1 and 1-2; and 0.83 (0.79–0.86; p < 0.001) between Examiner 3-1 and 3–2.

4. Discussion

This study showed the predictive ability of MRI to determine the outcome of conservative treatment. As hypothesized, patients with higher LE-MRI scores had worse outcomes under conservative treatment. The recovery rate of 6 months post-conservative treatment in the MRI-mild group was 88.5%, comparable to previous studies, while that of the MRI-severe group was only 35.3%, significantly worse than previously reported [10,11]. As MRI captures pathological changes, MRI-severe patients may have experienced severe pathological changes, and did not respond to conservative treatment.

Based on the treatment outcome, we demonstrated that MRI could guide treatment strategies for LE. In MRI-severe patients, the outcome of surgical treatment was superior to that of conservative treatment. Therefore, surgery should be considered for patients with refractory LE with MRI scores ≥6. Conversely, surgical indications should be considered carefully in patients with a slight MRI signal change. According to a previous study, MRI had a high diagnostic ability for LE, with a cut-off value of 3 [14]. Therefore, differential diagnoses should be considered when the MRI score is ≤ 2, despite severe and persistent pain. Furthermore, the MRI-mild group showed good outcomes after conservative treatment, and the outcomes were not significantly different between conservative and surgical treatments. Therefore, if LE patients with mild LE-MRI scores do not experience symptom relief despite conservative treatments, wrong initial diagnosis or poor compliance to conservative treatment should be considered, suggesting patient is not a good fit for surgery. Thus, we do not recommend surgery for patients with low LE-MRI scores.

We emphasize that the indication for MRI should be limited to refractory LE as 90% of patients with LE recover with conservative treatment [10,11]. Our study is significant as it shows that MRI could assist in treatment strategies for refractory LE rather than initial conservative treatment. When determining a treatment strategy for refractory LE, MRI is beneficial in also determining the surgical procedure. MRI can detect pathologies that develop from LE, such as extensive rupture of the CET/LCL complex and synovial fold disorders. Using positive MRI findings, we can select the appropriate surgical procedure for the pathology, e.g., Nirschl's

Table 2
Demographic and clinical characteristics of MRI-mild patients (LE-MRI scores ≤5).

	Group		
	With conservative treatment ($n = 26$)	With surgical treatment ($n = 15$)	p-value
Age (yrs)[a]	52 (48–57)	49 (44–50)	0.049
Sex	Male: 14 Female: 12	Male: 6 Female: 9	0.39
Duration of disease (month)[a]	3.2 (1.4–5.4)	11.4 (2.4–19.2)	0.035
Total number of previous steroid injections[a]	0 (0–1)	1 (1–4)	0.002

[a] Data are presented as the median (25–75 percentile); under line, $p < 0.05$; MRI, magnetic resonance imaging; LE, lateral epicondylitis.

Table 3
Demographic and clinical characteristics of MRI-severe patients (LE-MRI scores ≥6).

	Group		
	With conservative treatment ($n = 17$)	With surgical treatment ($n = 35$)	p-value
Age (yrs)[a]	49 (46–55)	49 (45–56)	$p = 0.67$
Sex	Male: 6 Female: 11	Male: 19 Female: 16	$p = 0.20$
Duration of disease (month)[a]	4.1 (2.0–9.2)	6.2 (2.8–12.9)	$p = 0.26$
A total number of previous steroid injections[a]	1 (0–2)	3 (1–4)	$p = 0.022$

[a] Data are presented as median (25–75 percentile); under line, $p < 0.05$; MRI, magnetic resonance imaging; LE, lateral epicondylitis.

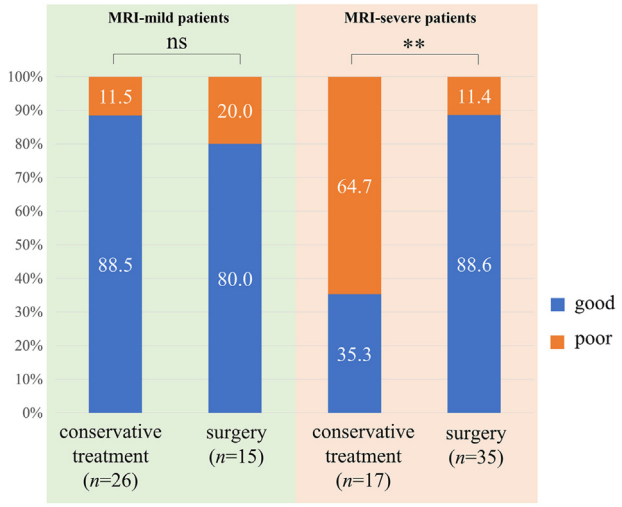


Fig. 4. Treatment outcomes for each LE-MRI severity groups. LE, lateral epicondylitis; MRI, magnetic resonance imaging; ns, not significant; **, $p < 0.01$.

procedure for severe degeneration of CET [23], reconstruction of the CET/LCL complex for massive tear [4,29], or arthroscopic surgery for the intraarticular disorder [25]. MRI is essential in determining the treatment strategy for refractory LE.

In MRI evaluation, the kappa statistics results were all excellent in this study. We believed we could provide reliable data using the highest resolution MRI and detailed MRI evaluation.

This study had some limitations. As this was a retrospective study, the conservative and surgical treatments were not standardized. Therefore, the possibility of confounding benefactors, including degeneration due to repeated steroid injection [30], cannot be excluded. Furthermore, since we determined treatment outcomes from limited medical records entries, we were limited to qualitative outcome evaluation to accurately assess. Future prospective study with standardized treatment protocols, comprehensive multivariate analysis of risk factors, and qualitative evaluation of treatment outcomes would reveal a more detailed prognostic potential of MRI.

5. Conclusion

The LE-MRI score was associated with conservative treatment outcomes. Patients with poor outcomes to conservative treatment had higher LE-MRI scores, with a cut-off value of 6. Patients in the MRI-severe group with LE-MRI scores ≥6 had better results with surgery than with conservative treatment. Conversely, in the MRI-mild group with MRI scores ≤5, the outcomes showed no significant difference between conservative and surgical treatments. LE-MRI scoring was beneficial for predicting conservative treatment outcomes and determining treatment strategies for LE.

Conflicts of interest

Any authors of this study have no COI.

Funding

None.

Acknowledgement

We would like to thank Naotaka Mamizuka for laying the foundation for this MRI study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jos.2023.03.014.

References

[1] Runge F. Zur Genese und Behandlung des schreibe Kranfes. Bed Klin Worchenschr 1873;10(1):245–8 (in German).
[2] Bredella MA, Tirman PF, Fritz RC, Feller JF, Wischer TK, Genant HK. MR imaging findings of lateral ulnar collateral ligament abnormalities in patients with lateral epicondylitis. AJR Am J Roentgenol 1999 Nov;173(5):1379–82.
[3] Cha YK, Kim SJ, Park NH, Kim JY, Kim JH, Park JY. Magnetic resonance imaging of patients with lateral epicondylitis: relationship between pain and severity of imaging features in elbow joints. Acta Orthop Traumatol Turcica 2019 Sep;53(5):366–71.
[4] Kalainov DM, Cohen MS. Posterolateral rotatory instability of the elbow in association with lateral epicondylitis. A report of three cases. J Bone Joint Surg Am 2005 May;87(5):1120–5.
[5] Ozone K, Kokubun T, Takahata K, Takahashi H, Yoneno M, Oka Y, et al. Structural and pathological changes in the enthesis are influenced by the

muscle contraction type during exercise. J Orthop Res 2022 Sep;40(9):2076–88.
[6] Bhabra G, Wang A, Ebert JR, Edwards P, Zheng M, Zheng MH. Lateral elbow tendinopathy: development of a pathophysiology-based treatment algorithm. Orthop J Sports Med 2016 Nov;4(11):2325967116670635.
[7] Kraushaar BS, Nirschl RP. Tendinosis of the elbow (tennis elbow). Clinical features and findings of histological, immunohistochemical, and electron microscopy studies. J Bone Joint Surg Am 1999 Feb;81(2):259–78.
[8] Regan W, Wold LE, Coonrad R, Morrey BF. Microscopic histopathology of chronic refractory lateral epicondylitis. Am J Sports Med 1992 Nov-Dec;20(6):746–9.
[9] Ahmad Z, Siddiqui N, Malik SS, Abdus-Samee M, Tytherleigh-Strong G, Rushton N. Lateral epicondylitis: a review of pathology and management. Bone Joint Lett J 2013 Sep;95-B(9):1158–64.
[10] Ahmed AF, Rayyan R, Zikria BA, Salameh M. Lateral epicondylitis of the elbow: an up-to-date review of management. Eur J Orthop Surg Traumatol 2023 Jan;33(2):201–6.
[11] Ma KL, Wang HQ. Management of lateral epicondylitis: a narrative literature review. Pain Res Manag 2020;2020 May:6965381.
[12] Zwerus EL, Somford MP, Maissan F, Heisen J, Eygendaal D, van den Bekerom MP. Physical examination of the elbow, what is the evidence? A systematic literature review. Br J Sports Med 2018 Oct;52(19):1253–60.
[13] Jeon JY, Lee MH, Jeon IH, Chung HW, Lee SH, Shin MJ. Lateral epicondylitis: associations of MR imaging and clinical assessments with treatment options in patients receiving conservative and arthroscopic managements. Eur Radiol 2018 Mar;28(3):972–81.
[14] Ikeda K, Ogawa T, Ikumi A, Yoshii Y, Kohyama S, Ikeda R, et al. Individual evaluation of the common extensor tendon and lateral collateral ligament improves the severity diagnostic accuracy of magnetic resonance imaging for lateral epicondylitis. Diagnostics 2022 Aug;12(8):1871.
[15] Pasternack I, Tuovinen EM, Lohman M, Vehmas T, Malmivaara A. MR findings in humeral epicondylitis. A systematic review. Acta Radiol 2001;42(5):434–40.
[16] Qi L, Zhang YD, Yu RB, Shi HB. Magnetic resonance imaging of patients with chronic lateral epicondylitis: is there a relationship between magnetic resonance imaging abnormalities of the common extensor tendon and the patient's clinical symptom? Med (Baltim) 2016 Feb;95(5):e2681.
[17] Shiri R, Viikari-Juntura E. Lateral and medial epicondylitis: role of occupational factors. Best Pract Res Clin Rheumatol 2011 Feb;25(1):43–57.
[18] Wolf JM, Mountcastle S, Burks R, Sturdivant RX, Owens BD. Epidemiology of lateral and medial epicondylitis in a military population. Mil Med 2010 May;175(5):336–9.
[19] Kessler RE, Day MS, Tyler TF, McHugh MP, Bedford BB, Lee SJ, et al. Predictive value of magnetic resonance imaging in outcomes of nonsurgical treatment of lateral epicondylitis. JSES Int 2022 Jan;6(2):305–8.
[20] Steinborn M, Heuck A, Jessel C, Bonel H, Reiser M. Magnetic resonance imaging of lateral epicondylitis of the elbow with a 0.2-T dedicated system. Eur Radiol 1999;9(7):1376–80.
[21] Walton MJ, Mackie K, Fallon M, Butler R, Breidahl W, Zheng MH, et al. The reliability and validity of magnetic resonance imaging in the assessment of chronic lateral epicondylitis. J Hand Surg Am 2011 Mar;36(3):475–9.
[22] Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957 Dec;16(4):494–502.
[23] Nirschl RP, Ashman ES. Elbow tendinopathy: tennis elbow. Clin Sports Med 2003 Oct;22(4):813–36.
[24] Plancher KD, Bishai SK. Open lateral epicondylectomy: a simple technique update for the 21st century. Tech Orthop 2006 Dec;21(4):276–82.
[25] Stiefel EC, Field LD. Arthroscopic lateral epicondylitis release using the "bayonet" technique. Arthrosc Tech 2014 Jan 31;3(1):e135–9.
[26] Smith-Forbes V, Howell DM, Willoughby J, Pitts DG, Uhl TL. A retrospective cohort study of QuickDASH scores for three hand therapy acute upper limb conditions. Mil Med 2018 Mar 1;183(suppl_1):522–9.
[27] Franchignoni F, Vercelli S, Giordano A, Sartorio F, Bravini E, Ferriero G. Minimal clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (QuickDASH). J Orthop Sports Phys Ther 2014 Jan;44(1):30–9.
[28] Nimura A, Fujishiro H, Wakabayashi Y, Imatani J, Sugaya H, Akita K. Joint capsule attachment to the extensor carpi radialis brevis origin: an anatomical study with possible implications regarding the etiology of lateral epicondylitis. J Hand Surg Am 2014 Feb;39(2):219–25.
[29] Noh YM, Kong GM, Moon SW, Jang HS, Kim S, Bak GG, et al. Lateral ulnar collateral ligament (LUCL) reconstruction for the treatment of recalcitrant lateral epicondylitis of the elbow: a comparison with open débridement of the extensor origin. JSES Int 2021 Feb;5(3):578–87.
[30] Dean BJF, Lostis E, Oakley T, Rombach I, Morrey ME, Carr AJ. The risks and benefits of glucocorticoid treatment for tendinopathy: a systematic review of the effects of local glucocorticoid on tendon. Semin Arthritis Rheum 2014 Feb;43(4):570–6.

Received: 29 March 2023 | Revised: 11 July 2023 | Accepted: 24 August 2023

DOI: 10.1002/jor.25684

RESEARCH ARTICLE

Journal of Orthopaedic Research®

Pathophysiology of sex difference in refractoriness in lateral epicondylitis: Biomechanical study of wrist torque

Kazuhiro Ikeda[1,2] | Yuichi Yoshii[3] | Sho Kohyama[1] | Akira Ikumi[2] | Takeshi Ogawa[4] | Reimi Ikeda[5] | Masashi Yamazaki[2]

[1]Department of Orthopedic Surgery, Kikkoman General Hospital, Noda City, Japan

[2]Department of Orthopedic Surgery, Faculty of Medicine, University of Tsukuba, Tsukuba City, Japan

[3]Department of Orthopedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ami Town, Japan

[4]Department of Orthopedic Surgery, Mito Medical Center, Ibarakimachi, Japan

[5]Department of Orthopedic Surgery, Moriya Daiichi General Hospital, Moriya City, Japan

Correspondence

Yuichi Yoshii, Department of Orthopedic Surgery, Tokyo Medical University Ibaraki Medical Center, 3-20-1 Chuo, Ami, Inashiki, Ibaraki 300-0395, Japan.
Email: yyoshii@tokyo-med.ac.jp

Abstract

Eccentric contractions of the wrist extensors worsen lateral epicondylitis (LE), whose pathophysiology may involve sex differences in wrist torque. This study aimed to investigate sex differences in wrist torque in patients with LE. The wrist extension and flexion torques of 22 patients with LE (11 males and 11 females) were measured. Maximum muscle output over time was measured for 20 s, initial torque was defined as muscle strength, and the degree of eccentric contraction was quantified and defined as the eccentric contraction index (ECI). The affected/unaffected side ratio of the wrist extensor, extensor/flexor ratio of muscle strength, and affected/unaffected side difference of ECI between sexes were statistically analyzed. Furthermore, correlations between wrist extensor torque, ECI, and Visual Analog Scale of pain during the examination were evaluated. Females were found to display lower affected/unaffected side ratios of the wrist extensor and wrist extension/flexion ratios for the affected side, compared with males; however, no differences were found in the wrist extension/flexion ratios for the unaffected side in both sexes. Additionally, females presented with larger differences between the affected and unaffected sides in the ECI. Based on correlations between wrist torques, ECI, and pain, females tended to suppress muscle output to prevent pain from eccentric contraction of wrist extensors more than males, which would induce an imbalance in muscle strength of the wrist extensors and flexors. This imbalance may result in chronic eccentric contraction of the wrist extensors with gripping, exacerbating LE.

KEYWORDS

biomechanics, elbow, tendon

1 | INTRODUCTION

1.1 | Background

Lateral epicondylitis (LE) of the humerus is the tendinopathy of the forearm extensors.[1,2] The prevalence of LE is 1%–3% of the general population[2,3] and 2%–14.5% of manual workers who perform repetitive gripping movements with the forearm in pronation.[4,5] Females have a higher prevalence of LE and a 2.7 times higher risk of refractoriness than males.[3,4,6] Pathophysiologically, repetitive eccentric contraction of wrist extensors is a factor in the development and exacerbation of LE, which induces micro rupture at the enthesis.[7,8] Uncontrolled conditions worsen histological severity with vascular fibroblast proliferation, mucoid degeneration, calcification, and ectopic ossification at the enthesis.[7,9,10] Decreased muscle strength and endurance of the wrist extensors are considered to be risk factors for the development of LE, because it induces an eccentric contraction of the wrist extensors in gripping with forearm

pronation.[11] Although LE is considered a self-limiting disease, about 10% of patients do not respond to conservative treatment and sometimes require surgery for refractory LE.[9]

As refractory LE produces a socioeconomic burden resulting from labor difficulties, overcoming it is a social issue; the economic loss reportedly reaches £27 million annually in England.[12] Although heavy labor, smoking, and females are epidemiologically known as risk factors for refractoriness,[4] the pathophysiology leading to refractoriness is poorly understood. Particularly, no literature provides pathophysiological evidence that females are more prone to develop refractory LE than males. The lack of pathophysiological knowledge leads to inappropriate treatment, which results in a difference in refractoriness rates between sex. Therefore, it is essential to understand the pathophysiology of the difference in refractoriness rates between sex to overcome refractory LE.

1.2 | Rationale

As LE exacerbates with repetitive eccentric contractions of wrist extensors,[7,8] female patients with LE would be expected to be more exposed to eccentric contractions than males. Muscles are prone to eccentric contraction when muscle performance declines, including low muscle strength, contraction velocity, and endurance.[11] In patients with LE, muscle contraction velocity of the extensor carpi radialis delays,[13] and the strength of the wrist extensor reduces relative to the strength of the flexor.[14,15] As wrist extensors and flexors co-contract in gripping,[16,17] this muscle strength imbalance of the wrist may result in repetitive eccentric contractions with each gripping motion. Therefore, we hypothesized that the loss of wrist extensor strength due to LE is more significant in females than in males, resulting in a more significant muscle imbalance between the wrist extensors and flexors. This study aimed to reveal sex differences in wrist extension and flexion torques in patients with LE at their initial examination.

2 | METHODS

2.1 | Study design

The study protocol conforms to the principles outlined in the 1964 Declaration of Helsinki. This study was approved by our Institutional Review Board. Written informed consent was obtained from all patients.

This case–control study investigated the wrist extension and flexion torques of patients with LE cross-sectionally (Level of evidence III).

2.2 | Participants

Participants were consecutive patients with unilateral LE whom the author examined in our hospital between February 2022 and January 2023. The exclusion criteria were patients with any shoulders or wrists symptoms, a history of elbow trauma, and arthritic diseases: osteoarthritis (Kellgren–Lawrence classification 2 or higher),[18] osteochondritis dissecans, and rheumatoid arthritis. Furthermore, we excluded mild patients to clarify the characteristics of wrist torque in patients with LE. As previous studies reported quick disabilities of the arm, shoulder, and hand (QuickDASH) scores of 11–15 points at the end of conservative treatment for LE,[19,20] we set a cutoff score of QuickDASH for exclusion criteria of patients with mild LE as 11.

LE diagnosis was based on physical examination findings, positive Thomsen or Maudsley's test, and tenderness at the lateral epicondyle. All patients underwent an X-ray and MRI to rule out elbow arthritis and trauma.

Thirty patients with LE received the author's examination during the study period. We excluded two patients with bilateral symptoms, four with a QuickDASH score of fewer than 11 points, one with a history of elbow trauma, and one who refused to participate in this study. Subsequently, this study included 22 patients with LE (11 males and females each, average age and SD: 55.6 ± 13.4 years, age range: 39–84 years).

2.3 | Physical examination and clinical evaluation

The following data were recorded on the subject background: height, weight, body mass index (BMI), forearm length and circumference, lever arm length for obtaining wrist torque, and grip strength. BMI is reported to be positively associated with grip strength[21] and wrist torque.[22] The grip strength was measured with a hand dynamometer in the upper limb hanging downward, elbow extension, and the forearm in a neutral position. The forearm length was from the lateral epicondyle of the humerus to the styloid process of the radius. The forearm circumference was around the thickest part of the forearm. These forearm lengths and circumferences were measured as indicators of the forearm muscle volume.[23] The lever arm length was from the tip of the radial styloid to the grip center. This measurement of the lever arm length was evaluated by gripping the torque measuring device with their wrist and forearm in a neutral position.

As the patient-based self-evaluation for LE, disability/symptom score of QuickDASH, Visual Analog Scale (VAS) for pain at rest, during daily activities, and wrist torque measurement were evaluated. Thereafter, we calculated the VAS difference as follows: VAS difference = (VAS during daily activity) − (VAS during torque measurement).

2.4 | Measurement of wrist torque

Wrist extension and flexion muscle strengths were measured with a wrist torque measuring device (Three-One Design Inc.) (Figure 1).[24] This apparatus measures wrist torques every 10 ms and averages

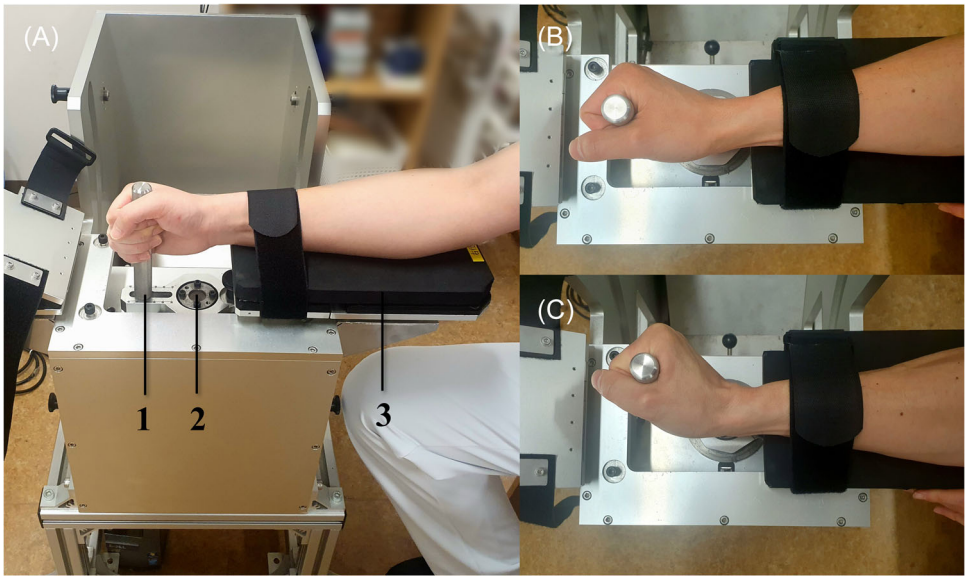


FIGURE 1 (A) The wrist torque measuring equipment consists of a handle (1), a rotation center to measure the wrist torque (2), and a table to place the forearms (3). (B) Participants adjusted the moving handle to ensure the rotation center of the wrist and the rotation center of the equipment together. (C) The handle rotates around the wrist axis in response to the input torque.

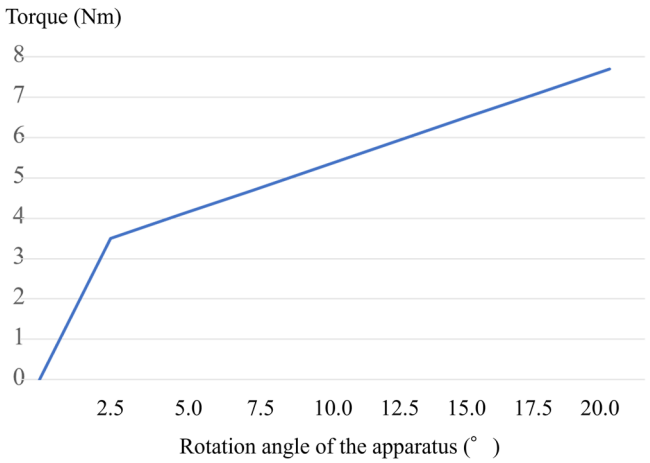


FIGURE 2 Input torque required to rotate the torque-measuring device, measured every 2.5°. The measuring machine required 1.5 Nm/° for the initial movement. Thereafter, the required force was 0.24 Nm/° to rotate the handle.

them every second, which enables us to evaluate reliable torque over time. Participants sat on a chair and placed their forearms on the apparatus table, with their elbows slightly flexed at 30°–45° and the forearm in a neutral position. The rotation centers of the wrist, which is the axis connecting the radial and ulnar styloid,[24] were adjusted to the rotational center of the device by moving the handle position. The gripping handle of this measurement device rotates around the wrist rotational axis in response to the output of wrist torque. Figure 2 shows the torque output required to rotate this torque-measuring device every 2.5°. Patients performed a maximal voluntary wrist extension or flexion contraction for 20 s. As wrist torque measurement in this study evaluates dynamic muscle contraction, decreased muscle output over time during examination represents eccentric contraction. On the affected side, the patient was encouraged to exert as much force as possible within the self-acceptable pain. Although we explained that the examination could be interrupted if the pain was severe, none of the patients were interrupted.

2.5 | Data analysis

We assessed maximum muscle strength and the degree of eccentric contraction that occurred during wrist torque examination.

The wrist torque over time was evaluated for 20 s from its first peak waveform. The wrist torque of the first peak waveform was defined as muscle strength at the starting point of the measurement (MS[S]). MS(S) was used for further evaluation as the maximum muscle strength of the patients. Muscle strength of 20 s after the start point was used as muscle strength at the end of the measurement (MS[E] [Nm]). Then, the muscle strength loss (MS[L]) per second was calculated with the following equation: MS(L) = (MS[S] − MS[E])/20 [Nm/s]). MS(L) reflects the degree of eccentric contraction that occurred during the measurement. Finally, we normalized MS(L) by dividing MS(S) and defined this as eccentric contraction index (ECI): ECI = MS(L)/MS(S).

The primary parameters in this study were the affected/unaffected side ratio of the MS(S), the wrist extension/flexion ratio of the MS(S), and the affected/unaffected side difference of the ECI.

2.6 | Statistical analysis

We performed the Shapiro–Wilk test for each evaluated item as a normality test. All collected parameters, including clinical characteristics and measured values, were compared. We used the χ^2 test for categorical parameters, the Student's *t* test for continuous parameters between sexes, and the paired *t* test to compare the affected and unaffected sides with a normal distribution. We used the Mann–Whitney *U* test for continuous parameters between sexes and Wilcoxon signed-rank test to compare the affected and unaffected sides with an irregular distribution.

We analyzed Spearman's correlation coefficients to assess the association between the clinical characteristics and wrist torque. Correlation coefficients of $\pm 0.3 < r < \pm 0.7$ were considered moderate correlations, and those of $r > \pm 0.7$ were considered strong correlations.

For statistical comparisons in the present study, we assumed that comparisons of the extension/flexion ratio of the muscle strength within each sex would have the most power, as previous studies have already shown significant differences.[15] Therefore, in determining sample size, we calculated the sample size that would satisfy an effect size of 0.8, a type I error of 0.05, and a power of 0.8 for the paired *t* test. Subsequently, we determined a sample size of 11 for each sex. $p < 0.05$ was considered to be significant. All statistical analyses were performed using Bellcurve for Excel version 3.20 (SSRI Co.n).

3 | RESULTS

3.1 | Demographic and clinical characteristics

Table 1 summarizes the demographic and clinical characteristics of males and females. There were no significant differences in the distribution of age. Concerning physical characteristics, height, weight, forearm length, and circumference, lever arm length to obtain wrist torque and grip strength were larger in males than in females. There was no significant difference in BMI or in any of the items related to clinical characteristics or severity of LE between sexes, that is, comparisons of dominant hand incidence, duration of disease, affected/unaffected side ratio of grip strength, QuickDASH score, or VAS.

TABLE 1 Demographic and clinical characteristics between males and females.

	Group		
	Males ($n = 11$)	Females ($n = 11$)	*p*
Age (years)[a]	54.1 ± 16.3	57.1 ± 10.4	0.61
Height (cm)[a]	169.3 ± 7.5	157.0 ± 5.4	<0.001
Weight (kg)[a]	71.6 ± 12.2	55.9 ± 13.5	0.010
BMI[a]	24.9 ± 2.9	22.7 ± 5.5	0.26
Forearm length (cm)[a]			
Affected side	26.3 ± 1.5	23.7 ± 1.5	<0.001
Unaffected side	26.3 ± 1.8	23.7 ± 1.5	0.001
Forearm circumference (cm)[a]			
Affected side	27.2 ± 2.2	23.5 ± 2.7	0.002
Unaffected side	27.2 ± 2.1	23.4 ± 2.9	0.002
Lever arm length for obtaining the wrist torque (mm)[a]	66.4 ± 3.2	60.0 ± 2.3	<0.001
Grip strength (kgf)[a]			
Affected side	27.2 ± 10.7	13.5 ± 7.6	0.002
Unaffected side	38.2 ± 10.4	23.5 ± 8.9	0.002
Affected/unaffected side ratio[a]	0.75 ± 0.28	0.57 ± 0.28	0.15
Affects the dominant hand	8 (72.7%)	9 (81.8%)	0.61
Duration of disease (days)[b]	57 (35–74.5)	62 (19.5–89.5)	0.53
QuickDASH score[a]	31.6 ± 18.3	42.0 ± 20.6	0.25
VAS during			
At rest[b]	0 (0–25.5)	0 (0–21.0)	0.97
Daily activity[a]	50.4 ± 18.9	54.2 ± 15.3	0.61
Wrist torque measurement[a]	27.1 ± 26.4	34.5 ± 21.6	0.48

Note: Underline, $p < 0.05$.

Abbreviations: BMI, body mass index; DASH, disabilities of arm, shoulder and hand; VAS, Visual Analog Scale.

[a]Data are normally distributed, described with average ± SD.

[b]Data are irregularly distributed, described with median (25–75 percentile).

3.2 | Wrist torque comparison

Figure 3 shows the results of the average values for wrist extension torque over time in males and females, with Table 2 summarizing the analysis. MS(S) was significantly increased in males than females on both the affected and unaffected sides. The affected/unaffected side ratio of MS(S) was significantly increased in males than in females. On the affected side, MS(L) and ECI were significantly increased in males than that in females. ECI difference between the affected and unaffected sides was more significantly increased in females than males. The ECI of the affected side was lower than that of the unaffected side in females ($p = 0.002$), whereas there was no significant difference in males ($p = 0.67$).

Figure 4 shows the results of the average values for wrist flexion torque over time, with Table 3 summarizing the analysis. MS(S) was increased in males than in females on both the affected and unaffected sides. There were no significant differences in the affected/unaffected side ratio of the MS(S), MS(L), ECI, and the

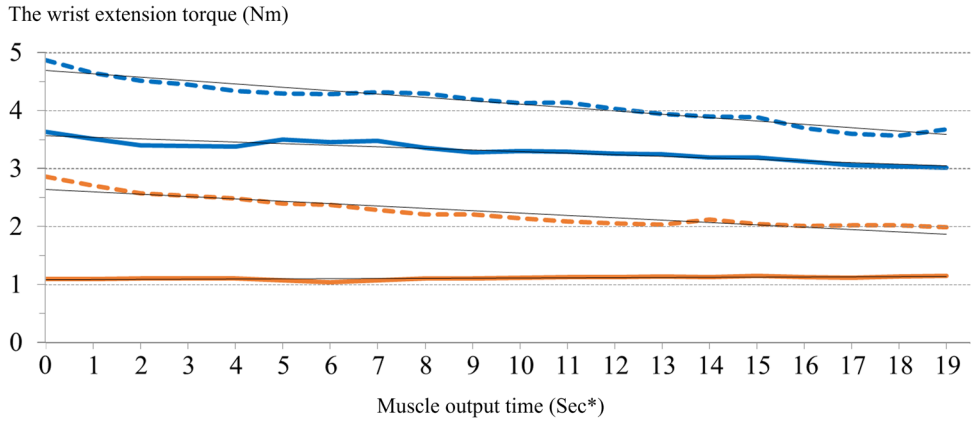


FIGURE 3 The average wrist extension torque for males and females over time. The intercept of the regression line approximates the muscle strength at the starting point of the measurement and the slope approximates the decrease in muscle strength per second. The regression line for each graph was as follows: Blue dotted line, the unaffected side in males ($y = -0.058x + 4.7$, $R^2 = 0.95$); blue solid line, the affected side in males ($y = -0.027x + 3.6$, $R^2 = 0.89$); orange dotted line, the unaffected side in females ($y = -0.041x + 2.7$, $R^2 = 0.89$); orange solid line; affected side in females ($y = -0.0031x + 1.1$, $R^2 = 0.49$). *The apparatus measures wrist torques every 10 ms and averages them every second; the torque at x seconds in the graph indicates the average torque of x to $x + 1$ s during the measurement.

TABLE 2 The result of wrist extension torques measurement.

	Males ($n = 11$)	Females ($n = 11$)	p
MS(S)			
Affected side (Nm)[b]	4.1 (2.6–4.3)	0.7 (0.5–1.5)	<0.001
Unaffected side (Nm)[a]	4.9 ± 1.5	2.7 ± 1.5	0.003
Affected/unaffected side ratio[a]	0.76 ± 0.24	0.40 ± 0.19	0.001
MS(L)			
Affected side ($\times 10^{-2}$ Nm/s)[b]	2.2 (0.2–6.5)	−0.4 (−0.9–0.0)	0.028
Unaffected side ($\times 10^{-2}$ Nm/s)[b]	4.8 (2.6–8.8)	2.2 (1.3–6.4)	0.18
ECI			
Affected side ($\times 10^{-3}$)[a]	9.6 ± 14	−9.2 ± 14	0.006
Unaffected side ($\times 10^{-3}$)[a]	11.8 ± 6.7	12.7 ± 7.8	0.77
Affected/unaffected side difference ($\times 10^{-3}$)[b]	−3.8 (−10.4 to 4.3)	−19.9 (−27.1 to −11.8)	0.016

Note: Underline, $p < 0.05$.

Abbreviations: ECI, eccentric contraction index; MS(L), muscle strength loss; MS(S), muscle strength at the starting point for the measurement.

[a]Data are normally distributed, described with average ± SD.

[b]Data are irregularly distributed, described with median (25–75 percentile).

affected/unaffected side difference of the ECI between males and females.

Figure 5 shows the results of the wrist extension/flexion ratio of the MS(S). In males, the wrist extension/flexion ratios of the MS(S) were 0.48 (0.43–0.67) and 0.60 (0.43–0.76) on the affected and unaffected sides, respectively. In females, the wrist extension/flexion ratios of the MS(S) were 0.29 (0.29–0.41) and 0.62 (0.44–0.72) on the affected and unaffected sides, respectively. On the affected side, the wrist extension/flexion ratio of the MS(S) was increased in males than in females ($p = 0.020$), whereas there was no significant difference between sexes on the unaffected side ($p = 0.95$). The wrist extension/flexion ratio of the MS(S) was increased on the unaffected side than that on the affected side in females ($p = 0.013$), whereas there was no significant difference in males ($p = 0.86$).

The resistible weight of the wrist extensors and flexors, which are calculated from the lever arm length and torque results, are shown in Supporting Information: Table S1.

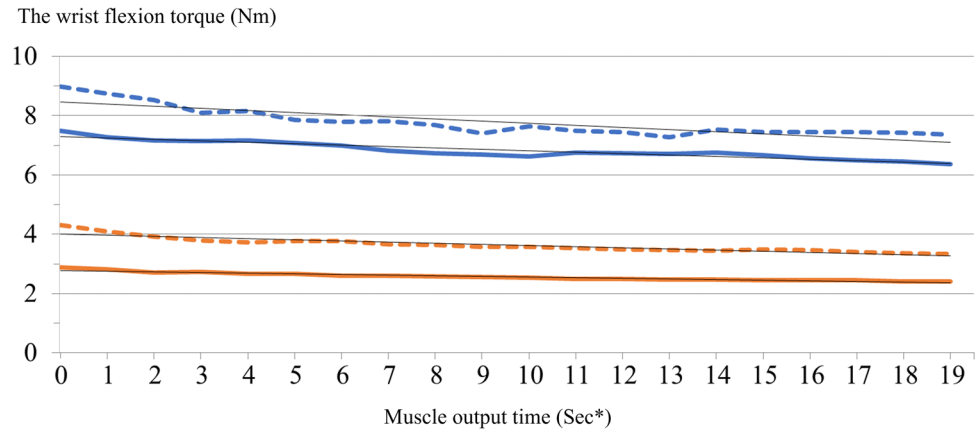


FIGURE 4 Wrist flexion torque for males and females over time. The regression line for each graph was as follows: Blue dotted line, the unaffected side in males ($y = -0.071x + 8.5$, $R^2 = 0.76$); blue solid line, the affected side in males ($y = -0.048x + 7.3$, $R^2 = 0.89$); orange dotted line, the unaffected side in females ($y = -0.039x + 4.0$, $R^2 = 0.85$); orange solid line; affected side in females ($y = -0.022x + 2.8$, $R^2 = 0.91$). *The torque at x seconds in the graph indicates the average torque of x to $x + 1$ s during the measurement.

TABLE 3 The result of wrist flexion torques and endurances in males and females.

	Males ($n = 11$)	Females ($n = 11$)	p
MS(S)			
Affected side (Nm)[b]	7.8 (4.2–9.0)	2.5 (1.7–4.5)	<0.001
Unaffected side (Nm)[b]	7.3 (6.5–9.6)	3.7 (3.0–5.3)	0.0014
Affected/Unaffected side ratio[a]	0.83 ± 0.30	0.69 ± 0.28	0.24
MS(L)			
Affected side ($\times 10^{-2}$ Nm/s)[b]	3.1(−0.1–11.0)	2.0 (−1.6–3.1)	0.49
Unaffected side ($\times 10^{-2}$ Nm/s)[b]	7.7 (3.4–11.4)	4.9 (2.6–6.4)	0.20
ECI			
Affected side ($\times 10^{-3}$)[b]	7.0 (−3.0–12.4)	8.5 (5.3–12.3)	0.34
Unaffected side ($\times 10^{-3}$)[a]	9.6 ± 6.3	11.1 ± 7.4	0.60
Affected/Unaffected side difference ($\times 10^{-3}$)[b]	−2.4 (−9.6–3.1)	−2.8 (−10.1–4.3)	0.81

Note: Underline, $p < 0.05$.

Abbreviations: ECI, eccentric contraction index; MS(L), muscle strength loss; MS(S), muscle strength at the starting point for the measurement.

[a]Data are normally distributed, described with average ± SD.

[b]Data are irregularly distributed, described with median (25–75 percentile).

3.3 | Correlations between body characteristics, clinical characteristics, and wrist extension torque

The correlations between body characteristics data and wrist extension torque of the unaffected side are shown in Supporting Information: Table S2. Table 4 shows correlations between clinical characteristics and wrist extension torque in males and females. The correlations between the affected/unaffected side ratio of MS(S) and each parameter were as follows: affected/unaffected side ratio of grip strength showed a moderate positive correlation in both sexes; VAS difference showed a moderate negative correlation in females; and QuickDASH score showed strong negative correlation in females. The correlations between the affected/unaffected side difference in ECI and each parameter were as follows: affected/unaffected side ratio of grip strength showed a moderate positive correlation in both sexes; VAS difference and QuickDASH score showed moderate negative correlations in both sexes; and affected/unaffected side ratio of MS (S) showed a strong positive correlation in females.

4 | DISCUSSION

The present study demonstrated that the affected/unaffected side ratio of the MS(S) of the wrist extensor was significantly lower in females than in males. This suggests that the wrist extensor strength in females decreased more remarkably than that of males because of

LE. In contrast, there was no difference in the loss of wrist flexor strength between the sexes. Consequently, the extension/flexion ratio of the MS(S) in females was lower on the affected side than on either their unaffected side or the affected side in males. As there was a co-contraction of wrist flexors and extensors during gripping in the wrist extension position,[16,17] this weakness of wrist extensors relative to flexors may induce chronic eccentric contraction of wrist extensors with gripping, which exacerbates the condition of LE. Therefore, this significant loss of wrist extensor strength in female patients with LE may be a risk of refractoriness.

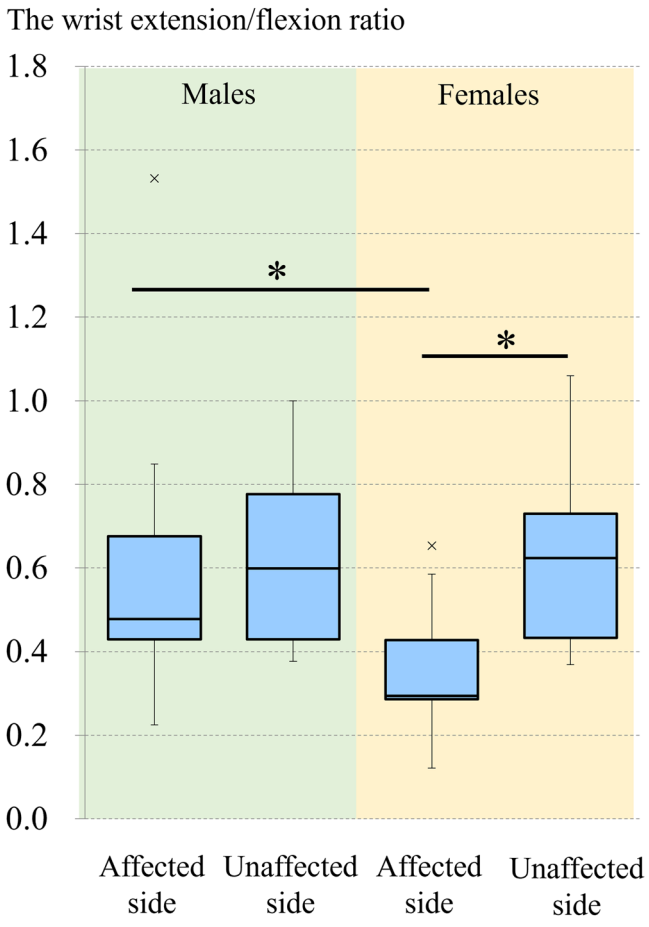


FIGURE 5 The wrist extension/flexion ratio of the muscle strength at the starting point for the measurement. The boxes show the median and the interquartile range, the whiskers with values within 1.5 times the interquartile range, and *x* marks of the outlier's values. *$p < 0.05$.

The present study identified the cause of remarkable muscle weakness in female patients with LE by quantifying the eccentric contractions and pain during the wrist torque examination. The correlation among MS(S), ECI, and VAS showed that patients with less muscle output experienced less pain and less eccentric contraction. Patients with LE experience pain due to eccentric contraction of the wrist extensors, which is clinically used in manual tests such as Thomsen and Maudsley's tests.[25] The pain caused by eccentric contraction activates the sensorimotor system and decreases muscle output through negative feedback.[13] This pain-induced suppression of muscle output was stronger in females. Although the muscle output of unaffected wrist extensors decreased almost linearly over time, the coefficient of determination of the regression line was lower in the affected wrist extensors of females. This indicates that there was unstable muscle output in the affected wrist extensors. Furthermore, the ECI of the wrist extensor in females was smaller than that in males and smaller ECI results in a smaller affected/unaffected side ratio of the MS(S) of the wrist extensor. These results indicate that females tended to suppress muscle output to prevent pain from eccentric contraction of wrist extensors more than males. Furthermore, patients who suppressed muscle output to prevent eccentric contractions had a lower affected/unaffected side ratio of grip strength in both males and females. Patients with markedly decreased grip strength should be treated with an understanding that they are at high risk for refractory LE. Aggressive conservative treatment should be considered for such patients, for example, extensive orthotic therapy of tennis elbow brace with a cock-up splint[26] and physical therapy, including controlled eccentric contraction exercises.[27]

Similar to the present study, previous studies of experimentally induced pain have consistently shown that females exhibit greater pain sensitivity than males.[28] These sex differences in pain are

TABLE 4 Correlation between clinical data and the affected wrist extension torque.

	Affected/unaffected side ratio of MS(S)		Affected/unaffected side difference in ECI	
	Males	Females	Males	Females
Affected/Unaffected side ratio of grip strength	0.34	0.37	0.52	0.42
VAS difference[a]	0.00	−0.61	−0.30	−0.39
QuickDASH score	−0.10	−0.72	−0.40	−0.47
Affected/unaffected side ratio of MS(S)			−0.19	0.83

Note: 0.3 < < 0.7; 0.7 < ; −0.7 < <−0.3; <−0.7.

Abbreviations: DASH score, disabilities of the arm, shoulder and hand score; ECI, eccentric contraction index; MS(S), muscle strength at the starting point for the measurement; VAS, visual analog scale.

[a]VAS difference = (VAS during daily activity) − (VAS during torque measurement).

reported as multiple biopsychosocial mechanisms, including sex hormones, endogenous opioid function, and genetic factors.[28] We have shown that severe pain in female patients with LE is not only the consequence of a severe condition but also the cause of morbid muscle output suppression, which leads to refractory LE. In LE refractoriness, severe pain triggers the vicious cycle of worsening LE and further reduces the muscle output of wrist extensors.

The limitation of the present study was that this was a cross-sectional study. We did not directly evaluate the prognosis of the patients. Resistance to LE treatment is multifactorial, including biological, environmental, and psychological factors.[5,11,29] A lower affected/unaffected side ratio of the MS(S) of the wrist extensor or an extension/flexion ratio of the MS(S) is only one of the biological risk factors for refractoriness. A longitudinal study with multivariate analysis would identify the risks more quantitatively. The present study provides basic research results on wrist torque characteristics in males and females of patients with LE, which may be able to use in future longitudinal studies.

In conclusion, wrist extensor strength on the affected side of LE was decreased more remarkably in females than in males. Females tended to suppress muscle output to prevent pain from eccentric contraction of wrist extensors more than males. This morbid condition resulted in wrist extensor/flexor strength imbalance. Female patients with LE are prone to eccentric contraction of the wrist extensors in each gripping, which is a risk of refractoriness of LE.

AUTHOR CONTRIBUTIONS

Conceptualization: Ikeda K. *Data curation*: Ikeda K. *Formal analysis*: Ikeda K., Ikeda R. *Investigation*: Ikeda K. *Methodology*: Ikeda K., Yoshii Y. *Project administration*: Ikumi A. *Resource*: Ikeda K., Kohyama S. *Software*: Yoshii Y. *Supervision*: Yoshii Y., Yamazaki M. *Validation*: Ikeda R., Ogawa T. *Visualization*: Ikeda K. *Writing—original draft*: Ikeda K. *Writing—review and editing*, Yoshii Y. All authors have read and approved the final submitted manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ETHICS STATEMENT

Our Institutional Review Board approved the present study (Approval No. KC-H26, Date 27th. May. 2022).

ORCID

Kazuhiro Ikeda http://orcid.org/0000-0002-7112-6511
Yuichi Yoshii http://orcid.org/0000-0003-1447-4664

REFERENCES

1. Maffulli N, Longo UG, Denaro V. Novel approaches for the management of tendinopathy. *J Bone Joint Surg Am*. 2010;92:2604-2613.
2. Shiri R, Viikari-Juntura E, Varonen H, Heliovaara M. Prevalence and determinants of lateral and medial epicondylitis: a population study. *Am J Epidemiol*. 2006;164:1065-1074.
3. Coombes BK, Bisset L, Vicenzino B. A new integrative model of lateral epicondylalgia. *Br J Sports Med*. 2009;43:252-258.
4. Fan ZJ, Silverstein BA, Bao S, et al. Quantitative exposure-response relations between physical workload and prevalence of lateral epicondylitis in a working population. *Am J Ind Med*. 2009;52:479-490.
5. Shiri R, Viikari-Juntura E. Lateral and medial epicondylitis: role of occupational factors. *Best Pract Res Clin Rheumatol*. 2011;25:43-57.
6. Wolf JM, Mountcastle S, Burks R, Sturdivant RX, Owens BD. Epidemiology of lateral and medial epicondylitis in a military population. *Mil Med*. 2010;175:336-339.
7. Bhabra G, Wang A, Ebert JR, Edwards P, Zheng M, Zheng MH. Lateral elbow tendinopathy: development of a pathophysiology-based treatment algorithm. *Orthop J Sports Med*. 2016;4:232596711667063.
8. Ozone K, Kokubun T, Takahata K, et al. Structural and pathological changes in the enthesis are influenced by the muscle contraction type during exercise. *J Orthop Res*. 2022;40:2076-2088.
9. Ikeda K, Ogawa T, Ikumi A, et al. Individual evaluation of the common extensor tendon and lateral collateral ligament improves the severity diagnostic accuracy of magnetic resonance imaging for lateral epicondylitis. *Diagnostics*. 2022;12:1871.
10. Kraushaar BS, Nirschl RP. Current concepts review - tendinosis of the elbow (Tennis Elbow). clinical features and findings of histological, immunohistochemical, and electron microscopy studies*. *J Bone Joint Surg*. 1999;81:259-278.
11. Ikeda K, Yoshii Y, Kohyama S, et al. Sex differences in wrist torque and endurance-biomechanical factors associated with developing lateral epicondylitis of the humerus. *J Orthop Res*. 2023;41(8):1670-1677.
12. Hopkins C, Fu SC, Chua E, et al. Critical review on the socio-economic impact of tendinopathy. *Asia Pac J Sports Med Arthrosc Rehabil Technol*. 2016;4:9-20.
13. Chourasia AO, Buhr KA, Rabago DP, Kijowski R, Irwin CB, Sesto ME. Effect of lateral epicondylosis on grip force development. *J Hand Ther*. 2012;25:27-37.; Quiz 37.
14. Pienimäki TT, Siira PT, Vanharanta H. Chronic medial and lateral epicondylitis: a comparison of pain, disability, and function. *Arch Phys Med Rehabil*. 2002;83:317-321.
15. Unyó C, Chaler J, Rojas-Martínez M, et al. A cross-sectional study comparing strength profile of dorsal and palmar flexor muscles of the wrist in epicondylitis and healthy men. *Eur J Phys Rehabil Med*. 2013;49:507-515.
16. Lee JA, Sechachalam S. The effect of wrist position on grip endurance and grip strength. *J Hand Surg*. 2016;41:e367-e373.
17. Mogk JPM, Keir PJ. Crosstalk in surface electromyography of the proximal forearm during gripping tasks. *J Electromyography Kinesiol*. 2003;13:63-71.
18. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16:494-502.
19. Guler T, Yildirim P. Comparison of the efficacy of kinesiotaping and extracorporeal shock wave therapy in patients with newly diagnosed lateral epicondylitis: a prospective randomized trial. *Niger J Clin Pract*. 2020;23:704-710.
20. Smith-Forbes EV, Howell DM, Willoughby J, Pitts DG, Uhl TL. A retrospective cohort study of QuickDASH scores for three hand therapy acute upper limb conditions. *Mil Med*. 2018;183:522-529.
21. Pasdar Y, Darbandi M, Mirtaher E, Rezaeian S, Najafi F, Hamzeh B. Associations between muscle strength with different measures of obesity and lipid profiles in men and women: results from RaNCD cohort study. *Clin Nutr Res*. 2019;8:148-158.
22. Seo NJ, Armstrong TJ, Ashton-Miller JA, Chaffin DB. 2008. Wrist strength is dependent on simultaneous power grip intensity. *Ergonomics*. 51: 1594-1605.
23. Vanmechelen IM, Shortland AP, Noble JJ. Lower limb muscle volume estimation from maximum cross-sectional area and muscle length in

cerebral palsy and typically developing individuals. *Clin Biomech*. 2018;51:40-44.
24. Yoshii Y, Yuine H, Kazuki O, Tung W, Ishii T. Measurement of wrist flexion and extension torques in different forearm positions. *Biomed Eng Online*. 2015;14:115.
25. Zwerus EL, Somford MP, Maissan F, Heisen J, Eygendaal D, van den Bekerom MP. Physical examination of the elbow, what is the evidence? A systematic literature review. *Br J Sports Med*. 2018;52:1253-1260.
26. Ma KL, Wang HQ. Management of lateral epicondylitis: a narrative literature review. *Pain Res Manag*. 2020;2020:1-9.
27. Lee J, Kim T, Lim K. Effects of eccentric control exercise for wrist extensor and shoulder stabilization exercise on the pain and functions of tennis elbow. *J Phys Ther Sci*. 2018;30:590-594.
28. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. *Br J Anaesth*. 2013;111:52-58.
29. Haahr JP. Physical and psychosocial risk factors for lateral epicondylitis: a population based case-referent study. *Occup Environ Med*. 2003;60:322-329.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ikeda K, Yoshii Y, Kohyama S, et al. Pathophysiology of sex difference in refractoriness in lateral epicondylitis: biomechanical study of wrist torque. *J Orthop Res*. 2024;42:277-285. doi:10.1002/jor.25684

Clin Orthop Relat Res (2024) 482:526-533
DOI 10.1097/CORR.0000000000002829

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Clinical Research

Is a Novel Fluoroscopic Intraoperative Reference System Superior to Conventional Management for Distal Radius Fracture Reduction? A Propensity-matched Comparative Study

Sho Kohyama MD, PhD[1], Yuichi Yoshii MD, PhD[2], Akira Ikumi MD, PhD[3], Takeshi Ogawa MD, PhD[4], Tomoo Ishii MD, PhD[2]

Received: 22 February 2023 / Accepted: 26 July 2023 / Published online: 6 September 2023

Abstract

Background Preoperative planning is generally performed to simulate the process of reduction as well as to determine the size and placement of implants in patients undergoing distal radius fracture surgery. We previously described a three-dimensional (3D) digital preoperative planning system for the osteosynthesis of distal radius fractures, and we have developed a novel intraoperative referencing system that superimposes preoperative planning (such as plate position and length) onto fluoroscopic images during surgery; however, its efficacy has not been evaluated compared with conventional planning and surgery.

Questions/purposes Does use of a novel intraoperative referencing system result in (1) better Mayo wrist scores at 3 and 6 months after surgery and (2) less loss of reduction in terms of ulnar variance, palmar tilt, and radial inclination on plain radiographs taken 1 week, 3 months, and 6 months after surgery compared with conventional preoperative planning?

Methods Between April 2014 and October 2021, we treated 294 patients with open reduction and volar plate fixation for distal radius fractures. Of 294 patients, 65% (191) underwent surgery using either conventional preoperative planning or a novel intraoperative referencing system. The remaining patients were excluded because they were younger than 18 years, they had some missing medical records related to the clinical outcomes, or they had a previous history of upper extremity injuries. During that time, we generally treated fractures with volar plates when there was: more than 2 mm of stepoff/gap in the articular surface, a dorsal tilt more than 15°, radial inclination less than 15°, or radial shortening more than 5 mm. Generally, we used a flexor carpi radialis approach. In some patients who had dorsal fragments, we added a dorsal approach. At that time, we were developing the new

Each author certifies that there are no funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article related to the author or any immediate family members.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research®* editors and board members are on file with the publication and can be viewed on request.
Clinical Orthopaedics and Related Research® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA approval status, of any drug or device before clinical use.
Ethical approval for this study was obtained from our institutional review board, registered as T2019-0178 and T2022-0041.
The work was performed at the Department of Orthopaedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ami, Japan.

[1]Department of Orthopaedic Surgery, Kikkoman General Hospital, Chiba, Japan

[2]Department of Orthopaedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ami, Japan

[3]Department of Orthopaedic Surgery, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan

[4]Department of Orthopedic Surgery, National Hospital Organization, Mito Medical Center, Mito, Japan

Y. Yoshii ✉, Department of Orthopaedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ami, 300-0395 Ibaraki, Japan, Email: yyoshii@tokyo-med.ac.jp

Wolters Kluwer

intraoperative referencing system, so it was not used consistently. To arrive at a fair assessment, we opted to perform propensity matching based on age, gender, and AO fracture type. During the period in question, 36% (69 of 191) of patients with distal radius fractures who received a volar plate were treated using our novel intraoperative referencing system, and 64% (122 of 191) had surgery using conventional preoperative planning (control group). Of those, 91% (63 of 69) of patients who were treated with the intraoperative referencing system and 89% (108 of 122) of those in the control group were available for follow-up with all imaging and Mayo wrist scores at least 6 months after surgery. After propensity matching, that left us with two groups of 39 patients, who were well matched in terms of age and fracture type; these were the study groups. We also tried to match them according to gender, but there were fewer patients in the intraoperative referencing group, and the percentage of women for each group differed: 70% (44 of 63) in the intraoperative referencing group and 76% (82 of 108) in the control group. Also, there were fewer men with C3 fractures in the control group. Therefore, 64% (25 of 39) of patients in the intraoperative referencing group were women and 77% (30 of 39) of patients in the control group were women. In the intraoperative referencing group, our novel intraoperative referencing system was used in combination with the 3D digital preoperative planning system for preoperative planning. In the control group, preoperative planning was performed manually in a conventional manner using tracing paper and implant templates or using a digital template. We compared the groups in terms of operative duration, the radiation dose used in surgery, and Mayo wrist scores at 3 and 6 months after surgery. We also compared the groups in terms of loss of reduction on ulnar variance, palmar tilt, and radial inclination on plain radiographs taken 3 months and 6 months after surgery. We considered the plain radiograph taken 1 week after surgery as a baseline. Each item was compared between the image fusion and control groups using a Welch t-test.

Results Mayo wrist scores were no different between the intraoperative referencing system and the control group at 3 months (71 ± 7 versus 72 ± 11, mean difference 1 [95% CI -3.7 to 5.7]; p = 0.07) or at 6 months after surgery (76 ± 6 versus 79 ± 11, mean difference 3 [95% CI -3.5 to 7.9]; p = 0.12). There were no differences in surgical duration or radiation doses between the intraoperative referencing and control groups. We found only a small advantage in favor of the intraoperative referencing system in terms of loss of reduction on ulnar variance (3 months after surgery: 0.2 ± 0.4 mm versus 0.6 ± 0.7 mm, mean difference 0.4 mm [95% CI 0.15 to 0.69]; p = 0.003, 6 months after surgery: 0.4 ± 0.6 mm versus 0.8 ± 0.8 mm, mean difference 0.4 mm [95% CI 0.05 to 0.73]; p = 0.02 for the intraoperative referencing system and the control group, respectively). This difference in radial shortening was so small that it was not likely to have been clinically important.

Conclusion We found no clinically important advantages from the use of our novel intraoperative referencing system except a slight improvement in ulnar variance. Therefore, we recommend against its use in everyday practice at this time. However, future improvements may lead to better clinical outcomes, so we plan further investigations.

Level of Evidence Level III, therapeutic study.

Introduction

Distal radius fractures are common, representing 44% of all fractures [7]; they occur most often in elderly patients after low-energy trauma and in young patients after high-energy trauma. Peak incidence is in Caucasian women older than 65 years [4]. Distal radius fractures can be treated either nonsurgically or with reduction and fixation. Although many patients treated nonsurgically report little pain or disability, there is evidence that the risk of redisplacement is up to 64% in those treated conservatively [9]. Generally, those patients with a dorsal angulation greater than 15°, radial shortening larger than 3 mm, or an intra-articular stepoff greater than 2 mm are indicated for surgical treatment to have better results, both in terms of achieving an anatomic reduction and clinical outcomes [11].

Anatomic fracture reduction provides a better chance of full functional restoration, since the risk of developing posttraumatic osteoarthritis is 20-fold higher for intra-articular fractures [1], and a residual articular stepoff of more than 2 mm results in posttraumatic osteoarthritis in 100% of patients [9]. Preoperative planning is performed to simulate the process of reduction as well as the size and placement of implants. This commonly is done using picture archiving and communication systems (PACS) [3]. However, this does not allow for planning in three dimensions (3D). We previously developed and reported a 3D digital preoperative planning system for the osteosynthesis of distal radius fractures [17]. The initial version of this system did not allow direct comparison of preoperative planning and intraoperative fluoroscopic images, so we developed a novel intraoperative referencing system that superimposes preoperative planning onto fluoroscopic images during surgery [19]. However, whether this new system provides benefits that patients can perceive has not been evaluated.

We therefore asked: Does use of our novel intraoperative referencing system result in (1) better Mayo wrist scores at 3 and 6 months after surgery and (2) less loss of reduction in terms of ulnar variance, palmar tilt, and radial inclination on plain radiographs taken 1 week, 3 months, and 6 months after surgery compared with conventional preoperative planning?

Patients and Methods

Study Design and Setting

This was a single-center, retrospective, comparative study that used propensity matching. Our institution is a core hospital in a city with a population of 50,000 that supports a wide variety of both community and referral patients from surrounding rural area.

Patients and Propensity Score Matching

Between April 2014 and October 2021, we treated 294 patients with open reduction and volar plate fixation for distal radius fractures. Of 294 patients, 65% (191) underwent surgery using either conventional preoperative planning or our novel intraoperative referencing system. The remaining patients were excluded because they were younger than 18 years, they had some missing medical records related to clinical outcomes, or they had a previous history of upper extremity injuries. During that time, we generally treated fractures with volar plates when there was more than 2 mm of stepoff/gap in the articular surface, a dorsal tilt more than 15°, radial inclination less than 15°, or radial shortening more than 5 mm. Usually, we used a flexor carpi radialis approach. In some patients with dorsal fragments, we added a dorsal approach.

As we were developing our novel intraoperative referencing system, which at this time is called the image fusion system, we used it inconsistently. We started consistently using the new system in July 2018.

We divided patients into two groups: the intraoperative referencing group and the control group. We used our novel intraoperative referencing system and the 3D digital preoperative planning system for preoperative planning in the intraoperative referencing group. Currently, this program is not commercially available. This program is in development with a software company (LEXI Co Ltd). In this study, we collected the data using the program to obtain a proof of concept. Preoperative planning was performed manually in the conventional manner using tracing paper and implant templates or using a digital template in the control group.

To arrive at a fair assessment, we opted to perform propensity matching based on age, gender, and AO fracture type (assessed on CT scans). During the time period in question, 36% (69 of 191) of patients with distal radius fractures who received a volar plate were treated using our novel intraoperative referencing system and 64% (122 of 191) had surgery using conventional preoperative planning (control group). Of those, 91% (63 of 69) of patients in the intraoperative referencing system group and 89% (108 of 122) of those in the control group were available for follow-up with all imaging and Mayo wrist scores at least 6 months after surgery. We used a nearest neighbor matching model to create a comparable control group. The matching variables of age and fracture type were inserted for a propensity score matching algorithm with one-to-one optimum matching between the two groups. Propensity matching left us with two groups each with 39 patients. We also tried to match patients based on gender, but because there were fewer patients in the intraoperative referencing group, the percentage of women for each group differed: 70% (44 of 63) for the intraoperative referencing group and 76% (82 of 108) for the control group. Also, there were fewer men with C3 fractures in the control group, with 21% (13 of 63) in the intraoperative referencing group and 9% (10 of 108) in the control group (Table 1). After propensity matching, the intraoperative referencing group was comprised of 64% (25 of 39) women and the control group included 77% (30 of 39) women.

Descriptive Data

Our two study groups did not differ in important ways after we applied propensity score matching in terms of age, gender, and fracture type (Table 2). The study population included 25 women and 14 men in the intraoperative referencing group with a mean age of 65 ± 14 years. In the control group, there were 30 women and 9 men with a mean age of 66 ± 12 years.

Preoperative Planning

We planned to use Stellar P/D locking plates (HOYA Technosurgical Inc) for all patients. This plate system is a monoaxial volar locking plate system and has three sizes for width (small, medium, and large) and two sizes for length (short and long). Distal screw lengths varied from 10 to 24 mm with a diameter of 2.4 mm, and proximal screw lengths varied from 10 to 20 mm with a diameter of 2.6 mm.

Details on preoperative planning using our novel intraoperative referencing system were previously reported [19]. Using the system, contours of bones and implants from the preoperative plan were extracted, and they could be superimposed to intraoperative fluoroscopic images. So the system enables the surgeons to refer to the overlapping images to determine implant placement. We performed preoperative planning in the control group by handwriting on tracing paper and adjusting the scale of the AP and lateral views of radiographs on PACS or using a digital template [17]. Radiographs of the contralateral side were used. The size and placement of the locking plate were planned in both groups. The surgical goals were identical in

Table 1. Proportion of patients before propensity matching

Fracture type	Intraoperative referencing group (n = 63)	Control group (n = 108)
A2		
Women	2 (1)	1 (1)
Men	0 (0)	0 (0)
A3		
Women	19 (12)	15 (16)
Men	0 (0)	3 (3)
B3		
Women	2 (1)	3 (3)
Men	2 (1)	1 (1)
C1		
Women	2 (1)	0 (0)
Men	0 (0)	2 (2)
C2		
Women	22 (14)	31 (34)
Men	8 (5)	9 (10)
C3		
Women	24 (15)	26 (28)
Men	21 (13)	9 (10)

Data presented as % (n).

both groups, in terms of anatomical reduction, plate placement, and prevention of screw protrusion either dorsally or intraarticularly.

Surgical Technique and Aftercare

After preoperative planning, we performed osteosynthesis in all patients under general anesthesia. In the intraoperative referencing system group, the outline of the planned image was displayed on a fluoroscopy monitor, and its size was calibrated and overlapped with the fluoroscopic image (Fig. 1). Surgeons made an effort to reproduce the planned reduction and implant position in accordance with the outline of the intraoperative referencing system, and they referred to the handwritten preoperative plan in the control group. Both groups underwent the same postoperative rehabilitation program. We immobilized the wrist with a splint for 1 to 2 weeks, then initiated ROM exercises of the wrist and forearm under the instruction of hand therapists. ROM exercises of fingers were initiated soon after surgery.

Table 2. Characteristics of included patients

	Intraoperative referencing group (n = 39)	Control group (n = 39)	p value
Women	64 (25)	77 (30)	0.21
Age in years	65 ± 14	66 ± 12	0.44
AO classification			0.63
A3	15 (6)	13 (5)	
B3	3 (1)	3 (1)	
C2	49 (19)	56 (22)	
C3	33 (13)	28 (11)	

Data presented as % (n) or mean ± SD.

Data Sources and Measurement

We assessed the surgical duration, radiation dose needed in surgery, and Mayo wrist scores 3 and 6 months after surgery. The presence or absence of complications in each group was recorded at the time of the final hospital visit. All data were abstracted from the medical records by one surgeon who participated in the study (YY); this surgeon was not blinded to the study groups. In addition, the number of distal screws protruding into the joint or dorsal compartment was assessed by radiograph and CT. We took postoperative CTs 1 month after surgery. Bearing in mind the patients' radiation exposure, we took CT scans with patients standing with both hands in front to reduce the absorbed radiation dose. We also used the Advanced intelligent Clear-IQ Engine application (Canon Medical Systems) for low-dose imaging during CT scans, which resulted in a reduction of the CT dose index from the conventional 6.9 mGy to 1.4 mGy. We assessed the loss of reduction on ulnar variance, palmar tilt, and radial inclination on plain radiographs taken 1 week, 3 months, and 6 months after surgery. Two raters (YY and one individual who was not a study author) independently assessed images. The second rater was blinded to the groups. Interrater reliabilities were excellent [2], with intraclass correlation coefficient values of 0.97, 0.88, and 0.91 for ulnar variance, palmar tilt, and radial inclination, respectively. After evaluating the reliability of the two raters' measurements, the mean values for each parameter were used in further analyses.

Primary and Secondary Study Outcomes

Our primary study goal was to evaluate the usefulness of our novel intraoperative referencing system on clinical outcomes. To achieve this, we evaluated Mayo wrist scores 3 and 6 months after surgery and compared them with the results of patients who underwent surgery with conventional preoperative planning.

Our secondary study goal was to clarify the effect of our novel intraoperative referencing system on radiological parameters. To achieve this, we assessed the loss of reduction in terms of ulnar variance, palmar tilt, and radial inclination

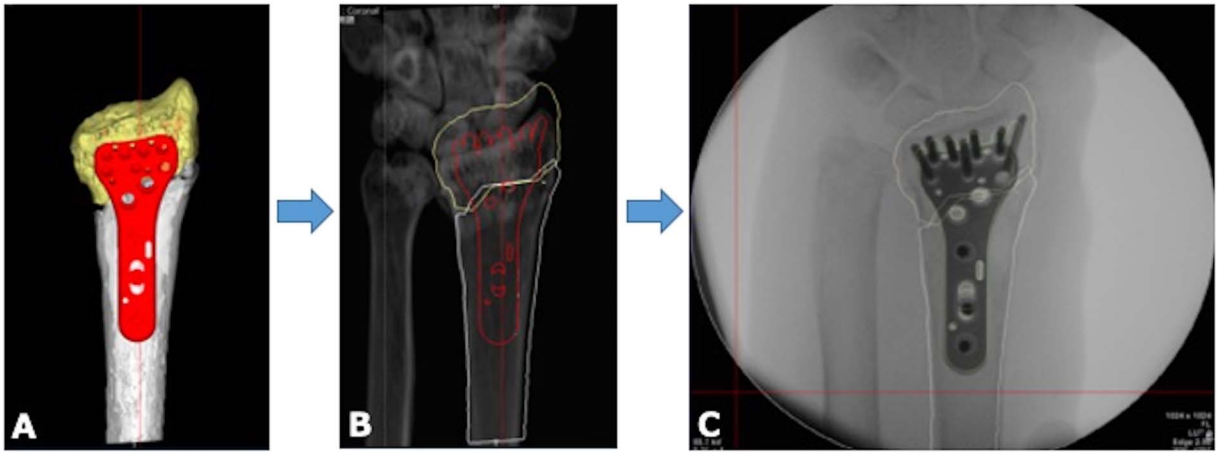


Fig. 1 The image fusion system is shown here. (**A**) Preoperative planning is done using a 3D digital preoperative planning system. (**B**) Contours were extracted from the preoperative 3D planning. (**C**) Extracted contours were superimposed onto intraoperative fluoroscopy images.

with the differences of the measurements between radiographs taken at 3 months and 6 months after surgery and radiographs taken at 1 week after surgery. We compared these findings with the results of patients who underwent surgery with conventional preoperative planning.

Ethical Approval

The study protocol was approved by the institutional review board of Tokyo Medical University (T2019-0178 and T2022-0041) and was performed in accordance with the ethical standards of the 1964 Declaration of Helsinki. Informed consent for surgery and for using the medical records was obtained from all patients included in the study at the time of admission. For access to the medical record in previous cases, the outline of research was published on the hospital's website to ensure that patients had the opportunity to decline the use of their data for the study.

Statistical Analysis

Results are expressed as the mean ± SD. A loss of reduction was defined as the difference between the values measured 1 week after surgery and those measured 3 and 6 months after surgery. Each item was compared between the intraoperative referencing and control groups using the Welch t-test. Screw protrusions were also compared between groups with the chi-square test. Differences were considered significant when p values were less than 0.05. All statistical analyses were performed using BellCurve for Excel version 2.12 (SSRI Co.) and SPSS 28 statistics (IBM Co.).

Results

Mayo Wrist Scores and Other Clinical Outcomes

Mayo wrist scores were no different between the intraoperative referencing system and the control group at 3 months (71 ± 7 versus 72 ± 11, mean difference 1 [95% CI -3.7 to 5.7]; p = 0.07) or at 6 months after surgery (76 ± 6 versus 79 ± 11, mean difference 3 [95% CI -3.5 to 7.9]; p = 0.12) (Fig. 2).

There were no differences in surgical duration or radiation dose. Surgical duration was 105 ± 35 and 94 ± 25 minutes in the intraoperative referencing system and control groups, respectively (p = 0.10). Radiation doses were 4.8 ± 2.3 mGy and 5.5 ± 3.1 mGy in the intraoperative referencing system and control groups, respectively (p = 0.33).

Three patients experienced nerve-related complications (one patient developed carpal tunnel syndrome and two patients had symptoms of median nerve palmar branch injury) in each group. Two patients in the intraoperative referencing system and three patients in the control group had wrist pain, and two patients in the intraoperative referencing system and one patient in the control group had symptoms related to wound scars. There were no between-group differences in terms of the number of distal screws protruding into the joint; there was one in the intraoperative referencing system group (total number of distal screws: 276) and five in the control group (total number of distal screws: 278; p = 0.10). Likewise, there were no between-group differences in the number of distal screws protruding into the dorsal extensor tendon compartment; there were 14 in the intraoperative referencing system group and 11 in the control group (p = 0.52).

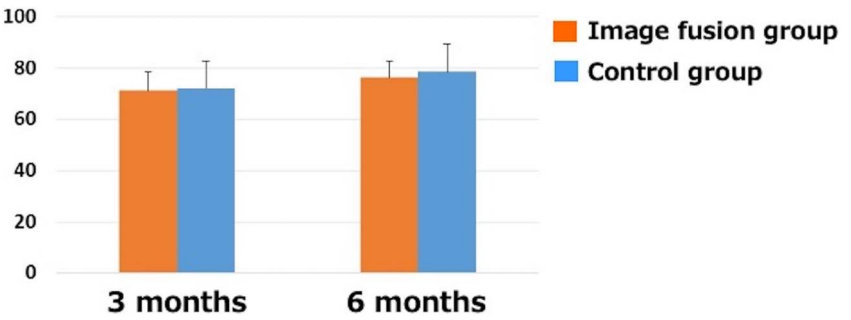


Fig. 2 Mayo wrist scores recorded at 3 and 6 months after surgery show no differences between the image fusion and control groups.

Loss of Reduction

We found no advantage to using the intraoperative referencing system in terms of loss of reduction in palmar tilt (3 months after surgery: 1.5° ± 1.5° versus 1.5° ± 1.3°, mean difference 0.0° [95% CI -0.6° to 0.6°]; p = 0.73, 6 months after surgery: 1.3° ± 1.0° versus 1.4° ± 1.4°, mean difference 0.1° [95% CI -0.4° to 0.6°]; p = 0.71 for the intraoperative referencing system and the control group, respectively) or radial inclination (3 months after surgery: 1.2° ± 0.8° versus 1.0° ± 0.8°, mean difference 0.2° [95% CI -0.2° to 0.5°]; p = 0.38, 6 months after surgery: 1.2° ± 1.2° versus 1.3° ± 1.0°, mean difference 0.1° [95% CI -0.4° to 0.6°]; p = 0.69 for the intraoperative referencing system and the control group, respectively). We found only a small advantage to the intraoperative referencing system in terms of loss of reduction on ulnar variance (3 months after surgery: 0.2 ± 0.4 mm versus 0.6 ± 0.7 mm, mean difference 0.4 mm [95% CI 0.15 to 0.69]; p = 0.003, 6 months after surgery: 0.4 ± 0.6 mm versus 0.8 ± 0.8 mm, mean difference 0.4 mm [95% CI 0.05 to 0.73]; p = 0.02 for the intraoperative referencing system and the control group, respectively) (Fig. 3). This difference in radial shortening was so small that it was not likely to have been clinically important.

Discussion

Volar locking plate fixation has been widely used for distal radius fractures since it was introduced in 2000 [12]. A plate must be placed properly to provide sufficient subchondral support [6, 8, 10]. Fragment-specific fixation is considered important when a fracture is comminuted [7]. We previously demonstrated that the reproducibility of the distance between the distal end of the volar locking plate and joint surface by our intraoperative referencing system was high [19]. We believe that precise 3D preoperative planning

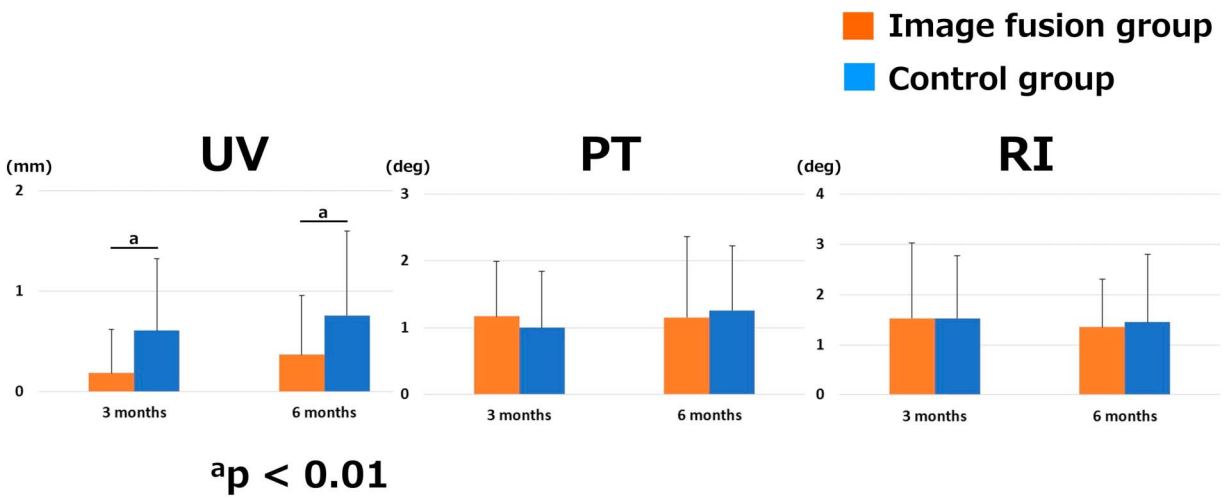


Fig. 3 In this graph, which shows the loss of reduction 3 and 6 months after surgery, the loss of ulnar variance was greater in the control group. However, no differences were observed in radial inclination or palmar tilt between the groups; UV = ulnar variance; RI = radial inclination; PT = palmar tilt.

and intraoperative reference to the planning may enable surgeons to achieve better anatomic fracture reduction, leading to a better chance of full functional restoration. We aimed to assess whether our novel intraoperative referencing system provides patient-perceived benefits. Our results suggest that the system prevented postoperative radial shortening; however, we observed no differences in clinical outcomes between the intraoperative referencing and control groups.

Limitations

There are several study limitations that must be addressed. Even though our novel intraoperative referencing system is easy to use, a specific application needs to be installed on the computer, which is an additional cost. Second, we did not subanalyze the patients by gender. The results drawn from the study may not be applied separately to men and women. In addition, we included patients from a wide range of ages, and the clinical results among these varying age groups may differ. We did not have enough patients to evaluate this, so we will work on it in the future. Third, since this is a retrospective study, we cannot completely exclude the possibility of selection bias. For example, the patients in the control group were enrolled earlier in the study period, whereas the patients in the intraoperative referencing group were recently treated cases. We could not avoid this selection bias because our intraoperative referencing system had not launched when we started collecting patients. In addition, we used the AO classification for the propensity matching. However, there can be a wide range of displacement and damage among the same AO fracture types. In the future, we need to take these into consideration. Fourth, we have used only a single type of volar locking plate system with the flexor carpi radialis approach. Our results may vary with different locking plate systems or surgical approaches. Fifth, our novel intraoperative referencing system is not yet commercially available. We are developing this program with a software company (LEXI Co Ltd). Some modifications may be necessary based on the results of this study. Finally, we believe that a follow-up period of 6 months would be sufficient to demonstrate a difference in treatment of distal radius fractures, but a longer follow-up period may affect clinical outcomes. We will continue to accumulate data for future investigations.

Mayo Wrist Scores and Other Clinical Outcomes

There was no clinical benefit to using the intraoperative referencing system as measured by the Mayo wrist score at either 3 or 6 months after surgery. There were no differences in operative duration, radiation dose, and the complication rate. Therefore, this system currently is unlikely to offer value to patients in terms of any endpoints they are likely to perceive.

Loss of Reduction

We found no difference in loss of reduction in terms of radial inclination or palmar tilt at 3 or 6 months, although there was slightly less loss of reduction in ulnar variance for the intraoperative referencing system group at both 3 and 6 months. This difference in ulnar variance was so small that it did not result in any differences in clinical outcomes.

Other Relevant Findings

Appropriate subchondral support is important not only for radial shortening but also for angular stability [6, 8, 10]. The results we obtained in this study were only a matter of several millimeters; therefore, we do not know whether overall stability was higher in the intraoperative referencing group. The present study included a wide variety of fractures. The results may have differed if we had evaluated only comminuted fractures, but this is speculative. Future studies might evaluate systems like ours in a study limited to comminuted fractures. Greater stability, if it can be achieved, might allow more aggressive postoperative rehabilitation.

Nerve-related complications, such as carpal tunnel syndrome after the osteosynthesis of distal radius fractures using volar locking plates, may have been associated with improper plate positioning [5, 15]. Carpal tunnel syndrome occurs in between 2% and 9% of patients in this setting [5, 16], and in this study, it occurred in 3% of both groups. Inappropriate plate positioning has also been reported to cause intra-articular screw protrusion, especially with monoaxial volar locking plate systems [13]. In the present study, there was no between-group difference in terms of the number of screws protruding into the joint, but we were underpowered on that endpoint. Future studies, perhaps multicenter collaborations, would be needed to determine whether there is any advantage to this system in terms of screw protrusion, and if there is, whether it is large enough to justify the costs of the system.

Our intraoperative referencing system can be used as an educational tool for young surgeons. It may help surgeons predict how each fragment should be reduced based on preoperative planning and perform fragment-specific fixation. For less-experienced surgeons, intraoperative referencing systems like this one may facilitate teaching and learning about fracture reduction and plate position. However, before it can be widely adopted for this purpose, its benefits would need to be validated in studies about surgical learning curves and accuracy of plate and screw position.

Conclusion

We found no clinically important advantages from the use of our novel intraoperative referencing system except a slight improvement in ulnar variance. No differences were observed in clinical outcomes, operative duration, or radiation doses between the intraoperative referencing system and control groups. We recommend against the use of this intraoperative referencing system in everyday practice in its current form. We now are modifying this system with 3D tracking and 3D reconstructions of the fluoroscopic images, and we plan to re-evaluate it after we have done so [14, 18].

Acknowledgment We thank Dr. Yusuke Eda for his cooperation in the study.

References

1. Andersson DD, Chubinskaya S, Guilak F, et al. Post-traumatic osteoarthritis: improved understanding and opportunities for elderly intervention. *J Orthop Res*. 2011;29:802-809.
2. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. *Psychol Rep*. 1966;19:3-11.
3. Choplin RH, Boehme JM 2nd, Maynard CD. Picture archiving and communication systems: an overview. *Radiographics*. 1992; 12:127-129.
4. Corsino BC, Reeves RA, Sieg RN. *Distal Radius Fractures.* StatsPearls Publishing; 2023.
5. Cooke ME, Gu A, Wessel LE, Koo A, Osei DA, Fufa DT. Incidence of carpal tunnel syndrome after distal radius fracture. *J Hand Surg Glob Online*. 2022;4:324-327.
6. Drobetz H, Bryant AL, Pokorny T, et al. Volar fixed-angle plating of distal radius extension fractures: influence of plate position on secondary loss of reduction-a biomechanical study in a cadaveric model. *J Hand Surg Am*. 2006;31:615-622.
7. Hozack BA, Tosti RJ. Fragment-specific fixation in distal radius fractures. *Curr Rev Musculoskelet Med*. 2019;12:190-197.
8. Inagaki K, Kawasaki K. Distal radius fractures–design of locking mechanism in plate system and recent surgical procedures. *J Orthop Sci.* 2016;21:258-262.
9. Leixnering M, Rosenauer R, Pezzel C, et al. Indications, surgical approach, reduction, and stabilization techniques of distal radius fractures. *Arch Orthop Trauma Surg*. 2020;140;611-621.
10. Lee SK, Chun YS, Shin HM, Kim SM, Choy WS. Double-tiered subchondral support fixation with optimal distal dorsal cortical distance using a variable-angle volar locking-plate system for distal radius fracture in the elderly. *Orthop Traumatol Surg Res.* 2018;104:883-891.
11. Lichtman DM, Bindra RR, Boyer MI, et al. Treatment of distal radius fractures. *J Am Acad Orthop Surg.* 2010;18:180-189.
12. Orbay JL. The treatment of unstable distal radius fractures with volar fixation. *Hand Surg.* 2000;5:103-112.
13. Piñal FD, Jupiter JB, Rozental T, et al. Distal radius fractures. *J Hand Surg Eur Vol*. 2021;47:12-23.
14. Shrestha P, Xie C, Shishido H, Yoshii Y, Kitahara I. 3D reconstruction of wrist bones from C-arm fluoroscopy using planar markers. *Diagnostics (Basel)*. 2023;13:330.
15. Siegerman D, Lutsky K, Fletcher D, et al. Complications in the management of distal radius fractures: how do we avoid them? *Curr Rev Musculoskelet Med.* 2019;12:204-212.
16. Soong M, Ring D. Ulnar nerve palsy associated with fracture of the distal radius. *J Orthop Trauma*. 2007;21:113-116.
17. Yoshii Y, Kusakabe T, Akita K, Tung WL, Ishii T. Reproducibility of three dimensional digital preoperative planning for the osteosynthesis of distal radius fractures. *J Orthop Res*. 2017;35:2646-2651.
18. Yoshii Y, Iwahashi Y, Sashida S, et al. An experimental study of a 3D bone position estimation system based on fluoroscopic images. *Diagnostics (Basel).* 2022;12:2237.
19. Yoshii Y, Totoki Y, Sashida S, Sakai S, Ishii T. Utility of an image fusion system for 3D preoperative planning and fluoroscopy in the osteosynthesis of distal radius fractures. *J Orthop Surg Res.* 2019;14:342.

JJCO Japanese Journal of Clinical Oncology

Japanese Journal of Clinical Oncology, 2024, 54(2), 153–159
https://doi.org/10.1093/jjco/hyad156
Advance Access Publication Date: 20 November 2023
Original Article

OXFORD

Original Article

Development and evaluation of a rapid one-step high sensitivity real-time quantitative PCR system for minor *BCR-ABL* (e1a2) test in Philadelphia-positive acute lymphoblastic leukemia (Ph+ ALL)

Michihiro Hidaka[1,*], Koiti Inokuchi[2], Nobuhiko Uoshima[3], Naoto Takahashi[4], Nao Yoshida[5], Shuichi Ota[6], Hirohisa Nakamae[7], Hiromi Iwasaki[8], Kenichiro Watanabe[9], Yoshiyuki Kosaka[10], Norio Komatsu[11], Kuniaki Meguro[12], Yuho Najima[13], Tetsuya Eto[14], Takeshi Kondo[15], Shinya Kimura[16], Chikashi Yoshida[17], Yuichi Ishikawa[18], Masashi Sawa[19], Tomoko Hata[20], Keizo Horibe[21], Hiroatsu Iida[22], Takeshi Shimomura[23], Nobuaki Dobashi[24], Isamu Sugiura[25], Junya Makiyama[26], Naoyuki Miyagawa[27], Asuka Sato[28], Ryuta Ito[28], Itaru Matsumura[29], Yuzuru Kanakura[30] and Tomoki Naoe[31]

[1]Department of Hematology, National Hospital Organization Kumamoto Medical Center, Kumamoto, Japan, [2]Department of Hematology, Nippon Medical School, Tokyo, Japan, [3]Department of Hematology, Japanese Red Cross Kyoto Daini Hospital, Kyoto, Japan, [4]Department of Hematology, Nephrology and Rheumatology, Akita University Graduate School of Medicine, Akita, Japan, [5]Department of Hematology and Oncology, Children's Medical Center, Japanese Red Cross Aichi Medical Center Nagoya First Hospital, Aichi, Japan, [6]Department of Hematology, Sapporo Hokuyu Hospital, Hokkaido, Japan, [7]Department of Hematology, Graduate School of Medicine, Osaka City University, Osaka, Japan, [8]Department of Hematology, National Hospital Organization Kyushu Medical Center, Fukuoka, Japan, [9]Department of Hematology and Oncology, Shizuoka Children's Hospital, Shizuoka, Japan, [10]Department of Hematology and Oncology, Center of Childhood Cancer, Hyogo Prefectural Kobe Children's Hospital, Hyogo, Japan, [11]Department of Hematology, Juntendo University School of Medicine, Tokyo, Japan, [12]Department of Hematology, National Hospital Organization Sendai Medical Center, Miyagi, Japan, [13]Hematology Division, Tokyo Metropolitan Cancer and Infectious Diseases Center, Komagome Hospital, Tokyo, Japan, [14]Department of Hematology, Hamanomachi Hospital, Fukuoka, Japan, [15]Blood Disorders Center, Aiiku Hospital, Hokkaido, Japan, [16]Department of Hematology, Respiratory Medicine and Oncology, Saga University, Saga, Japan, [17]Department of Hematology, National Hospital Organization Mito Medical Center, Ibaraki, Japan, [18]Department of Hematology and Oncology, Nagoya University Graduate School of Medicine, Aichi, Japan, [19]Department of Hematology and Oncology, Anjo Kosei Hospital, Aichi, Japan, [20]Department of Hematology, Nagasaki University Hospital, Nagasaki, Japan, [21]Department of Pediatrics, National Hospital Organization Nagoya Medical Center, Aichi, Japan, [22]Department of Hematology, National Hospital Organization Nagoya Medical Center, Aichi, Japan, [23]Department of Hematology, National Hospital Organization Hiroshimanishi Medical Center, Hiroshima, Japan, [24]Department of Clinical Oncology and Hematology, The Jikei University Daisan Hospital, Tokyo, Japan, [25]Department of Hematology and Oncology, Toyohashi Municipal Hospital, Aichi, Japan, [26]Department of Hematology/Oncology, Institute of Medical Science, The University of Tokyo Hospital, Tokyo, Japan, [27]Division of Hematology/Oncology, Kanagawa Children's Medical Center, Kanagawa, Japan, [28]Diagnostic Division, Otsuka Pharmaceutical Co. Ltd., Tokushima, Japan, [29]Department of Hematology and Rheumatology, Kindai University

Faculty of Medicine, Osaka, Japan, [30]Sumitomo Hospital, Osaka, Japan and [31]National Hospital Organization Nagoya Medical Center, Aichi, Japan

*For reprints and all correspondence: Michihiro Hidaka, National Hospital Organization Kumamoto Medical Center, 1-5 Ninomaru, Chuo-ku, Kumamoto 860-0008 Japan. E-mail: hidaka.michihiro.rp@mail.hosp.go.jp

Received 19 June 2023; Editorial Decision 23 October 2023; Accepted 25 October 2023

Abstract

Objective: Minimal residual disease assessment of *BCR-ABL* messenger ribonucleic acid levels is crucial in Philadelphia chromosome-positive acute lymphoblastic leukemia for prognosis and treatment planning. However, accurately quantifying minor *BCR-ABL* transcripts, which comprise 70% of Philadelphia chromosome-positive acute lymphoblastic leukemia cases, lacks a national-approved method.

Methods: We developed the "Otsuka" minor BCR-ABLmessenger ribonucleic acid assay kit with exceptional precision (0.00151%). Minor BCR-ABL messenger ribonucleic acid levels were analyzed in 175 adults, 36 children with acute lymphoblastic leukemia and 25 healthy individuals to evaluate the kit's performance.

Results: The "Otsuka" kit showed high concordance with a commonly used chimeric gene screening method, indicating reliable detection of positive cases. Quantitative results demonstrated a robust correlation with both a laboratory-developed test and a diagnostic research product. The "Otsuka" kit performs comparably or even surpass to conventional products, providing valuable insights into Philadelphia chromosome-positive acute lymphoblastic leukemia pathology.

Conclusions: The 'Otsuka" minor BCR-ABL messenger ribonucleic acid assay kit exhibits excellent performance in quantifying minor *BCR-ABL* transcripts in Philadelphia chromosome-positive acute lymphoblastic leukemia patients. Our results align well with established screening methods and show a strong correlation with laboratory-developed tests and diagnostic research products. The "Otsuka" kit holds great promise as a valuable tool for understanding Philadelphia chromosome-positive acute lymphoblastic leukemia pathology and guiding effective treatment strategies.

Key words: minor *BCR-ABL*, Philadelphia-positive acute lymphoblastic leukemia, minimal residual disease

Introduction

Acute lymphoblastic leukemia (ALL) is a hematopoietic tumor characterized by neoplastic changes in immature lymphoid cells and bone marrow (BM) infiltration. The reported annual incidence is ~1 in 100 000 adults and ~3 in 100 000 children (1). The Philadelphia chromosome is present in 30–40% of adult patients (2) and 3–5% of pediatric patients (3), making its confirmation crucial for diagnosis and treatment. Philadelphia chromosome-positive (Ph+) ALL is classified into two types of *Breakpoint cluster region-abelson* (*BCR-ABL*) fusion genes: minor *BCR-ABL* and major *BCR-ABL*. Minor *BCR-ABL* accounts for ~70% of adult patients (4) and ~90% of pediatric patients (5).

The standard treatment for Ph+ ALL patients now involves chemotherapy with tyrosine kinase inhibitors (TKIs), followed by allogeneic hematopoietic stem cell transplantation. The combination of chemotherapy with TKIs has shown significant improvement in survival rates and high rates of complete remission with negative minimal residual disease (MRD) (6–9). Negative MRD is the most significant prognostic factor in Ph+ ALL (10–13). Detecting MRD after 3 months of remission induction therapy indicates a high likelihood of relapse and reduced survival (11). Furthermore, it has been observed that patients with negative MRD before transplantation have lower post-transplant relapse rates (12). Negative MRD is now widely used as a marker for treatment strategies. However, standardized evaluation methods for MRD in minor *BCR-ABL*, which accounts for 70% of adults and 90% of children with Ph+ ALL, have not been established.

Until recently, Japan lacked approved diagnostic agents for determining minor *BCR-ABL* positivity and MRD levels, resulting in the absence of a standardized measurement method. Measurement of minor *BCR-ABL* levels relied on various diagnostic reagents, including laboratory-developed tests from registered laboratories. Typically, a sample would be considered MRD negative if quantitative polymerase chain reaction (PCR) did not detect BCR-ABL.

We have recently developed an *in vitro* diagnostic system called the "Otsuka" minor BCR-ABL messenger ribonucleic acid (mRNA) assay kit. This kit is designed to specifically target the e1a2 breakpoint that is predominantly found in Ph+ ALL. In addition, this kit utilizes *ABL* mRNA instead of *GAPDH* mRNA as the reference gene. This adaptation facilitates comparison of MRD data with results obtained in other countries. The assay procedure is remarkably simple, requiring only 2 hours and a small amount of BM and peripheral blood (PB). The reverse transcription quantitative PCR (RT-qPCR) reaction is performed in a single tube, facilitating comparison and evaluation of measurements across different laboratories.

In this study, we validated the kit using individual patients' samples and compared it with currently used diagnostics in research. The data obtained clearly demonstrated the assay kit's efficacy as a powerful tool for diagnosing and monitoring the pathological conditions of Ph+ ALL patients.

Patients and methods

Ethical conduct of the study

This study (Protocol No. ODK-1601-CLN-001) adhered to the ethical principles outlined in the Declaration of Helsinki, Ethical Guidelines for Medical and Health Research Involving Human Subjects, Japan's Act on Securing Quality, Efficacy and Safety of Products Including Pharmaceuticals and Medical Devices, and other relevant ministerial ordinances, notifications and the study protocol. No instances of ethical misconduct were reported throughout the study.

The National Hospital Organization (NHO) Central Review Board approved the implementation of this study for the 12 NHO-operated institutions. In addition, each hospital's institutional review board reviewed and approved the study based on the protocol, informed consent form and case report forms for the remaining 41 institutions.

Before participating in the study, the principal investigator or subinvestigator provided a detailed explanation to each patient using the informed consent form, allowing sufficient time for the patient to make an informed decision. Written voluntary informed consent was obtained from each patient, ensuring their comprehensive understanding of the information. For pediatric patients, written informed consent was obtained from either the patients themselves or their legally authorized representatives.

Patients and samples

This multicenter, open-label study was conducted from June 2017 to March 2019, involving 53 participating institutions. A total of 236 participants were enrolled, including 175 adults and 36 pediatric patients with suspected ALL who met the inclusion criteria and none of the exclusion criteria. In addition, 25 healthy individuals were included. Due to the limited number of children with Ph+ ALL, pediatric patients undergoing remission induction therapy or consolidation therapy were eligible for enrollment. Furthermore, existing samples stored at the institutions were utilized.

Among the adult patients with suspected ALL (n = 175), there were 33 patients with Ph+ ALL harboring minor *BCR-ABL*, 8 patients with Ph+ ALL harboring major *BCR-ABL*, 46 patients with Ph- ALL and 88 patients with other diseases. The pediatric patients with suspected ALL (n = 36) comprised 9 patients with Ph+ ALL harboring minor *BCR-ABL*, 22 patients with Ph- ALL and 5 patients with other diseases.

All adult patients with suspected ALL were enrolled before initiating treatment, and samples of BM and PB were collected. Patients diagnosed with Ph- ALL or non-ALL malignancies based on the test results were withdrawn from the study. Patients diagnosed with Ph+ ALL remained in the study, and BM and/or PB samples were collected at nine time points: Days 0, 8, 15, 22 and 29 after remission induction therapy, as well as weeks 0, 2, 4 and at the end of consolidation therapy.

Study design

RNA was extracted from BM and PB samples for analysis. The levels of minor *BCR-ABL* mRNA were measured using the minor BCR-ABL mRNA Assay Kit "Otsuka" (referred to as "Otsuka"), an *in vitro* diagnostic product by Otsuka Pharmaceutical Co., Japan. To evaluate the performance of Otsuka, "a screening test for leukemia-related chimeric genes", one of the tests used to diagnose Ph+ ALL, was used as a control for the pre-treatment samples (Sample No. 1). Using the screening test as a control, the positive concordance rate, negative concordance rate and overall concordance rate for Otsuka were calculated.

We then evaluated the correlation between Otsuka and two control reagents to validate the performance of Otsuka in monitoring treatment efficacy (Sample Nos. 1–9). The two control reagents were a laboratory-developed test from a registered laboratory (minor *BCR/ABL* assay (14), hereafter referred to as control reagent A) and the Ipsogen *BCR-ABL1* mbcr kit (15), hereafter referred to as control reagent B). In addition, minor *BCR-ABL* levels in patients with minor *BCR-ABL* and major *BCR-ABL* were compared, as well as other hematologic disorders. Results from BM samples were compared with PB samples, and the changes in minor *BCR-ABL* mRNA levels over the clinical course were graphically plotted. Statistical analyses were performed using SAS system version 9.4 (SAS Institute Japan).

Diagnostic kits

Otsuka is designed to detect minor *BCR-ABL* mRNA fusion transcript e1a2. The kit can perform reverse transcription reaction and quantitative PCR of minor *BCR-ABL* mRNA and *ABL* mRNA in a single reaction solution simultaneously and continuously. The components and detailed procedure of this kit are attached as Supplement. Briefly, in assays using the kit, 15 μL of PCR mix was added to 10 μL of RNA sample, and measurement was performed using an ABI™ 7500 Fast Dx system (Applied Biosystems, Foster City, CA, USA). Based on the obtained results, the ratio of minor *BCR-ABL* mRNA copies to *ABL* mRNA copies was calculated and normalized.

The limit of detection for this kit was determined according to Clinical and Laboratory Standards Institute guideline EP17-A2. Based on the results, for an *ABL* mRNA copy number of $\geq$10 000, the minimum detectable sensitivity was reported as a minor *BCR::ABL1* mRNA copy number of 13.58 copies/test and a minor *BCR::ABL1* mRNA to *ABL* mRNA ratio of 0.00151%.

Fusion gene transcripts

At initial diagnosis, fusion gene transcripts were screened for in patient BM samples using RT-qPCR, including minor *BCR-ABL* and major *BCR-ABL*. Once detected, the fusion gene transcripts were assayed, and minor *BCR-ABL* mRNA levels in BM or PB samples were determined.

Results

Concordance of test results between Otsuka and existing screening test for chimeric genes related to leukemia

Concordance between Otsuka and the existing screening test for leukemia-related chimeric genes was examined using samples collected from patients with suspected ALL (Sample No. 1), obtained before treatment initiation. Positive results in the control test, at or above 250 copies/μg RNA (minimum detectable sensitivity), were considered positive, while results below this threshold were considered negative. Similarly, positive results obtained with Otsuka at or above the minimum detectable sensitivity were classified as positive, and results below that threshold were considered negative.

Table 1A presents the results for 180 (153 adult and 27 pediatric) BM samples: the positive agreement rate was 97.6% (40/41), negative agreement rate was 95.0% (132/139) and overall agreement rate was 95.6% (172/180). The kappa coefficient, which measures diagnostic agreement, was 0.88, with a 95% confidence interval (CI) of 0.80–0.96.

Table 1. Consistency rate between Otsuka and control test

			Control test				
			Positive	Negative	Total		
(A) BM							
	Otsuka	Positive	40	7	47	Positive consistency rate:	97.6% (40/41)
		Negative	1	132	133	Negative consistency rate:	95.0% (132/139)
		Total	41	139	180	Total consistency rate:	95.6% (172/180)
(B) PB							
	Otsuka	Positive	38	8	46	Positive consistency rate:	97.4% (38/39)
		Negative	1	114	115	Negative consistency rate:	93.4% (114/122)
		Total	39	122	161	Total consistency rate:	94.4% (152/161)

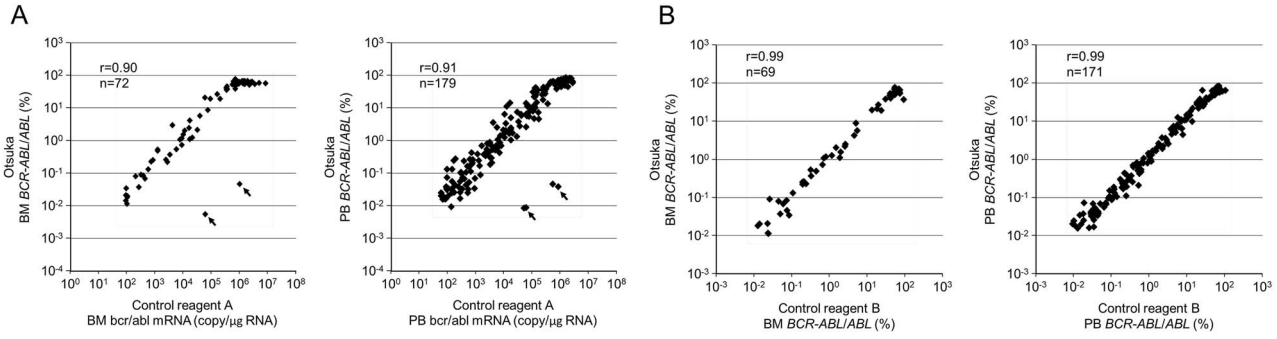


Figure 1. Correlation between Otsuka and control reagent A or control reagent B. (A) A strong correlation was observed between *BCR-ABL* levels measured by Otsuka and those by control reagent A for both BM and PB samples from Ph+ ALL patients. Samples with *BCR-ABL* subtype e1a3 are indicated by arrows. (B) A strong correlation was observed between *BCR-ABL* levels measured by Otsuka and those by control reagent B for both BM and PB samples from Ph+ ALL patients.

Table 1B displays the results for 161 adult PB samples (pediatric PB was not collected): the positive agreement rate was 97.4% (38/39), negative agreement rate was 93.4% (114/122) and overall agreement rate was 94.4% (152/161). The kappa coefficient was 0.86, with a 95% CI of 0.77–0.95.

To evaluate performance in monitoring therapeutic effects, the correlation between measurements obtained by Otsuka and control reagent A or B was assessed. All samples above the detection threshold were examined for correlation: 72 (63 adult and 9 pediatric) BM and 179 (155 adult and 24 pediatric) PB samples were tested with Otsuka and control reagent A, showing a strong correlation with a correlation coefficient of r = 0.90 for BM samples and r = 0.91 for PB samples (Fig. 1A).

Similarly, using Otsuka and control reagent B, 69 (61 adult and 8 pediatric) BM and 171 (149 adult and 22 pediatric) PB samples above the detection threshold were evaluated, demonstrating a strong correlation with a correlation coefficient of r = 0.99 for both BM and PB samples (Fig. 1B).

Stratified analyses by disease

Figure 2 presents elegant scatter plots showing the minor *BCR-ABL* mRNA/ABL mRNA ratio (%) for BM samples (n = 153) in Fig. 2A and PB samples (n = 193) in Fig. 2B. These samples were collected from patients with ALL, non-ALL malignancies and healthy individuals (PB only) before treatment initiation. Measurements below the limit of detection were represented as 0.0001%.

In Fig. 2A, all patients with Ph+ ALL and minor BCR-ABL (n = 29) tested positive (maximum: 78.2%, mean: 55.0%), whereas patients with Ph- ALL (n = 42) tested negative. Interestingly, although patients with Ph+ ALL and major *BCR-ABL* (n = 6) tested positive, the minor *BCR-ABL* mRNA/ABL mRNA ratios (%) were low (maximum: 0.0822%, mean: 0.0356%). Among patients with non-ALL malignancies (n = 76), positive cases included one with acute myeloid leukemia (AML) harboring minor *BCR-ABL* (7.64%) and one with chronic myeloid leukemia (CML) harboring minor *BCR-ABL* (37.9%). The remaining cases involved patients with CML

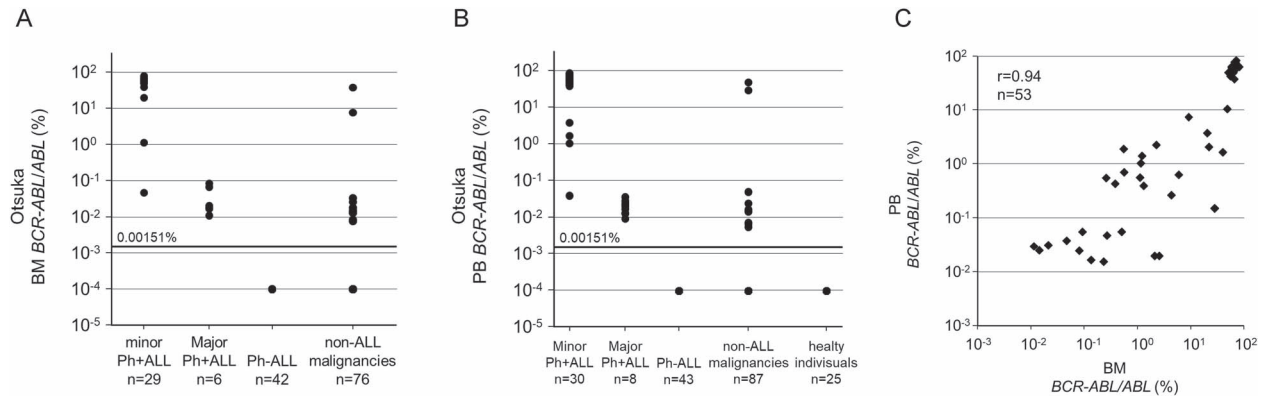


Figure 2. Distribution of minor *BCR-ABL/ABL* (%) measured by Otsuka (BM). (A) Comparison of *BCR-ABL* levels in BM samples from patients with Ph+ ALL harboring minor *BCR-ABL*, Ph+ ALL harboring major *BCR-ABL*, Ph− ALL and non-ALL malignancies. (B) Comparison of *BCR-ABL* levels in PB samples from patients with Ph+ ALL harboring minor *BCR-ABL*, Ph+ ALL harboring major *BCR-ABL*, Ph− ALL and non-ALL malignancies and healthy individuals. (C) Correlation of *BCR-ABL* levels between BM and PB samples.

and major *BCR-ABL*, where minor *BCR-ABL* mRNA/*ABL* mRNA ratios (%) were low (maximum: 0.0326%, mean: 0.0166%).

In Fig. 2B, all patients with Ph+ ALL and minor *BCR-ABL* (n = 30) tested positive (maximum: 83.8%, mean: 53.5%), whereas patients with Ph- ALL (n = 43) and healthy individuals (n = 25) tested negative. Similarly, although patients with Ph+ ALL and major *BCR-ABL* (n = 8) tested positive, the minor *BCR-ABL* mRNA/*ABL* mRNA ratios (%) were low (maximum: 0.0355%, mean: 0.0199%). Among patients with non-ALL malignancies (n = 87), positive cases included one with AML harboring minor *BCR-ABL* (28.4%) and one with CML harboring minor *BCR-ABL* (46.9%). The remaining cases involved patients with CML and major *BCR-ABL*, where minor *BCR-ABL* mRNA/*ABL* mRNA ratios (%) were low (maximum: 0.0490%, mean: 0.0173%).

Comparison of BM and PB

A total of 82 sample pairs were available for patients with Ph+ ALL harboring minor *BCR-ABL*, allowing measurement comparison in both BM and PB samples before treatment initiation and before/after consolidation therapy (Sample Nos. 1, 6 and 9, respectively). Out of these pairs, 53 exhibited BM and PB data exceeding the limit of detection. Correlation analysis between BM and PB data, specifically for BCR-ABL/ABL (%) (Fig. 2C), revealed a strong correlation with r = 0.94. However, three pairs displayed a ≥2-log difference, consistently showing higher values in BM compared with PB.

Eighteen of the 82 sample pairs had both BM and PB data below the detection limit. In the remaining 11 pairs, BM data were above the detection limit and PB data were below the detection limit. Notably, all these samples were collected after treatment initiation (Samples Nos. 6 and 9). This disparity is likely attributed to the treatment response, which eliminates leukemia cells from PB earlier than from the BM (16) (data not shown, Supplementary Fig. 1).

Time course of *BCR-ABL*/*ABL* in BM and PB during therapy in individual patients

We continued sample collection from 31 out of 33 patients with Ph+ ALL harboring minor *BCR-ABL*. Figure 3A presents representative data from the 31 sample pairs. Measurements below the minimum detectable sensitivity were plotted as 0.0001%.

The minor *BCR-ABL/ABL* (%) results obtained using Otsuka throughout the study period ranged from 0.00485 to 78.2% for BM and from 0.00761 to 83.8% for PB.

Comparison of the time course of minor *BCR-ABL* mRNA levels measured by Otsuka and existing research diagnostic kits for research use

By comparing the time course of minor *BCR-ABL* mRNA levels measured by Otsuka with those obtained using control reagents A and B, we analyzed 30 patients with Ph+ ALL harboring minor *BCR-ABL* who had data available for at least 2 consecutive time points. Figure 3B presents data from BM and PB samples of two patients as representative examples. The time course results obtained using the Otsuka kit closely aligned with those obtained using control reagents A and B.

Data measured by control reagent A were plotted as 1 copy/μg RNA if they fell below the minimum detectable sensitivity (50 copies/μg RNA). If data were below the minimum detectable sensitivity of Otsuka or were undetectable or not calculable using control reagent B, they were plotted as 0.0001%.

Discussion

In Japan, there was no approved diagnostic agent to identify minor *BCR-ABL* MRD levels, and no standard measurement method had been established. In this study, we evaluated the efficacy of Otsuka, a novel assay kit focusing on the e1a2 breakpoint of *BCR-ABL* mRNA, which is frequently found in Ph+ ALL. Notably, the kit uses *ABL* mRNA as the reference gene instead of *GAPDH* mRNA. As a result, this kit makes it possible to compare MRD data with those from other countries. The study revealed a significant concordance between Otsuka and the conventional control test in detecting minor *BCR-ABL* mRNA in both BM and PB samples across multiple centers. Quantitative analysis also established a strong correlation between Otsuka and the two control reagents. Otsuka specifically detected minor *BCR-ABL* mRNA and exhibited comparable time course data to the control reagents during the therapeutic process, indicating its favorable performance.

Although the results demonstrated a high agreement between Otsuka and the control assay, a few false positive cases were reported (7/115 in BM samples and 8/122 in PB samples). This was thought

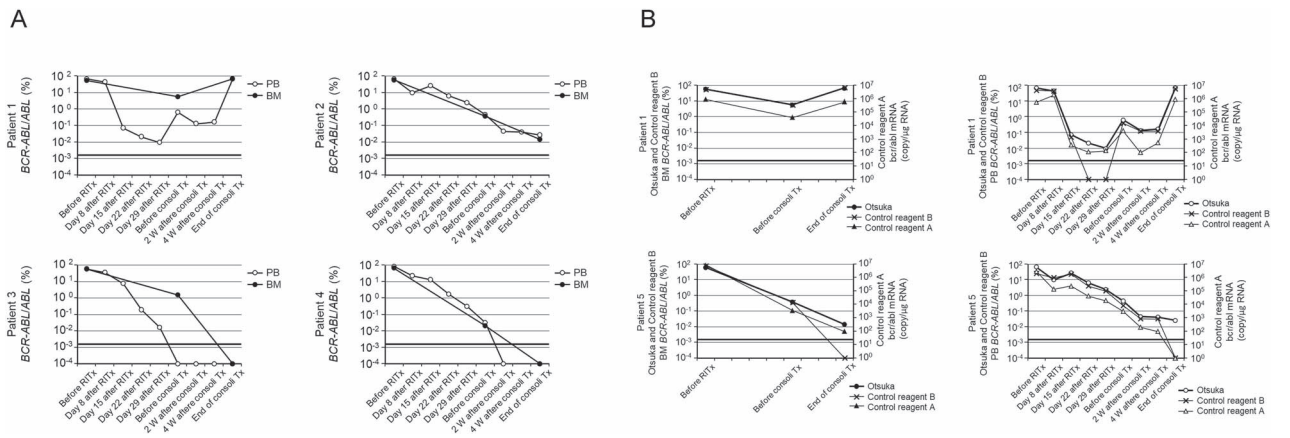


Figure 3. Time course of *BCR-ABL* levels in BM and PB samples over the therapeutic process. (A) Patient 1; The *BCR-ABL* level in PB was lowest 29 days after remission induction therapy but did not decrease to at or below the minimum detectable sensitivity. Later, this patient experienced hematological relapse. Patient 2; In both BM and PB samples, *BCR-ABL* levels declined with treatment, but did not decrease to at or below the minimum detectable sensitivity. Patient 3; *BCR-ABL* levels in PB samples were at or below the minimum detectable sensitivity before the start of consolidation therapy. After the completion of consolidation therapy, *BCR-ABL* levels in BM samples also decreased to at or below the minimum detectable sensitivity. Patient 4; *BCR-ABL* levels in PB samples decreased to at or below the minimum detectable sensitivity within 2 weeks of the start of consolidation therapy. After the completion of consolidation therapy, *BCR-ABL* levels in BM samples also decreased at or below the minimum detectable sensitivity. (B) Time course of *BCR-ABL* levels in Patient 1 (BM and PB) and Patient 5 (BM and PB), measured by Otsuka, control reagent A and control reagent B. Time courses of *BCR-ABL* levels were similar regardless of diagnostic reagents.

to be due to the difference in minimum detectable sensitivity between Otsuka and the control assay [13.58 copies/test (1 test is equivalent to 1 μgRNA) and 250 copies/μgRNA, respectively]. All of these cases occurred in patients with ALL or CML who had major *BCR-ABL*. This can be attributed to infrequent alternative splicing events that occur when major BCR-ABL mRNA is abundantly transcribed, a phenomenon often observed in Ph+ leukemia patients with major BCR-ABL (17). Stratified analyses by disease revealed positive test results for samples from patients with Ph+ ALL or CML harboring major BCR-ABL, but with a low minor BCR-ABL mRNA/ABL mRNA ratio, which can also be explained by alternative splicing (Fig. 2A and B).

In the correlation plots between Otsuka and control reagent A (Fig. 1A), a few samples displayed inconsistent results between the two methods (indicated by arrows in Fig. 1A), all originating from a single patient (two BM and four PB samples). Sanger sequencing analysis determined the presence of *BCR-ABL* e1a3, a rare variant subtype accounting for 1 to 2% of Ph+ ALL cases (18). In our study, the e1a3 subtype occurred in 1 out of 33 adult patients and 1 out of 3 pediatric patients. Otsuka was unable to detect e1a3 due to its primer design targeting the second exon of the *ABL1* gene (data not shown). Consequently, after excluding data from this patient, we analyzed the correlation between Otsuka and control reagent A in BM and PB samples, revealing a strong correlation with r = 0.98 (n = 61) for BM samples and r = 0.97 (n = 151) for PB samples.

The correlation analysis of *BCR-ABL/ABL* (%) between BM and PB (Fig. 2C) revealed that BM values exceeded PB data in some samples collected post-treatment initiation. However, Fig. 3A demonstrated good agreement between BM and PB data in samples collected pre-treatment (Sample No. 1), suggesting that either BM or PB samples can be used for measurements before treatment commencement. As mentioned previously, the 11 points where the detection results differed between PB and BM were all patterns where BM exceeded the detection sensitivity and PB was below the detection sensitivity. Conversely, there was no pattern where PB sensitivity exceeded the detection sensitivity and BM sensitivity was below the detection sensitivity. In other words, when PB exceeded the detection sensitivity or showed an increasing trend, the sensitivity of BM exceeded the detection sensitivity without exception. Therefore, the PB measurement of this assay is at least useful for predicting relapse or non-remission in BM. (19).

While major *BCR-ABL* testing has been standardized for clinical use, the evaluation of minor *BCR-ABL* fusion genes lacks such standardization, particularly in Japan where in-house or laboratory-developed reagents are employed. The Europe Against Cancer program has emphasized the importance of sensitive and accurate MRD quantification, leading to efforts to standardize qRT-PCR analysis. In the CML community, international standardization efforts have been ongoing since 2003. Similar endeavors have been reported for Ph+ ALL, where the e1a2 breakpoint, prevailing in this disease, was prioritized for qRT-PCR standardization, differing from the e13a2 and e14a2 breakpoints typical of CML (20).

To summarize, Otsuka represents the first clinical assay kit in Japan that accurately detects minor *BCR-ABL* mRNA levels, matching or surpassing the performance of conventionally used research reagents. Based on these findings, the assay kit obtained approval from the Ministry of Health, Labour, and Welfare in Japan in June 2021, with health insurance coverage commencing in November 2021. Regular monitoring of minor *BCR-ABL* mRNA in Ph+ ALL BM or PB samples at clinics will provide valuable insights to healthcare professionals for assessing individual patients' disease status.

Acknowledgements

We thank all participating institutions and physicians, not only the authors, but also the following investigators, for their support for the study.

The following investigators participated in this study: Kazutaka Sunami (National Hospital Organization Okayama Medical Center), Kenichi Ishizawa (Yamagata University Hospital), Yasunori Ueda (Kurashiki Central Hospital), Emiko Sakaida (Chiba University Hospital), Hirokazu Tanaka (Kindai University Faculty of Medicine), Junya Kuroda (Kyoto Prefectural University of Medicine), Yuki Yuza (Tokyo Metropolitan Children's Medical Center), Shingo Yano (The Jikei University School of Medicine), Mitsutoshi Kurosawa (National Hospital Organization Hokkaido Cancer Center), Koichi Miyamura (Japanese Red Cross Nagoya First Hospital), Jiro Fujita (Osaka

University Graduate School of Medicine), Hironori Ueno (National Hospital Organization Tokyo Medical Center), Tomohiko Kamimura (Harasanshin Hospital), Kaichi Nishiwaki (The Jikei University Kashiwa Hospital), Kazuma Oyashiki (Tokyo Medical University Hospital), Katsuyoshi Koh (Saitama Children's Medical Centre), Shinichiro Yoshida (National Hospital Organization Nagasaki Medical Center), Takuo Ito (National Hospital Organization Kure Medical Center), Yasushi Takamatsu (Fukuoka University Hospital), Masaya Okada (Hyogo College of Medicine Hospital), Tsuyako Iwai (Shikoku Medical Center for Children and Adults) and Kayo Taira (Hakata Clinic).

Supplementary material

Supplementary material is available at *Japanese Journal of Clinical Oncology* online.

Funding

This study was funded by Otsuka Pharmaceutical Co., Ltd., Tokyo, Japan.

Conflict of interest statement

Michihiro Hidaka, Masashi Sawa, Nobuaki Dobashi and Tomoki Naoe received research funding from Otsuka. Nobuaki Dobashi also received honoraria and payment for expert testimony from Otsuka. Naoto Takahashi received research funding and honoraria from Otsuka. Itaru Matsumura, Yuzuru Kanakura and Tomoki Naoe received personal fees from Otsuka during the conduct of the study. Asuka Sato and Ryuta Ito are employees of Otsuka Pharmaceutical Co., Ltd.

References

1. Redaelli A, Laskin L, Stephens M, Botteman F, Pashos L. A systematic literature review of the clinical and epidemiological burden of acute lymphoblastic leukaemia (ALL). *Eur J Cancer Care* 2005;14:53–62.
2. Liu-Dumlao T, Kantarjian H, Thomas A, O'Brien S, Ravandi F. Philadelphia-positive acute lymphoblastic leukemia: current treatment options. *Curr Oncol Rep* 2012;14:387–94.
3. Kaczmarska A, Śliwa P, Zawitkowska J, Lejman M. Genomic analyses of pediatric acute lymphoblastic leukemia Ph+ and Ph-like-recent progress in treatment. *Int J Mol Sci* 2021;22:6411–27.
4. Kantarjian M, Talpaz M, Dhingra K, et al. Significance of the P210 versus P190 molecular abnormalities in adults with Philadelphia chromosome-positive acute leukemia. *Blood* 1991;78:2411–8.
5. Bernt M, Hunger P. Current concepts in pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia. *Front Oncol* 2014;4:1–21.
6. Yanada M, Takeuchi J, Sugiura I, et al. High complete remission rate and promising outcome by combination of imatinib and chemotherapy for newly diagnosed BCR-ABL-positive acute lymphoblastic leukemia: a phase II study by the Japan Adult Leukemia Study Group. *J Clin Oncol* 2006;24:460–6.
7. Hatta Y, Mizuta S, Matsuo K, et al. Final analysis of the JALSG Ph+ALL202 study: tyrosine kinase inhibitor-combined chemotherapy for Ph+ALL. *Ann Hematol* 2018;97:1535–45.
8. Sugiura I, Doki N, Hata T, et al. Dasatinib-based 2-step induction for adults with Philadelphia chromosome–positive acute lymphoblastic leukemia. *Blood Adv* 2022;6:624–36.
9. Rousselot P, Coudé M, Gokbuget N, et al. Dasatinib and low-intensity chemotherapy in elderly patients with Philadelphia chromosome-positive ALL. *Blood* 2016;128:774–82.
10. Berry A, Zhou S, Higley H, et al. Association of minimal residual disease with clinical outcome in pediatric and adult acute lymphoblastic leukemia: a meta-analysis. *JAMA Oncol* 2017;3:e170580.
11. Ravandi F, Jorgensen L, Thomas A, et al. Detection of MRD may predict the outcome of patients with Philadelphia chromosome-positive ALL treated with tyrosine kinase inhibitors plus chemotherapy. *Blood* 2013;122:1214–21.
12. Nishiwaki S, Imai K, Mizuta S, et al. Impact of MRD and TKI on allogeneic hematopoietic cell transplantation for Ph+ALL: a study from the adult ALL WG of the JSHCT. *Bone Marrow Transplant* 2016;51:43–50.
13. Akahoshi Y, Arai Y, Nishiwaki S, et al. Minimal residual disease (MRD) positivity at allogeneic hematopoietic cell transplantation, not the quantity of MRD, is a risk factor for relapse of Philadelphia chromosome-positive acute lymphoblastic leukemia. *Int J Hematol* 2021;113:832–9.
14. Osumi K, Fukui T, Kiyoi H, et al. Rapid screening of leukemia fusion transcripts in acute leukemia by real-time PCR. *Leuk Lymphoma* 2002;43:2291–9.
15. *ipsogen BCR-ABL1 mbcr Controls Kit [Internet]*. Venlo, The Netherlands; [cited 2022 Aug 27]. ipsogen BCR-ABL1 mbcr Kit; Available from: https://www.qiagen.com/us/products/diagnostics-and-clinical-research/oncology/ipsogen-leukemia/ipsogen-bcr-abl1-mbcr-p190-kit.
16. Scheuring J, Pfeifer H, Wassmann B, et al. Early minimal residual disease (MRD) analysis during treatment of Philadelphia chromosome/Bcr-Abl–positive acute lymphoblastic leukemia with the Abl-tyrosine kinase inhibitor imatinib (STI571). *Blood* 2003;101:85–90.
17. Lee S, LeMaistre A, Kantarjian M, et al. Detection of two alternative BCR/ABL mRNA junctions and minimal residual disease in Philadelphia chromosome positive chronic myelogenous leukemia by polymerase chain reaction. *Blood* 1989;73:2165–70.
18. Burmeister T, Schwartz S, Taubald A, et al. Atypical BCR-ABL mRNA transcripts in adult acute lymphoblastic leukemia. *Haematologica* 2007;92:1699–702.
19. Muffly L, Sundaram V, Chen C, et al. Concordance of peripheral blood and bone marrow measurable residual disease in adult acute lymphoblastic leukemia. *Blood Adv* 2021;5:3147.
20. Pfeifer H, Cazzaniga G, van der Velden VHJ, et al. Standardisation and consensus guidelines for minimal residual disease assessment in Philadelphia-positive acute lymphoblastic leukemia (Ph + ALL) by real-time quantitative reverse transcriptase PCR of e1a2 BCR-ABL1. *Leukemia* 2019;33:1910–22.

Received: 5 July 2023 | Revised: 12 October 2023 | Accepted: 12 November 2023

DOI: 10.1002/cnr2.1938

ORIGINAL ARTICLE

Cancer Reports | WILEY

Female and preserved platelet count subgroups of myelodysplastic syndrome patients benefit from standard-dose azacitidine

Shinichi Ogawa[1] | Tatsuhiro Sakamoto[2] | Ryota Matsuoka[3] | Kantaro Ishitsuka[4] | Yasuko Ogino[1] | Ayano Sootome[2] | Kenichi Makishima[4] | Chikashi Yoshida[5] | Yufu Ito[6] | Seiichi Shimizu[6] | Takuya Suyama[7] | Atsushi Shinagawa[7] | Takayoshi Ito[1] | Naoshi Obara[2] | Manabu Kusakabe[2] | Mamiko Sakata-Yanagimoto[2] | Yasushi Miyazaki[8] | Yasuhito Nannya[9] | Shigeru Chiba[2]

[1]Division of Hematology, JA Toride General Medical Center, Toride, Ibaraki, Japan

[2]Department of Hematology, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan

[3]Department of Pathology, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan

[4]Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

[5]Division of Hematology, National Hospital Organization Mito Medical Center, Mito, Ibaraki, Japan

[6]Division of Hematology, Tsuchiura Kyoudou General Hospital, Tsuchiura, Ibaraki, Japan

[7]Division of Hematology, Hitachi General Hospital, Hitachi, Ibaraki, Japan

[8]Department of Hematology, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki, Japan

[9]Department of Hematology, Institute of Medical Science, University of Tokyo, Tokyo, Japan

Correspondence
Shinichi Ogawa, Division of Hematology, JA Toride General Medical Center, 302-0022, Hongo 2-1-1, Toride, Ibaraki, Japan.
Email: shin199901598ogawa@hotmail.co.jp

Funding information
Japan Agency for Medical Research and Development, Grant/Award Number: cm0106505h; Japan Society for the Promotion of Science, Grant/Award Number: 21K16261

Abstract

Background: Hypomethylating agents, including azacytidine (AZA), are standard therapeutics for patients with high-risk myelodysplastic syndromes (MDS), a group of myeloid neoplasms. However, treatment schedules are not unified in real-world practice; in addition to the standard 7-day (standard-dose) schedule, shortened (reduced-dose) schedules are also used.

Aims: The aim of this study was to discover the patient group(s) which show differential efficacy between standard-and reduced-dose AZA to MDS.

Methods and Results: The outcome of different AZA doses in a cohort of 151 MDS patients were retrospectively analyzed. Overall survival (OS) was not significantly different between standard- and reduced-dose AZA groups by multivariate analysis. However, an interaction was found between either the sex (female vs. male), the platelet counts (< 40 $\times$ 10^3/μl vs. ≥ 40 $\times$ 10^3/μl), or the karyotype risk (< poor vs. ≥ poor) and standard-dose AZA for longer OS. Subgroup analyses revealed better OS with standard- over reduced-dose AZA in female patients (HR, 0.27 [95% CI,

0.090-0.79]; $p = 0.011$), and those with platelet counts ≥ 40 × 10^3/μl (HR, 0.51 [95% CI, 0.26-0.99]; $p = 0.041$). The union of female and preserved platelet count subgroups also benefited from standard-dose AZA. With this as a test cohort, we next analyzed patients registered in the JALSG MDS212 study, for whom 7-day and 5-day AZA treatment strategies were prospectively compared, as a validation cohort ($N = 172$). That cohort showed the same tendency as the retrospective results.

Conclusion: We identified the union of female and preserved platelet count subgroups which benefited from standard-dose AZA, imparting crucial information to physicians planning treatment regimens in MDS patients.

KEYWORDS
azacitidine, dose, myelodysplastic syndrome, platelet counts, sex

1 | INTRODUCTION

Myelodysplastic syndromes (MDS) are a group of chronic myeloid neoplasms characterized by pancytopenia, dysplasia, and predisposition to acute myeloid leukemia (AML).[1,2] The mainstay of therapy includes hypomethylating agents (HMAs), including azacitidine (AZA), and hematopoietic stem cell transplantation (HSCT) if eligible.[3–6] In the standard protocol, AZA is given at 75 mg/m^2 per day for 7 consecutive days every 28 days based on a Phase III study demonstrating prolonged overall survival (OS).[5] High-risk MDS patients are also reported to respond to shortened schedules (reduced-doses) of AZA although OS benefit differences between standard- and reduced-doses are controversial.[7–11] Based on these reports, the 5-day protocol is often clinically used because of convenience and better tolerability.[12] If the standard protocol is superior to the reduced protocol, the patients who receive the reduced protocol may lengthen their OS by changing their administration protocol. On the other hand, if the reduced protocol is equal to the standard protocol, it needs to be considered whether the standard protocol is reconsidered to lighten the adverse effects on patients and reduce economic burden to both patients and the society. Therefore, it is important to exhaustively compare the efficacy and difference between the protocols. Because AZA is a backbone of new combinatorial therapies for MDS and AML with venetoclax, magrolimab, APR246, and so on,[13–17] detailed data between the standard and the reduced doses may influence clinical studies and resulting new therapeutic regimens. Although a Phase III clinical trial was conducted to prove the superiority of the 7-day over the 5-day protocol, it was never completed and statistically significant OS differences between the 7- and 5-day protocols were not proven.[18]

We had a community-based information that reduced-dose AZA is prescribed to a significant proportion of MDS patients in Ibaraki Prefecture in Japan. Thus, we conducted a multicenter retrospective study to disclose real-world dosing schedules and investigate any potential differences in OS between patients receiving AZA at standard- or reduced-doses. Furthermore, we intended to identify subcohorts in which AZA dose delineated OS. To define such subcohorts, interaction analyses between cumulative AZA dose and each clinical parameter were performed as screening before subgroup analyses were performed for selected parameters. Sex and platelet count were each related to AZA dose dependency.

After validation with a prospective cohort registered in the JALSG MDS212 study,[18] it was suggested that the standard- or near standard-dose of AZA, in comparison with the reduced-dose, improved OS in female patients and those with preserved platelet counts.

2 | METHODS

2.1 | Patients and inclusion criteria for clinical analyses

One-hundred and eighty-six patients were enrolled, all diagnosed with MDS according to either the FAB[19] or the WHO 2016 criteria,[1] and treated by AZA from March 2011 to May 2019 at 5 hospitals in Ibaraki Prefecture. Two patients with a history of HSCT before AZA administration and one with a shortage of clinical data were removed (Table S1).

To investigate the influence of differences in the dose of AZA on hematological improvement (HI) and OS, we further removed 32 patients who died sooner than day 112 after the commencement of AZA. Then, resulting 151 patients who survived for 112 (28 days × 4 courses) days or longer (survivor112) were determined as a main target of our analysis. This was because we planned to exclude the short survivors dying sooner than day 112 based on our understanding represented by the following reports. First, the median number of courses required for the initial response was three, and 90% of responses were seen by 6 courses in MDS.[20] Second, AZA should be continued for at least 4–6 courses to judge whether the patients respond to AZA or not in AML patients.[21] Therefore, we collected the cumulative AZA dose at day 112, as well as data on the total number of AZA treatment course and the mean period of AZA administration in each course (6 days or shorter, or longer than 6 days). This retrospective study was approved

by the institutional review board in each hospital. This retrospective study was based on the medical records. Obeying the approval of each institutional review board, we performed opt-out in each hospital instead of written informed consent.

2.2 | Definitions of hematological improvement, survival, and cumulative AZA dose

Hematological improvement (HI) to AZA was defined according to the revised IWG 2019 hematological response criteria.[22] OS was defined as the time from the day of the first administration of AZA to the day of death caused by any reasons. Living patients were censored at the last contact and those patients receiving stem cell transplantation were censored at the day of the stem cell infusion. Cumulative doses of AZA (mg/m^2) in the first 4 courses were calculated by dividing the sum of AZA given on or before day 112 by the body surface area at the first administration of AZA. If AZA was administered at 75 mg/m^2/day for 5 days and body surface area was unchanged, the cumulative AZA dose was considered to be 1500 mg/m^2. Based on this calculation, cumulative AZA doses equal to or less than 1500 mg/m^2 were defined as reduced-dose while over 1500 mg/m^2 was defined as the standard-dose.

2.3 | Statistics

Fisher's exact test was used for univariate analyses of binary variables for response to AZA while the Mann–Whitney U test was used for univariate analysis of continuous variables. Logistic regression modeling was used for multivariate analyses of binary variables for response to AZA. OS was evaluated using the Kaplan–Meier method. The log-rank test was used to compare the survival curves between the patient groups of interest. The Cox proportional hazard model was used to estimate hazard ratios (HR) and 95% confidence intervals (95% CI) of HR in univariate and multivariate analyses of OS. In multivariate analyses of hematological improvement rate and OS, age (<75 vs. ≥75), sex, karyotype risk defined by the revised international prognostic scoring system criteria (IPSS-R) (<poor vs. ≥poor),[23] bone marrow blast percentage (<10% vs. ≥10%), neutrophile counts (<800/μl vs. ≥800/μl), hemoglobin levels (<8 g/dL vs. ≥8 g/dL), platelet counts (<40 × 10^3/μl vs. ≥40 × 10^3/μl), and cumulative AZA doses (reduced-dose vs. standard-dose) were included as explanatory variables, irrespective of p values. Factors with p values <.05 were additionally included in explanatory variables. To obtain the propensity score (PS), the probability to receive the standard-dose was calculated using a logistic regression model in which explanatory variables were age, sex, bone marrow blast percentage, WHO 2016 diagnosis, karyotype-risk defined by IPSS-R, with or without transplantation, neutrophile counts (Neu; /μL), hemoglobin levels (Hb; g/dL), and platelet counts (Plt; × 10^4/μL) at the first administration of AZA. The PS matching was performed using 1:1 caliper matching (caliper 0.2). Statistical analyses were performed using EZR.[24]

3 | RESULTS

3.1 | Patient characteristics

Characteristics of the 183 patients are shown in Table S1. The median age was 72 years (range, 29–90) with a male/female ratio of 2.05. Myelodysplastic syndrome with excess blasts 1 (MDS-EB1) and MDS-EB2 were the most prevalent (62.8%), followed by MDS with multiple lineage dysplasia (MDS-MLD; 19.1%) and AML with myelodysplasia-related changes (AML-MRC; 10.4%), according to the WHO 2016 criteria. All AML-MRC cases corresponded to refractory anemia with excess blasts (RAEB) in transformation (RAEB-t) according to the FAB classification. Based on IPSS-R, 71.1% were judged to have high- or very high-risk prognosis. The median number of AZA courses was 6 (range, 1–61).

Of Survivors112, the standard- and reduced-doses were given to 91 and 60 patients, respectively (Table 1). Median cumulative AZA doses at day 112 were 2074 mg/m^2 (10–90 percentile, 1575–2100 mg/m^2) and 1232 mg/m^2 (853–1500 mg/m^2) in the standard- and reduced-dose groups, respectively (Figure S1). The median Hb concentrations at the first administration of AZA were significantly higher in the standard-dose group than the reduced-dose group (8.6 g/dL vs. 7.8 g/dL, $p = .04$, Table 1), which potentially influenced the choice of the AZA dose. All other factors, including age, sex, diagnosis, IPSS-R-risk, karyotype-risk, Neu and Plt at the first administration of AZA, and bone marrow blast percentage at diagnosis or within 3 months before AZA start, were not significantly different between the two groups.

3.2 | Hematological improvement

The hematological improvement (HI) rate in any parameter by AZA was 54.1% in Survivors112 (95% CI, 45.7%–62.4%) (Table 1). In univariate analyses, cumulative AZA dose and sex significantly affected the HI rate; these rates were greater with regard to standard-dose and male sex (Tables 1 and S2). All other factors, such as age (<75 or ≥75), bone marrow blast percentage (<10% or ≥10%), or IPSS-R-risk (< high or ≥high), karyotype-risk (< poor or ≥poor), Neu, Hb, and Plt, did not significantly affect the HI rate. In our multivariate analysis, sex and the cumulative AZA dose were again the significant parameter affecting the HI rate (Table S2). Response to AZA based on bone marrow evaluations could not be investigated because these data at the appropriate time points after AZA initiation were missing in a substantial number of patients.

3.3 | Survival in the entire Survivors112 cohort

Median survival time (MST) was 509 days (95% CI, 445–640 days), while 1-year OS was 72.7% (95% CI, 64.2%–79.6%) in the Survivors112 cohort (Table 1). MST and 1-year OS in the whole cohort (183 patients) were described in Table S1. In univariate analyses, high or very high IPSS-R (HR, 1.85 [95% CI, 1.10–3.12]; $p = .019$), poor or

TABLE 1 Patient characteristics divided by cumulative dose of AZA at day 112.

	All patients (%)	Cumultive dose of AZA at day 112		*p* value
		≤1500 mg/m^2 (%)	>1500 mg/m^2 (%)	
N	151	60	91	
Age, median [range]	72 [29, 90]	74 [29, 90]	72 [42, 86]	.47
Sex				1
Male	102 (67.5)	41 (68.3)	61 (67.0)	
Female	49 (32.5)	19 (31.7)	30 (33.0)	
WHO 2016 criteria				.30
MDS-SLD	4 (2.6)	1 (1.7)	3 (3.3)	
MDS-MLD	27 (17.9)	10 (16.7)	17 (18.7)	
MDS-EB1	50 (33.1)	25 (41.7)	25 (27.5)	
MDS-EB2	46 (30.4)	15 (25.0)	31 (34.1)	
AML-MRC	17 (11.3)	6 (10.0)	11 (12.1)	
MDS with isolated del(5q)	1 (0.7)	1 (1.7)	0 (0.0)	
MDS-RS	2 (1.3)	0 (0.0)	2 (2.2)	
CMML	1 (0.7)	0 (0.0)	1 (1.1)	
tMN	2 (1.3)	2 (3.3)	0 (0.0)	
MDS-U	1 (0.7)	0 (0.0)	1 (1.1)	
IPSS-R risk group				.30
Very low	2 (1.3)	0 (0.0)	2 (2.2)	
Low	16 (10.6)	7 (11.7)	9 (9.9)	
Intermediate	26 (17.2)	6 (10.0)	20 (22.0)	
High	50 (33.1)	22 (36.7)	28 (30.8)	
Very high	5 (37.1)	25 (41.7)	31 (34.1)	
NA	1 (0.7)	0 (0.0)	1 (1.1)	
IPSS-R karyotype group				.17
Very good	3 (2.0)	0 (0.0)	3 (3.3)	
Good	61 (40.4)	18 (30.0)	43 (47.3)	
Intermediate	33 (21.9)	15 (25.0)	18 (19.8)	
Poor	11 (7.3)	6 (10.0)	5 (5.5)	
Very poor	39 (25.8)	19 (31.7)	20 (22.0)	
NA	4 (2.6)	2 (3.3)	2 (2.2)	
Transplantation				1
No	137 (90.7)	55 (91.7)	82 (90.1)	
Yes	14 (9.3)	5 (8.3)	9 (9.9)	
Bone marrow blast %, median [range]	7.6 [0.0, 29.8]	7.0 [0.4, 26.5]	8.4 [0.0, 29.8]	.56
Neutrophile count (/μL), median [range]	888 [47, 22 243]	904 [110, 22 243]	880 [47, 19 757]	.76
Hemoglobin (g/dL), median [range]	8.1 [2.4, 12.9]	7.8 [4.2, 12.9]	8.6 [2.4, 12.9]	.040
Platelet count ($\times 10^3$/μL), median [range]	63 [5, 629]	64 [10, 629]	62 [5, 364]	.32
Hematological improvement rate, % (95% CI)	54.1 (45.7–62.4)	36.2 (24.0–49.9)	65.9 (55.0–75.7)	.00064
Median follow up time, days (95% CI)	427 (364–449)	349.5 (263–428)	445 (400–522)	.35
Median survival time, days (95% CI)	509 (445–640)	427 (321–584)	623 (482–850)	.010
OS at 1 year, % (95% CI)	72.7 (64.2–79.6)	59.1 (44.6–71.0)	82.0 (71.4–88.9)	
OS at 2 year, % (95% CI)	35.9 (26.8–45.0)	24.3 (13.3–37.2)	44.7 (31.9–56.7)	

Note: Patients were included whoose overall survival was 112 days or longer.

Abbreviations: 95% CI, 95% confidence interval; AML-MRC, acute myeloid leukemia with myelodysplasia-related changes; CMML, chronic myelomonocytic leukemia; IPSS-R, revised international prognostic scoring system; MDS-EB1, myelodysplastic syndrome with excess blasts 1; MDS-EB2, myelodysplastic syndrome with excess blasts 2; MDS-MLD, myelodysplastic syndrome with multilineage dysplasia; MDS-SLD, myelodysplastic syndrome with single lineage dysplasia; MDS-RS, myelodysplastic syndrome with ring sideroblasts; MDS-U, myelodysplastic syndrome, unclassifieable; OS, overall survival; NA, not available; tMN, therapy related myeloid neoplasms.

very poor karyotype-risk (HR, 3.29 [95% CI, 2.04–5.31]; $p = 2.5 \times 10^{-7}$), no HI by AZA (HR, 2.21 [95% CI, 1.43–3.42]; $p = 2.5 \times 10^{-4}$), Hb <8 g/dL (HR, 0.62 [95% CI, 0.40–0.95]; $p = .025$), and the reduced-dose (HR, 0.58 [95% CI, 0.38–0.88]; $p = .010$) were significant factors for poor prognosis (Table S3 and Figure S2). In the multivariate Cox proportional hazards model, poor or very poor karyotype-risk (HR, 3.14 [95% CI, 1.88–5.23]; $p = 1.2 \times 10^{-5}$) and no HI (HR, 1.89 [95% CI, 1.13–3.18]; $p = .016$) significantly shortened the OS. The cumulative AZA dose was not an independent significant prognostic factor (Table S3).

Consequently, poor or very poor karyotype risk and no HI from AZA were negative prognostic factors in both univariate and multivariate analyses for the entire Survivors112 cohort, similarly to previously verified reports.

TABLE 2 The *p* values of interaction with cumulative AZA dose in the propensity score-matched analysis.

Age (<75 vs. ≥75)	.91
Sex (female vs. male)	.040
Bone marrow blast percentage (<10% vs. ≥10%)	.51
Neutrophile count (<800/μL vs. ≥800/μL)	.83
Hemoglobin (<8 g/dL vs. ≥8 g/dL)	.98
Platelet counts (<40 × 10^3/μL vs. ≥40 × 10^3/μL)	.28
Karyotype-risk (<poor vs. ≥poor)	.25

3.4 | Survival in the subcohorts

Because the univariate analysis showed OS differences between the AZA doses, we hypothesized that the benefit of the standard-dose AZA would be clearer if confounding factors were excluded. To remove such confounding factors and delineate subcohorts in which the standard-dose AZA prolonged OS than the reduced-dose AZA, we selected 94 patients by propensity score matching from Survivors112 (Table S4, Figure S3) and performed interaction analyses between each clinical parameter and cumulative AZA dose. We picked up any interactions providing *p* values of interaction less than .30. Sex (female vs. male), platelet counts (<40 × 10^3/μl vs. ≥40 × 10^3/μl), and karyotype risk (<poor vs. ≥poor) matched the criteria and were selected as

TABLE 3 Subgroup analyses of overall survival in the propensity score-matched analysis.

	Cumulative dose of AZA at day 112	*N*	Median survival time (95% CI) (Day)	HR (95% CI)	*p* value
Female					.011
	≤1500 mg/m^2	13	321 (129–422)		
	>1500 mg/m^2	14	850 (479–1737)	0.27 (0.090–0.79)	
Male					.78
	≤1500 mg/m^2	34	484 (371–707)		
	>1500 mg/m^2	33	482 (438–812)	0.91 (0.48–1.72)	
Platelet count <40 × 10^3/μL					.84
	≤1500 mg/m^2	15	429 (164–NA)		
	>1500 mg/m^2	15	473.5 (192–850)	0.91 (0.36–2.30)	
Platelet count ≥40 × 10^3/μL					.041
	≤1500 mg/m^2	32	412 (222–594)		
	>1500 mg/m^2	32	682 (458–940)	0.51 (0.26–0.99)	
Karyotype-risk < poor					.66
	≤1500 mg/m^2	27	640 (422–1072)		
	>1500 mg/m^2	28	812 (482–1737)	0.85 (0.41–1.76)	
Karyotype-risk ≥ poor					.062
	≤1500 mg/m^2	20	222 (141–427)		
	>1500 mg/m^2	19	448 (248–NA)	0.48 (0.22–1.07)	
Female or Platelet counts ≥40 × 10^3/μL					.0078
	≤1500 mg/m^2	36	371 (222–505)		
	>1500 mg/m^2	36	682 (458–940)	0.43 (0.23–0.82)	
Male and Platelet counts <40 × 10^3/μL					.38
	≤1500 mg/m^2	11	640 (132–NA)		
	>1500 mg/m^2	11	443 (167–NA)	1.60 (0.55–4.64)	

Abbreviations: 95% CI, 95% confidence interval; HR, hazard ratio; NA, not available.

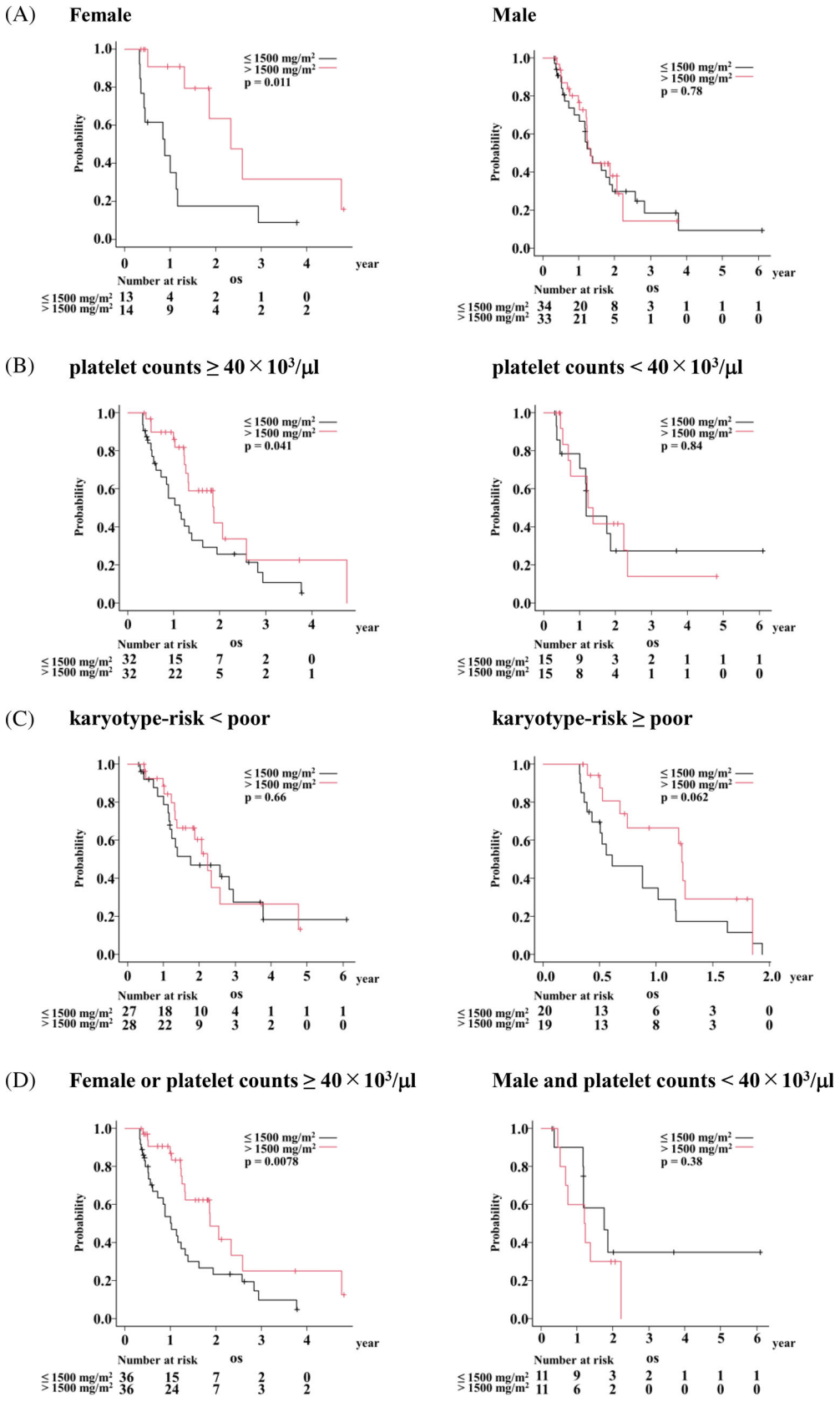


FIGURE 1 Subgroup analyses by sex, platelet counts, or karyotype in the propensity score-matched analysis. Female and male (A), platelet counts <40 × 10^3/μl and ≥40 × 10^3/μl (B), and karyotype-risk <poor and ≥poor (C), and the union of female and platelet counts ≥40 × 10^3/μl, and patients other than the union (D).

candidates for the subcohorts (Table 2). Then, univariate analyses were performed in each subcohort to investigate whether cumulative AZA dose influenced OS. The standard-dose significantly prolonged OS in the female (HR, 0.27 [95% CI, 0.090–0.79]; $p = .011$) and platelet counts ≥40 × 10^3/μl (HR, 0.51 [95% CI, 0.26–0.99]; $p = .041$) subcohorts (Table 3 and Figure 1). In the karyotype-risk ≥ poor

TABLE 4 Subgroup analyses of overall survival in 172 patients from the JALSG MDS212 cohort.

	Dose of AZA	*N*	Median survival time (95% CI) (Day)	HR (95% CI)	*p* value
Female					.26
	5 days	28	484 (438–732)		
	7 days	26	756 (420–1104)	0.68 (0.35–1.34)	
Male					.57
	5 days	61	497 (443–652)		
	7 days	57	537 (349–710)	0.88 (0.56–1.37)	
Karyotype-risk < poor					.49
	5 days	54	652 (483–848)		
	7 days	50	756 (569–1028)	0.84 (0.5–1.39)	
Karyotype-risk ≥ poor					.45
	5 days	35	438 (307–474)		
	7 days	33	378 (309–463)	0.81 (0.47–1.4)	
Platelet counts <40 × 10^3/μL					.85
	5 days	27	457 (288–695)		
	7 days	24	378 (238–455)	1.07 (0.55–2.07)	
Platelet counts ≥40 × 10^3/μL					.10
	5 days	62	509 (458–652)		
	7 days	59	710 (489–911)	0.69 (0.44–1.08)	
Female or platelet counts ≥40 × 10^3/μL					.067
	5 days	71	509 (458–652)		
	7 days	67	673 (463–868)	0.68 (0.44–1.03)	
Male and platelet counts <40 × 10^3/μL					.31
	5 days	18	457 (288–NA)		
	7 days	16	329 (169–455)	1.52 (0.68–3.4)	

Abbreviations: 95% CI, 95% confidence interval; HR, hazard ratio; NA, not available.

subcohort, there was a tendency that the standard-dose prolonged OS but without statistical significance (HR, 0.48 [95% CI, 0.22–1.07]; $p = .062$). The union of the female and platelet counts ≥40 × 10^3/μl subcohorts, in other words, the patients other than the male with platelet counts <40 × 10^3/μl, was delineated as the subcohort in which the standard-dose AZA improved OS than the reduced-dose AZA (HR, 0.43 [95% CI, 0.23–0.82]; $p = .0078$) (Table 3 and Figure 1).

To validate the results of our retrospective cohort, the OS of 172 patients who were prospectively treated with 7- and 5-day AZA (which correlates to the standard-and the reduced-doses, respectively, of the retrospective analysis) and survived 112 days or longer in the JALSG MDS212 study was compared.[18] In this entire JALSG day 112 survivor cohort, OS was not significantly different between 7- and 5-day AZA groups (HR, 0.80 [95% CI, 0.55–1.16]; $p = .24$). We then compared the OS between 7- and 5-day AZA arms of the following three subcohorts: female patients, those with platelet counts ≥40 × 10^3/μl, and those with karyotype-risk ≥ poor. In the female and the platelet counts ≥40 × 10^3/μl subcohorts, there was a tendency that OS was better in the 7-day AZA arm (HR, 0.68 [95% CI 0.35–1.34], $p = .26$; and 0.69 [95% CI, 0.44–1.08], $p = .10$; respectively). These OS differences were not observed in the male patients and those with platelet counts <40 × 10^3/μl. The difference between the two dose groups was marginal irrespective of karyotype-risk in the JALSG day 112 survivor cohort (Table 4 and Figure 2). In the patients other than the male with platelet counts <40 × 10^3/μl, there was a strong tendency that 7-day AZA prolonged OS (Table 4 and Figure 2; $p = .067$).

Taken together, our results suggested that the standard-dose AZA provided female patients and those with preserved platelet counts with better OS.

4 | DISCUSSION

By a retrospective analysis of 151 MDS patients who survived 112 days or longer after the starting of AZA, we found that OS in the female and the platelet counts ≥40 × 10^3/μl subcohorts significantly benefitted from the standard- rather than the reduced-dose. In the cohort of the Phase III clinical trial comparing the 7- and 5-day AZA scheduling,[18] the OS tended to be better with 7-day scheduling in the female patients and those with platelet counts ≥40 × 10^3/μl.

In real-world practice, either the standard- (7-day) or the reduced-dose (5-day) regimen is chosen without prognostic stratification. Our results showed that both regimens may have equal efficiency for OS prolongation in the male MDS patients with platelet

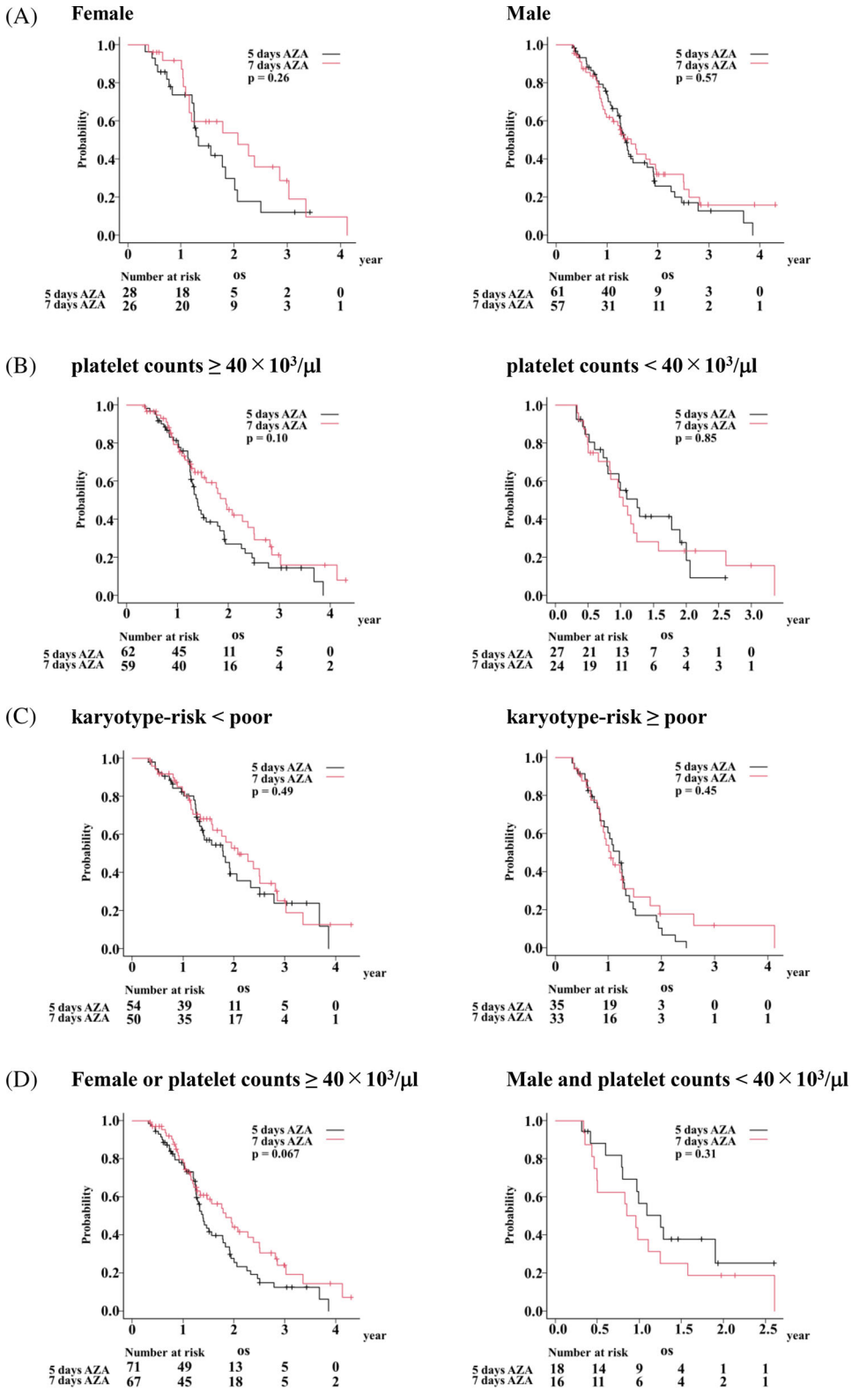


FIGURE 2 Subgroup analyses by sex, platelet counts, or karyotype in the validation cohort. Female and male (A), platelet counts <40 × 10^3/μl and ≥40 × 10^3/μl (B), karyotype-risk <poor and ≥poor (C), and the union of female and platelet counts ≥40 × 10^3/μl, and patients other than the union (D).

counts <40 × 10^3/μl. On the other hand, the standard-dose regimen reduced the risk of mortality by 57% and prolonged OS in the patients other than the male with platelet counts <40 × 10^3/μl (Table 3 and Figure 1). According to these results, AZA is recommended to be administered as the standard-dose, if the patients are not the male with platelet counts <40 × 10^3/μl. This provides novel and crucial information for the physicians treating high-risk MDS patients in choosing a treatment protocol, contributing to better quality of life and health economics.

A trend of shortened OS in patients receiving the reduced-dose, compared to the standard-dose, was previously described in a retrospective large cohort study.[25] Such a trend was also described in a

prospective study, albeit in a small number ($N = 22$) of patients based on a comparison with the AZA-001 study.[9] In a Phase III, JALSG MDS212 trial comparing 7- and 5-day AZA for RAEB and RAEB-t, although prematurely terminated because of poor recruitment, the 7-day protocol showed a statistically insignificant but visible trend of better OS (MST 538 [95% CI, 396–711] days) than the 5-day protocol (MST 477 [95% CI, 456–554] days).[18] Failure to demonstrate statistical OS differences in that study was attributed by the authors to insufficient statistical power, given that time to leukemia trasnformation was significantly longer with the 7-day protocol by multivariate analysis if only the centrally reviewed patients were investigated.[18]

In other previous reports comparing 5- and 7-day protocols, the conclusions have been controversial. Laribi et al. introduced relative dose intensity (RDI) of AZA (the relative dose intensity is the percentage of the dose received by the patient on the dose that theoretically should have been administered) to investigate how the dose of AZA influence the outcome of 93 high-risk MDS patients retrospectively. The OS and PFS were not different significantly with or without <80% RDI. The patients who responded to AZA were retrospectively divided into two groups; one group with <80% RDI and the other group without <80% RDI at the time when response was achieved. Dose reduction after the response was not considered. Then, they concluded that the group without <80% RDI showed significantly longer OS than the group with <80% RDI.[8] The time of response could be approximated by the day112,[20,21] and the doses administered to standard-dose group patients in our analysis resembled the group without <80% RDI in the report by Laribi, et al. Thus, the result of univariate analysis, but not multivariate analysis, in the current study may recapitulate the conclusion of Laribi et al. García-Delgado et al. retrospectively compared three regimens, 5 days (AZA5), 7 days including 2-day break (AZA 5-2-2), and 7 days (AZA7) in 200 patients with both high- and low-risk MDS patients, with majority with the latter. In this analysis, AZA 5-2-2 had significantly better response rate than AZA 5 or AZA 7, but the OS was not different among three regimens.[11] Fujimaki et al. compared the HI rate of their high-risk MDS patients on the 5-day AZA protocol with the HI rate of the high-risk MDS patients on the 7-day AZA protocol, and concluded that the HI rate was similar in both protocol.[10]

In the present study, 32 patients who died sooner than 112 days were removed from the landmark analysis for the Survivors112, but this removal could affect the conclusion. We explored potential differences between these short survivors and the Survivors112 by comparing the characteristics of patients. Although information on performance status and comorbidities was missing, we did not detect significant differences in other characteristics of patients between the two groups such as age, sex, diagnosis, risks on IPSS-R and karyotype, hematological parameters at the first administration of AZA, and bone marrow blast percentage at diagnosis or within 3 months before AZA start (Table S1).

In the Survivors112, the response to AZA by standard-dose was significantly better than reduced-dose in univariate and multivariate analyses. OS differences within Survivors112 were found between the two cumulative AZA dose groups in the univariate analysis. However, multivariate analysis did not show a significant difference. This could be due to biases or the dilution effect, which obscured findings in a specific subgroup of patients by other patients, according to the results of subgroup analyses in our study.

We reported that the standard-dose improved OS in female and the platelet counts ≥40 $\times$ 10^3/μl subcohorts within Survivors112. While we found an association between standard-dose and longer OS in specific subcohorts, there might be factors that influenced the results other than the standard-dose, given that the nature of retrospective analysis.

Preserved platelet count is an important component for prognosis prediction in IPSS-R[23] and, thus, should be selective for a subcohort with better OS. While this could be correlated to tolerability, clear explanations on why platelet, but not other blood cell lineages, affect the AZA dose preference remain elusive.

It was unexpected that the OS advantage imparted by the standard-dose was seen in female but not male patients. The activity of cytidine deaminase that inactivates AZA is known to be lower in females than males in a murine model.[26] Likewise, as cytarabine clearance from blood is known to be faster in males than females (as reported in a clinical trial),[27] AZA metabolism could differ between females and males and activity could persist in females if the dose is the same. It is, however, unclear whether and how this knowledge can explain the differences in observed outcomes.

There were several limitations in our study. First, the patients of our cohort came from 5 hospitals in Ibaraki prefecture, Japan, thus, there might be a geographical bias. Second, performance status and comorbidities were missing in our study. Third, the cumulative dose of AZA was surrogate index of 5-day dose or 7-day dose, but not equal to those. Fourth, the inclusion criteria between the current retrospective study and the JALSG MDS212 study were different. Of note, 26.8% of patients in the current retrospective cohort were MDS with low blasts, in contrast to the JALSG MDS212 cohort that included only RAEB and RAEB-t patients. Fifth, availability of the mutation profiles was incomplete in our retrospective cohort and not useful for the analysis.

Prospective study including large numbers of MDS patients is ideal to confirm results from our retrospective cohort. Given the premature termination of JALSG MDS212 prospective study due to poor recruitment, however, it might not be easy to perform a new prospective study comparing AZA doses in the future when new drugs would be equipped. In another way, better-designed retrospective analysis which takes the limitations of our cohort into account may be feasible. Simultaneously, it is warranted to elucidate the mechanism how the gender and platelet count influence OS under different AZA doses.

In conclusion, we identified by retrospective analysis that female and platelet counts ≥40 $\times$ 10^3/μl subcohorts of MDS (including oligoblastic AML), receive OS benefits from standard-dose rather than reduced-dose AZA. The same tendency was observed in the validation cohort independent of our cohort, although statistical significance was not seen.

AUTHOR CONTRIBUTIONS

Shinichi Ogawa: Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); resources

(lead); writing – original draft (lead). **Tatsuhiro Sakamoto:** Conceptualization (equal); data curation (equal); formal analysis (equal); methodology (equal); resources (equal). **Ryota Matsuoka:** Formal analysis (equal). **Kantaro Ishitsuka:** Formal analysis (supporting). **Yasuko Ogino:** Resources (equal). **Ayano Sootome:** Resources (equal). **Kenichi Makishima:** Formal analysis (supporting). **Chikashi Yoshida:** Resources (equal). **Yufu Ito:** Resources (equal). **Seiichi Shimizu:** Resources (equal). **Takuya Suyama:** Resources (equal). **Atsushi Shinagawa:** Resources (equal). **Takayoshi Ito:** Resources (equal). **Naoshi Obara:** Validation (equal). **Manabu Kusakabe:** Resources (equal). **Mamiko Sakata-Yanagimoto:** Validation (equal). **Yasushi Miyazaki:** Validation (equal). **Yasuhito Nannya:** Validation (lead). **Shigeru Chiba:** Supervision (lead); writing – review and editing (lead).

ACKNOWLEDGMENTS

We thank all patients who participated in the study. We also thank physicians participating in the JALSG MDS212 study. This work was supported in part by AMED under Grant Number cm0106505h (to SC). This work was supported in part by a grant from JSPS to TS (KAKENHI #21K16261). The authors would also like to thank Dr. Bryan J. Mathis of the University of Tsukuba Hospital International Medical Center for language revision.

CONFLICT OF INTEREST STATEMENT

S.Ogawa received honorarium from Novartis, Janssen Pharmaceutical, Nippon Shinyaku, Chugai Pharmaceutical, Daiichi Sankyo, and Bristol-Myers Squibb outside the submitted work. The husband of A. Sootome has stock of Otsuka Holdings. K.Makishima received honorarium from Ono Pharmaceutical outside the submitted work. C. Yoshida received honorarium from Novartis, Otsuka Pharmaceutical, AbbVie GK, Janssen Pharmaceutical, Nippon Shinyaku, Chugai Pharmaceutical., and honoraria and research funding from Bristol-Myers Squibb outside the submitted work. Y. Ito received honorarium from Nippon Shinyaku outside the work. T. Suyama received honorarium from Abbvie, Meiji-Seika, Sanofi, Janssen Pharmaceutical, Kyowa-Kirin and Nippon Shinyaku outside the work, and has stock of Takeda Pharmaceutical. T. Ito received honorarium from Janssen Pharmaceutical, Ono Pharmaceutical, Nippon Shinyaku, Chugai Pharmaceutical, Sanofi, Meiji-Seika, Bristol-Myers Squibb, Takeda Pharmaceutical, Novartis, and Otsuka Pharmaceutical outside the work. N.Obara received research funding from Kyowa-Kirin outside the submitted work. M.Kusakabe received honorarium from Janssen Pharmaceutical, Astellas, Takeda Pharmaceutical, Chugai Pharmaceutical, Astrazeneca, Meiji-Seika outside the submitted work. M.Sakata-Yanagimoto received research funding from Eisai, Bristol-Myers Squibb, Otsuka and Mundipharma outside the submitted work, and honorarium from Nippon Shinyaku, Kyowa-Kirin, Chugai Pharmaceutical, Astellas, Meiji-Seika, Zenyaku Kogyo, Mundipharma outside the submitted work. Y.Miyazaki received honorarium from Nippon-Shinyaku, Bristol-Myers Squibb, Novartis, Sumitomo Pharma, Kyowa-Kirin, Abbvie, Daiichi-Sankyo, Takeda Pharmaceutical, Janssen Pharmaceutical, Astellas, Pfizer, Chugai Pharmaceutical, SymBio, Otsuka Pharmaceutical, and Research funding from Sumitomo-Dainippon outside the submitted work. Y.Nannya received honorarium from Takeda Pharmaceutical, Pfizer, Chugai Pharmaceutical, Sumitomo Pharma, Astrazeneca, Kyowa-Kirin, Asahi Kasei Pharma, Nippon Shinyaku, Fuji Pharma, Janssen Pharmaceutical, Filgen, Otsuka Pharmaceutical and Bristol Myers Squibb outside the submitted work, and research funding from Daiichi Sankyo RD Novare outside the submitted work. Y. Nannya participates in advisory boards in Novartis and Otsuka Pharmaceutical outside the submitted work. S.Chiba reports research funding from Ono Pharmaceutical, Chugai Pharmaceutical, Eisai, Kyowa Kirin, Astellas, Bayer, and Thyas outside the submitted work, and honorarium from Nippon Shinyaku, Chugai Pharmaceutical, Eisai, Kyowa-Kirin, Astellas, Janssen Pharmaceutical, Asahi Kasei, Sanofi, Meiji-seika, Takeda Pharmaceutical, AMGEN, Ono Pharmaceutical, MSD, Astrazeneca, Abbvie, Daiichisankyo, Pfizer, and Otsuka Pharmaceutical outside the submitted work. Any of the authors of this manuscript is not a current Editor or Editorial Board Member of Cancer Science. All authors had full access to all the data in this study and had final responsibility for the decision to submit for publication.

DATA AVAILABILITY STATEMENT

The raw data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

Approval of research protocol by an institutional review board. This retrospective study was approved by the institutional review board in each hospital.

INFORMED CONSENT

N/A. This retrospective study was based on the medical records. Obeying the approval of each institutional review board, we performed opt-out in each hospital instead of written informed consent.

ORCID

Shinichi Ogawa https://orcid.org/0000-0002-7416-0926

REFERENCES

1. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, WHO Classification of Tumours*. Vol 2. 4th ed. International Agency for Research on Cancer; 2017.
2. Cazzola M. Myelodysplastic syndromes. *N Engl J Med*. 2020;383(14):1358-1374.
3. Kantarjian H, Issa JP, Rosenfeld CS, et al. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. *Cancer*. 2006;106(8):1794-1803.
4. Silverman LR, Demakos EP, Peterson BL, et al. Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. *J Clin Oncol*. 2002;20(10):2429-2440.
5. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol*. 2009;10(3):223-232.
6. Palacios-Berraquero ML, Alfonso-Piérola A. Current therapy of the patients with MDS: walking towards personalized therapy. *J Clin Med*. 2021;10(10).

7. Lyons RM, Cosgriff TM, Modi SS, et al. Hematologic response to three alternative dosing schedules of azacitidine in patients with myelodysplastic syndromes. *J Clin Oncol*. 2009;27(11):1850-1856.
8. Laribi K, Bolle D, Alani M, et al. Impact of the relative dose intensity on survival of patients with high-risk myelodysplastic syndromes treated with Azacitidine. *Cancer Med*. 2019;8(5):2188-2195.
9. Martin MG, Walgren RA, Procknow E, et al. A phase II study of 5-day intravenous azacitidine in patients with myelodysplastic syndromes. *Am J Hematol*. 2009;84(9):560-564.
10. Fujimaki K, Miyashita K, Kawasaki R, Tomita N. Efficacy and safety of a 5-day regimen of azacitidine for patients with high-risk myelodysplastic syndromes. *Eur J Haematol*. 2016;97(3):228-231.
11. García-Delgado R, de Miguel D, Bailén A, et al. Effectiveness and safety of different azacitidine dosage regimens in patients with myelodysplastic syndromes or acute myeloid leukemia. *Leuk Res*. 2014;38(7):744-750.
12. Grinblatt DL, Sekeres MA, Komrokji RS, Swern AS, Sullivan KA, Narang M. Patients with myelodysplastic syndromes treated with azacitidine in clinical practice: the AVIDA registry. *Leuk Lymphoma*. 2015;56(4):887-895.
13. DiNardo CD, Jonas BA, Pullarkat V, et al. Azacitidine and Venetoclax in previously untreated acute myeloid leukemia. *N Engl J Med*. 2020;383(7):617-629.
14. Sallman DA, Al Malki MM, Asch AS, et al. Magrolimab in combination with Azacitidine in patients with higher-risk myelodysplastic syndromes: final results of a phase Ib study. *J Clin Oncol*. 2023;41(15):2815-2826.
15. Cluzeau T, Sebert M, Rahmé R, et al. Eprenetapopt plus Azacitidine in TP53-mutated myelodysplastic syndromes and acute myeloid leukemia: a phase II study by the Groupe francophone des Myélodysplasies (GFM). *J Clin Oncol*. 2021;39(14):1575-1583.
16. Sallman DA, DeZern AE, Garcia-Manero G, et al. Eprenetapopt (APR-246) and Azacitidine in TP53-mutant myelodysplastic syndromes. *J Clin Oncol*. 2021;39(14):1584-1594.
17. Garcia-Manero G, Goldberg AD, Winer ES, et al. Eprenetapopt combined with venetoclax and azacitidine in TP53-mutated acute myeloid leukaemia: a phase 1, dose-finding and expansion study. *Lancet Haematol*. 2023;10(4):e272-e283.
18. Miyazaki Y, Kiguchi T, Sato S, et al. Prospective comparison of 5- and 7-day administration of azacitidine for myelodysplastic syndromes: a JALSG MDS212 trial. *Int J Hematol*. 2022;116:228-238.
19. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol*. 1982;51(2):189-199.
20. Silverman LR, McKenzie DR, Peterson BL, et al. Further analysis of trials with azacitidine in patients with myelodysplastic syndrome: studies 8421, 8921, and 9221 by the cancer and leukemia group B. *J Clin Oncol*. 2006;24(24):3895-3903.
21. Santini V, Ossenkoppele GJ. Hypomethylating agents in the treatment of acute myeloid leukemia: a guide to optimal use. *Crit Rev Oncol Hematol*. 2019;140:1-7.
22. Platzbecker U, Fenaux P, Adès L, et al. Proposals for revised IWG 2018 hematological response criteria in patients with MDS included in clinical trials. *Blood*. 2019;133(10):1020-1030.
23. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*. 2012;120(12):2454-2465.
24. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. *Bone Marrow Transplant*. 2013;48(3):452-458.
25. Itzykson R, Thépot S, Quesnel B, et al. Prognostic factors for response and overall survival in 282 patients with higher-risk myelodysplastic syndromes treated with azacitidine. *Blood*. 2011;117(2):403-411.
26. Mahfouz RZ, Jankowska A, Ebrahem Q, et al. Increased CDA expression/activity in males contributes to decreased cytidine analog half-life and likely contributes to worse outcomes with 5-azacytidine or decitabine therapy. *Clin Cancer Res*. 2013;19(4):938-948.
27. Fleming RA, Capizzi RL, Rosner GL, et al. Clinical pharmacology of cytarabine in patients with acute myeloid leukemia: a cancer and leukemia group B study. *Cancer Chemother Pharmacol*. 1995;36(5):425-430.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ogawa S, Sakamoto T, Matsuoka R, et al. Female and preserved platelet count subgroups of myelodysplastic syndrome patients benefit from standard-dose azacitidine. *Cancer Reports*. 2023;e1938. doi:10.1002/cnr2.1938

Cureus

Open Access Original Article

DOI: 10.7759/cureus.52025

The Effect of Axial Traction MRI on the Articular Cartilage Visibility in Thumb Carpometacarpal Arthritis

Review began 12/21/2023
Review ended 01/05/2024
Published 01/10/2024

Akira Ikumi [1, 2], Yuichi Yoshii [3], Sho Kohyama [4], Sho Iwabuchi [1], Takeo Mammoto [2], Takeshi Ogawa [5], Masashi Yamazaki [1]

1. Department of Orthopedic Surgery, Institute of Medicine, University of Tsukuba, Ibaraki, JPN 2. Department of Orthopedic Surgery and Sports Medicine, Tsukuba University Hospital Mito Kyodo General Hospital, Ibaraki, JPN 3. Department of Orthopedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ibaraki, JPN 4. Department of Orthopedic Surgery, Kikkoman General Hospital, Chiba, JPN 5. Department of Orthopedic Surgery, National Hospital Organization, Mito Medical Center, Ibaraki, JPN

Corresponding author: Akira Ikumi, ikumi@tsukuba-seikei.jp

Abstract

Objectives: Thumb carpometacarpal arthritis has a high incidence. However, the degree of damage to the cartilage has not been accurately assessed. The purpose of this study was to examine the effects of axial traction of the thumb carpometacarpal joint during magnetic resonance imaging (MRI) on the visibility of articular cartilage in patients with thumb carpometacarpal arthritis and to evaluate the articular cartilage defect using MRI findings.

Materials and methods: Forty-four patients with thumb carpometacarpal arthritis (14 males, 30 females) and a mean age of 67.3±8.6 years were classified according to Eaton Stages 1, 2, 3, and 4 in 2, 14, 24, and 4 patients, respectively. Axial traction MRI was performed with and without traction (3 kg) using 3-Tesla MRI (Siemens Magnetom Skyra) with a 3D T2* multiecho data imaging combination. The effectiveness of traction was verified using the joint space width before and after traction at five points (central, volar, dorsal, radial, and ulnar margins) and the original articular cartilage outline visibility classification (poor, intermediate, complete). The rate of remaining cartilage on each joint surface was also evaluated. Statistical significance was set at $p<0.05$ in this study.

Results: Joint space width increased significantly at all points with traction ($P<0.01$). The grade of articular cartilage outline visibility significantly improved from seven intermediate and 37 poor cases to 15 complete, 23 intermediate, and six poor cases ($P<0.01$). Significantly more articular cartilage remained in Stages 1-2 compared with Stages 3-4 arthritis of both articular surfaces ($P<0.01$ in first metacarpal, $P=0.01$ in trapezium).

Conclusion: Axial traction of the thumb increased the joint space width and improved articular cartilage visibility in the thumb carpometacarpal joint. Our results suggested that axial traction MRI can be used for noninvasive evaluation of articular cartilage defects in patients with thumb carpometacarpal arthritis and aid in selecting the optimal surgical procedure.

Categories: Orthopedics
Keywords: cartilage defect, articular cartilage, magnetic resonance imaging, axial traction, thumb carpometacarpal arthritis

Introduction

The thumb carpometacarpal joint is a saddle joint that connects the first metacarpal (MC1) and the trapezium and can be moved in multiple directions during daily activities such as pinching or grasping because of its anatomical characteristics [1,2]. The incidence of thumb carpometacarpal arthritis is high, occurring in >15% of adults aged >30 years and one-third of postmenopausal women, despite it being a non-weight-bearing joint [3-6]. Thumb carpometacarpal arthritis is usually diagnosed based on patient history, physical examination, and radiographs. The Eaton classification of thumb carpometacarpal arthritis is widely used to determine the staging and severity of this type of arthritis [7,8]. However, the degree of damage to the articular cartilage has not been accurately assessed, because the Eaton classification is based only on radiographs. One study reported that the intra- and inter-examiner reliabilities of this classification are low [9], whereas other studies have reported a poor correlation between clinical symptoms and intraoperative articular cartilage findings [10,11].

Magnetic resonance imaging (MRI) is widely used to evaluate articular cartilage damage. However, accurate articular cartilage evaluation of the thumb carpometacarpal joint is challenging because of its anatomical complexity and relatively small size compared to those of large joints such as the hip and knee. Although some reports have evaluated the articular cartilage of the thumb carpometacarpal joint using MRI [12-14], an

accurate method of evaluation has not yet been established because of the contact of the articular cartilage between the MC1 and the trapezium, as well as the underestimation of the degree of cartilage damage compared with other pathological findings [15].

To enhance the visibility of the articular cartilage, we performed an MRI while applying axial traction to the thumb carpometacarpal joint in healthy volunteers [16]. The joint space width of the thumb carpometacarpal joint was significantly increased in accordance with the traction weight, and the articular cartilage visibility of the thumb carpometacarpal joint was significantly improved by axial traction. This method of applying axial traction to improve the visualization of articular cartilage has also been previously used to observe other joints, such as the elbow and knee [17,18].

The aim of this study was to examine the effects of axial traction during MRI of the thumb carpometacarpal joint on the visibility of articular cartilage and to evaluate the remaining articular cartilage in patients with thumb carpometacarpal arthritis.

Materials And Methods

Study population

This study was approved by the institutional review board of the Tsukuba University Hospital Mito Clinical Education and Training Center, Ibaraki, Japan (No. R04-01), and was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

We enrolled 44 patients who visited our facility between April 2021 and October 2022 and were diagnosed with thumb carpometacarpal arthritis. Written informed consent was obtained from each patient after a thorough explanation of the objectives, methods, and expected complications.

Image acquisition

We used a 3-Tesla (3T) whole-body MRI system (Magnetom Skyra, Siemens Healthneers AG©, Munich, Germany) with a 4-channel 3T special purpose coil (Siemens Healthneers AG©, Munich, Germany). For the MRI sequence, a three-dimensional T2* multiecho data imaging combination (MEDIC) scan was used with the following parameters: slice thickness, 0.2 mm; slice gap, 0.15 mm; field of vision, 130 × 130 × 78 mm; matrix, 384 × 292; time to repeat, 20 ms; echo time, 11.0 ms; and flip angle, 25°. The required time, according to the protocol, was 5 min and 43 sec for each image. The patients were asked to lay supine on a table with their arms outstretched and their forearms pronated at the side of the body. The wrist and thumb were fixed with custom-made splints to standardize the limb position during the MRI examination. The hand under observation was centered parallel to the long axis of the gantry (Figure *1a*).

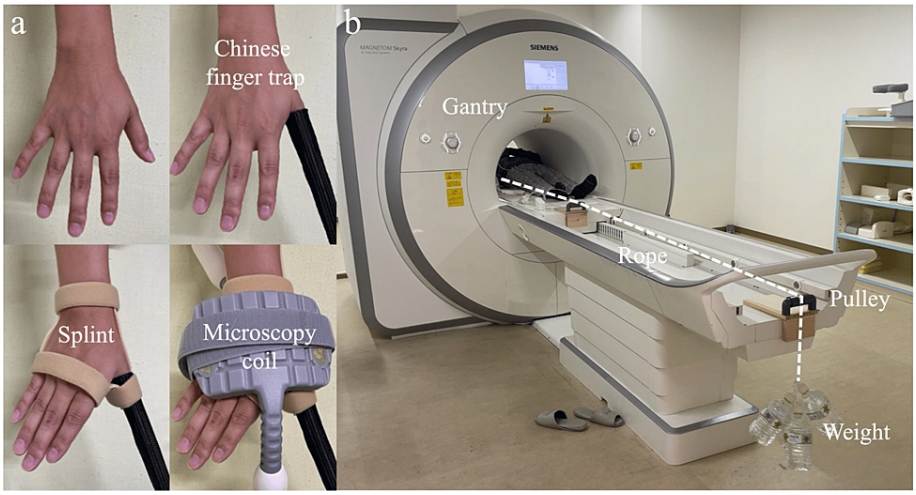


FIGURE 1: Application of axial traction during thumb carpometacarpal joint MRI

(a) The thumb is enclosed within a Chinese finger trap. Then, the wrist and thumb are fixed using a custom-made splint. Finally, the microscopy coil is placed around the wrist. (b) The Chinese finger trap is connected to the traction weight using a nonelastic rope routed through the pulley system.

Application of axial traction during MRI

The patient's thumb was enclosed within a Chinese finger trap (Allen® Sterile Mesh Finger Traps, AliMed, Inc., Massachusetts, USA) using a rope. After fixing the wrist and thumb using a splint to keep the thumb position at 40-degree abduction during traction, the other end of the rope was hung over the edge of the MR table via a pulley system and attached to non-magnetic traction weights (Figure *1b*). The MRI was initially performed without traction (no weight was used), followed by an MRI with traction. A traction weight of 3 kg was used based on our previous research [16].

Image analysis

We evaluated the effects of traction on the joint space width and articular cartilage outline visibility. In this study, the joint space width was defined as the space between opposing articular cartilages within the target joint. All MR images were independently evaluated by two orthopedic surgeons (with 15 and 10 years of clinical experience, respectively). All study images were interpreted on a workstation (Materialise Mimics, version 20.0; Materialise©, Leuven, Belgium), which was used to obtain the multiplanar reconstructed (MPR) images. Specifically, coronal and sagittal images were reconstructed parallel to the longitudinal axis of the first metacarpal region. The plane connecting the depressed portions of the distal metacarpal condyles with respect to the long axis was defined as coronal, and the plane perpendicular to the coronal plane was defined as sagittal. This procedure was performed by the first author for all the images. The images were initially enlarged, and the grayscale contrast was adjusted to optimize the visualization of the assessed structure. The images were then randomly numbered to minimize bias.

Measurement of the joint space width

Joint space width was measured on sagittal and coronal images at the center of the proximal articular surface of the first metacarpal bone as previously described [16]. On the sagittal image, the AB line, the line through both the volar (point A) and dorsal (point B) borders at the proximal articular surface of the first metacarpal bone; and the CD line, the line through both the volar (point C) and dorsal (point D) borders of the distal articular surface of the trapezium were drawn first. Subsequently, a perpendicular line was drawn at the center of the AB line and the intersection point of the articular surface of the first metacarpal bone was labeled as point E and the intersection point of the articular surface of the trapezium was labeled as point F. Furthermore, the intersection points of the perpendicular line drawn from points A and B to the CD line were labeled as points G and H. On the sagittal image, the following two lines were drawn: the IJ line, the line through the radial (point I) and ulnar (point J) borders of the proximal articular surface of the first metacarpal bone, and the KL line, the line through the radial (point K) and ulnar (point L) borders at the distal articular surface of the trapezium. Subsequently, the intersection points of the perpendicular lines drawn from points K and L to the IJ line were defined as points M and N. The distance between points E and F was defined as the center of the joint space width, those between points A and G and B and H were defined as the volar and dorsal joint space widths, and those between points I and M and J and N were defined as the radial and ulnar joint space widths, respectively (Figure *2*).

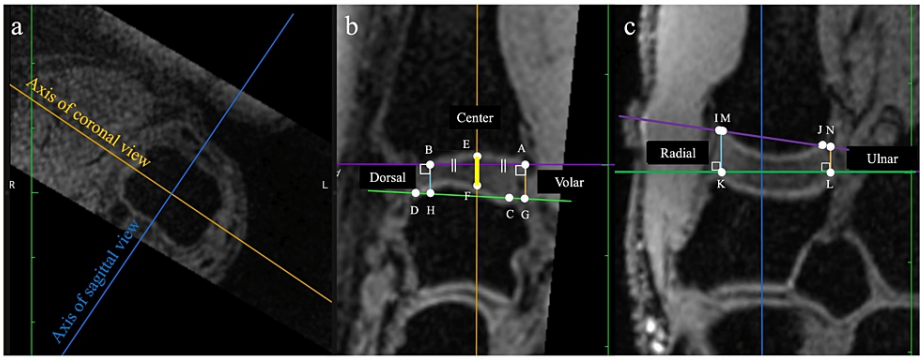


FIGURE 2: Definition of measurement points.

(a) The coronal and sagittal axes are defined using the axial plane at the first metacarpal bone head. (b) Sagittal image of first carpometacarpal joint. Distances E-F, A-G, and B-H were defined as the center, volar, and dorsal distances, respectively. (c) Coronal image at the first carpometacarpal joint. Distances K-M and L-N were defined as the radial and ulnar distances, respectively.

Assessment of the articular cartilage outline visibility

The articular cartilage outline visibility was graded using a three-grade scale as previously described [16]. Grade 2 (complete) occurred when 100% of the articular cartilage outline was clearly visible in the entire range when facing the opposing articular cartilage. Grade 1 (intermediate) occurred when ≥50% but <100% of the articular cartilage outline was clearly visible in the range when facing the opposing articular cartilage. Grade 0 (poor) occurred when the articular cartilage outline was visible in <50% of the entire range when

facing the opposing articular cartilage (Figure 3).

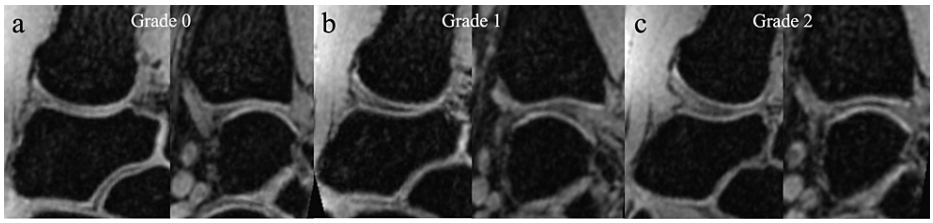


FIGURE 3: Articular cartilage outline visibility grade.

(a) Grade 0 (poor), (b) Grade 1 (intermediate), (c) Grade 2 (complete).

Measurement of the remaining cartilage rate

The rate of the remaining cartilage was evaluated using the same images on which joint space width was measured. On the sagittal and coronal images, the distance between the subchondral bone of the articular surface (line A-D) and the remaining cartilage (line A'-D') was measured both on the first metacarpal bone and trapezium. The rate of the remaining cartilage (%) was calculated by dividing the distance of the remaining cartilage by subchondral bone, and the average value of the sagittal and coronal images was used as the remaining cartilage rate of each bone (Figure 4).

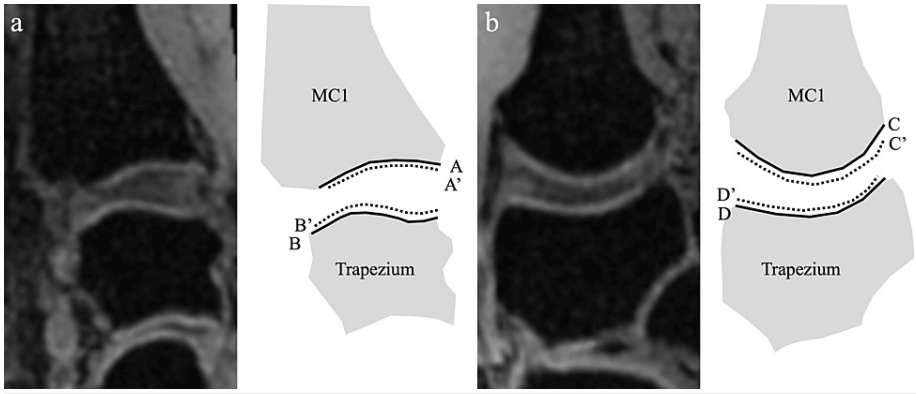


FIGURE 4: Calculation method of the remaining cartilage rate.

(a) Sagittal image, (b) coronal image. The remaining cartilage (%) was calculated by dividing the distance of remaining cartilage (A' to D') by subchondral bone (A to D). A, B, C, D: the distance of subchondral bone, A', B', C', D': the distance of remaining cartilage, MC1: first metacarpal.

Statistical analyses

GraphPad Prism 8 (GraphPad Software, La Jolla, CA, USA) was used for all statistical analyses. All data were tested for normal distribution using the Shapiro-Wilk test. None of the data on articular cartilage outline visibility, joint space width, and remaining cartilage rate followed a normal distribution, owing to the small sample size. Therefore, the Mann-Whitney and Friedman tests were used to assess the differences in joint space widths and articular cartilage outline visibility with and without traction and the differences in the remaining cartilage rate in Stages 1-2 and Stages 3-4 patients. Spearman's rank correlation coefficient was used to evaluate the correlation between the Eaton classification and articular cartilage visibility and changes in joint space width. Statistical significance was set at $p < 0.05$.

Results

The study population comprised 44 patients (men: 14, women: 30), with a mean age of 67.3 ± 8.6 (range, 50-82) years. According to the Eaton classification, 2 patients were in Stage 1, 14 in Stage 2, 24 in Stage 3, and 4 in Stage 4. There is no patient who suspended the MRI examination due to discomfort relating to traction.

The grades of articular cartilage outline visibility were intermediate in 7 and poor in 37 cases without traction, complete in 15, intermediate in 23, and poor in 6 cases with traction. The visibility of the articular cartilage outlines significantly improved with traction ($P<0.01$) (Figure 5). There was a significant correlation between the Eaton classification stage and the grade of articular cartilage visibility with traction ($r = -0.3096$, $P = 0.0409$), but no correlation was observed without traction ($r = -0.2890$, $P = 0.0571$).

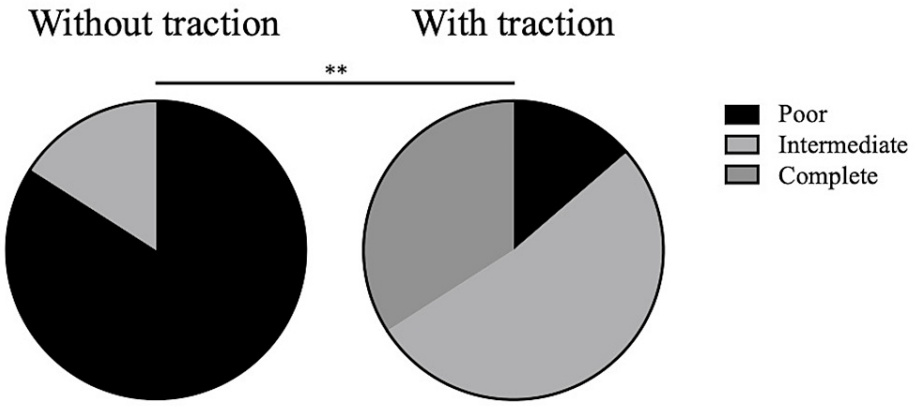


FIGURE 5: Visibility of the articular cartilage outlines.

The visibility is significantly improved by traction. ** P < 0.01.

The joint space width without/with traction was 1.02 ± 0.85/2.25 ± 1.04 mm in the center, 2.28 ± 1.41/3.83 ± 1.43 mm in the volar edge, 1.60 ± 1.77/2.67 ± 1.75 mm in the dorsal edge, 3.70 ± 1.61/4.93 ± 1.66 mm in the radial edge, and 1.43 ± 1.10/2.18 ± 1.39 mm in the ulnar edge. The joint space width increased significantly at all points with traction (P < 0.01) (Figure 6).

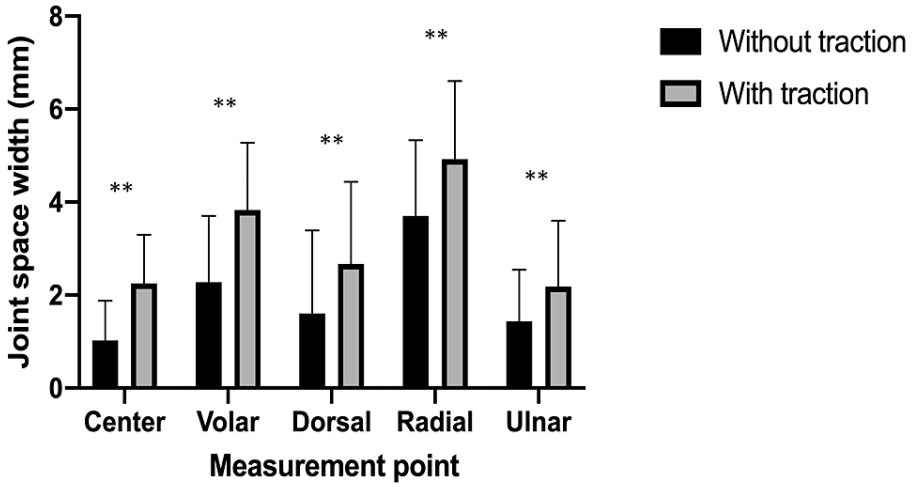


FIGURE 6: Joint space widths at each point (mean ± standard deviation).

Joint space widths significantly widened after axial traction at all points. ** P < 0.01.

The amount of change in joint space width before and after traction at each point was 1.23 ± 0.75 mm in the center, 1.55 ± 1.13 mm in the volar edge, 1.06 ± 1.11 mm in the dorsal edge, 1.22 ± 1.21 mm in the radial edge, and 0.75 ± 1.07 mm in the ulnar edge. No correlation between the Eaton classification stage and the change in joint space width was observed at any point (P=0.8801 at the center, 0.4884 at the volar, 0.6458 at the dorsal, 0.6080 at the radial, and 0.2983 at the ulnar edges). After comparing each point, the change in the ulnar edge was significantly smaller than that in the center (P<0.01), volar edge (P<0.01), and radial edge (P=0.02).

The rate of remaining cartilage on the first metacarpal surface was 51.0 ± 22.8% in Stages 1-2 and 32.5 ± 16.8% in Stages 3-4 (P<0.01). The rate of remaining cartilage on the trapezium surface was 46.5 ± 23.9% in Stages 1-2 and 26.8 ± 18.7% in Stages 3-4 (P=0.01) (Figure 7). However, when considering individual cases, large cartilage defect (>50%) was observed in four Stages 1-2 patients (25.0% of all Stages 1-2 patients), and >50% cartilage remained intact in three Stages 3-4 patients (10.7% of all Stage 3-4 patients).

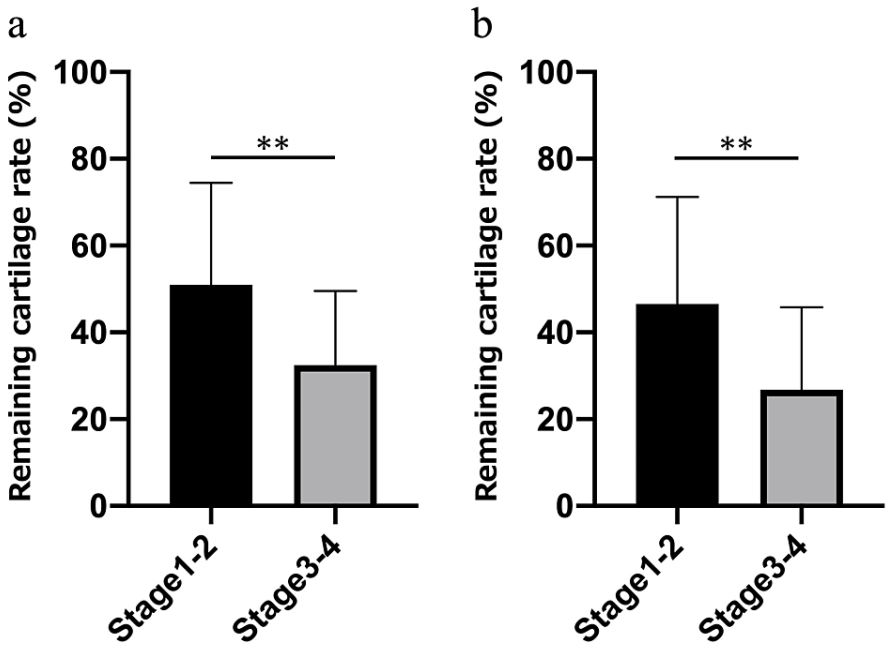


FIGURE 7: Remaining cartilage rate on sagittal and coronal images.

(a) Sagittal image; (b) coronal image. The presence of cartilage remained significant in Eaton Stages 1-2. ** P<0.01.

Discussion

Our results showed that the visibility of the cartilage outline of the thumb carpometacarpal joint on MRI was significantly better when axial traction was applied to the thumb in patients with thumb carpometacarpal arthritis, similar to that in healthy volunteers. The rate of remaining cartilage was significantly higher in patients in the early Eaton stages.

Various surgical procedures have been reported on thumb carpometacarpal arthritis [19]. For patients with mild osteoarthritic changes (Eaton classification Stages 1-2), ligament repair, arthroscopic synovectomy, and first metacarpal osteotomy are generally selected for joint-sparing surgery; whereas for patients with advanced osteoarthritic changes (Eaton classification Stages 3-4), ligament reconstruction with or without tendon interposition, arthrodesis, and artificial joint replacement are generally selected as non-joint-sparing surgery. Although the clinical outcomes of each surgical procedure are generally good [20], some joint preservation surgeries have achieved good symptomatic improvement even in patients with advanced osteoarthritic changes on plain radiographs [21,22]. In the present study, although the articular cartilage defect was significantly larger in patients with Eaton Stages 3-4, the cartilage remained >50% intact in 10% of those patients. The change in load distribution due to the first metacarpal osteotomy is one of the mechanisms used to reduce pain in patients with thumb carpometacarpal arthritis [23,24]. We believe that if the residual articular cartilage can be evaluated preoperatively, joint-sparing surgery can be performed more aggressively even in progressive cases. Thus, articular cartilage evaluation using axial traction MRI has the potential to be useful for surgeons in selecting less-invasive surgical procedures. We also believe that joint-sparing surgery is preferable for younger patients or those involved in heavy labor, although the risk of osteoarthritic progression exists even after long-term joint-sparing surgery.

In this study, the widening of the articular surface due to axial traction was significantly smaller on the ulnar side. The anterior oblique (or volar beak) and dorsoradial ligaments have been reported as key ligaments in thumb carpometacarpal arthritis [25-27]. These ligaments contribute to the stability of the radiodorsal side of the thumb carpometacarpal joint. According to the progression of thumb carpometacarpal arthritis, loosening of these ligaments related to radiodorsal instability of the thumb carpometacarpal joint may have caused the difference in joint space width between the measurement points by traction in this study. Although our results suggest that the degree of traction-induced widening of the joint space can be used to evaluate joint instability, which reflects ligament dysfunction, the sample size was insufficient for evaluating joint instability (ligament dysfunction). Future research efforts should aim to increase the sample size and verify whether axial traction MRI can evaluate ligament function and articular cartilage.

Badia et al. reported a treatment algorithm based on the intraoperative thumb carpometacarpal arthroscopy findings in 2006 [11]. Depending on the degree of articular cartilage damage or loss, joint-sparing or non-

joint-sparing surgery was selected for this algorithm. This algorithm is more innovative than the Eaton classification because it assesses articular cartilage damage and loss with higher accuracy before deciding on the surgical technique. However, several limitations have been encountered, including the inability to evaluate the articular cartilage preoperatively, invasiveness, and the need to decide the surgical technique intraoperatively. Our results demonstrate that axial traction MRI of the thumb carpometacarpal joint could be used to preoperatively evaluate the articular cartilage condition, which would allow the selection of the optimal surgical technique that reflects the articular cartilage condition, rather than depending on the Eaton classification.

This study has several limitations. First, the assessment of the rate of remaining cartilage was limited to a single slice of sagittal and coronal images in this study. Developing an evaluation method encompassing the whole joint surface, possibly using techniques such as 3D reconstruction, is necessary for a more accurate assessment of the articular cartilage condition in future studies. Second, the optimal traction weight was not evaluated in this study. While increasing the axial traction on the thumb with heavier weights might further widen the joint space and improve the articular cartilage outline visibility, this added traction may induce pain in these individuals. Compared with our previous result of the change in joint space width in healthy volunteers [16], there was no significant difference between 3 kg traction in patients with thumb carpometacarpal arthritis and 2 kg and 5 kg traction in healthy volunteers (Figure *8*). Thus, a 3 kg traction was considered sufficient to evaluate articular cartilage defects in the thumb carpometacarpal joint. Thirdly, this study did not verify whether articular cartilage was correctly detected by MRI. To evaluate the precision of detecting articular cartilage through axial traction MRI, a study comparing the consistency between MRI and arthroscopy findings is necessary for future research. Finally, the axial traction system used in this study could not control the rotational force on the carpometacarpal joint of the thumb. A slight twist caused by axial traction may have affected the measurement of the joint space widths at each point. Therefore, the addition of a system to control the rotational force during axial traction becomes imperative.

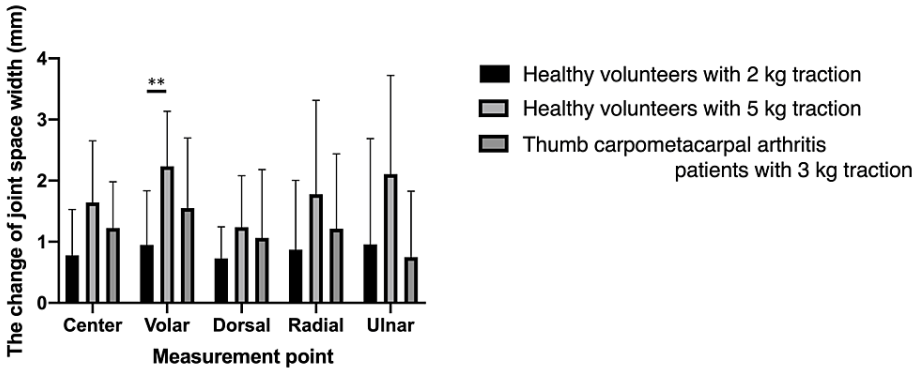


FIGURE 8: The change of joint space width between with and without traction.

There is no significant difference between 3 kg traction in thumb carpometacarpal arthritis patients and both 2 kg and 5 kg traction in healthy volunteers at any measurement point by two-way ANOVA with the Geissler-Green house correction. ** P < 0.01 between 2 kg and 5 kg traction in healthy volunteers by Tukey's multiple comparisons test.

Conclusions

Axial traction of the thumb increased the joint space width and improved the visibility of the articular cartilage of the thumb carpometacarpal joint on the MRIs of patients with thumb carpometacarpal arthritis.

Our results suggest that axial traction MRI can be used to noninvasively evaluate articular cartilage defects in patients with thumb carpometacarpal arthritis.

Axial traction MRI holds potential as an evaluation index for selecting the optimal surgical method in patients with thumb carpometacarpal arthritis.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Akira Ikumi, Sho Kohyama

Acquisition, analysis, or interpretation of data: Akira Ikumi, Sho Iwabuchi, Takeo Mammoto, Takeshi Ogawa, Yuichi Yoshii, Masashi Yamazaki

Drafting of the manuscript: Akira Ikumi

Critical review of the manuscript for important intellectual content: Sho Iwabuchi, Sho Kohyama, Takeo Mammoto, Takeshi Ogawa, Yuichi Yoshii, Masashi Yamazaki

Supervision: Yuichi Yoshii

Disclosures

Human subjects: Consent was obtained or waived by all participants in this study. Tsukuba University Hospital Mito Clinical Education and Training Center issued approval No.21-01. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** This work was supported by the JA co-commissioned research business of the Japanese Association of Rural Medicine (grant No. 2022-5). **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

We would like to thank Editage (www.editage.com) for English language editing.

References

1. Cooney WP 3rd, Chao EY: Biomechanical analysis of static forces in the thumb during hand function . J Bone Joint Surg Am. 1977, 59:27-36.
2. Strauch RJ, Behrman MJ, Rosenwasser MP: Acute dislocation of the carpometacarpal joint of the thumb: an anatomic and cadaver study. J Hand Surg Am. 1994, 19:93-8. 10.1016/0363-5023(94)90229-1
3. Armstrong AL, Hunter JB, Davis TR: The prevalence of degenerative arthritis of the base of the thumb in post-menopausal women. J Hand Surg Br. 1994, 19:340-1. 10.1016/0266-7681(94)90085-x
4. Dillon CF, Hirsch R, Rasch EK, Gu Q: Symptomatic hand osteoarthritis in the United States: prevalence and functional impairment estimates from the third U.S. National Health and Nutrition Examination Survey, 1991-1994. Am J Phys Med Rehabil. 2007, 86:12-21. 10.1097/phm.0b013e31802ba28e
5. Haara MM, Heliövaara M, Kröger H, et al.: Osteoarthritis in the carpometacarpal joint of the thumb. Prevalence and associations with disability and mortality. J Bone Joint Surg Am. 2004, 86:1452-7. 10.2106/00004623-200407000-00013
6. Sodha S, Ring D, Zurakowski D, Jupiter JB: Prevalence of osteoarthrosis of the trapeziometacarpal joint . J Bone Joint Surg Am. 2005, 87:2614-8. 10.2106/JBJS.E.00104
7. Eaton RG, Littler JW: Ligament reconstruction for the painful thumb carpometacarpal joint . J Bone Joint Surg Am. 1973, 55:1655-66.
8. Eaton RG, Glickel SZ: Trapeziometacarpal osteoarthritis. Staging as a rationale for treatment . Hand Clin. 1987, 3:455-71.
9. Berger AJ, Momeni A, Ladd AL: Intra- and interobserver reliability of the Eaton classification for trapeziometacarpal arthritis: a systematic review. Clin Orthop Relat Res. 2014, 472:1155-9. 10.1007/s11999-013-3208-z
10. Hoffler CE 2nd, Matzon JL, Lutsky KF, Kim N, Beredjiklian PK: Radiographic stage does not correlate with symptom severity in thumb basilar joint osteoarthritis. J Am Acad Orthop Surg. 2015, 23:778-82. 10.5435/JAAOS-D-15-00329
11. Badia A: Trapeziometacarpal arthroscopy: a classification and treatment algorithm . Hand Clin. 2006, 22:153-63. 10.1016/j.hcl.2006.02.006
12. Kroon FP, Conaghan PG, Foltz V, et al.: Development and reliability of the OMERACT thumb base osteoarthritis magnetic resonance imaging scoring system. J Rheumatol. 2017, 44:1694-8. 10.3899/jrheum.161099
13. Connell DA, Pike J, Koulouris G, van Wettering N, Hoy G: MR imaging of thumb carpometacarpal joint ligament injuries. J Hand Surg Br. 2004, 29:46-54. 10.1016/s0266-7681(03)00170-0
14. Cardoso FN, Kim HJ, Albertotti F, Botte MJ, Resnick D, Chung CB: Imaging the ligaments of the trapeziometacarpal joint: MRI compared with MR arthrography in cadaveric specimens. AJR Am J Roentgenol. 2009, 192:W13-9. 10.2214/AJR.07.4010
15. Saltzherr MS, Coert JH, Selles RW, et al.: Accuracy of magnetic resonance imaging to detect cartilage loss in severe osteoarthritis of the first carpometacarpal joint: comparison with histological evaluation. Arthritis Res Ther. 2017, 19:55. 10.1186/s13075-017-1262-8
16. Ikumi A, Kohyama S, Okuwaki S, et al.: Effects of magnetic resonance imaging with axial traction of the thumb carpometacarpal joint on articular cartilage visibility: a feasibility study. Cureus. 2022, 14:e22421. 10.7759/cureus.22421
17. Kohyama S, Tanaka T, Shimasaki K, Kobayashi S, Ikumi A, Yanai T, Ochiai N: Effect of elbow MRI with axial traction on articular cartilage visibility-a feasibility study. Skeletal Radiol. 2020, 49:1555-66. 10.1007/s00256-020-03455-3
18. Kikuchi N, Kohyama S, Kanamori A, Taniguchi Y, Okuno K, Ikeda K, Yamazaki M: Improving visualization of

the articular cartilage of the knee with magnetic resonance imaging under axial traction: a comparative study of different traction weights. Skeletal Radiol. 2022, 51:1483-91. 10.1007/s00256-021-03971-w

19. Wilkens SC, Meghpara MM, Ring D, Coert JH, Jupiter JB, Chen NC: Trapeziometacarpal Arthrosis. JBJS Rev. 2019, 7:e8. 10.2106/JBJS.RVW.18.00020
20. Gottschalk MB, Patel NN, Boden AL, Kakar S: Treatment of basilar thumb arthritis: a critical analysis review . JBJS Rev. 2018, 6:e4. 10.2106/JBJS.RVW.17.00156
21. Ogawa T, Tanaka T, Asakawa S, Tatsumura M, Mammoto T, Hirano A: Arthroscopic synovectomy for the treatment of stage II to IV trapeziometacarpal joint arthritis. J Rural Med. 2018, 13:76-81. 10.2185/jrm.2962
22. Spielman AF, Sankaranarayanan S, Lessard AS: Joint preserving treatments for thumb CMC arthritis . Hand Clin. 2022, 38:169-81. 10.1016/j.hcl.2022.01.002
23. Molitor PJ, Emery RJ, Meggitt BF: First metacarpal osteotomy for carpo-metacarpal osteoarthritis . J Hand Surg Br. 1991, 16:424-7. 10.1016/0266-7681(91)90018-j
24. Pellegrini VD Jr: Pathomechanics of the thumb trapeziometacarpal joint . Hand Clin. 2001, 17:175-84, vii-viii. 10.1016/S0749-0712(21)00238-9
25. Pagalidis T, Kuczynski K, Lamb DW: Ligamentous stability of the base of the thumb . Hand. 1981, 13:29-36. 10.1016/s0072-968x(81)80026-5
26. Doerschuk SH, Hicks DG, Chinchilli VM, Pellegrini VD Jr: Histopathology of the palmar beak ligament in trapeziometacarpal osteoarthritis. J Hand Surg Am. 1999, 24:496-504. 10.1053/jhsu.1999.0496
27. Bettinger PC, Berger RA: Functional ligamentous anatomy of the trapezium and trapeziometacarpal joint (gross and arthroscopic). Hand Clin. 2001, 17:151-68, vii. 10.1016/S0749-0712(21)00236-5

Online January 13, 2024
Journal of Neuroendovascular Therapy 2024; 18: 29–36
DOI: 10.5797/jnet.oa.2023-0069

Original Article

Efficacy and Safety of Lumbar Drainage before Endovascular Treatment for Ruptured Intracranial Aneurysms

Toshitsugu Terakado,[1,2] Yoshiro Ito,[3] Koji Hirata,[4] Masayuki Sato,[3] Tomoji Takigawa,[5] Aiki Marushima,[3] Mikito Hayakawa,[4] Wataro Tsuruta,[6] Noriyuki Kato,[7] Yasunobu Nakai,[1] Kensuke Suzuki,[5] Yuji Matsumaru,[3,4] and Eiichi Ishikawa[3]

Objective: Intraoperative rebleeding during endovascular treatment for ruptured intracranial aneurysms is associated with poor prognosis. Lumbar drainage is performed preoperatively to control intracranial pressure; however, it is associated with a risk of brain herniation or rebleeding because intracranial pressure may change rapidly. Therefore, this study aimed to examine the efficacy and safety of preoperative lumbar drainage.

Methods: This retrospective study enrolled 375 patients who underwent endovascular treatment of ruptured intracranial aneurysms at our institution between April 2013 and March 2018. The incidence of rebleeding and clinical outcomes were compared between patients who did and did not undergo preoperative lumbar drainage.

Results: Among the 375 patients with ruptured intracranial aneurysms, 324 (86.0%) and 51 (14.0%) patients did and did not undergo lumbar drainage, respectively. The incidence of rebleeding was 11/324 (3.4%) and 2/51 (3.9%) in lumbar drainage and nonlumbar drainage groups, respectively, with no statistical differences (p = 0.98). Of the rebleeding cases, 9/11 (81%) and 2/2 (100%) in lumbar drainage and nonlumbar drainage groups, respectively, were due to intraoperative bleeding, and 2/11 (19%) in the lumbar drainage group, the causes of the rebleeding were undetermined. The incidence of symptomatic vasospasm did not differ significantly between the groups (13.2% vs. 11.8%, P = 0.776), while the incidence of hydrocephalus (24.6% vs. 11.8%, P = 0.043) and meningitis (15.2% vs. 5.9%, P = 0.075) were slightly higher in the lumbar drainage group. Favorable clinical outcomes (modified Rankin Scale score <2) at discharge were less frequent in the lumbar drainage group (55.3% vs. 70.0%, P = 0.051). No significant differences were observed in the propensity score-matched analysis.

Conclusion: Lumbar drainage before endovascular treatment for ruptured intracranial aneurysms is a safe procedure that does not increase the incidence of rebleeding.

Keywords ▶ cerebrospinal fluid, drainage, subarachnoid hemorrhage, endovascular procedures

Introduction

There are many reports on the use of lumber drainage during the perioperative period of endovascular treatment of ruptured intracranial aneurysms. Many of these studies have reported that lumbar drainage after aneurysm treatment reduces delayed vasospasm and cerebral infarction.[1–5] Nowadays, there is a consensus on lumbar drainage management after ruptured intracranial aneurysm treatment. However, although intraoperative rebleeding is

[1]*Department of Neurosurgery, Tsukuba Medical Center Hospital, Tsukuba, Ibaraki, Japan*
[2]*Department of Neurosurgery, Koyama Memorial Hospital, Kashima, Ibaraki, Japan*
[3]*Department of Neurosurgery, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan*
[4]*Division of Stroke Prevention and Treatment, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan*
[5]*Department of Neurosurgery, Saitama Medical Center, Dokkyo Medical University, Koshigaya, Saitama, Japan*
[6]*Department of Neuro-Endovascular Therapy, Toranomon Hospital, Tokyo, Japan*
[7]*Department of Neurosurgery, National Hospital Organization Mito Medical Center, Higashi-Ibaraki, Ibaraki, Japan*

Received: October 2, 2023; Accepted: December 11, 2023
Corresponding author: Yoshiro Ito. Department of Neurosurgery, Faculty of Medicine, University of Tsukuba, Tennodai 1-1-1, Tsukuba, Ibaraki 305-8575, Japan
Email: yoshiro@md.tsukuba.ac.jp

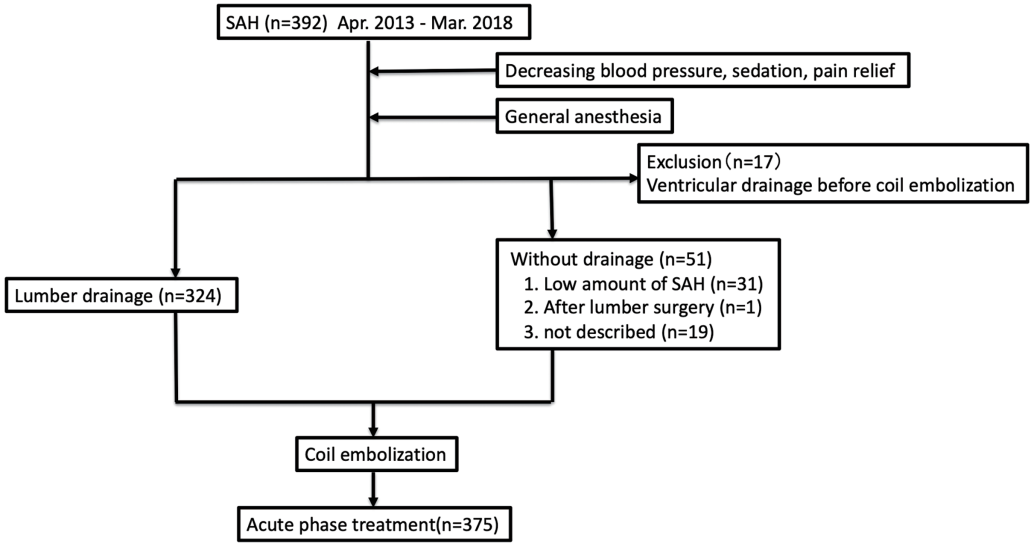


Fig. 1 Patient inclusion flowchart. SAH: subarachnoid hemorrhage

rare in this disease (5.0%–7.7%), 63% of intraoperative rebleeding cases have poor outcomes.[6,7] The sudden increase in intracranial pressure caused by an aneurysm rupture is associated with poor outcomes. Lumbar drainage in the preoperative period has been shown to control intracranial pressure rapidly and may also decrease secondary brain damage.[8] However, rebleeding or brain herniation may occur during insertion of a lumbar drainage catheter,[9–12] though the effectiveness and safety of lumbar drainage before endovascular treatment of ruptured intracranial aneurysms remain unclear. We hypothesize that lumbar drainage before endovascular treatment would not increase perioperative rebleeding and would reduce intracranial pressure in the event of intraoperative rebleeding, preventing poor functional outcomes. Therefore, the present study examined the effectiveness and safety of preoperative lumbar drainage during the endovascular treatment of ruptured cerebral aneurysms.

Materials and Methods

Study design and patients

This retrospective study was approved by the Institutional Review Committee of the Neuroendovascular and Surgical Management from Multicenter Observation to build a PHILosophical Approach (NEMMOPHILA) study (Reference number: H30-137) and was conducted according to the tenets of the Declaration of Helsinki. This retrospective observational study was conducted using the opt-out method on our department's website.

Three hundred and seventy-five patients who underwent endovascular treatment for ruptured intracranial aneurysms between April 2013 and March 2018 at five associated institutions were included (**Fig. 1**). Patients with acute hydrocephalus that required ventricular drainage were excluded.

Endovascular technique and perioperative management

Endovascular treatment of ruptured intracranial aneurysms was performed within 48 hours of onset. All patients with a longer treatment duration were re-evaluated using head CT scans. Lumbar drainage was basically performed on patients who underwent endovascular treatment, but the final decision was left to the surgeon. Reasons for not performing lumbar drainage included low hematoma volume and postoperative lumbar surgery (**Fig. 1**). A lumbar drainage catheter was inserted after general anesthesia, and cerebrospinal fluid (CSF) was continuously drained during the procedure. Systemic heparinization (80 IU/kg) was performed prior to insertion of the guiding catheter, and antiplatelet drugs were not administered. However, dual antiplatelet therapy loading doses were administered in cases that required stent placement. Head CT scans were performed within 24 hours after the treatment to evaluate complications. When rebleeding occurred, the blood pressure was decreased as soon as possible. In cases of intraoperative rebleeding, bleeding was stopped with a balloon catheter or coil. After endovascular treatment, the drain rate was set at 5 to 10 mL/hour. Drainage was continued

until the hematoma resolved on head CT, if an adverse event occurred, or at 7 days postoperatively. Fasudil and ozagrel sodium were administered intravenously for the prevention of vasospasm. Treatment-resistant vasospasm was managed with intra-arterial fasudil or percutaneous transluminal angioplasty.

Data collection

The following data were collected: age, sex, aneurysm location, aneurysm size (largest dimension of the aneurysm), bleb, pretreatment World Federation of Neurological Surgeons (WFNS) grade; Fisher's group as determined using the first head CT, treatment methods, adverse events associated with lumbar drainage (e.g., brain herniation due to excessive drainage, and a piece of the drainage catheter remained in the patient at the time of its removal, which required additional treatment), perioperative rebleeding, symptomatic vasospasm, hydrocephalus requiring shunt surgery, meningitis, and the modified Rankin Scale (mRS) score at discharge. Fisher's group was determined as follows: Group 1, with no blood detected; Group 2, with diffuse deposition or a thin layer with all vertical layers of blood <1 mm in thickness; Group 3, with localized clots and/or a vertical layer of blood ≥1 mm in thickness; Group 4, with diffuse or no subarachnoid blood but with intracerebral or intraventricular clots.

Study endpoints

A comparative study was conducted both with and without lumbar drainage. Lumbar drainage insertion was defined as LD (+), and noninsertion was defined as LD (−). The primary endpoint was the frequency of favorable clinical outcomes at discharge, defined as an mRS score <2. The secondary endpoints were the frequency of perioperative rebleeding, symptomatic vasospasm, hydrocephalus, and meningitis. Perioperative rebleeding was defined as bleeding that occurred up to the time of the postoperative CT. Therefore, bleeding during lumbar drainage insertion, intraoperative bleeding, and enlarged subarachnoid hemorrhage (SAH) on postoperative CT were included.[7] Symptomatic vasospasm was defined using the following criteria: (1) newly developed neurological deficit; (2) no explanation for neurological deficits, such as hyponatremia, infection, hypoxia, and epilepsy; and (3) evidence of vasospasm on magnetic resonance imaging, CT angiography, and transcranial Doppler. Hydrocephalus was defined as the need for shunt surgery. Meningitis was defined as increased cell count and hypoglycorrhachia in the CSF with fever.

Statistical analysis

Data were compared between patients who did and did not undergo preoperative lumbar drainage. Continuous data were expressed as the mean ± standard deviation and categorical data were presented as the counts and percentages. Between-group differences were assessed using Fisher's exact test for discrete data and two-sample Student's t-test for continuous data. The propensity score was calculated using a multivariable logistic regression model with the two groups as dependent variables and sex, WFNS score, Fisher's group, rebleeding, symptomatic vasospasm, hydrocephalus, meningitis, and mRS at discharge as independent variables. The inverse probability of treatment weighting (IPTW) was determined using propensity score-matched analysis. Briefly, IPTW uses weights based on the propensity score to create a synthetic sample in which the distribution of measured covariates is independent of the treatment assignment. All statistical analyses were performed using SPSS statistics software (version 25.0; IBM, Armonk, NY, USA). Statistical significance was set at $P < 0.05$.

Results

Of the 375 patients with ruptured intracranial aneurysms, 324 (86.0%) and 51 (14.0%) did and did not undergo lumbar drainage. There were no significant between-group differences in age, sex, aneurysm location, and aneurysm size. The WFNS grade was significantly lower in the LD (−) group ($P = 0.013$). Fisher group was lower in the LD (−) group ($P < 0.001$), and aneurysms with blebs were more frequent in the LD (+) group (65.7% vs. 46.0%, $P = 0.007$). A double catheter or balloon-assisted technique was used more frequently in the LD (+) group ($P = 0.034$) (**Table 1**).

The frequency of an mRS ranging between 0–2 at discharge was lower in the LD (+) group than in the LD (−) group (55.3% vs. 70.0%, $P = 0.051$), but there were no significant differences in rebleeding (4.0% vs. 3.9%, $P = 0.975$). Furthermore, symptomatic vasospasm did not differ significantly between the LD (+) and LD (−) groups (13.2% vs. 11.8%, $P = 0.776$). The incidences of hydrocephalus (24.6% vs. 11.8%, $P = 0.043$) and meningitis (15.2% vs. 5.9%, $P = 0.075$) were slightly higher in the LD (+) group (**Table 2**).

The characteristics of the patients with rebleeding are summarized in **Table 3**. Re-bleeding occurred in 13 patients in the LD (+) group, and intraoperative rebleeding in 11 of these patients was perforated by a coil,

Table 1 Patient characteristics and adverse events

Patients	Drainage (+)	Drainage (–)	P value
	324	51	
Age (years)	63.0 ± 14.4	60.3 ± 14.5	0.211
Male	98 (30.2)	13 (25.5)	0.489
WFNS grade	324	50	0.013
1	69 (21.3)	18/50 (36.0)	
2	96 (29.6)	15/50 (30.0)	
3	42 (13.0)	5/50 (10.0)	
4	80 (24.7)	3/50 (6.0)	
5	37 (11.4)	9/50 (18.0)	
Fisher's grade	324	49	<0.001
1	3 (0.9)	5/49 (10.2)	
2	78 (24.1)	18/49 (36.7)	
3	160 (49.4)	15/49 (30.6)	
4	83 (25.6)	11/49 (22.4)	
Aneurysm location	323	51	0.176
ICA proximal	15 (4.6)	5 (9.8)	
ICA distal	97 (30.0)	9 (17.6)	
AcomA	77 (23.8)	11 (21.6)	
MCA	23 (7.1)	4 (7.8)	
VA	59 (18.3)	8 (15.7)	
BA apex	36 (11.1)	8 (15.7)	
Others	16 (5.0)	6 (11.8)	
Fusiform	37 (11.4)	1 (2.0)	0.037
Size	6.45 ± 3.25	6.62 ± 4.98	0.75
Bleb	205/312 (65.7)	23/50 (46.0)	0.007
Treatment method	305	47	0.034
Simple catheter	115 (57.4)	30 (58.8)	
Double catheter	22 (7.2)	0 (0.0)	
Balloon assist	145 (47.5)	18 (38.3)	
Stent assist	23 (7.5)	2 (4.3)	

Data are presented as the mean ± standard deviation or n (%). AcomA: anterior communicating artery; BA: basilar artery; ICA: internal carotid artery; MCA: middle cerebral artery; VA: vertebral artery; WFNS: World Federation of Neurological Surgeons

microcatheter, or microguidewire. In the remaining 2 patients, no adverse events occurred during the intraoperative period; however, postoperative CT revealed an enlarged SAH. No rebleeding occurred between the lumbar drainage procedure and the endovascular treatment. In contrast, rebleeding occurred in two patients in the LD (−) group during the intraoperative period. The adverse events that occurred in 2 patients (1.0%) who underwent lumbar drainage included brain herniation due to excessive drainage and a piece of the drainage catheter that remained in the patient at the time of its removal, which required additional surgery.

Table 2 Clinical outcomes

	Drainage (+) (n = 324)	Drainage (–) (n = 51)	P value
Primary endpoint			
Modified Rankin Scale score 0–2	176/318 (55.3)	35/50 (70.0)	0.051
Secondary endpoint			
Rebleeding	13 (4.0)	2 (3.9)	0.975
Symptomatic vasospasm	42/318 (13.2)	6 (11.8)	0.776
Hydrocephalus	78/317 (24.6)	6 (11.8)	0.043
Meningitis	48/316 (15.2)	3 (5.9)	0.075

Given that patient characteristics differed between the LD (+) and LD (−) groups, we matched 42 patients in both groups and compared the rebleeding and outcomes at discharge (**Table 4**). There was no significant difference between the LD (+) and LD (−). The odds ratio of good outcomes at discharge were 1.378 (95% confidence interval [CI]: 0.555–3.421; P = 0.489) by propensity score matching and 1.824 (95% CI: 0.546–6.094; P = 0.329) with IPTW for the LD (+) group relative to the LD (−) group. The odds ratio of rebleeding was 1.00 (95% CI: 0.134–7.451; P = 1.000) by propensity score matching and 1.089 (95% CI: 0.210–5.652; P = 0.919) with IPTW (**Tables 5** and **6**). Preoperative lumbar drainage was not associated with intraoperative rebleeding or poor neurological outcome.

Discussion

Perioperative lumbar drainage is commonly performed during endovascular treatment to prevent cerebral vasospasms. The volume of SAH was previously shown to be associated with cerebral vasospasm.[13] The drainage of SAH from ruptured aneurysms is important for preventing cerebral vasospasm. Lumbar drainage is known to reduce SAH more rapidly than ventricular drainage.[3] Previous studies showed that 17%–29% of symptomatic vasospasm cases occurred with lumbar drainage, whereas 27%–45% occurred without lumbar drainage.[2,4,5,14,15] However, some studies have reported occasional rebleeding and herniation associated with lumbar drainage.[9,10] The most critical complication of endovascular treatment of ruptured cerebral aneurysms is intraoperative rebleeding. Previous studies reported an intraoperative rebleeding rate of 5.0%–7.7%, which is lower than surgical clipping.[6,7,16] Intraoperative

Table 3 Summary of characteristics of patients with rebleeding

	Lumbar drainage	Age (years)	Sex	Location	WFNS	Fisher's grade	Details of rebleeding	Symptomatic vasospasm	Hydrocephalus	mRS at discharge
1	+	83	F	AcomA	2	4	Intraoperative	–	+	4
2	+	60	F	IC-PC	5	3	Intraoperative	–	–	1
3	+	67	F	MCA	3	2	Intraoperative	Dead (day 9)		6
4	+	72	M	IC-PC	4	3	Intraoperative	Dead (day 3)		6
5	+	60	F	MCA	4	3	Intraoperative	–	–	3
6	+	62	F	VA	4	3	SAH increased at postoperative CT	–	–	3
7	+	63	F	IC-PC	1	2	Intraoperative	–	–	0
8	+	41	F	IC-PC	1	2	Intraoperative	–	–	1
9	+	80	F	ACA	1	3	Intraoperative	–	+	5
10	+	92	F	IC-PC	2	3	Intraoperative	–	–	3
11	+	84	M	VA	1	2	Intraoperative	–	–	2
12	+	59	F	IC-PC	2	2	Intraoperative	–	+	2
13	+	47	F	AcomA	1	3	SAH increased at postoperative CT	Dead (day 8)		6
14	–	56	F	ICA	1	3	Intraoperative	–	–	0
15	–	80	F	AcomA	2	3	Intraoperative	–	+	5

ACA: anterior cerebral artery; AcomA: anterior communicating artery; ICA: internal carotid artery; IC-PC: internal carotid-posterior communicating artery; MCA: middle cerebral artery; SAH: subarachnoid hemorrhage; VA: vertebral artery; WFNS: World Federation of Neurological Surgeons

rebleeding during surgical clipping does not yield a poor outcome because a sudden increase in intracranial pressure is prevented by craniotomy and the cessation of bleeding as soon as possible.[17,18] Meanwhile, intraoperative rebleeding during endovascular treatment is associated with poor outcomes due to a sudden increase in intracranial pressure because craniotomy is not performed.[6,7] A previous study showed that lumbar drainage after SAH effectively controlled intracranial pressure and may contribute to preventing secondary brain damage.[8] If the risk of rebleeding or complications associated with the insertion of a lumbar drainage catheter in the preoperative period is low, it may contribute to controlling intracranial pressure during the acute phase and preventing a sudden increase in intracranial pressure due to rebleeding.

In the present study, the rebleeding rate was 5.8% in the LD (+) group, which is consistent with previous findings (5.0%–7.7%).[6,7] Moreover, the rebleeding rate associated with lumbar drainage was 0.6% (2/324 patients). In the propensity score-matched analysis to adjust for patient backgrounds between the LD (+) and LD (−) groups, the findings confirmed that lumbar drainage before endovascular treatment did not contribute to aneurysm rebleeding. Connolly et al. previously reported a series of 314 patients who underwent lumbar drainage before surgical clipping for ruptured aneurysms.[19] Rebleeding due to lumbar drainage occurred in only 1 patient (0.3%). In the study by Ochiai et al., although no patient had rebleeding associated with lumbar drainage among the 31 patients, 1 (9.1%) had rebleeding while awaiting treatment.[20] Furthermore, Ruijs et al. showed that, in 11 patients with SAH who underwent lumbar drainage during the acute phase, rebleeding occurred in 5 patients (45.4%) while waiting for treatment (several hours to 6 days).[21] These findings support that rebleeding associated with a lumbar drainage catheter insertion is extremely low. However, the risk of rebleeding is high while awaiting treatment after the insertion of a lumbar drainage catheter. Ruptured aneurysms need to be treated early when a lumbar drainage catheter is inserted.

Most patients in this study underwent lumbar drainage under general anesthesia. Lumbar drainage was safely performed with pain relief and strict blood pressure control. Endovascular treatment was administered immediately after lumbar drainage. If the aneurysm is treated soon after lumber drainage, the risk of rebleeding is extremely low, and lumbar drainage before endovascular treatment is considered a safe procedure. Lumbar drainage prior to endovascular treatment has an additional advantage. If a lumbar

Table 4 Propensity score matching between the lumbar drainage and no lumbar drainage groups

Patients	Drainage (+)	Drainage (–)	P value
	42	42	
Age (years)	59.7 ± 14.0	61.6 ± 15.0	0.554
Male	15 (35.7)	11 (26.2)	0.345
WFNS grade			0.871
1	14 (33.3)	15 (35.7)	
2	13 (31.0)	12 (28.6)	
3	3 (7.1)	5 (11.9)	
4	4 (9.5)	2 (4.8)	
5	8 (19.0)	8 (19.0)	
Fisher's grade			0.616
1	2 (4.8)	3 (7.1)	
2	12 (28.6)	17 (40.5)	
3	17 (40.5)	13 (31.0)	
4	11 (26.2)	9 (21.4)	
Aneurysm location			0.673
ICA proximal	4 (9.5)	3 (7.1)	
ICA distal	9 (21.4)	8 (19.0)	
AcomA	14 (33.3)	11 (26.2)	
MCA	6 (14.3)	4 (9.5)	
VA	5 (11.9)	5 (11.9)	
BA apex	2 (4.8)	8 (19.0)	
Others	2 (4.8)	3 (7.1)	
Fusiform	0 (0.0)	0 (0.0)	–
Size	5.79 ± 3.32	6.30 ± 4.93	0.585
Bleb	18 (42.9)	20 (47.6)	0.661
Treatment method			0.801
Simple catheter	28 (66.7)	26 (61.9)	
Double catheter	0 (0.0)	0 (0.0)	
Balloon assist	13 (31.0)	14 (33.3)	
Stent assist	1 (2.4)	2 (4.8)	

Data are presented as the mean ± standard deviation or n (%). AcomA: anterior communicating artery; BA: basilar artery; ICA: internal carotid artery; MCA: middle cerebral artery; WFNS: World Federation of Neurological Surgeon; VA: vertebral artery

Table 5 Clinical outcomes and odds ratio of adverse events by propensity score matching and inverse probability of treatment weighting

Patients	Drainage (+)	Drainage (–)	P value
	42	42	
Primary endpoint			
Modified Rankin Scale Score 0–2	25 (59.5)	28 (66.7)	0.488
Secondary endpoint			
Rebleeding	2 (4.8)	2 (4.8)	1.000
Symptomatic vasospasm	5 (11.9)	5 (11.9)	1.000
Hydrocephalus	10 (23.8)	6 (14.3)	0.243
Meningitis	8 (19.0)	3 (7.1)	0.097

drainage catheter is inserted before endovascular treatment when antithrombotic therapy is administered during the perioperative period, hemorrhagic complications associated with its insertion may be prevented.

We suspected that lumbar drainage during intraoperative rebleeding may prevent the deterioration of patient outcomes. However, lumbar drainage did not affect patient outcomes in the propensity score-matched or ITPW analyses in this study. The sample size of patients with rebleeding in the LD (−) group was considered too small to prove its effects on the clinical outcomes.

Incidences of meningitis and hydrocephalus were slightly higher in the LD (+) group than in the LD (−) group before propensity score matching. One reason for this was the more severe SAH grade in the LD (+) group than in the LD (−) group. Long-term lumbar drainage was needed to wash out thick SAH. However, lumbar drainage for more than four days increased the risk of meningitis.[22] Moreover, meningitis is associated with an increased risk of hydrocephalus (odds ratio: 5.90).[23] Therefore, the duration of lumbar drainage warrants careful consideration, and the drainage catheter needs to be removed as soon as the hemorrhage is washed out.

The present study has several limitations. This was a nonrandomized retrospective study, and the clinical backgrounds of the two groups were different. Therefore, further randomized controlled trials involving a larger number of patients are required to confirm the safety of lumbar drainage. Furthermore, the LD (−) group included only 14% (51/375) of all patients, and the decision to perform spinal drain insertion was completely based on individual judgment; thus, statistical analyses may have been inadequate because there were too few patients in this group. It remains unknown whether lumbar drainage controls intracranial pressure in rebleeding cases because it was not calculated during treatment. Spinal drainage may promote bleeding when an intraoperative rebleeding occurs. However, massive bleeding did not occur in most cases, and we considered this to indicate that spinal drainage controlled the intracranial pressure until the bleeding stopped. The reason for not inserting the lumbar drainage catheter was not described in 19 patients in the LD (−) group. Patients who had rebleeding between lumbar drainage and endovascular

Table 6 Odds ratio of adverse events by propensity score matching and inverse probability of treatment weighting

	Propensity matched score		IPTW	
	OR (95% CI)	P value	OR (95% CI)	P value
Modified Rankin Scale 0–2 at discharge	1.378 (0.555–3.421)	0.489	1.824 (0.546–6.094)	0.329
Rebleeding	1.000 (0.134–7.451)	1.000	1.089 (0.210–5.652)	0.919
Symptomatic vasospasm	1.028 (0.274–3.854)	0.968	1.545 (0.457–5.217)	0.484
Hydrocephalus	1.935 (0.631–5.934)	0.248	1.267 (0.314–5.122)	0.740
Meningitis	3.152 (0.773-12.851)	0.109	3.560 (0.898–14.113)	0.071

IPTW: inverse probability of treatment weighting; OR: odds ratio

treatment were not suitable for endovascular treatment, and thus, they were not enrolled in the present study. Further randomized controlled trials involving a larger number of patients are required to confirm the safety of lumbar drainage.

Conclusion

Lumbar drainage before endovascular treatment for ruptured intracranial aneurysms is a safe procedure that does not increase the incidence of rebleeding.

Acknowledgments

The authors thank Editage (www.editage.jp) for editing the English language.

Funding

This work was supported by JSPS KAKENHI Grant Number JP20242763.

Author Contributions

TT (1st), YI, MS, MH, and YM contributed to the study conception and design. TT (1st) and YI, KH, MS, TT (5th), AM, MH, WT, NK, YN, KS, YM, and EI acquired the data. TT (1st) and YI analyzed and interpreted the data. TT (1st) and YI drafted the manuscript. KH, MS, TT (5th), AM, MH, WT, NK, YN, and KS revised and edited the manuscript.

Disclosure Statement

The authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article.

References

1) Kwon OY, Kim YJ, Kim YJ, et al. The utility and benefits of external lumbar CSF drainage after endovascular coiling on aneurysmal subarachnoid hemorrhage. *J Korean Neurosurg Soc* 2008; 43: 281–287.
2) Wolf S, Mielke D, Barner C, et al. Effectiveness of lumbar cerebrospinal fluid drain among patients with aneurysmal subarachnoid hemorrhage. A randomized clinical trial. *JAMA Neurol* 2023; 80: 833–842.
3) Al-Tamimi YZ, Bhargava D, Feltbower RG, et al. Lumbar drainage of cerebrospinal fluid after aneurysmal subarachnoid hemorrhage: A prospective, randomized, controlled trial (LUMAS). *Stroke* 2012; 43: 677–682.
4) Klimo P Jr., Kestle JRW, MacDonald JD, et al. Marked reduction of cerebral vasospasm with lumbar drainage of cerebrospinal fluid after subarachnoid hemorrhage. *J Neurosurg* 2004; 100: 215–224.
5) Panni P, Fugate JE, Rabinstein AA, et al. Lumbar drainage and delayer cerebral ischemia in aneurysmal subarachnoid hemorrhage: a systematic review. *J Neurosurg Sci* 2017; 61: 665–672.
6) Elijovich L, Higashida RT, Lawton MT, et al. Predictors and outcomes of intraprocedural rupture in patients treated for ruptured intracranial aneurysms: The CARAT Study. *Stroke* 2008; 39: 1501–1506.
7) Stapleton CJ, Walcott BP, Butler WE, et al. Neurological outcomes following intraprocedural rerupture during coil embolization of ruptured intracranial aneurysms. *J Neurosurg* 2015; 122: 128–135.
8) Murad A, Ghostine S, Colohan ART. A case for further investigating the use of controlled lumbar cerebrospinal fluid drainage for the control of intracranial pressure. *World Neurosurg* 2012; 77: 160–165.
9) Bloch J, Regli L. Brain stem and cerebellar dysfunction after lumbar spinal fluid drainage: case report. *J Neurol Neurosurg Psychiatry* 2003; 74: 992–994.
10) Motoyama Y, Nakajima T, Takamura Y, et al. Risk of brain herniation after craniotomy with lumbar spinal drainage: a propensity score analysis. *J Neurosurg* 2018; 130: 1710–1720.

11) Hellingman C, van den Bergh WM, Beijer I, et al. Risk of rebleeding after treatment of acute hydrocephalus in patients with aneurysmal subarachnoid hemorrhage. *Stroke* 2007; 38: 96–99.
12) Panni P, Donofrio C, Barzaghi L, et al. Safety and feasibility of lumbar drainage in the management of poor grade aneurysmal subarachnoid hemorrhage. *J Clin Neurosci* 2019; 64: 64–70.
13) Weir B, Macdonald RL, Stoodley M. Etiology of cerebral vasospasm. *Acta Neurochir Suppl* 1999; 72: 27–46.
14) Yong CI, Hwang SK, Kim SH. The role of lumbar drainage to prevent shunt-dependent Hydrocephalus after coil embolization for aneurysmal subarachnoid Hemorrhage in good-grade patients. *J Korean Neurosurg Soc* 2010; 48: 480–484.
15) Maeda Y, Shirao S, Yoneda H, et al. Comparison of lumbar drainage and external ventricular drainage for clearance of subarachnoid clots after Guglielmi detachable coil embolization for aneurysmal subarachnoid hemorrhage. *Clin Neurol Neurosurg* 2013; 115: 965–970.
16) Horie N, Sato S, Kaminogo M, et al. Impact of perioperative aneurysm rebleeding after subarachnoid hemorrhage. *J Neurosurg* 2019; 133: 1401–1410.
17) Sandalcioglu IE, Schoch B, Regel JP, et al. Does intraoperative aneurysm rupture influence outcome? Analysis of 169 patients. *Clin Neurol Neurosurg* 2004; 106: 88–92.
18) Schramm J, Cornelia Cedzich C. Outcome and management of intraoperative aneurysm rupture. *Surg Neurol* 1993; 40: 26–30.
19) Connolly ES Jr., Kader AA, Frazzini VI, et al. The safety of intraoperative lumbar subarachnoid drainage for acutely ruptured intracranial aneurysm: Technical note. *Surg Neurol* 1997; 48: 338–342; discussion, 342–344.
20) Ochiai H, Yamakawa Y. Continuous lumbar drainage for the preoperative management of patients with aneurysmal subarachnoid hemorrhage. *Neurol Med Chir (Tokyo)* 2001; 41: 576–580; discussion, 581.
21) Ruijs ACJ, Dirven CMF, Algra A, et al. The risk of rebleeding after external lumbar drainage in patients with untreated ruptured cerebral aneurysms. *Acta Neurochir (Wien)* 2005; 147: 1157–1161 discussion, 1161–1162.
22) Liang H, Zhang L, Gao A, et al. Risk factors for infections related to lumbar drainage in spontaneous subarachnoid hemorrhage. *Neurocrit Care* 2016; 25: 243–249.
23) Xie Z, Hu X, Zan X, et al. Predictors of shunt-dependent hydrocephalus after aneurysmal subarachnoid hemorrhage? A systematic review and meta-analysis. *World Neurosurg* 2017; 106: 844–860.e6.

Cureus. 2024 Jan 25;16(1):e52959. doi: 10.7759/cureus.52959. eCollection 2024 Jan.

Effect of Preoperative Oral Antibiotics and Mechanical Bowel Preparations on the Intestinal Flora of Patients Undergoing Laparoscopic Colorectal Cancer Surgery: A Single-Center Prospective Pilot Study

Sho Fujiwara [1 2], Kenji Kaino [1], Kazuki Iseya [1 3], Nozomi Koyamada [1], Tatsuya Nakano [1 4]

Affiliations

PMID: 38406026 PMCID: PMC10894073 DOI: 10.7759/cureus.52959

Abstract

Introduction: In the last few decades, considerable progress has been made in controlling surgical site infections (SSIs) using a combination of mechanical and oral antibiotic bowel preparation. However, the number of bacteria present after bowel preparation has not been clarified. In this study, we investigated the bacterial cultures of intestinal fluid samples from patients undergoing laparoscopic surgery for colorectal cancer after preoperative bowel preparation.

Methods: This prospective observational study was designed as a pilot study at a single center. We enrolled 25 consecutive patients who underwent laparoscopic surgery for colorectal cancer between March 2021 and February 2022 at our institution.

Results: The rate of bacterial culture positivity was 56.0%. The most abundant bacterium was *Escherichia coli* (44.0%). The positivity rates for *E. coli* on the right and left sides were 54.5% and 35.7%, respectively ($P = 0.60$). Moreover, there was a significant relationship between a low American Society of Anesthesiologists Physical Status score and *E. coli* positivity on the right side ($P = 0.031$). In the left-sided group, female sex and large tumor size were significantly associated with *E. coli* positivity ($P = 0.036$ and 0.049, respectively). Superficial SSI occurred in the patient in the left-sided group, but *E. coli* was negative.

Conclusion: This study emphasizes the importance of understanding intestinal fluid contamination and its relationship to infection risk. Future prospective multicenter studies should be conducted to determine the association between intestinal bacteria and different types of preoperative preparation.

Keywords: bowel preparation; colorectal cancer; intracorporeal anastomosis; laparoscopic surgery; surgical site infection.

Figures

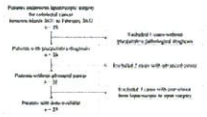


Figure 1. Flow diagram depicting the selection...

Related information

MedGen

LinkOut - more resources

Full Text Sources

Europe PubMed Central

PubMed Central

Heliyon 10 (2024) e25751

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Exploring the relationship between plasma substance P and glottal incompetence in the elderly☆

Koichi Tsunoda [a,b,*], Toyota Ishii [c], Hiroyuki Kuroda [d], Hiroaki Nakatani [e], Masaru Tateda [f], Sawako Masuda [g], Tetsuya Takiguchi [h], Fujinobu Tanaka [i], Hayato Misawa [j], Masamitsu Senarita [k], Mihiro Takazawa [b], Kenji Itoh [b], Thomas Baer [l]

[a] Department of Otolaryngology, NHO (National Hospital Organization) Tokyo Medical Center, Tokyo, Japan
[b] Department of Artificial Organs & Medical Creations, NHO (National Hospital Organization) Tokyo Medical Center, Tokyo, Japan
[c] Department of Otolaryngology, NHO (National Hospital Organization) Sagamihara Hospital, Kanagawa, Japan
[d] Department of Otolaryngology, NHO (National Hospital Organization) Kobe Medical Center, Hyogo, Japan
[e] Department of Otolaryngology, NHO (National Hospital Organization) Fukuyama Medical Center, Hiroshima, Japan
[f] Department of Otolaryngology, NHO (National Hospital Organization) Sendai Medical Center, Miyagi, Japan
[g] Department of Otolaryngology, NHO (National Hospital Organization) Mie Hospital, Mie, Japan
[h] Department of Otolaryngology, NHO (National Hospital Organization) Kanazawa Medical Center, Ishikawa, Japan
[i] Department of Otolaryngology, NHO (National Hospital Organization) Nagasaki Medical Center, Nagasaki, Japan
[j] Department of Otolaryngology, NHO (National Hospital Organization) Nagoya Medical Center, Aichi, Japan
[k] Department of Otolaryngology, NHO (National Hospital Organization) Mito Medical Center, Ibaraki, Japan
[l] University of Cambridge, UK

A B S T R A C T

We speculated that increased blood-plasma levels of Substance P may serve as an indicator of glottal incompetence, which is usually indicated by reduced maximum phonation time. We performed an initial study to test the plausibility of this hypothesis. Patients with dysphonia caused by glottal incompetence were asked to perform vocal exercises for six months to reduce glottal incompetence and we compared the plasma concentration of Substance P before and after the vocal exercise to detect correlation between maximum phonation time and plasma concentration of Substance P. Based on the results, we further hypothesized that patients exhibiting dysphonia with maximum phonation time less than 14 s, in particular less than 10 sec, caused by glottal incompetence may have increased plasma concentration of Substance P with the results of elevated thresholds of cough reflex associated with subclinical aspiration in airways. Further study is needed on patients with decreased Substance P levels, with low scores on Activities of Daily Living and who are hospitalized with aspiration pneumonia.

1. Introduction

Coughing results from a reflex to protect against airway irritation [1]. There is a marked depression of cough reflex in patients with aspiration pneumonia [2]. In the elderly with glottal incompetence, cough effectiveness is reduced at the time in life of greatest risk of

☆ The research was approved by NHO (National Hospital Organization) Research Ethics Committee on Nov. 11, 2018 (approval #H30-1116001).

* Corresponding author. Chairman of Artificial Organs & Medical Creations, Department of Otolaryngology NHO (National Hospital Organization) Tokyo Medical Center 2-5-1 Higashigaoka, Meguro-ku, Tokyo, 152-8902, Japan.
E-mail address: koichi.tsunoda@kankakuki.jp (K. Tsunoda).

https://doi.org/10.1016/j.heliyon.2024.e25751
Received 13 September 2023; Received in revised form 28 January 2024; Accepted 1 February 2024
Available online 8 February 2024

pneumonia [3]. Glottal incompetence reduces ability to achieve effective glottal closure, resulting in frequent laryngeal ingress and difficulty in expectoration owing to subglottic pressure that is insufficient to remove aspirated substances through coughing [4,5] It also causes breathy hoarseness [3].

In the population at risk of aspiration pneumonia, it is necessary to screen for glottal incompetence, especially because of its effect on cough effectiveness. Measurement of maximum phonation time (MPT), which is how long an individual can maintain phonation of the vowel 'a' without taking a breath, provides a straightforward bedside indication for this purpose [6]. This was exemplified in the work of our clinical research group. A novel self-controlled exercise regime (NHO-exercise) was developed to strengthen the laryngeal adduction muscles and improve glottal competence. In a randomized-control-trial study with 543 patients, 6-months of NHO-exercise was found to significantly increase MPT and reduce the incidence of aspiration pneumonia [3].

Screening for glottal competence using MPT or direct fiberoptic observation of glottal closure is somewhat subjective. It is necessary to establish an objective indicator. We considered that blood sampling of Substance P (SP) may meet this need.

SP was the first neuropeptide found to be a sensory neurotransmitter [7] in the laryngeal afferent system [8]. SP is released by sensory nerves and contributes to the cough reflex. The SP released in response to stimuli may mediate cough and neutral endopeptidase may have a role in modulating SP-induced effects [9]. Coughing during ACE-inhibitor therapy to prevent aspiration is due to an increased inflammatory state in the airways for preventing reduction of SP [10]. Some reports suggested that a reduction of SP concentration in serum [10–14], sputum [15], or saliva [16,17] may be a useful predictive marker for the increased risk of developing aspiration or aspiration pneumonia. Therefore, a relationship between SP, cough reflex, and the risk of aspiration pneumonia has been established. We speculate that increased blood-plasma levels of SP may serve more generally as an indicator of glottal incompetence. This study was designed to test the plausibility of this hypothesis.

In this study, patients with breathy hoarseness (dysphonia) caused by glottal incompetence were asked to perform NHO-exercise for six months. We measured the plasma concentration of SP and the MPT before and after the NHO-exercise to test for correlation between their changes. It was hoped that glottal incompetence, which could lead to early-stage aspiration pneumonia and is expected to occur in more patients in the future, could be reliably detected by a passive indicator using blood sampling of SP in combination with MPT [14]. This would greatly help to identify cases at risk of aspiration.

2. Material and method

This study was focused on elderly patients over the age of 65. To establish the distribution of normal MPT values in this population, MPT measurements were obtained from 131 patients from this age group who had no glottal incompetence according to diagnosis by screening tests. The distribution was found to have a mean of 22.66 and standard deviation of 6.84. Approximating the distribution as Gaussian, its cumulative probability exceeds 2.5% at MPT of 10, 10% at MPT of 14, and 33% at MPT of 20. These values were used to categorize the level of glottal incompetence of study participants according to their MPT: severe below 10, moderate from 10 to 13, and mild from 14 to 20.

Based on data in the protocol paper [14] associated with this study, the recommended sample size of participants was calculated using G*Power (version 3.1) [18] with power = 0.95, $\alpha = 0.05$ and $d = 0.5$, producing an estimate of $n = 105$. The study included 122 patients (outpatients), all over the age of 65, each of whom provided informed written consent. Each of them was examined by laryngeal fiberscopy for diagnosis.

The test group comprised 86 patients who complained of dysphonia (with breathy hoarseness, cough, or dysphagia) and had age-related glottal incompetence due to vocal-fold atrophy with MPT of 4–13 s (sec). Patients unable to perform the NHO exercise and to schedule a follow-up appointment 6 months after the first session were excluded. Any patient with glottal incompetence caused by pathology other than age, including heart failure or pulmonary diseases and rheumatism, was also excluded as was any patient receiving angiotensin-converting-enzyme (ACE) inhibitor therapy, which affects the cough reflex and increases the plasma level of SP. The remaining 36 patients formed the control group. Each had mild MPT in the range of 14–19 s but complained of subjectively experienced dysphonia. Inclusion criteria were similar to those for the test group, except that previous history of otolaryngological conditions including glottal incompetence were exclusionary. The patients with low MPT were divided into two groups based on MPT (severe and moderate). Thus, the respective patient groups were; *severe* [4–9 s] ($n = 41$, mean age 78.55 ± 6.57 years, 68.29% male), *moderate* [10–13 s] ($n = 45$, mean age 79.71 ± 6.91 years, 57.78% male) and *mild* [14–19 s] ($n = 36$, mean age 79.83 ± 5.23 years, 52.78% male). The number in each of the three groups was considered to be large enough to provide statistically reliable results.

During the study, participants performed NHO-exercise every day for 6 months. The exercise regime has been reported in the literature [3]. Before beginning the exercises, patients are given a brochure explaining the anatomy and physiology of the glottis and the role of glottal closure and the cause of failure of glottal closure due to aging, how age-related vocal fold atrophy leads to hoarseness and aspiration, recommendations on how to modulated the voice while speaking and singing, and how to prevent aspiration, They are then shown a video on how to perform NHO-exercise with simultaneous explanation by a physician or speech therapist and they are instructed to perform the exercise in the following manner:

1 Sit on a chair and grip the sides of the seat with both hands.
2. While saying each number from 1 to 10 aloud, pull firmly on both sides of the seat and then relax and inhale naturally before saying the next number.
3. Repeat this exercise for a total of 2 sets in both the morning and evening, for a total of 4 sets per day.

Before and after the 6-month NHO-exercise, MPT was measured using standard procedures and blood sampling was performed for

measurement of plasma concentration of SP. Blood sampling was performed with EDTA-2Na + aprotinin (NP-EA0305-123D, SRL, Tokyo, Japan). Samples were centrifuged at 0 °C. Plasma SP was stored at −80 °C until assay and analyzed within 2-weeks. Age, disease, MPT, and the difference of measured plasma concentration of SP before and after the NHO-exercise of those who performed all the steps were analyzed at SRL Kitakanto Laboratory Inc., Tokyo, Japan, who were certified to perform these procedures. Relationships of plasma concentration of SP with MPT and the NHO-exercise were analyzed using a two-factor repeated measures mixed ANOVA. The significant differences were specified by post hoc t tests if needed. The statistical analysis was performed in R Studio v.2023.06.1–524 using R v.4.3.2.

The research was approved by National Hospital Organization Research Ethics Committee (approval #H30-1116001). This study was registered in the Clinical Trial Registry (UMIN-CTR) as UMIN000035080. The detail was published as a protocol [14].

3. Result

The effects of the NHO-exercise on SP and MPT are shown in Figure-1. In the two groups (severe and moderate) with glottal incompetence and MPT of less than 14 sec. with dysphonia and choking/cough, plasma concentration of SP decreased after the NHO-exercise and moved below 3000 pg/ml. The plasma concentration of SP never increased and MPTs always increased after the exercise. The results showed significant relation and correlation between decrease of plasma concentration of SP and increase of MPT (Fig. 1a and b). Furthermore, the severe and moderate groups of patients with initial MPTs less than 14 s had higher average levels of plasma concentration of SP than those of the mild group with initial MPTs between 14 and 19 s (3099 vs 2539, $p = 0.011$, $d = 0.415$). Both severe + moderate and mild MPT groups, especially the combination group showed decreased average plasma concentration of SP that moved toward 2500 pg/ml after the NHO-exercise (2706 vs 2378, $p = 0.188$, $\eta p2 = 0.235$).

As shown in Fig. 1a, the plasma concentration of SP was over 3000 pg/ml in both the severe and moderate groups with less-than-14-sec MPTs but significantly fell below 3000 pg/ml after the exercise ($\Delta = -393$ pg/ml; $p = 0.012$; $d = 0.283$). This effect of exercise was more prominent in the moderate group with 10-to-13-sec MPTs ($\Delta = -563$ pg/ml; $p = 0.004$; $d = 0.435$) compared to the severe group with 4-to-9-sec MPTs ($\Delta = -207$ pg/ml; $p = 0.468$; $d = 0.121$). In the mild group with 14-to-19-sec MPTs, the SP value was around 2500 pg/ml before and after the NHO-exercise ($\Delta = -161$ pg/ml; $p = 0.420$; $d = 0.162$). Using the results of our study for the short and moderate cases whose MPTs recovered to more than 14 s after NHO-exercise as representative of those for healthy elderly individuals, the normal level of plasma concentration of SP may be estimated to be 2512 ± 1117 pg/ml (95% CL: 2269–2755 pg/ml).

4. Discussion

Average SP levels are a bit above 3000 pg/ml before exercise regardless of how bad is the average pre-exercise MPT, as long as that MPT is below the normal range (14 s or above), at least without pneumonia. The pattern of changes due to the 6-month NHO-exercise for the three different MPT groups is very different for the SP levels compared to the MPT values. This makes it clear that the information given by SP level is different to that given by MPT so, the utility of SP is not only that it is more objective than measuring MPT.

Increased levels of SP in plasma are associated with persistent cough in humans and might be related to airway sensitivity in asthmatic cough [13]. Clinically, a reduction of SP concentration in the sputum in the elderly is associated with a reduced cough reflex [10].

We could not get the results we expected in the initial study based on our speculation. After this, we further hypothesized that patients exhibiting dysphonia with MPTs less than 14 s, in particular less than 10 sec., caused by glottal incompetence may have

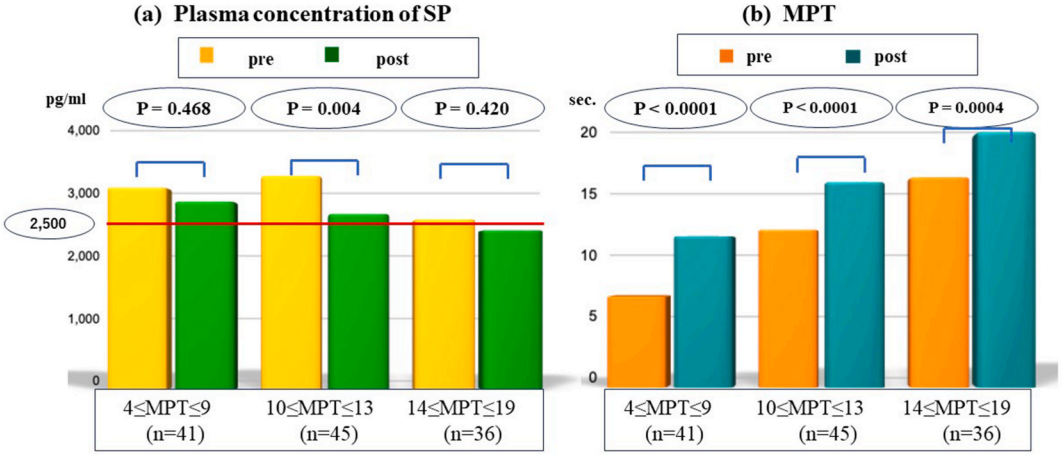


Fig. 1. Mean of plasma concentration of SP and MPT of pre/post NHO-exercise. a) Mean of plasma concentration of SP b) Mean of plasma concentration of MPT.

increased plasma concentration of SP as an inflammatory change with the results of elevated thresholds of cough reflex associated with subclinical aspiration in airways. However, in some cases, patients who have subclinical aspiration never exhibit the aspiration with/without cough. Therefore, if patients whose MPT has decreased to less than 10 sec complain of hoarseness and have elevation of SP to around 3000 pg/ml, or who cough frequently, this might be a sign of the beginning of aspiration (Fig. 2).

It is necessary to maintain satisfactory plasma concentration of SP as neurotransmitter to induce the cough reflex, even with elevated threshold, to prevent aspiration. That is why plasma concentration of SP increases physiologically to compensate for pathological elevated threshold. Also, we hypothesize that improvement of glottal closure with NHO-exercise helps to decrease pathological elevated threshold for cough reflex. Then, physiologically elevated SP based on homeostasis, might return to normal. In cases of severe breathy hoarseness with frequent cough, and with increased plasma concentration of SP, appropriate treatment would be necessary to prevent glottal incompetence. That is the appropriate timing for preventing aspiration pneumonia.

5. Conclusion

Our finding is that raised SP is a sign that the cough is weak and the threshold for the cough reflex needs to be reduced while lowering of the SP level is a sign that the cough has become more effective. Glottal closure is necessary for that cough. In other words, the measurement of substance P in combination with fiberscopic laryngeal examination and MPT is the first step in a study that may lead to an objective biomarker to detect the risk of aspiration pneumonia and prevent it in advance.

Further study is needed on patients with decreased SP levels, with low scores on Activities of Daily Living indexes, and those who are hospitalized for the treatment of aspiration pneumonia to develop our hypothesis, also, comparing the SP before and after intervention with ACE-inhibitor and Pharyngeal Electrical Stimulation (PES) [16] for such cases, is awaited.

Additional information

No additional information is available for this paper.

Data availability

Data will be made on request.

CRediT authorship contribution statement

Koichi Tsunoda: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization. **Toyota Ishii:** Data curation. **Hiroyuki Kuroda:** Data curation. **Hiroaki Nakatani:** Data curation. **Masaru Tateda:** Data curation. **Sawako Masuda:** Data curation. **Tetsuya Takiguchi:** Data curation. **Fujinobu Tanaka:** Data curation. **Hayato Misawa:** Data curation. **Masamitsu Senarita:** Data curation. **Mihiro**

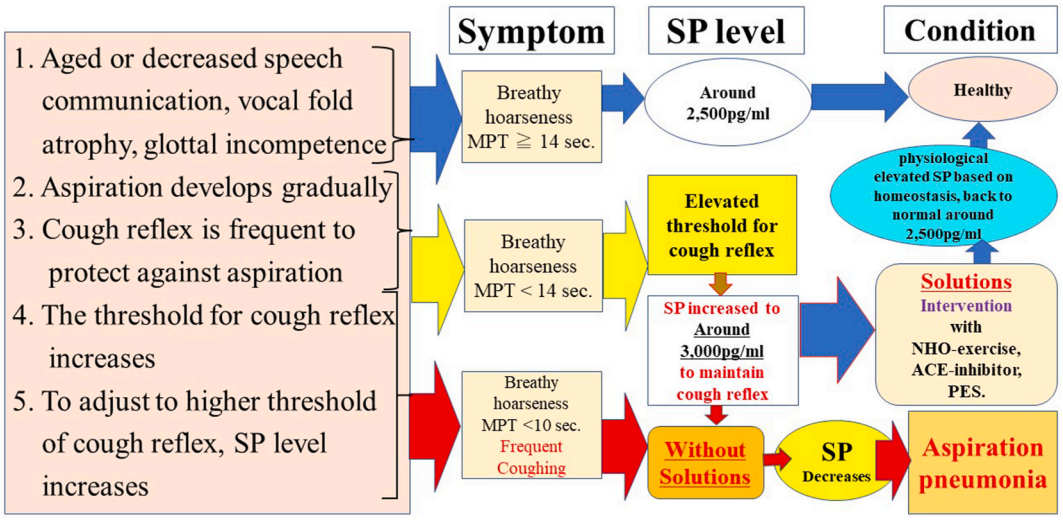


Fig. 2. Relation of aspiration and SP level. Roles of increase or decrease of SP, 1. By aged or decrease speech communication, glottal incompetence due to vocal fold atrophy occurs, 2. Aspiration gradually occurs, 3. To prevent aspiration, the cough reflex occurs frequently, 4. The threshold for cough reflex might be increasing pathologically, 5. To induce cough reflex, the plasma concentration of SP increases physiologically as a neurotransmitter, *With the suitable solution, SP would decrease and moved toward around 2500 pg/ml * Without any solution, SP decreases and aspiration pneumonia would be occurred caused by subclinical aspiration.

Takazawa: Writing – review & editing, Investigation. **Kenji Itoh:** Writing – review & editing, Software, Methodology, Formal analysis. **Thomas Baer:** Writing – review & editing, Software, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank to National Hospital Organization Japan Network Joint Research, and special thanks to Dr. Kazuhiro Araki and Dr. Yukihiko Momiyama. We would like to express our sincere admiration to Dr. Masanori Otsuka, who was the first to discover the substance P as neurotransmitter.

References

[1] SB Mazzone, An overview of the sensory receptors regulating cough, Cough. 1:2 (2005 Aug 4), https://doi.org/10.1186/1745-9974-1-2.
[2] K. Sekizawa, Y. Ujiie, S. Itabashi, H. Sasaki, T. Takishima, Lack of cough reflex in aspiration pneumonia, Lancet 335 (8699) (1990) 1228–1229, https://doi.org/10.1016/0140-6736(90)92758-a.
[3] Y. Fujimaki, K. Tsunoda, R. Kobayashi, et al., Research Group for Aspiration Pneumonia, National Hospital Organization, Japan. Independent exercise for glottal incompetence to improve vocal problems and prevent aspiration pneumonia in the elderly: a randomized controlled trial, Clin Rehabil (8) (2017) 1049–1056, https://doi.org/10.1177/0269215516673208.
[4] D. Britton, J.O. Benditt, A.L. Merati, et al., Associations between laryngeal and cough dysfunction in motor neuron disease with bulbar involvement, Dysphagia 29 (6) (2014) 637–646, https://doi.org/10.1007/s00455-014-9554-5.
[5] T. Pitts, Airway protective mechanisms, Lung 192 (1) (2014) 27–31, https://doi.org/10.1007/s00408-013-9540-y.
[6] Z. Zhou, F. Vincent, J.Y. Salle, M.T. Antonini, V. Aliamus, J.C. Daviet, Acute stroke phase voluntary cough and correlation with maximum phonation time, Am J Phys Med Rehabil 91 (6) (2012) 494–500, https://doi.org/10.1097/PHM.0b013e31824fa66a.
[7] M. Otsuka, S. Konishi, Release of substance P-like immunoreactivity from isolated spinal cord of newborn rat, Nature 264 (5581) (1976) 83–84, https://doi.org/10.1038/264083a0.
[8] Y. Hisa, S. Koike, N. Tadaki, H. Bamba, K. Shogaki, T. Uno, Neurotransmitters and neuromodulators involved in laryngeal innervation, Ann Otol Rhinol Laryngol Suppl 178 (1999) 3–14, https://doi.org/10.1177/00034894991080s702.
[9] K. Sekizawa, Y.X. Jia, T. Ebihara, Y. Hirose, Y. Hirayama, H. Sasaki, Role of substance P in cough, Pulm Pharmacol 9 (5–6) (1996) 323–328, https://doi.org/10.1006/pulp.1996.0042.
[10] B.R. Lindgren, U. Rosenqvist, T. Ekström, R. Grönneberg, B.E. Karlberg, R.G. Andersson, Increased bronchial reactivity and potentiated skin responses in hypertensive subjects suffering from coughs during ACE-inhibitor therapy, Chest 95 (6) (1989) 1225–1230, https://doi.org/10.1378/chest.95.6.1225.
[11] T. Nagamine, Serum substance P levels in patients with chronic schizophrenia treated with typical or atypical antipsychotics, Neuropsychiatr Dis Treat 4 (1) (2008) 289–294, https://doi.org/10.2147/ndt.s2367.
[12] Y. Nakamori, T. Yasuda, H. Imamoto, et al., Assessment of risk of aspiration pneumonia after esophageal cancer surgery by measuring substance P concentration in blood saliva and citrate-induced cough reflex threshold test, Med J Kinki Univ 35 (1) (2010) 31–40 (in Japanese).
[13] K. Otsuka, A. Niimi, H. Matsumoto, et al., Plasma substance P levels in patients with persistent cough, Respiration 82 (5) (2011) 431–438, https://doi.org/10.1159/000330419.
[14] K. Tsunoda, S. Hashimoto, H. Kuroda, T. Ishii, M. Takazawa, Exploring the relation between glottal closure and plasma substance P: a study protocol, Tohoku J Exp Med 249 (4) (2019) 237–240, https://doi.org/10.1620/tjem.249.237.
[15] T. Nakagawa, T. Ohrui, K. Sekizawa, H. Sasaki, Sputum substance P in aspiration pneumonia, Lancet 345 (8962) (1995) 1447, https://doi.org/10.1016/s0140-6736(95)92638-0.
[16] A. Muhle P, S. Suntrup-Krueger, S. Bittner, et al., Increase of substance P concentration in saliva after pharyngeal electrical stimulation in severely dysphagic stroke patients - an indicator of decannulation success? Neurosignals 25 (1) (2017) 74–87, https://doi.org/10.1159/000482002.
[17] J.B. Schröder, T. Marian, I. Claus, et al., Substance P saliva reduction predicts pharyngeal dysphagia in Parkinson's disease, Front Neurol 10 (2019) 386, https://doi.org/10.3389/fneur.2019.00386.
[18] F. Faul, E. Erdfelder, A. Buchner, A.G. Lang, Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses, Behav Res Methods 41 (4) (2009) 1149–1160, https://doi.org/10.3758/BRM.41.4.1149.

Optimal Limb Position for the Stress Ultrasound Evaluation of Elbow Valgus Laxity in Baseball Players

Ryuhei Michinobu,* MD, Takeshi Ogawa,† MD, PhD, Yuichi Yoshii,‡§ MD, PhD, Akira Ikumi,* MD, PhD, Kazuhiro Ikeda,* MD, Hiromitsu Tsuge,* MD, Shotaro Teruya,* MD, Yuki Hara,‖ MD, PhD, and Masashi Yamazaki,* MD, PhD

Investigation performed at Faculty of Medicine, Orthopedic Surgery, Tsukuba University, Ibaraki, Japan

Background: The optimal limb position during stress ultrasound (SUS) evaluation of elbow valgus laxity has not been standardized.

Purpose: To compare 2 elbow positions (at 90° and 30° of flexion) and report which position method better represents the increased valgus laxity characteristics of baseball players.

Study Design: Controlled laboratory study.

Methods: Eighteen college baseball players with no history of elbow pain or elbow disorders who belonged to a college baseball club between April and November 2021 participated in this study. The medial elbow joint space (MEJS) was recorded by ultrasonography at rest and under valgus stress, and the difference in MEJS between the conditions was considered the valgus laxity. For all participants, the MEJS was recorded at 90° and 30° of elbow flexion. In the 90° of flexion position, the participant was positioned in the supine position with abduction and external rotation of the shoulder, and 2.5 kgf of valgus stress was applied proximally to the wrist. In the 30° of flexion position, the participant was positioned in the sitting position with abduction and external rotation of the shoulder, and 3.0 kgf of valgus stress was applied to the ulnar head. Valgus laxity on the throwing and nonthrowing sides was compared between the 2 elbow positions using paired *t* tests or Wilcoxon signed-rank tests after checking the normality.

Results: There was a significant difference in valgus laxity on the throwing side between the 90° and 30° of flexion positions (1.9 vs 1.1 mm, respectively; $P = .002$), whereas no significant difference between positions was seen on the nonthrowing side ($P = .06$).

Conclusion: SUS with the elbow flexed at 90° more clearly detected valgus laxity in the study participants than the 30° of flexion position.

Clinical Relevance: The quantitative evaluation of valgus laxity is important for baseball players to assess the risk of ulnar collateral ligament injury.

Keywords: stress ultrasound; evaluation; elbow valgus laxity; baseball players

In the baseball throwing motion, a large valgus stress is loaded on the throwing elbow joint from the late cocking phase to the acceleration phase.[7] This valgus stress is repeatedly loaded on the elbow with each throw, resulting in valgus laxity.[5,6,8,9,11,13,20] Excessive increase in valgus laxity is a reported risk factor for ulnar collateral ligament (UCL) injury among baseball players.[22] Hence, quantitative evaluation of valgus laxity is important.

Stress radiography has been commonly used to evaluate valgus laxity; however, the use of stress ultrasound (SUS) is being increasingly reported.[2,4,10,11,18,20,22,23] Nevertheless, the limb position for SUS has not been standardized. Ciccotti et al[4] evaluated valgus laxity in baseball players in the sitting position with 30° of elbow flexion and 90° of forearm supination and reported that valgus laxity was

The Orthopaedic Journal of Sports Medicine, 12(2), 23259671231221523
DOI: 10.1177/23259671231221523

greater on the throwing side. Several reports also applied the same limb position method for SUS.[2,18] On the other hand, Sasaki et al[20] evaluated valgus laxity among baseball players in the supine position, 90° of elbow flexion, and forearm neutral rotation; they reported that valgus laxity was greater on the throwing side. This limb position method for SUS was applied in similar studies.[10,11,22,24] However, it is unclear which limb position method more clearly detects valgus laxity in baseball players.

The purpose of this study was to determine the method that can detect valgus laxity in baseball players (90° or 30° of elbow flexion) and to directly compare the 2 positions to report which elbow position method better represents the increased valgus laxity characteristics of baseball players. We hypothesized that the 90° of flexion method, with an elbow joint flexion angle similar to that of the throwing motion, is superior to the 30° of flexion method.

METHODS

Study Population

Eighteen asymptomatic college baseball players (4 pitchers and 14 fielders) who belonged to a college baseball club between April and November 2021 were included in this study. All participants were young men with no history of elbow pain or elbow surgery. The overall age of the participants was 18 to 20 years (mean ± SD, 18.8 ± 0.5 years), and they had 9 to 12 years (mean ± SD, 10.6 ± 0.9 years) of baseball experience. Seventeen were right-handed throwers and 1 was a left-handed thrower. All participants provided written informed consent. This study was conducted in accordance with the Declaration of Helsinki, and the study protocol was approved by the ethics committee of our institution.

Assessment

For all participants, the medial elbow joint space (MEJS), defined as the distance between the medial distal end of the humeral trochlea and the proximal end of the ulnar sublime tubercle, was measured on rest days when they had no athletic training. We used an ultrasound system with an 11-MHz linear array transducer (SONIMAGE MX1; Konica Minolta Japan Inc) and a standard transducer gel to capture images at 90° and 30° of elbow flexion, with the elbow at rest and under valgus stress (Figure 1). Valgus laxity was then calculated as the difference between MEJS values under valgus stress and at rest.

The elbow limb positions were based on previous studies.[4,22] In the 90° of flexion position, the participant was positioned supine on the examination table, with 90° of shoulder abduction, 90° of elbow flexion, and neutral forearm rotation. At rest, the forearm was placed on the examination table to avoid valgus stress of the weight of the forearm. If the shoulder joint external rotation was below 90°, a towel was placed under the forearm. Under valgus stress, the position of the participant was adjusted as shown in Figure 1B. Additionally, a handheld dynamometer (μTas F-2; ANIMA Co, Ltd) adjusted to 2.5 kgf of valgus stress was attached to the proximal wrist joint. In the 30° of flexion position, the participant was positioned in the sitting position, with 90° of shoulder abduction, 30° of elbow flexion, and 90° of forearm supination. A processed acrylic plate was used to maintain the 30° of flexion position of the elbow. Under valgus stress, the same handheld dynamometer adjusted to 3.0 kgf of valgus stress was attached to the ulnar head. An assistant supported the upper arm of the participant with a wooden block to ensure that the valgus stress was properly loaded on the elbow (Figure 2). Forearm rotation and amount of valgus stress in the 2 limb positions were based on previous studies.[4,22]

The fiber direction of the UCL was identified, and a probe was placed parallel to it. Ultrasonographic images of the medial epicondyle, UCL, medial surface of the humeral trochlea, and coronoid process of the ulna depicted in the same field of view were used for MEJS measurements (Figure 3). ImageJ Version 1.53t (National Institutes of Health) was used as the image analysis software.

Three orthopaedic surgeons who use ultrasound in their daily practice (R.M., T.O., and Y.H.) participated in the acquisition of the images, and MEJS measurements were performed by a single orthopaedic surgeon with 10 years of experience (R.M.). We evaluated the intraobserver and interobserver reliability of the MEJS and valgus laxity measurements by using the intraclass correlation coefficient (ICC) on a different set of study participants. Intraobserver reliability data were obtained twice by a single orthopaedic surgeon (R.M.) from 10 elbows with a 1-month interval between measurements. Interobserver reliability data were obtained independently by 2 orthopaedic surgeons (R.M. and S.T.) from 10 elbows each. ICC values were interpreted according to the criteria of Landis and Koch[15]: 0.00 to 0.20, slight; 0.21 to 0.40, fair; 0.41 to 0.60, moderate; 0.61 to 0.80, substantial; and 0.81 to 1.00, almost perfect.

[§]Address correspondence to Yuichi Yoshii, MD, PhD, Tokyo Medical University Ibaraki Medical Center, 3-20-1 Chuo, Ami, Inashiki, Ibaraki 300-0395, Japan (email: yyoshii@tokyo-med.ac.jp).

*Faculty of Medicine, Orthopedic Surgery, Tsukuba University, Ibaraki, Japan.

[†]Orthopedic Surgery, National Hospital Organization Mito Medical Center, Ibaraki, Japan.

[‡]Orthopedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ibaraki, Japan.

[‖]Orthopedic Surgery, National Center of Neurology and Psychiatry, Tokyo, Japan.

Final revision submitted May 21, 2023; accepted July 31, 2023.

The authors declared that they have no conflicts of interest in the authorship and publication of this contribution. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from the University of Tsukuba (reference No. 1517-3).

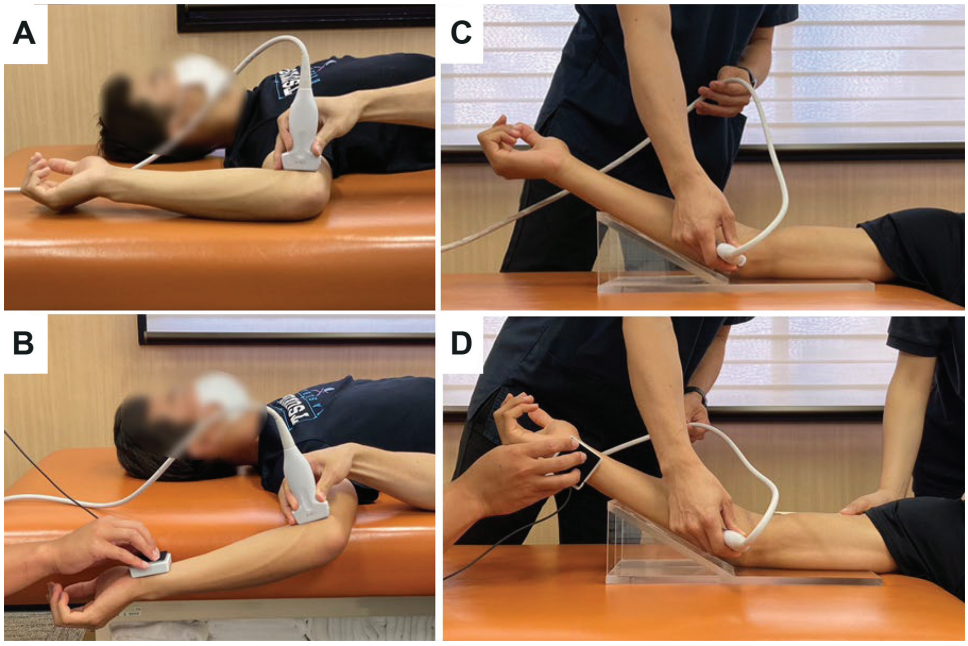


Figure 1. Measurement positions of a study participant: (A) 90° of flexion position with the elbow at rest, (B) 90° of flexion position with the elbow under valgus stress, (C) 30° of flexion position with the elbow at rest, and (D) 30° of flexion position with the elbow under valgus stress.

Figure 2. Photograph showing the 30° of flexion position under valgus stress. An assistant supported the participant's upper arm to ensure that valgus stress was properly loaded on the elbow.

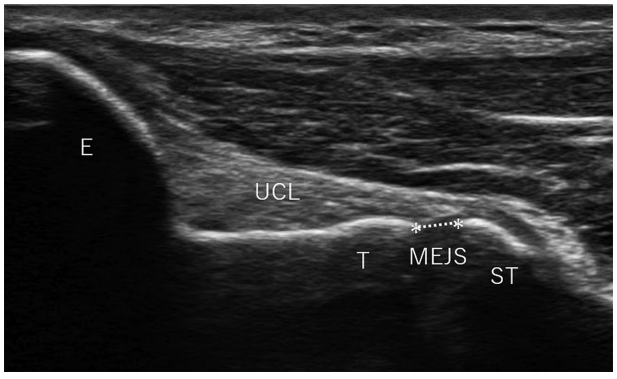


Figure 3. Long-axis image demonstrating the measurement of medial elbow joint space (MEJS; represented by asterisks). E, medial humeral epicondyle; ST, proximal end of the ulnar sublime tubercle; T, medial distal end of the humeral trochlea; UCL, ulnar collateral ligament.

Statistical Analysis

All data were analyzed using SPSS Statistics Version 27 (IBM Corp). First, the normality of each variable was checked using the Shapiro-Wilk test. If the P values of the 2 variables to be compared were $>.05$, they were considered to follow a normal distribution and a paired t test was performed. If either variable had a P value $<.05$, the Wilcoxon signed-rank test was performed. The paired t test or Wilcoxon signed-rank test was performed for comparisons of the measurements between the throwing side and nonthrowing side in both the 30° and 90° of flexion positions. Significant differences were set at a level of 5%. The effect size (ES) was calculated using the Cohen d.

RESULTS

The intraobserver reliability (ICC[1,1]) of the MEJS and valgus laxity measurements ranged from 0.75 to 0.94,

TABLE 1
Intraobserver and Interobserver Reliability of MEJS and Valgus Laxity[a]

	Intraobserver (ICC[1,1])		Interobserver (ICC[2,1])	
	90° of Flexion Position	30° of Flexion Position	90° of Flexion Position	30° of Flexion Position
MEJS				
At rest	0.937	0.75	0.851	0.788
Under valgus stress	0.935	0.872	0.873	0.888
Valgus laxity	0.824	0.843	0.748	0.728

[a]ICC, intraclass correlation coefficient; MEJS, medial elbow joint space.

and the interobserver reliability (ICC[2,1]) ranged from 0.73 to 0.89, indicating substantial to almost perfect agreement for all measurements (Table 1).

Figure 4 shows the MEJS results for each elbow position. At rest, MEJS was 3.9 ± 0.9 mm on the throwing side and 3.8 ± 1.0 mm on the nonthrowing side in the 90° of flexion position. In the 30° of flexion position, it was 4.5 ± 0.8 mm on the throwing side and 4.7 ± 0.7 mm on the nonthrowing side. There was no significant difference between the 2 sides for both limb positions in the resting condition (90° of flexion position: $P = .62$; 30° of flexion position: $P = .56$). Under valgus stress, MEJS was 5.8 ± 1.2 mm on the throwing side and 5.1 ± 1.5 mm on the nonthrowing side in the 90° of flexion position and 5.6 ± 0.7 mm on the throwing side and 5.4 ± 0.7 mm on the nonthrowing side in the 30° of flexion position. In the 90° of flexion position, the MEJS was significantly larger on the throwing side than on the nonthrowing side ($P = .019$; ES = 0.61); whereas in the 30° of flexion position, there was no significant difference between the throwing and nonthrowing sides ($P = .28$).

Valgus laxity was 1.9 ± 0.7 mm on the throwing side and 1.3 ± 0.8 mm on the nonthrowing side in the 90° of flexion position and 1.1 ± 0.7 mm on the throwing side and 0.8 ± 0.6 mm on the nonthrowing side in the 30° of flexion position. In the 90° of flexion position, the difference was significantly larger on the throwing side than on the nonthrowing side ($P = .022$; ES = 0.60), whereas in the 30° of flexion position, there was no difference between the throwing and nonthrowing sides ($P = .13$).

Valgus laxity was significantly larger in the 90° of flexion position than in the 30° of flexion position on the throwing side ($P = .002$; ES = 0.87). Ultrasound images of a representative case are shown in Figure 5. In contrast, there was no significant difference in the valgus laxity between the 30° and 90° of flexion positions on the nonthrowing side ($P = .06$).

DISCUSSION

The major findings from our study were that the elbow valgus laxity on the throwing side was significantly larger (by 0.8 mm) in the 90° of flexion position than in the 30° of flexion position (1.9 vs 1.1 mm; $P = .002$; ES = 0.87), whereas there was no significant difference on the nonthrowing side. Therefore, the valgus laxity of the throwing elbow joint was more clearly detected with a static force in the 90° of flexion position than in the 30° of flexion position among baseball players. This indicates that the 90° of flexion position is the more optimal limb position for SUS evaluation of valgus laxity.

The evaluation of valgus instability at the elbow joint in stress radiography indicated that instability occurs when the medial joint opening exceeds 1 mm.[14] Generally, joint laxity is less severe than joint instability. Considering these factors, a difference of 0.8 mm is a clinically significant value for valgus laxity despite the difference between SUS and stress radiography. Furthermore, the difference between MEJS on the throwing and nonthrowing sides could be detected in the 90° of flexion position but not in the 30° of flexion position. These results lend further support that the 90° of flexion position is the more optimal limb position for the SUS evaluation of valgus laxity.

There are 2 important points regarding this result from a biomechanics perspective. First, the UCL is the primary static stabilizer for valgus stress of the elbow joint.[12,17] Previous research on cadavers indicated that the maximum valgus instability occurs in 90° flexion position when the UCL is dissected.[3] This indicates that the maximum contribution of the UCL as a static stabilizer to valgus stress is in the 90° of flexion position. Morrey and An[16] also reported that in the 90° of flexion position, the contribution of the UCL as a static stabilizer to valgus stress is greater than that of the bone and joints. Second, the maximum valgus stress on the elbow joint in the throwing motion occurs immediately before the maximum external rotation of the shoulder joint and around 90° of elbow flexion.[8] At this instant, the elbow joint is loaded with a large varus torque comparable to the breaking strength of the UCL.[1,8,25] With each pitch, the elbow joint is repeatedly loaded with this stress, resulting in valgus laxity over time.[5,6,8,9,11,13,20] These 2 statements support the hypothesis that the valgus laxity in baseball players is best detected in the 90° of flexion position.

Regarding the difficulty of the procedure in different limb positions, in actual clinical practice, we consider that a reproducible and simple technique such as the 90° of flexion position is visually clear and is considered highly reproducible. In contrast, we consider that the 30° of flexion position is not visually clear and requires time, effort, and measurement with an angle meter to ensure

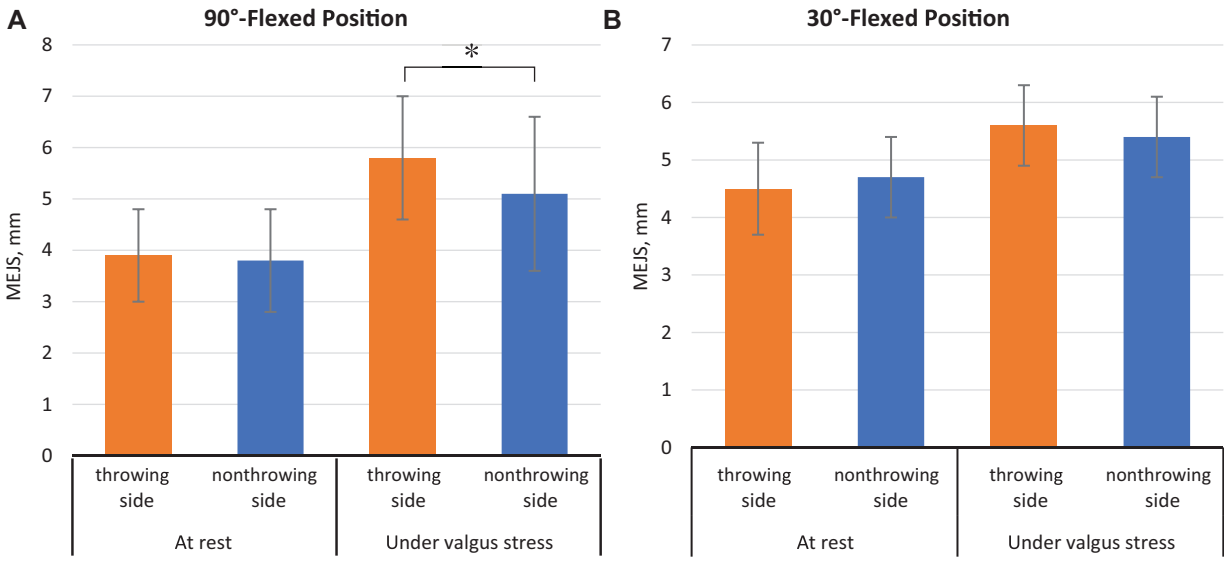


Figure 4. Bar graph illustrating the MEJS for the (A) 90° of flexion position and (B) 30° of flexion position. Error bars indicate standard deviation. *Statistically significant difference between throwing and nonthrowing sides ($P < .05$). MEJS, medial elbow joint space.

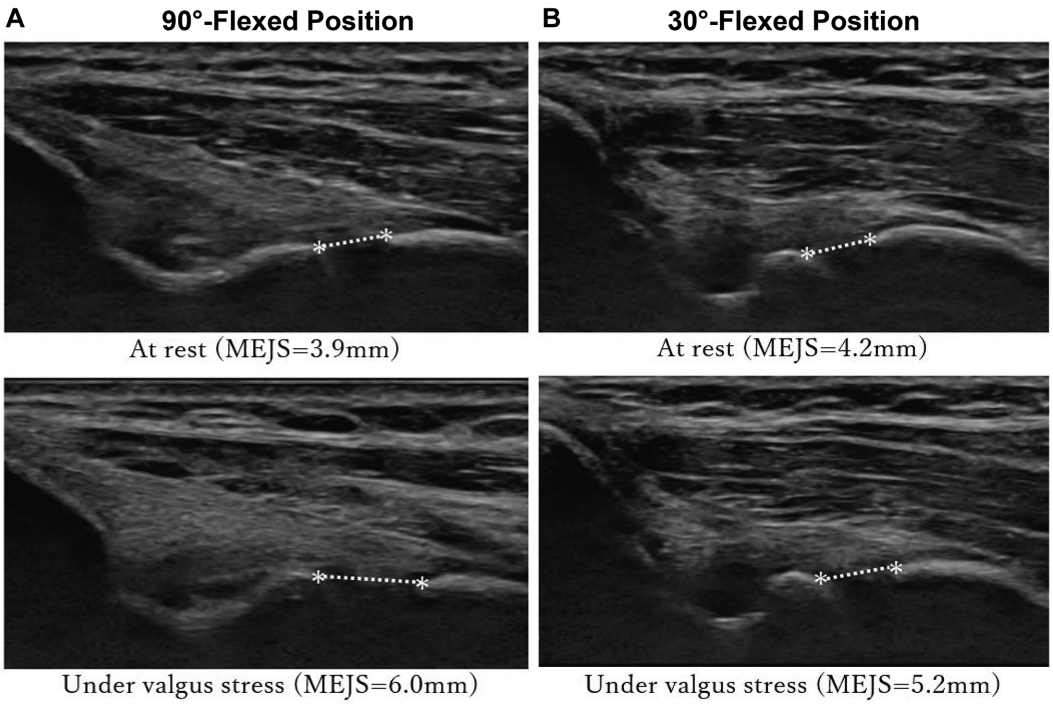


Figure 5. Ultrasound images from a 19-year-old pitcher showing the medial elbow joint space (MEJS; represented by asterisks) on the throwing side for the (A) 90° of flexion and (B) 30° of flexion positions. Valgus laxity was greater at the 90° of flexion position than at the 30° of flexion position (2.1 vs 1.0 mm).

reproducibility. In terms of simplicity, the 90° of flexion position does not require any special equipment. However, in the 30° of flexion position, equipment is needed to properly load valgus stress while maintaining the flexion angle. In this study, a simply processed acrylic plate was used to maintain the 30° of flexion position, but multiple assistants were required to maintain the limb position, such as upper arm support when loading valgus stress. Some studies have indicated using dedicated devices; however, they are expensive and not practical in clinical or sports fields.[2,4,6,10,21,23] We consider that the same can be stated for limb positions other than 30°, such as 45° and 60°.

Therefore, the evaluation method using the 90° of flexion position is easier to use clinically than the evaluation method using the 30° of flexion position.

In this study, the difference between the 2 limb positions is the forearm rotation in addition to the elbow flexion angle. In the 90° of flexion position, the forearm was in neutral rotation, and in the 30° of flexion position, the forearm was in supination. Safran et al[19] investigated the effect of forearm rotation on elbow valgus laxity at 30°, 50°, and 70° of elbow flexion in a cadaveric study. Results showed that when the UCL was intact, valgus laxity was greater in the neutral forearm position compared with forearm pronation or supination at all elbow flexion angles. In this study, forearm rotation in the 30° of flexion position was defined as supination on the basis of previous studies, but it is unclear whether the forearm rotation affected the magnitude of valgus laxity.

Strengths and Limitations

The strengths of this study include the uniformity of the technique, the precision of high-resolution ultrasonography, and the methods being conducted by experienced orthopaedic surgeons. It is possible that previous studies conducted with the 30° of flexion elbow position adopted in this study have not adequately assessed the valgus laxity characteristics of baseball players.

There are several limitations of this study. First, the study is limited to college baseball players. It is not known whether the same results could be obtained for other age or sex groups. Second, the number of participants in this study is small. Third, subgroup analyses by age and years of baseball experience could not be performed because of the small individual differences in these parameters in the participants. Fourth, it did not consider the differences in valgus laxity by forearm rotation within the same elbow flexion angle. Fifth, the 30° of flexion position was not measured under conditions where gravity was applied to the forearm, and this approach may have resulted in differences from the measurements in the 90° of flexion position. Sixth, ultrasound effectiveness depends on the observer. Although intraobserver and interobserver reliability was high in this study, results can differ depending on the observer's experience with SUS. Seventh, this study did not investigate the association between valgus laxity and injuries, such as UCL injuries.

CONCLUSION

In the current study, the 90° of flexion position of the elbow more clearly detected valgus laxity than the 30° of flexion position, indicating that the 90° of flexion position is more optimal for SUS evaluation of elbow valgus laxity in baseball players.

ACKNOWLEDGMENT

The authors thank the university baseball team and team officials who assisted in the study. They also thank Editage (www.editage.jp) for their English-language editing.

REFERENCES

1. Ahmad CS, Lee TQ, ElAttrache NS. Biomechanical evaluation of a new ulnar collateral ligament reconstruction technique with interference screw fixation. *Am J Sports Med*. 2003;31(3):332-337. doi:10.1177/03635465030310030201
2. Atanda A Jr, Buckley PS, Hammoud S, Cohen SB, Nazarian LN, Ciccotti MG. Early anatomic changes of the ulnar collateral ligament identified by stress ultrasound of the elbow in young professional baseball pitchers. *Am J Sports Med*. 2015;43(12):2943-2949. doi:10.1177/0363546515605042
3. Callaway GH, Field LD, Deng XH, et al. Biomechanical evaluation of the medial collateral ligament of the elbow. *J Bone Joint Surg Am*. 1997;79(8):1223-1231. doi:10.2106/00004623-199708000-00015
4. Ciccotti MG, Atanda A Jr, Nazarian LN, Dodson CC, Holmes L, Cohen SB. Stress sonography of the ulnar collateral ligament of the elbow in professional baseball pitchers: a 10-year study. *Am J Sports Med*. 2014;42(3):544-551. doi:10.1177/0363546513516592
5. Conway JE, Jobe FW, Glousman RE, Pink M. Medial instability of the elbow in throwing athletes. Treatment by repair or reconstruction of the ulnar collateral ligament. *J Bone Joint Surg Am*. 1992;74(1):67-83. doi:10.2106/00004623-199274010-00009
6. Ellenbecker TS, Mattalino AJ, Elam EA, Caplinger RA. Medial elbow joint laxity in professional baseball pitchers. A bilateral comparison using stress radiography. *Am J Sports Med*. 1998;26(3):420-424. doi:10.1177/03635465980260031301
7. Fleisig GS, Andrews JR, Dillman CJ, Escamilla RF. Kinetics of baseball pitching with implications about injury mechanisms. *Am J Sports Med*. 1995;23(2):233-239. doi:10.1177/036354659502300218
8. Fleisig GS, Barrentine SW, Escamilla RF, Andrews JR. Biomechanics of overhand throwing with implications for injuries. *Sports Med*. 1996;21(6):421-437. doi:10.2165/00007256-199621060-00004
9. Hamilton CD, Glousman RE, Jobe FW, Brault J, Pink M, Perry J. Dynamic stability of the elbow: electromyographic analysis of the flexor pronator group and the extensor group in pitchers with valgus instability. *J Shoulder Elbow Surg*. 1996;5(5):347-354. doi:10.1016/s1058-2746(96)80065-6
10. Harada M, Takahara M, Maruyama M, Nemoto T, Koseki K, Kato Y. Assessment of medial elbow laxity by gravity stress radiography: comparison of valgus stress radiography with gravity and a Telos stress device. *J Shoulder Elbow Surg*. 2014;23(4):561-566. doi:10.1016/j.jse.2014.01.002
11. Hattori H, Akasaka K, Otsudo T, Hall T, Amemiya K, Mori Y. The effect of repetitive baseball pitching on medial elbow joint space gapping associated with 2 elbow valgus stressors in high school baseball players. *J Shoulder Elbow Surg*. 2018;27(4):592-598. doi:10.1016/j.jse.2017.10.031
12. Hotchkiss RN, Weiland AJ. Valgus stability of the elbow. *J Orthop Res*. 1987;5(3):372-377. doi:10.1002/jor.1100050309
13. Jobe FW, Stark H, Lombardo SJ. Reconstruction of the ulnar collateral ligament in athletes. *J Bone Joint Surg Am*. 1986;68(8):1158-1163. doi:10.2106/00004623-198668080-00004
14. Karbach LE, Elfar J. Elbow instability: anatomy, biomechanics, diagnostic maneuvers, and testing. *J Hand Surg Am*. 2017;42(2):118-126. doi:10.1016/j.jhsa.2016.11.025
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
16. Morrey BF, An KN. Articular and ligamentous contributions to the stability of the elbow joint. *Am J Sports Med*. 1983;11(5):315-319. doi:10.1177/036354658301100506
17. Morrey BF, Tanaka S, An KN. Valgus stability of the elbow. A definition of primary and secondary constraints. *Clin Orthop Relat Res*. 1991;(265):187-195.
18. Nazarian LN, McShane JM, Ciccotti MG, O'Kane PL, Harwood MI. Dynamic US of the anterior band of the ulnar collateral ligament of the elbow in asymptomatic Major League Baseball pitchers. *Radiology*. 2003;227(1):149-154. doi:10.1148/radiol.2271020288
19. Safran MR, McGarry MH, Shin S, Han S, Lee TQ. Effects of elbow flexion and forearm rotation on valgus laxity of the elbow. *J Bone Joint Surg Am*. 2005;87(9):2065-2074. doi:10.2106/JBJS.D.02045

20. Sasaki J, Takahara M, Ogino T, Kashiwa H, Ishigaki D, Kanauchi Y. Ultrasonographic assessment of the ulnar collateral ligament and medial elbow laxity in college baseball players. *J Bone Joint Surg Am*. 2002;84(4):525-531. doi:10.2106/00004623-200204000-00003
21. Seiber K, Bales C, Wörner E, Lee T, Safran MR. Assessment of the reliability of a non-invasive elbow valgus laxity measurement device. *J Exp Orthop*. 2020;7(1):74. doi:10.1186/s40634-020-00290-2
22. Shanley E, Smith M, Mayer BK, et al. Using stress ultrasonography to understand the risk of UCL injury among professional baseball pitchers based on ligament morphology and dynamic abnormalities. *Orthop J Sports Med*. 2018;6(8):2325967118788847. doi:10.1177/2325967118788847
23. Singh H, Osbahr DC, Wickham MQ, Kirkendall DT, Speer KP. Valgus laxity of the ulnar collateral ligament of the elbow in collegiate athletes. *Am J Sports Med*. 2001;29(5):558-561. doi:10.1177/03635465010290050601
24. Watanabe H, Masuma H, Kenmoku T, et al. Increased medial laxity of the elbow in preadolescent baseball players with or without medial elbow apophysitis. *JSES Int*. 2021;5(6):1119-1124. doi:10.1016/j.jseint.2021.07.010
25. Werner SL, Fleisig GS, Dillman CJ, Andrews JR. Biomechanics of the elbow during baseball pitching. *J Orthop Sports Phys Ther*. 1993;17(6):274-278. doi:10.2519/jospt.1993.17.6.274

American Journal of Transplantation 24 (2024) 1495–1508

Contents lists available at ScienceDirect

American Journal of Transplantation

journal homepage: www.amjtransplant.org

Original Article

Excess mortality in COVID-19-affected solid organ transplant recipients across the pandemic

Shigeyoshi Yamanaga [1], Keita Shimata [2], Satoko Ohfuji [3], Mikiko Yoshikawa [4], Yoichiro Natori [5,6], Taizo Hibi [2,*], Kenji Yuzawa [7], Hiroto Egawa [8], on behalf of The Japan Society for Transplantation COVID-19 Registry Study Group

[1] *Department of Surgery, Japanese Red Cross Kumamoto Hospital, Nagamine-Minami, Higashi Ward, Kumamoto, Japan*
[2] *Department of Pediatric Surgery and Transplantation, Kumamoto University Graduate School of Medical Sciences, Honjō, Chuo Ward, Kumamoto, Japan*
[3] *Department of Public Health, Osaka Metropolitan University Graduate School of Medicine, Asahimachi, Abeno Ward, Osaka, Japan*
[4] *Organ Transplantation and General Surgery, Kyoto Prefectural University of Medicine, Kajiichō, Kamigyo Ward, Kyoto, Japan*
[5] *Solid Organ Transplant Infectious Diseases, Miami Transplant Institute, Jackson Health System, Miami, Florida, USA*
[6] *Division of Infectious Diseases, Department of Medicine, University of Miami Miller School of Medicine, Miami, Florida, USA*
[7] *Department of Transplantation Surgery, National Hospital Organization Mito Medical Center, Sakuranosato, Ibaraki, Higashiibaraki District, Ibaraki, Japan*
[8] *Department of Surgery, Tokyo Women's Medical University, Kawadacho, Shinjuku Ward, Tokyo, Japan*

ARTICLE INFO

Keywords:
COVID-19
organ transplantation
excess mortality
standardized mortality ratio

ABSTRACT

The excess mortality of coronavirus disease 2019 (COVID-19) solid organ transplant recipients (SOTRs) throughout the pandemic remains unclear. This prospective cohort study based on the Japanese nationwide registry included 1632 SOTRs diagnosed with COVID-19 between February 1, 2020, and July 31, 2022, categorized based on dominant phases of variants of concern (VOCs): Waves 1 to 3 (Beta), 4 (Alpha), 5 (Delta), 6 (Omicron BA.1/BA.2), and 7 (Omicron BA.5). Excess mortality of COVID-19-affected SOTRs was analyzed by calculating standardized mortality ratios (SMRs). Overall, 1632 COVID-19-confirmed SOTRs included 1170 kidney, 408 liver, 25 lung, 20 heart, 1 small intestine, and 8 multi-organ recipients. Although disease severity and all-cause mortality decreased as VOCs transitioned, SMRs of SOTRs were consistently higher than those of the general population throughout the pandemic, showing a U-shaped gap that peaked toward the Omicron BA.5 phase; SMR (95% CI): 6.2 (3.1-12.5), 4.0 (1.5-10.6), 3.0 (1.3-6.7), 8.8 (5.3-14.5), and 21.9 (5.5-87.6) for Waves 1 to 3 (Beta), Wave 4 (Alpha), Wave 5 (Delta), Wave 6 (Omicron BA.1/

Abbreviations: CI, confidence interval; COVID-19, coronavirus disease 2019; ECMO, extracorporeal membrane oxygenation; HER-SYS, Health Center Real-Time Information-sharing System; HLA, human leukocyte antigen; ICU, intensive care unit; JST, Japan Society for Transplantation; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SMRs, standardized mortality ratios; SOTRs, solid organ transplant recipients; VOC, variants of concern.

* Corresponding author. Taizo Hibi, Department of Pediatric Surgery and Transplantation, Kumamoto University Graduate School of Medical Sciences, 1-1-1 Honjō, Chuo Ward, Kumamoto, Kumamoto 860-0811, Japan.
E-mail address: taizohibi@gmail.com (T. Hibi).

https://doi.org/10.1016/j.ajt.2024.03.016
Received 8 September 2023; Received in revised form 5 March 2024; Accepted 12 March 2024
Available online 20 March 2024

2), and Wave 7 (Omicron BA.5), respectively. In conclusion, COVID-19 SOTRs had greater SMRs than the general population across the pandemic. Vaccine boosters, immunosuppression optimization, and other protective measures, particularly for older SOTRs, are paramount.

1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has severely affected organ transplantation worldwide, considerably reducing patient life-years in solid organ transplant recipients (SOTRs) due to the increased risks of severe disease and mortality.[1-3] The in-hospital mortality rate of COVID-19–related pneumonia was 2.5-fold higher than that of non-COVID-19–related pneumonia among SOTRs.[4] In Japan, in the early period of the COVID-19 pandemic, the number of deceased donor organ donations and transplantations dropped to 61% and 69%, respectively, of the year-on-year average (2019 vs 2020),[5] and living donor kidney transplantation was particularly affected due to its nonurgent nature and lack of deceased donors.[6] Despite measures and therapeutics being implemented to restore transplant activities, cautionary measures are still required due to the suboptimal immunogenicity of SARS-CoV-2 vaccines in SOTRs[7,8] and the high infectivity of the Omicron variant of concern (VOC).[9]

Previous studies have delineated the short-term outcomes of COVID-19-affected SOTRs, but mostly in the early phase and with a limited follow-up period.[2,10] Furthermore, although guidelines published by transplant societies, including the Japan Society for Transplantation (JST), have recommended treatment options, most pivotal trials for therapeutics have excluded SOTRs.[3] Accordingly, the excess mortality due to COVID-19 among SOTRs compared to the general populations across the pandemic has yet to be elucidated. In the present study, we investigated the detailed epidemiology, interventions, and prognoses over a year-long follow-up period to reveal the excess mortality due to COVID-19 among SOTRs throughout the pandemic period.

2. Methods

This study was conducted by the JST COVID-19 Registry Study Group (The author list of the JST COVID-19 Registry Study Group is available in the Supplementary material) involving multiple centers nationwide. The study participants were SOTRs diagnosed with COVID-19 between February 1, 2020, and July 31, 2022. In Japan, SOTRs are actively monitored by the transplant centers and encouraged to contact the centers when they are diagnosed with COVID-19. The inclusion criteria for the participants are shown in Figure 1. The primary survey was sent to 122 programs, including 11, 10, and 33 programs for the heart, lung, and liver, respectively. For the kidney, the primary survey was sent to the programs performing >10 cases/year. Of these,

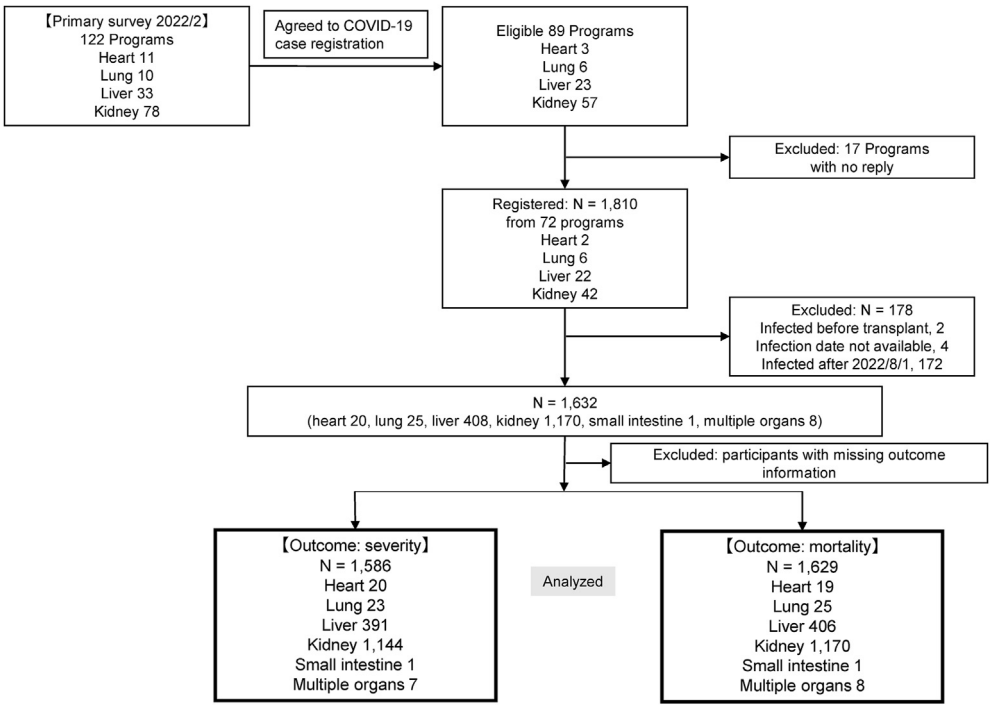


Figure 1. Outline of the study enrolment process. COVID-19, coronavirus disease 2019.

72 programs (2 heart, 6 lung, 22 liver, and 42 kidney programs) were enrolled, with 1632 cases in total. These 72 programs are geographically dispersed high-volume centers responsible for >50% of all de novo organ transplant cases in the country. For example, in 2019, the attending kidney transplant programs performed 56.5% of de novo kidney transplantations in Japan.[11] The questionnaire of this study is available in Supplementary Table S1. The prospectively collected clinical data were analyzed to comprehensively investigate the outcomes of COVID-19 during the pandemic.

2.1. Ethical approval

Ethical approval was obtained from the Institutional Review Board of Kumamoto University, Kumamoto, Japan (approval No. 2504). The requirement for informed consent was waived owing to the noninvasive nature of the study.

2.2. Outcomes

The primary outcome was death up to 1 year after the COVID-19 diagnosis. To capture COVID-19–associated death, deaths within 30, 60, 180, and 365 days after diagnosis with COVID-19 were also considered. Secondary outcomes included hospital admission, intensive care unit (ICU) admission, oxygen demand, mechanical ventilation, extracorporeal membrane oxygenation (ECMO), graft loss within 1 year, and the severity. Other outcomes included acute kidney injury, need for hemodialysis, pneumonia (diagnosed by chest imaging), secondary infection (bacterial, viral, and fungal), high-flow cannula, noninvasive positive pressure ventilation, cardiovascular events, cerebrovascular events, deep vein thrombosis/pulmonary embolism, acute rejection, development of de novo antihuman leukocyte antigen (HLA) antibody, and long COVID. The severity of COVID-19 was defined as follows: asymptomatic, mild (no oxygen demand and signs of pneumonia on chest imaging), moderate (O_2 demand with a nasal cannula or mask and/or positive for pneumonia on chest imaging), and severe (severe hypoxia requiring a high-flow cannula, noninvasive positive pressure ventilation, mechanical ventilation, ECMO, and/or ICU admission).[12] We defined acute kidney injury according to the Kidney Disease Improving Global Outcomes criteria.[13] Long COVID was defined as symptoms persisting for over 2 months that could not be explained by an alternative diagnosis according to the World Health Organization criteria.[14]

2.3. Statistical analysis

The background characteristics of the study participants were compared according to the VOC-dominant phases. The definitions of the VOC-dominant phases and corresponding waves in Japan were as follows: Wave 1: original strain and Beta VOC (B.1.1) January 29, 2020, to June 9, 2020, Wave 2: Beta VOC (B.1.1.284) June 10, 2020, to October 5, 2020, Wave 3: Beta VOC (B.1.1.214) October 6, 2020, to March 2, 2021, Wave 4: Alpha VOC March 3, 2021 to June 22, 2021, Wave 5: Delta VOC June 23, 2021 to December 14, 2021, Wave 6: Omicron-1 VOC (BA.1/BA.2) December 15, 2021 to June 21, 2022, and Wave 7: Omicron-2 VOC (BA.5) June 22, 2022 to July 31, 2022 (Supplementary Fig. S1). The distribution of variables was compared between groups using the chi-square or Fisher exact tests for categorical data and the Mann–Whitney U and Kruskal–Wallis tests for continuous data. Unless otherwise specified, all continuous data are expressed as medians (interquartile range).

The standardized mortality ratio (SMR) and 95% confidence interval (CI) were calculated using an indirect method with the age-sex structure of the COVID-19 general population to analyze the excess mortality of COVID-19-affected SOTRs. In Japan, all patients diagnosed with COVID-19 were registered into the Health Center Real-Time Information-sharing System (HER-SYS) on COVID-19, and the weekly number of newly diagnosed patients with COVID-19 and deaths by 10-year age categories and sex were available as Japanese open data.[15] The case-fatality rates of the general population during each VOC-dominant phase, classified according to 10-year age categories and sex, were calculated by using the information from HER-SYS and were defined as the reference for SMR regarding COVID-19–affected SOTRs. The Japanese government tracked all COVID-19 cases until September 2022, which falls within the observation period of the present study. We chose to incorporate 60-day mortality into SMR analysis, aligning with the previous studies investigating short-term outcomes related to acute infectious disease including COVID-19, wherein 60-day mortality is often designated as a primary outcome.[16-18] We conducted sensitivity analyses for validating SMRs using 1-year mortality data after COVID-19 diagnosis or COVID-19–related deaths. Because Japanese open data included reinfected patients among the general population, we included the reinfected SOTR patients in SMR analyses.

Survival of the study participants according to the VOC-dominant phases, transplanted organs, or age categories was estimated using the Kaplan–Meier method, and statistical differences between the curves were assessed using the log-rank test. The follow-up period was calculated from the date of the COVID-19 diagnosis to the date of the last visit to the study hospital or death. If the dates of diagnosis and the last visit to the study hospital were unavailable, the date of admission and the discharge date were used, respectively.

Univariate and multivariate proportional hazard models were employed to calculate crude and adjusted hazard ratios and 95% CIs to investigate the risk factors for death within 60 days after COVID-19 diagnosis. Risk factors for composite outcomes, including moderate or severe COVID-19 or death, were examined by odds ratios and 95% CIs in univariate and multivariate logistic regression models because the date of disease progression to moderate or severe condition was not fully collected. Multivariate models included variables statistically correlating with the outcome in univariate analyses ($P < .05$).

Statistical significance was set at $P < .05$. IBM SPSS Statistics for Windows version 28 (IBM Corp, Armonk, NY, USA) was used for all statistical analyses.

All results were reanalyzed, excluding SOTRs with recurrent infections, for validation.

3. Results

3.1. Baseline characteristics

A total of 21 923 outpatient SOTRs were identified, of which 7.3% (n = 1600) were confirmed cases of COVID-19. Recurrent infections (≥2) were observed in 2.0% (n = 32) of the patients. The baseline characteristics of the patients are summarized in Table 1. The median follow-up period was 125 (33–233) days. The median age of the participants was 48 (35–59) years, and most participants were male (61%). The most common comorbidities included hypertension (56%) and diabetes mellitus (22%). tacrolimus and mycophenolate mofetil were commonly used, and everolimus was used in 19% of the patients. Steroids were used in 76%. Notably, some kidney transplant recipients were maintained using the 4 immunosuppressive drugs. Less than 1% of patients had a history of acute rejection (within 3 months) or administration of either rituximab (within 6 months) or antithymoglobulin (within 3 months) before COVID-19 diagnosis. Approximately >50% received at least 1 vaccine during Wave 5 (Delta), which increased to 84% during Wave 7 (Omicron BA.5).

3.2. Disease severity, mortality, and other outcomes

The major outcomes are summarized in Table 2. The all-cause case-fatality rates at 30 days, 60 days, 6 months, 1 year, and overall were 1.4% (n = 22), 2.1% (n = 35), 2.4% (n = 39), 2.6% (n = 43), and 2.7% (n = 44), respectively. The number of COVID-19–related mortalities was 31 (76%), and fatal events occurred in 61% of patients within 30 days and 97% of cases within 60 days. The rates of hospital admission, ICU admission, pneumonia, oxygen demand, mechanical ventilation, and ECMO decreased over time. Disease severity decreased from 19% in Waves 1 to 3 (Beta) to 2% in Wave 7 (Omicron BA.5).

The 1-year all-cause mortality corresponding to different VOC-dominant phases gradually declined over time, with Kaplan-Meier mortality estimates of 11.2%, 10.6%, 7.4%, 2.7%, and 0.6% for Waves 1 to 3 (Beta), Wave 4 (Alpha), Wave 5 (Delta), Wave 6 (Omicron BA.1/2), and Wave 7 (Omicron BA.5), respectively (Fig. 2A for patient survival by waves, and Fig. 2B for patient survival by age categories). Despite this decline, the SMR remained persistently high in the population of SOTRs who expired within 60 days following COVID-19 diagnosis (Fig. 3), widening the U-shaped gap after the Delta VOC phase that peaked toward the Omicron BA.5 VOC phase. Specifically, SMR (95% CI) values of 6.2 (3.1-12.5), 4.0 (1.5-10.6), 3.0 (1.3-6.7), 8.8 (5.3-14.5), and 21.9 (5.5-87.6) were obtained for Waves 1 to 3 (Beta), Wave 4 (Alpha), Wave 5 (Delta), Wave 6 (Omicron BA.1/2), and Wave 7 (Omicron BA.5), respectively (Supplementary Table S2). The same SMR trends were obtained for the sensitivity analysis limited to the 1-year overall or COVID-19–related mortality. Although the COVID-19–related case-fatality rates for severe cases declined over time, they remained consistently high (>25%) during each phase of COVID-19 infection dominated by a VOC, with rates of 38.9%, 38.5%, 46.2%, 29.7%, and 25.0% during Waves 1 to 3 (Beta), Wave 4 (Alpha), Wave 5 (Delta), Wave 6 (Omicron BA.1/2), and Wave 7 (Omicron BA.5), respectively.

Details of the short- and long-term outcomes of the pandemic are available in Supplementary Table S3. Both short- and long-term outcomes exhibited similar gradual declining trends across all phases of COVID-19. Specifically, long-term outcomes included acute rejection (1%), de novo anti-HLA antibodies (6%), and long COVID (3%). Among the 64 recipients positive for de novo anti-HLA antibodies, 32.8% (n = 21) were confirmed to have donor-specific antibodies.

3.3. Treatment and clinical courses

Changes in therapeutic approaches during the pandemic are shown in Supplementary Figure S2. Remdesivir and molnupiravir were the gold-standard antiviral drugs. Anti-SARS-CoV-2 monoclonal antibodies, tocilizumab, and antithrombotic therapy were used in 27%, 2%, and 13% of patients, respectively.

The evolving patterns of immunosuppression adjustments throughout the pandemic are presented in Supplementary Figure S3. Notably, a considerable decrease in the frequency of calcineurin inhibitors and antimetabolites adjustments was observed over time. Conversely, there was an increase in steroid dosage (including escalated-pulse and dexamethasone-switch regimens) in 15% of cases, but this decreased in the Omicron era.

3.4. Risk factors for moderate/severe disease and mortality

The results of the independent risk factors for moderate or severe COVID-19 or deaths are summarized in Table 3 (details are available in Supplementary Table S4). The multivariate analysis revealed that harmful factors, including age ≥ 60 years or having underlying illnesses, mTORi use, steroid use, or antithymoglobulin use within 3 months before COVID-19 diagnosis, were independent risk factors for moderate/severe COVID-19 or deaths; protective factors included vaccination, Waves 6 (Omicron BA.1/2) and 7 (Omicron BA.5) infection, and age <20 years old. The independent risk factors for mortality within 60 days of COVID-19 diagnosis are summarized in Table 4 (details are available in Supplementary Table S5). In the multivariate analysis, independent risk factors impacting mortality were diagnosis during Wave 7 (Omicron BA.5, protective) and age ≥ 60 years (harmful). Of note, all results obtained were consistent after excluding SOTRs with recurrent infections (n = 32, 2.0%).

4. Discussion

This comprehensive and detailed analysis of nationwide prospective registry data provided valuable insights into the real-world experiences of Japanese COVID-19 SOTRs from the beginning to the final stages of the pandemic. To the best of our knowledge, this is the first study to present the SMR trends of COVID-19 in the SOTR population. The strength of the present

Table 1
Characteristics of study participants.

		Total (N = 1632)	Waves 1-3 (Beta) (N = 96)	Wave 4 (Alpha) (N = 75)	Wave 5 (Delta) (N = 124)	Wave 6 (Omicron BA.1/2) (N = 884)	Wave 7 (Omicron BA.5) (N = 453)	*P* value
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Transplant organ	Heart	20 (1)	2 (2)	1 (1)	3 (2)	10 (1)	4 (1)	.28
	Lungs	25 (2)	2 (2)	1 (1)	3 (2)	11 (1)	8 (2)	
	Liver	408 (25)	14 (15)	23 (31)	33 (27)	241 (27)	97 (21)	
	Kidneys	1170 (72)	77 (80)	50 (67)	84 (68)	616 (70)	343 (76)	
	Small intestine	1 (0.1)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.2)	
	Multiple organs	8 (0.5)	1 (1)	0 (0)	1 (1)	6 (1)	0 (0)	
Duration from transplantation (mo)	Median (IQR)	80 (38-145)	72 (36-143)	94.5 (48.5-181.5)	69.5 (31-143)	81 (40-146)	83 (37-141)	.21
Follow-up period (d)	Median (IQR)	125 (33-233)	606.5 (31-688)	497 (48-554)	380.5 (32-432)	167 (43.5-232)	70 (24-98)	<.01
Age (y)	Median (IQR)	48 (35-59)	49.5 (38.5-62.5)	52 (35-62)	50 (39.5-60)	47 (33-58)	48 (34-59)	.06
Sex	Male	1001 (61)	68 (71)	49 (65)	79 (64)	528 (60)	277 (61)	.25
Body mass index/ Laurel index/Kaup index	Underweight	193 (12)	12 (13)	5 (7)	16 (13)	110 (13)	50 (11)	.01
	Normal	915 (57)	48 (50)	40 (53)	72 (59)	471 (54)	284 (64)	
	Obese	510 (32)	36 (38)	30 (40)	35 (28)	297 (34)	112 (25)	
Smoking	Never	936 (69)	38 (51)	37 (61)	62 (60)	530 (72)	269 (70)	<.01
	Current/Former	419 (31)	37 (49)	24 (39)	42 (40)	203 (28)	113 (30)	
Underlying illnesses	Any	1055/1535 (69)	73/87 (84)	44/68 (65)	92/117 (79)	568/832 (68)	278/431 (65)	<.01
	Diabetes mellitus	362/1631 (22)	30 (31)	21 (28)	41 (33)	184/883 (21)	86 (19)	<.01
	Hypertension	617/1631 (56)	59 (61)	38 (51)	77 (62)	499 (56)	244/452 (54)	.32
	Cardiovascular diseases	114 (7)	7 (7)	5 (7)	12 (10)	60 (7)	30 (7)	.82
	Chronic respiratory diseases	37/1630 (2)	2/95 (2)	2 (3)	4/123 (3)	24 (3)	5 (1)	.38
	Chronic liver diseases	21/1630 (1)	8/95 (8)	0 (0)	2 (2)	8/883 (1)	3 (1)	<.01
	Chronic kidney diseases	30/1499 (2)	2/86 (2)	3/65 (5)	3/112 (3)	17/810 (2)	5/426 (1)	.39
	Cerebrovascular diseases	49 (3)	4 (4)	1 (1)	7 (6)	25 (3)	12 (3)	.35
	Malignant neoplasm	69 (4)	3 (3)	3 (4)	5 (4)	41 (5)	17 (4)	.92
HLA A24 typing	Yes	749 (59)	37 (53)	28 (51)	60 (59)	419 (60)	205 (58)	.55

(*continued on next page*)

Table 1 (*continued*)

		Total (N = 1632)	Waves 1-3 (Beta) (N = 96)	Wave 4 (Alpha) (N = 75)	Wave 5 (Delta) (N = 124)	Wave 6 (Omicron BA.1/2) (N = 884)	Wave 7 (Omicron BA.5) (N = 453)	*P* value
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Donor type	Deceased	161 (10)	11 (12)	8 (11)	19 (15)	87 (10)	36 (8)	.18
Acute rejection <3 months before COVID-19 diagnosis	Yes	11 (1)	0 (0)	1 (1)	2 (2)	5 (1)	3 (1)	.57
Calcineurin inhibitors	No	23 (1)	0 (0)	1 (1)	4 (3)	14 (2)	4 (1)	.11
	Tac	460 (28)	29 (31)	26 (35)	33 (27)	266 (30)	106 (23)	
	TacER	904 (55)	55 (58)	34 (45)	67 (54)	471 (53)	277 (61)	
	CsA	243 (15)	11 (12)	14 (19)	20 (16)	132 (15)	66 (15)	
Antimetabolites	No	301 (19)	15 (17)	11 (16)	22 (18)	166 (19)	87 (19)	.43
	MMF	1239 (77)	70 (78)	53 (78)	95 (79)	682 (78)	339 (75)	
	AZ	23 (1)	2 (2)	2 (3)	0 (0)	7 (1)	12 (3)	
	MZ	46 (3)	3 (3)	2 (3)	4 (3)	22 (3)	15 (3)	
mTOR inhibitors	No	1314 (81)	67 (72)	64 (85)	98 (80)	720 (82)	365 (81)	.16
	Yes	305 (19)	26 (28)	11 (15)	25 (20)	158 (18)	85 (19)	
Steroids	No	385 (24)	18 (20)	14 (19)	26 (21)	231 (26)	96 (21)	.02
	PSL	449 (28)	30 (33)	32 (44)	30 (24)	236 (27)	121 (27)	
	mPSL	781 (48)	44 (48)	26 (36)	67 (54)	411 (47)	233 (52)	
Combination of medications	None	5 (0.3)	0 (0)	0 (0)	1 (1)	3 (0.4)	1 (0.2)	.49
	1	170 (11)	4 (5)	4 (6)	10 (8)	101 (12)	51 (11)	
	2	217 (14)	12 (14)	14 (22)	17 (14)	125 (14)	49 (11)	
	3	1021 (65)	57 (67)	39 (60)	79 (66)	550 (64)	296 (66)	
	4	167 (11)	12 (14)	8 (12)	12 (10)	85 (10)	50 (11)	
Rituximab <6 mo before COVID-19 diagnosis	Yes	12 (1)	1 (1)	0 (0)	1 (1)	9 (1)	1 (1)	.5

(*continued on next page*)

Table 1 (*continued*)

		Total (N = 1632)	Waves 1-3 (Beta) (N = 96)	Wave 4 (Alpha) (N = 75)	Wave 5 (Delta) (N = 124)	Wave 6 (Omicron BA.1/2) (N = 884)	Wave 7 (Omicron BA.5) (N = 453)	*P* value
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Antithymoglobulin <3 mo before COVID-19 diagnosis	Yes	6 (0.4)	1 (1)	0 (0)	0 (0)	3 (0.3)	2 (0.4)	.73
Number of infections	1	1600 (98)	94 (98)	74 (99)	123 (99)	866 (98)	443 (98)	.89
	2	31 (2)	2 (2)	1 (1)	1 (1)	18 (2)	9 (2)	
	3	1 (0.1)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.2)	
Vaccination before infection	No	420 (29)	96 (100)	50 (88)	54 (49)	157 (20)	63 (16)	<.01
	Yes	1004 (71)	0 (0)	7 (12)	57 (51)	615 (80)	325 (84)	
Number of vaccinations	1	35 (4)		2 (33)	19 (35)	12 (2)	2 (1)	<.01
	2	486 (52)		2 (33)	35 (64)	377 (66)	72 (23)	
	3	394 (42)		2 (33)	1 (2)	182 (32)	209 (68)	
	4	26 (3)		0 (0)	0 (0)	2 (0.4)	24 (8)	
Period from onset to diagnosis (d)	Four days or longer	150 (13)	25 (40)	15 (25)	18 (19)	69 (12)	23 (7)	<.01
Symptoms at diagnosis	Yes	1479 (92)	88 (93)	66 (88)	116 (96)	793 (92)	416 (93)	.32

AZ, azathioprine; COVID-19, coronavirus disease 2019; CsA, cyclosporine; HLA, human leukocyte antigen; IQR, interquartile range; MMF, mycophenolate mofetil; mTOR, mammalian target of rapamycin; mPSL, methylprednisolone; MZ, mizoribine; PSL, prednisolone; Tac, tacrolimus; TacER, tacrolimus extended-release.

Table 2
Major outcomes of study participants.

		Total (N = 1632)	Waves 1-3 (Beta) (N = 96)	Wave 4 (Alpha) (N = 75)	Wave 5 (Delta) (N = 124)	Wave 6 (Omicron BA.1/2) (N = 884)	Wave 7 (Omicron BA.5) (N = 453)	*P* value
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	
Hospital admission		808 (50)	88 (92)	63 (85)	105 (85)	415 (47)	137 (31)	<.01
ICU admission		63 (5)	14 (21)	10 (20)	11 (11)	23 (3)	5 (1)	<.01
Pneumonia		338 (22)	47 (52)	30 (43)	57 (48)	152 (19)	52 (12)	<.01
Oxygen demand		220 (14)	38 (41)	31 (44)	40 (35)	82 (10)	29 (7)	<.01
Nasal cannula or mask		150 (10)	24 (23)	23 (33)	29 (26)	53 (6)	21 (5)	<.01
High-flow nasal cannula		39 (3)	3 (4)	5 (8)	8 (7)	18 (2)	5 (1)	<.01
NPPV		5 (0.3)	0 (0)	0 (0)	2 (2)	0 (0)	3 (0.7)	.02
Mechanical ventilation		49 (3)	12 (14)	5 (8)	8 (7)	19 (2)	5 (1)	<.01
ECMO		5 (0.3)	1 (1)	1 (2)	2 (2)	1 (0.1)	0 (0)	<.01
Graft loss		14 (0.9)	2 (2)	0 (0)	1 (1)	9 (1)	2 (0.5)	.45
Severity	Mild	1,214 (77)	37 (39)	26 (35)	53 (45)	707 (82)	391 (89)	<.01
	Moderate	283 (18)	40 (42)	35 (47)	52 (44)	114 (13)	42 (10)	
	Severe (including death)	89 (6)	18 (19)	13 (18)	13 (11)	37 (4)	8 (2)	
Prognosis	Death	44 (3)	10 (10)	7 (9)	7 (6)	18 (2)	2 (0.4)	<.01
	Death within 30 d	22 (1)	5 (5)	4 (5)	2 (2)	10 (1)	1 (0.2)	<.01
	Death within 60 d	35 (2)	8 (8)	4 (5)	6 (5)	15 (2)	2 (0.4)	<.01
	Death within 180 d	39 (2)	8 (8)	6 (8)	6 (5)	17 (2)	2 (0.4)	<.01
	Death within 1 y	43 (3)	9 (9)	7 (9)	7 (6)	18 (2)	2 (0.4)	<.01

COVID-19, coronavirus disease 2019; ECMO, extracorporeal membrane oxygenation; ICU, intensive care unit; NPPV, noninvasive positive pressure ventilation.

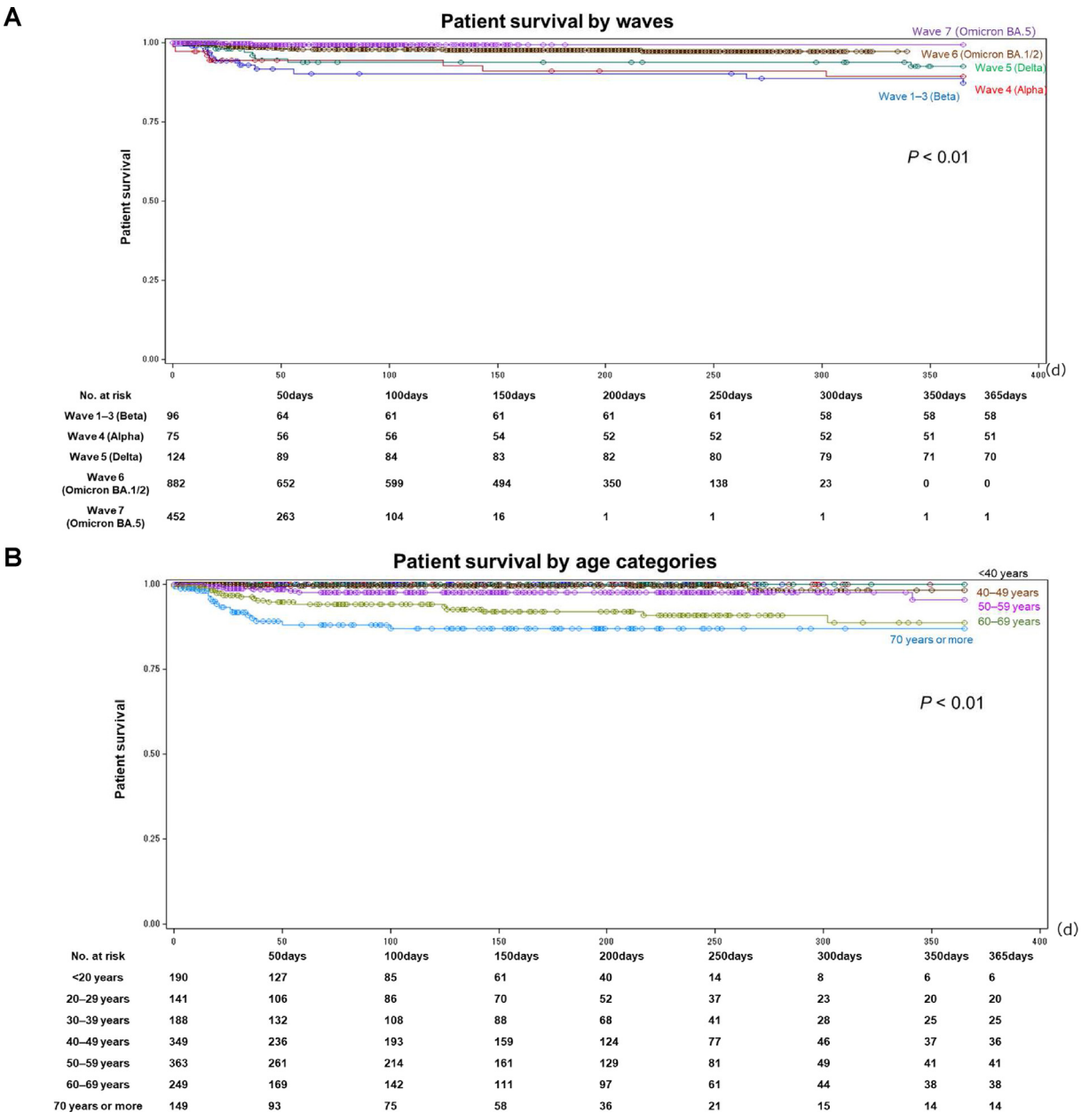
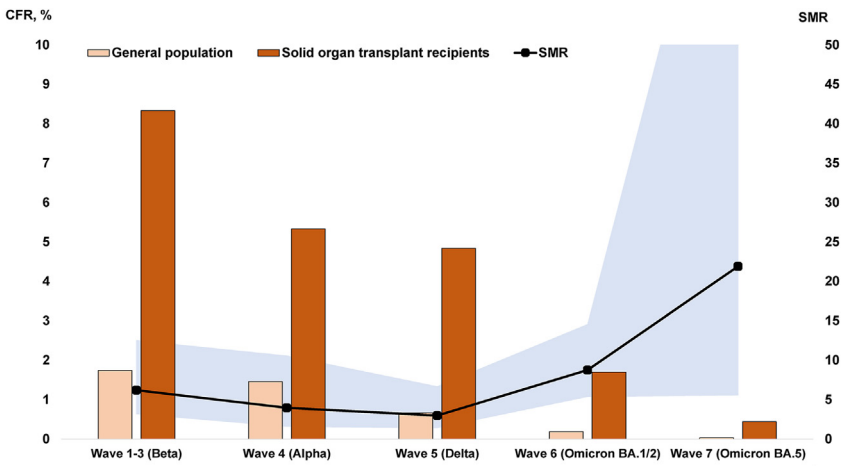


Figure 2. Patient survival categorized by (A) wave and (B) age categories. *P* value for the log-rank test.

study lies in the homogenous population of Japanese SOTRs, who generally have universal insurance and immunosuppression regimens and demonstrate strict compliance (87.1%–94.4%) with the safety measures or therapeutics published by the JST COVID-19 guidelines.[6] Consequently, this minimized the possible impact of unexpected confounders.

In the present study, we observed that disease severity and mortality rates decreased over time, particularly after the Omicron era, which is consistent with the findings of the general population.[19] However, the case-fatality rates for severe disease and SMR were consistently high, despite the availability of vaccines and booster shots in the late phase of the pandemic.[20] Specifically, in the present study, SMR showed a U-shaped pattern, with a nadir in the Delta era followed by a subsequent increase thereafter.

The suboptimal immune response to vaccines and boosters in SOTRs compared to the general population could explain this widened SMR gap toward the Omicron era.[21-25] We speculated that the original vaccines were effective until the Delta era but their efficacy declined drastically, which led to increased breakthrough infection in the Omicron era.[25,26] This phenomenon might also apply to future pandemics. A multicenter study from Spain conducted from February 28, 2020, to April 7, 2020, with an observation period of 23 days (median), revealed a low SMR of 0.96 (95% CI: 0.94-0.97, converted from the percent expression) in 111 liver transplant recipients compared to that in the general population.[27] Although the author concluded that this was possibly due to the protective effect of chronic immunosuppression on COVID-19 exacerbation, this has now been refuted and was regarded as a risk in a recent meta-analysis and



	Wave 1-3 (Beta)	Wave 4 (Alpha)	Wave 5 (Delta)	Wave 6 (Omicron BA.1/2)	Wave 7 (Omicron BA.5)
CFR, solid organ transplant recipients (death/infected)	8.33% (8/96)	5.33% (4/75)	4.84% (6/124)	1.70% (15/882)	0.44% (2/452)
CFR, general population (death/infected) *	1.74% (7,231/146,081)	1.46% (4,693/321,823)	0.66% (5,663/854,309)	0.19% (11,908/6,422,824)	0.03% (610/2,409,783)
SMR (95% CI)	**6.2 (3.1–12.5)**	**4.0 (1.5–10.6)**	**3.0 (1.3–6.7)**	**8.8 (5.3–14.5)**	**21.9 (5.5–87.6)**

Figure 3. Standardized mortality ratio (SMR) of solid organ transplant recipients after coronavirus disease 2019 (COVID-19) compared to the general population. The blue shaded area represents the 95% confidence intervals (CIs) of the SMR. *Open data published by the Ministry of Health, Labour, and Welfare. https://covid19.mhlw.go.jp/extensions/public/index.html.[15] CFR, case-fatality rate.

in this study.[28] Furthermore, we chose the 60-day mortality for the SMR calculation because >95% of COVID-19–related deaths occurred within 60 days of the COVID-19 diagnosis. Additionally, the year-long follow-up period in this study also revealed a trend in deaths during the pandemic.

However, within the SOTR population, the present study showed that vaccines and boosters still provided benefits in terms of reducing disease severity or mortality, despite uncertainties in their real-world effectiveness beyond reducing the immunological response and increasing the risk of breakthrough infections by 82-fold.[29-31] In Japan, BNT162b2 vaccines became available since February 17, 2021, just before the beginning of the Alpha era, followed by mRNA1273 and ChAdOx1 nCoV-19 on May 23, 2021, and August 16, 2021, respectively, though the vaccine most administered was BNT162b2 followed by mRNA1273.[12] As of July 31, 2022 (the end of enrolment of this study), the first, second, and third doses were completed in 77.4%, 76.9%, and 63.7% of cases, respectively.[32] Over 50% of cases in the Delta era received the vaccine, and there is universal access to additional booster shots, provided every 3 months by the government, which resulted in >80% vaccination coverage during the Omicron era in the present study. Consequently, this may have contributed to the low mortality rates observed in Japanese SOTRs. A study from Colombia revealed the benefits of full vaccination and boosters on prevention, hospitalization, and death in SOTRs. However, it did not specify the baseline immunosuppression or transplanted organs.[29] Our study confirmed the independent impact of vaccinations in reducing disease severity and mortality.

Moreover, this study makes a novel contribution to the existing literature by directly comparing the VOCs across the pandemic and revealing that the Omicron VOC had a more favorable impact on the severity of COVID-19 in SOTRs than other VOCs, independent of vaccination and the progress of therapeutics.[3,33] As expected, the Omicron VOC still warrants further attention because it has a higher SMR in COVID-19 SOTRs than that in the general population, as found in the present study, and still poses a high risk.[20] Of note, the vaccine efficacy will decline over time, and with the emergence of a new VOC; additional boosters of updated vaccine are recommended along with updated prophylactic monoclonal antibodies.[31]

Additionally, steroids and mTORi were identified as independent risk factors. This was possibly due to some centers adopting 4 immunosuppressive drug regimens to avoid rejection and anti-HLA antibodies, which may induce overimmunosuppression that could have been avoided during the COVID-19 pandemic.[34] Moreover, the present study did not find any considerable differences in disease severity and mortality rates among different transplanted organs, which may be attributed to the relatively small sample size of lung transplant recipients and lower overall mortality rates than those reported in other studies.[2,33,35] Other risk factors, including older age and underlying illnesses, were in line with the previous studies.[2,33]

The potential impact of immunosuppression adjustment on COVID-19 and other infections remains to be fully elucidated.[3] In a previous meta-analysis, adjustments were made in 76.2% and 38.7% of cases for antimetabolites and calcineurin inhibitors, respectively, resulting in a 1.0% incidence of acute rejection, which is in line with our data.[36] The present study could not definitively establish the benefits of immunosuppression adjustment because the disease severity was retrospectively assigned to the highest degree. However, because disease severity requiring withdrawal of immunosuppression would evoke immunological damage to allografts, attention should be focused on

Table 3
Independent risk and protective factors for moderate or severe COVID-19 or deaths per multivariate analysis.

		Outcome incidence (N = 1586[a])	Multivariate model	
		n/N (%)	OR (95% CI)	*P* value
Wave	1–3 (Beta)	58/95 (61)	1.00	
	4 (Alpha)	48/74 (65)	1.65 (0.59-4.57)	.34
	5 (Delta)	65/118 (55)	1.22 (0.50-2.95)	.66
	6 (Omicron BA.1/2)	151/858 (18)	0.23 (0.10-0.52)	<.01
	7 (Omicron BA.5)	50/441 (11)	0.11 (0.05-0.28)	<.01
Age (y)	<20	4/178 (2)	0.1 (0.02-0.50)	<.01
	20-29	22/135 (16)	0.93 (0.42-2.04)	.85
	30-39	33/186 (18)	0.8 (0.42-1.50)	.48
	40-49	74/344 (22)	1.00	
	50-59	87/359 (24)	0.84 (0.51-1.38)	.49
	60-69	93/242 (38)	1.96 (1.15-3.34)	.01
	70+	59/142 (42)	2.61 (1.40-4.86)	<.01
			(Trend $P < .01$)	
Underlying illnesses	No	45/467 (10)	1.00	
	Yes	315/1029 (31)	2.33 (1.37-3.97)	<.01
mTOR inhibitors	No	275/1273 (22)	1.00	
	Yes	91/300 (30)	1.67 (1.06-2.64)	.03
Steroids	No	39/367 (11)	1.00	
	PSL/mPSL	322/1202 (27)	2.4 (1.24-4.63)	<.01
Antithymoglobulin <3 mo before	No	363/1571 (23)	1.00	
COVID-19 diagnosis	Yes	5/6 (83)	23.47 (2.07-265.92)	.01
Vaccination before infection	No	146/409 (36)	1.00	
	Yes	189/984 (19)	0.46 (0.28-0.76)	<.01
Number of vaccinations	0	146/409 (36)	1	
	1-2	120/412 (23)	0.49 (0.29-0.84)	<.01
	3-4	59/412 (14)	0.29 (0.15-0.55)	<.01
			(Trend $P < .01$)	

Logistic regression model.
The multivariate model included variables associated significantly with disease severity in univariate analysis ($P < .05$). Details are available in Supplementary Table S4.
CI, confidence interval; COVID-19, coronavirus disease 2019; mTOR, mammalian target of rapamycin; mPSL, methylprednisolone; OR, odds ratio; PSL, prednisolone.
[a] Forty-six participants without disease severity and outcome status were excluded.

reducing the risk of severe disease.[37] Trends in therapeutics suggest that remdesivir and molnupiravir remain the gold standards for treating inpatient and outpatient SOTRs, respectively.

Long COVID is a multimodal disease occurring in 10% of the global population (>65 million cases). The incidence is increased in severe cases: 50%–70%, 10%–30%, and 10%–12% in hospitalized, nonhospitalized, and vaccinated cases, respectively.[38] Evidence regarding long COVID in SOTRs is limited. A recent prospective study via phone survey including 780 kidney transplant recipients with a prior COVID-19 infection revealed that 27% had experienced long COVID and 17% had been unable to return to work within 3 months.[39] The risk of long COVID was related to the number of symptoms, similar to that in the general population.[38,39] The reported incidence of long COVID was much lower (3%) in this study, reflecting a difference in the perspectives of the respondents. However, the decline in long COVID prevalence over time was in line with the previous study, possibly reflecting the intensity of acute illness.[39] Further follow-up is warranted.

The limitations of this study include its retrospective assignment of the disease severity, which was defined as the highest degree during the disease course, as the initial grading was not collected in this study. Because this registry was designated only for COVID-19–positive SOTRs, we could not analyze whether

Table 4
Independent risk factors for mortality within 60 days of COVID-19 diagnosis.

		n/Observation periods (Person-months)	Mortality rate per 1000 person-months	Multivariate model HR (95% CI)	*P* value
Total (N = 1629[a])		35/2578.3	13.57		
Wave	1-3	8/151.8	52.70	1.00	
	4	4/124.6	32.10	0.63 (0.12-3.37)	.59
	5	6/198.0	30.30	0.86 (0.19-3.94)	.85
	6	15/1444.6	10.38	0.25 (0.05-1.39)	.11
	7	2/659.4	3.03	0.05 (0.00-0.55)	.02
Age (y)	<20	0/315.4	0.00	NA	
	20-29	0/238.4	0.00	NA	
	30-39	0/293.2	0.00	NA	
	40-49	1/536.0	1.87	1.00	
	50-59	7/579.4	12.08	4.48 (0.52-38.60)	.17
	60-69	12/389.8	30.79	10.65 (1.32-86.07)	.03
	70+	15/226.1	66.34	24.34 (3.06-193.42) (Trend $P < .01$)	<.01

Cox proportional hazard model.
The multivariate model included variables associated significantly with death within 60 days in univariate analysis ($P < .05$). Details are available in Supplementary Table S5.
CI, confidence interval; COVID-19, coronavirus disease 2019; NA, not applicable; OR, odds ratio.
[a] Three participants without outcome status were excluded.

SMR was higher in the COVID-positive subgroup than comparing COVID-19-negative SOTRs to the COVID-19–negative general population, as well as SMR for SOTRs prior to the COVID-19 pandemic. We have collected data from the transplant centers on whether SOTRs were initially treated at transplant centers or at different institutions. There is a possibility that some data from different institutions might have been missing. Although the number of nonabdominal SOT programs is low, the majority of the nonabdominal SOTs were performed in the participating programs. Therefore, the results of this study sufficiently represent the national trend for COVID-19–affected nonabdominal SOTRs.

Regarding the observation period, patient enrolment ended in July 2022, in the middle of Wave 7 (Omicron BA.5), and the Japanese government stopped tracking all COVID-19 cases in September 2022, 2 months after the completion of the study. Thus, the SMR calculated in this study is based on accurate statistics and the maximum observation period available in our patient cohort. The standardized incidence ratio was difficult to calculate from the present study. However, the prevalence of confirmed cases was 7.3% in SOTRs in the present study, whereas it was 10.2% (calculated from open data as of July 31, 2023) in the general population.[15] The reason for the discrepancy in the data might be explained by the recommendation of self-quarantine by the transplant societies, including JST, after the COVID-19 pandemic.[6] Although the evidence was scarce on whether transplant recipients were more likely to seek testing and care in COVID-19, a study from Finland on seasonal influenza before the COVID-19 era revealed that the risk of laboratory-confirmed influenza was significantly higher among kidney transplant recipients than the general population, possibly due to lower threshold for acquiring laboratory testing in transplanted patients.[40] Because accurate data on the type of vaccine were not available, we were unable to investigate whether the difference in vaccine affected outcomes. Because the Japanese open data for COVID-19 mortality rates in the general population are only available for the overall case-fatality rate, comparing the granular mortality rates, such as at 30, 60, 180, and 365 days after diagnosis, was difficult in the present study. The demographic data for the general population stem from confirmed COVID-19 daily cases and deaths recorded in the HER-SYS in Japan.[15] Unfortunately, the HER-SYS only records simple outcomes and lacks detailed baseline data (such as diabetes and immunosuppression) beyond age and sex. Thus, tracking individual outcomes in the general population across the waves was difficult. Furthermore, our registry of transplant recipients may be encompassed within this general population reference for SMR analysis. However, alongside SOTRs,[27] individuals with immunosuppressed conditions such as autoimmune rheumatic diseases, inflammatory bowel diseases, and hematopoietic stem cell transplantation may be included.[41-43] These limitations of the SMR analysis, consistent with other reports, were disregarded due to the substantial disparity between the number of COVID-19–affected SOTRs and that of COVID-19 in the general population.[27,41-43] The generalization of the results of this study in another geographical setting may be difficult due to genetic differences, overall risk factors such as obesity, and management

of SOTRs during the COVID-19 pandemic. In the early phase of the pandemic, there were insufficient supplies of remdesivir, and hospital admission itself was difficult because of the exponential increase of COVID-19 patients. Thus, various nonauthorized drugs, such as favipiravir and ivermectin, were used globally for outpatients as there was no choice at that time, which may have affected the outcomes of COVID-19–positive SOTRs.

Conclusively, in this large-scale national registry study, we showed the epidemiology of the real-world experience in Japan, including the efficacy of vaccines, risk factors of moderate/severe disease, including older age and intense immunosuppression, and the consistently higher mortality risk than that in the general population, even in the final phase of the pandemic. Therefore, SOTRs need to undergo protective measures and receive additional booster shots to mitigate the risk of moderate/severe disease and death.

Funding

This work was supported by MHLW Special Research Program (#JPMH20CA2046) and MHLW Research Program on Emerging and Reemerging Infectious Diseases (#JPMH21HA2011). The funding source for this study played no role in the design, data accumulation, analysis, interpretation, or writing of this article.

A completed STROBE statement checklist for cohort studies is provided in Supplementary Table S6.

Author contributions

Concept and design: S.Y., K.S., M.Y., Y.N., T.H., K.Y., H.E.
Data curation: S.Y., K.S.
Analysis and interpretation of data: S.Y., K.S., S.O., M.Y., Y.N., T.H., K.Y., H.E.
Drafting of the manuscript: S.Y., S.O., T.H.
Critical revision of the manuscript for important intellectual content: S.Y., K.S., M.Y., Y.N., T.H., K.Y., H.E.
Direct access to the data: S.Y., K.S., S.O., T.H., H.E.
Statistical analysis: S.O.
Obtained funding: H.E.

Data availability

The data are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors of this manuscript have no conflicts of interest to disclose as described by the American Journal of Transplantation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajt.2024.03.016.

ORCiD

Shigeyoshi Yamanaga https://orcid.org/0000-0002-4372-0011
Keita Shimata https://orcid.org/0000-0002-8391-5593
Satoko Ohfuji https://orcid.org/0000-0003-3239-5249
Mikiko Yoshikawa https://orcid.org/0000-0002-8625-7389
Yoichiro Natori https://orcid.org/0000-0002-4938-125X
Taizo Hibi https://orcid.org/0000-0002-6867-228X
Kenji Yuzawa https://orcid.org/0000-0002-8357-5664
Hiroto Egawa https://orcid.org/0000-0003-2573-4548

References

1. Aubert O, Yoo D, Zielinski D, et al. COVID-19 pandemic and worldwide organ transplantation: a population-based study. *Lancet Public Health*. 2021;6(10):e709–e719. https://doi.org/10.1016/S2468-2667(21)00200-0.
2. Kates OS, Haydel BM, Florman SS, et al. Coronavirus disease 2019 in solid organ transplant: a multicenter cohort study. *Clin Infect Dis*. 2021;73(11):e4090–e4099. https://doi.org/10.1093/cid/ciaa1097.
3. Nimmo A, Gardiner D, Ushiro-Lumb I, Ravanan R, Forsythe JLR. The global impact of COVID-19 on solid organ transplantation: two years into a pandemic. *Transplantation*. 2022;106(7):1312–1329. https://doi.org/10.1097/TP.0000000000004151.
4. Jering KS, McGrath MM, Mc Causland FR, Claggett B, Cunningham JW, Solomon SD. Excess mortality in solid organ transplant recipients hospitalized with COVID-19: a large-scale comparison of SOT recipients hospitalized with or without COVID-19. *Clin Transplant*. 2022;36(1):e14492. https://doi.org/10.1111/ctr.14492.
5. Ito T, Kenmochi T, Ota A, et al. National survey on deceased donor organ transplantation during the COVID-19 pandemic in Japan. *Surg Today*. 2022;52(5):763–773. https://doi.org/10.1007/s00595-021-02388-1.
6. Kuramitsu K, Yamanaga S, Osawa R, et al. Impact of COVID-19 on living donor liver and kidney transplantation programs in Japan in 2020. *Transpl Infect Dis*. 2022;24(3):e13845. https://doi.org/10.1111/tid.13845.
7. Lee ARYB, Wong SY, Chai LYA, et al. Efficacy of COVID-19 vaccines in immunocompromised patients: systematic review and meta-analysis. *BMJ*. 2022;376:e068632. https://doi.org/10.1136/bmj-2021-068632.
8. Li J, Ayada I, Wang Y, et al. Factors associated with COVID-19 vaccine response in transplant recipients: a systematic review and meta-analysis. *Transplantation*. 2022;106(10):2068–2075. https://doi.org/10.1097/TP.0000000000004256.
9. Cochran W, Shah P, Barker L, et al. COVID-19 clinical outcomes in solid organ transplant recipients during the Omicron surge. *Transplantation*. 2022;106(7):e346–e347. https://doi.org/10.1097/TP.0000000000004162.
10. Akalin E, Azzi Y, Bartash R, et al. COVID-19 and kidney transplantation. *N Engl J Med*. 2020;382(25):2475–2477. https://doi.org/10.1056/NEJMc2011117.
11. Japanese Society for Clinical Renal Transplantation, The Japan Society for Transplantation. Annual Progress Report from the Japanese Renal Transplant Registry: number of renal transplantations in 2019 and follow-up survey. *Jpn J Transplant*. 2020;55(3):225–243. https://doi.org/10.11386/jst.55.3_225.
12. Yamakawa K, Yamamoto R, Terayama T, et al. Japanese rapid/living recommendations on drug management for COVID-19: updated guidelines (July 2022). *Acute Med Surg*. 2022;9(1):e789. https://doi.org/10.1002/ams2.789.
13. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract*. 2012;120(4):c179–c184. https://doi.org/10.1159/000339789.
14. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. WHO Clinical Case Definition Working Group on Post-COVID-19 Condition. A clinical case definition of post-COVID-19 condition by a Delphi consensus.

Lancet Infect Dis. 2022;22(4):e102–e107. https://doi.org/10.1016/S1473-3099(21)00703-9.
15. Ministry of Health, Labour and Welfare. Japanese COVID-19 Open data. https://covid19.mhlw.go.jp/extensions/public/index.html. Accessed April 29, 2023.
16. Ely EW, Ramanan AV, Kartman CE, et al. Efficacy and safety of baricitinib plus standard of care for the treatment of critically ill hospitalised adults with COVID-19 on invasive mechanical ventilation or extracorporeal membrane oxygenation: an exploratory, randomised, placebo-controlled trial. *Lancet Respir Med*. 2022;10(4):327–336. https://doi.org/10.1016/S2213-2600(22)00006-6.
17. Bouadma L, Mekontso-Dessap A, Burdet C, et al. High-dose dexamethasone and oxygen support strategies in intensive care unit patients with severe COVID-19 acute hypoxemic respiratory failure: the COVIDICUS randomized clinical trial. *JAMA Intern Med*. 2022;182(9):906–916. https://doi.org/10.1001/jamainternmed.2022.2168.
18. Trøseid M, Arribas JR, Assoumou L, et al. Efficacy and safety of baricitinib in hospitalized adults with severe or critical COVID-19 (Bari-SolidAct): a randomised, double-blind, placebo-controlled phase 3 trial. *Crit Care*. 2023;27(1):9. https://doi.org/10.1186/s13054-022-04205-8.
19. Hyams C, Challen R, Marlow R, et al. Severity of Omicron (B.1.1.529) and Delta (B.1.617.2) SARS-CoV-2 infection among hospitalised adults: a prospective cohort study in Bristol, United Kingdom. *Lancet Reg Health Eur*. 2023;25:100556. https://doi.org/10.1016/j.lanepe.2022.100556.
20. Anjan S, Khatri A, Viotti JB, et al. Is the Omicron variant truly less virulent in solid organ transplant recipients? *Transpl Infect Dis*. 2022;24(6):e13923. https://doi.org/10.1111/tid.13923.
21. Manothummetha K, Chuleerarux N, Sanguankeo A, et al. Immunogenicity and risk factors associated with poor humoral immune response of SARS-CoV-2 vaccines in recipients of solid organ transplant: a systematic review and meta-analysis. *JAMA Netw Open*. 2022;5(4):e226822. https://doi.org/10.1001/jamanetworkopen.2022.6822.
22. Yang ZR, Jiang YW, Li FX, et al. Efficacy of SARS-CoV-2 vaccines and the dose-response relationship with three major antibodies: a systematic review and meta-analysis of randomised controlled trials. *Lancet Microbe*. 2023;4(4):e236–e246. https://doi.org/10.1016/S2666-5247(22)00390-1.
23. Efros O, Anteby R, Halfon M, Meisel E, Klang E, Soffer S. Efficacy and safety of third dose of the COVID-19 vaccine among solid organ transplant recipients: a systemic review and meta-analysis. *Vaccines (Basel)*. 2022;10(1):95. https://doi.org/10.3390/vaccines10010095.
24. Chen X, Luo D, Mei B, et al. Immunogenicity of COVID-19 vaccines in solid organ transplant recipients: a systematic review and meta-analysis. *Clin Microbiol Infect*. 2023;29(4):441–456. https://doi.org/10.1016/j.cmi.2022.12.004.
25. Wei Z, He J, Wang C, Bao J, Leng T, Chen F. The importance of booster vaccination in the context of Omicron wave. *Front Immunol*. 2022;13:977972. https://doi.org/10.3389/fimmu.2022.977972.
26. Fiolet T, Kherabi Y, MacDonald C-J, Ghosn J, Peiffer-Smadja N. Comparing COVID-19 vaccines for their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: a narrative review. *Clin Microbiol Infect*. 2022;28(2):202–221. https://doi.org/10.1016/j.cmi.2021.10.005.
27. Colmenero J, Rodríguez-Perálvarez M, Salcedo M, et al. Epidemiological pattern, incidence, and outcomes of COVID-19 in liver transplant patients. *J Hepatol*. 2021;74(1):148–155. https://doi.org/10.1016/j.jhep.2020.07.040.
28. Giovinazzo F, Vaccaro A, Pascale MM, et al. SARS-CoV-2 infection in adult liver transplantation recipients: a systematic review of risk factors for mortality and immunosuppression role. *Eur Rev Med Pharmacol Sci*. 2023;27(4):1695–1707. https://doi.org/10.26355/eurrev_202302_31413.
29. Pinto-Álvarez M, Fernández-Niño JA, Arregocés-Castillo L, et al. Real-world evidence of COVID-19 vaccines effectiveness in solid-organ transplant recipient population in Colombia: a study nested in the Esperanza cohort. *Transplantation*. 2023;107(1):216–224. https://doi.org/10.1097/TP.0000000000004411.
30. Qin CX, Moore LW, Anjan S, et al. Risk of breakthrough SARS-CoV-2 infections in adult transplant recipients. *Transplantation*. 2021;105(11):e265–e266. https://doi.org/10.1097/TP.0000000000003907.
31. Solera JT, Ierullo M, Arbol BG, et al. Bivalent COVID-19 mRNA vaccine against omicron subvariants in immunocompromised patients. *Lancet Infect Dis*. 2023;23(8):e266–e267. https://doi.org/10.1016/S1473-3099(23)00357-2.
32. Digital Agency. Digital Agency, Government of Japan: Vaccine Recording System. https://info.vrs.digital.go.jp/dashboard/. Accessed November 12, 2023.
33. Hall VG, Solera JT, Al-Alahmadi G, et al. Severity of COVID-19 among solid organ transplant recipients in Canada, 2020-2021: a prospective, multicentre cohort study. *CMAJ*. 2022;194(33):E1155–E1163. https://doi.org/10.1503/cmaj.220620.
34. Bae S, Alejo JL, Chiang TPY, et al. mTOR inhibitors, mycophenolates, and other immunosuppression regimens on antibody response to SARS-CoV-2 mRNA vaccines in solid organ transplant recipients. *Am J Transplant*. 2022;22(12):3137–3142. https://doi.org/10.1111/ajt.17158.
35. Coll E, Fernández-Ruiz M, Padilla M, et al. COVID-19 in solid organ transplant recipients in Spain throughout 2020: catching the wave? *Transplantation*. 2021;105(10):2146–2155. https://doi.org/10.1097/tp.0000000000003873.
36. Raja MA, Mendoza MA, Villavicencio A, et al. COVID-19 in solid organ transplant recipients: a systematic review and meta-analysis of current literature. *Transplant Rev (Orlando)*. 2021;35(1):100588. https://doi.org/10.1016/j.trre.2020.100588.
37. Lefaucheur C, Louis K, Morris AB, et al. Clinical recommendations for posttransplant assessment of anti-HLA (human leukocyte antigen) donor-specific antibodies: a sensitization in transplantation: assessment of risk consensus document. *Am J Transplant*. 2023;23(1):115–132. https://doi.org/10.1016/j.ajt.2022.11.013.
38. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol*. 2023;21(3):133–146. https://doi.org/10.1038/s41579-022-00846-2.
39. Amorim CEN, Gomes VLT, Cristelli MP, et al. High prevalence of long-COVID among kidney transplant recipients: a longitudinal cohort study. *Transplantation*. 2022;106(12):2408–2415. https://doi.org/10.1097/tp.0000000000004359.
40. Helanterä I, Gissler M, Rimhanen-Finne R, et al. Epidemiology of laboratory-confirmed influenza among kidney transplant recipients compared to the general population—a nationwide cohort study. *Am J Transplant*. 2021;21(5):1848–1856. https://doi.org/10.1111/ajt.16421.
41. Passamonti F, Cattaneo C, Arcaini L, et al. Clinical characteristics and risk factors associated with COVID-19 severity in patients with haematological malignancies in Italy: a retrospective, multicentre, cohort study. *Lancet Haematol*. 2020;7(10):e737–e745. https://doi.org/10.1016/S2352-3026(20)30251-9.
42. Brenner EJ, Ungaro RC, Gearry RB, et al. Corticosteroids, but not TNF antagonists, are associated with adverse COVID-19 outcomes in patients with inflammatory bowel diseases: results from an international registry. *Gastroenterology*. 2020;159(2):481–491.e3. https://doi.org/10.1053/j.gastro.2020.05.032.
43. Rutter M, Lanyon PC, Grainge MJ, et al. COVID-19 infection, admission and death and the impact of corticosteroids among people with rare autoimmune rheumatic disease during the second wave of COVID-19 in England: results from the RECORDER Project. *Rheumatology (Oxford)*. 2023;62(12):3828–3837. https://doi.org/10.1093/rheumatology/kead150.

ANTICANCER RESEARCH *44*: 1751-1757 (2024)
doi:10.21873/anticanres.16974

Patient Age and *EGFR*-positive Non-small Cell Lung Cancer: A Multicenter Retrospective Study

YOSUKE MAEZAWA[1], MANATO TAGUCHI[2,3], TAKESHI KAWAKAMI[2,3], TOSHIHIDE INUI[3,4], SHINICHIRO OKAUCHI[1], TAKESHI NUMATA[5], TOSHIHIRO SHIOZAWA[3], KUNIHIKO MIYAZAKI[6], RYOTA NAKAMURA[5], KESATO IGUCHI[1], TAKEO ENDO[5], TOHRU SAKAMOTO[4], HIROAKI SATOH[1] and NOBUYUKI HIZAWA[3]

[1]*Divisions of Respiratory Medicine and Thoracic Surgery, Mito Medical Center, University of Tsukuba-Mito Kyodo General Hospital, Mito, Japan;*
[2]*Division of Respiratory Medicine, Kobari General Hospital, Noda, Japan;*
[3]*Division of Respiratory Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan;*
[4]*Division of Respiratory Medicine, Tsukuba Memorial Hospital, Tsukuba, Japan;*
[5]*Departments of Respiratory Medicine and Surgery, National Hospital Organization Mito Medical Center, Ibarakimachi, Japan;*
[6]*Division of Respiratory Medicine, Ryugasaki Saiseikai Hospital, Ryugasaki, Japan*

Abstract. *Background/Aim: The median age of subjects in many clinical trials of epidermal growth factor receptor (EGFR)-tyrosine kinase inhibitor conducted to date has been approximately 60 years. However, it is not uncommon to encounter EGFR gene-positive patients in their 70s or 80s. Based on information obtained from these clinical trials, EGFR gene-positive non-small cell lung cancer (NSCLC) patients are considered to be younger than EGFR-negative patients. In this study, we analyzed clinical data to identify whether this assumption is true. Patients and Methods: We retrospectively reviewed the medical records of NSCLC patients diagnosed in a multicenter clinical practice from 2009 to 2023. Patients included all cases of non-advanced and advanced NSCLC. Results: Information on 2,540 patients, including 605 EGFR gene-positive patients, was collected. The median age of EGFR-positive and EGFR-negative patients was 72 years and 71 years, respectively, and there was no significant difference in the age of patients between these two groups (p=0.7887). The most common age in these two groups was 70 years. Among the EGFR gene subtypes, the frequency of exon 19 deletion decreased with age, whereas that of EGFR L858R increased. Conclusion: Patients in their 70s and 80s with non-small cell lung cancer were relatively frequently EGFR gene-positive. To avoid missing out on treatment opportunities, EGFR gene testing should also be performed on patients in this age group.*

Correspondence to: Professor Hiroaki Satoh, MD, Ph.D., Division of Respiratory Medicine, Mito Medical Center, University of Tsukuba, 3-2-7 Miya-machi, Mito, Ibaraki, 3100015, Japan. Tel: +81 292312371, Fax: +81 292215137, e-mail: hirosato@md.tsukuba.ac.jp

Key Words: EGFR, epidermal growth factor receptor, non-small cell lung-cancer, age factors, aged.

The epidermal growth factor receptor (*EGFR*) gene was the first driver gene identified in non-small cell lung cancer (NSCLC) (1). The introduction of tyrosine kinase inhibitors (TKIs) that act on this gene has dramatically improved treatment outcomes for patients with *EGFR* gene-positive advanced NSCLC (1). Initially, *EGFR*-positive NSCLC patients were more likely reported to be female, nonsmokers, had adenocarcinoma, and be of Asian ethnicity, and these characteristics remain uncontested (2, 3). As for other characteristics, *EGFR* positivity has been reported to more frequently characterize younger patients (2). In fact, the median age of patients in many clinical trials of TKIs in advanced *EGFR*-positive patients was between 50 and 65 years (4-6). In a recent clinical trial of osimertinib, a third-generation EGFR-TKI, as postoperative adjuvant therapy, the median age of the patients was similar (7-9). Likewise, in retrospective "real-practice" studies of *EGFR*-mutated patients, the patients tended to be in their 50s to mid-60s (10-20). However, most of these studies were small-scale (21-25), and few included hundreds of *EGFR* gene-positive patients (10-20, 26). In contrast, in recent years, a number of elderly patients have presented with *EGFR* mutations (27-29). Additionally, questions have emerged regarding the applicability of TKI doses determined in clinical trials, which mostly evaluated younger patients, to older patients (30-32).

Table I. *Comparison of clinical features in non-small cell lung cancer patients with or without EGFR mutation.*

	Patients with *EGFR* mutation	Patients without *EGFR* mutation	*p*-Value
Number of patients	605	1.935	
Sex, male:female	228:377	1,513:422	0.0001
PS, 0-1:2-4	566:39	1,666:269	0.0001
Age, median (range) years	72 (35-82)	71 (29-95)	0.7887
Stage, IA-IIIA:IIIB-IVB	344:261	928:1,007	0.0001
Pathology, adenocarcinoma:others than adenocarcinoma	586:19	1,171:764	0.0001
Type of EGFR: Exon 19 deletion, Exon 21 L858R, uncommon mutations	290:268:47		

PS: Performance status.

In view of this background, we conducted a large-scale retrospective study of *EGFR*-positive and *EGFR*-negative patients with the aim of determining the age distribution of *EGFR*-positive patients in clinical practice. This study included resectable and unresectable *EGFR*-mutated patients. We also compared the characteristics of *EGFR*-positive and *EGFR*-negative patients.

Patients and Methods

We investigated the medical charts of all NSCLC patients diagnosed from May 2018 to April 2023 at 11 medical institutions affiliated with Tsukuba University. The pathological diagnosis of each patient was based on the WHO classification (33). Using head computed tomography or magnetic resonance imaging, bone scans, and ultrasonography and/or computed tomography of the abdomen prior to initiation of any treatment, all patients underwent TNM classification (34). The following information on patients at the time of NSCLC diagnosis was analyzed; sex, age, Eastern Cooperative Oncology Group performance status (PS), clinical stage, and *EGFR* gene mutation status and subtype.

The chi-squared test was used to test for differences in proportions. Mann–Whitney *U*-test was used to compare values between two unpaired groups, such as patient age. Spearman correlation coefficient was used to test the correlation between age and the number of patients with *EGFR* gene types. Statistical analysis was performed using SPSS Statistics software (version 11.5; SPSS, Chicago, IL, USA). A *p*-value less than 0.01 was considered to indicate a significant difference.

This research was approved by the University of Tsukuba Mito Medical Center-Mito Kyodo General Hospital (NO-23-53) and each Institutional Review Board.

Results

Patient characteristics. Information was collected on a total of 2,540 patients during the study period, including 605 *EGFR*-positive NSCLC patients and 1,935 *EGFR*-negative patients, resulting in an *EGFR* positivity rate of 23.8%. The *EGFR* positivity rate in adenocarcinoma patients was 33.3%. Table I shows the characteristics of *EGFR*-positive and *EGFR*-negative patients. There were significant differences in the proportion of women, good performance status, and non-advanced stage, and adenocarcinoma between *EGFR*-positive and *EGFR*-negative patients, whereas PS and clinical stage at the time of diagnosis of NSCLC did not differ significantly between these two groups.

The median age of *EGFR*-positive patients was 72 years, with the most frequent age in the age group 70 to 80 years, whereas the median age of *EGFR* gene-negative patients was 71 years. Patient age between these two groups did not differ significantly (p=0.7887).

Age distribution of EGFR-positive and EGFR-negative patients. When comparing the distribution of patients in both *EGFR*-positive and *EGFR*-negative groups, the largest number of patients were in their 70s (Figure 1A). Equally, the distribution of the proportion of patients in *EGFR*-positive and *EGFR*-negative groups was comparable (Figure 1B). There were 102 (16.9%) of 597 *EGFR*-positive patients aged 80 years or older (Figure 1B). The proportion of patients aged 80 years or more was not significantly different from that of patients aged 60 years or less (89 of 605 patients: 14.7%; p=0.3441), but was significantly higher than that of patients aged 50 years or less (22 of 605 patients: 3.6%; p=0.0001).

Age distribution of EGFR gene subtypes. The results for all *EGFR*-positive 605 patients are shown in Figure 2A and B. The highest number of patients for all subtypes were in their 70s. Uncommon mutations were found at a higher rate in patients under 50 years and over 80 years. The frequency of exon 19 deletion decreased with age (p=0.0003), and that of EGFR L858R increased with age (p=0.0035). However, there was no significant correlation in patients with uncommon *EGFR* mutations (p=0.6341).

Figure 3A and B shows the results for female patients. The distribution of exon 19 deletion and exon 21 L858R patients was similar to that of all 605 *EGFR*-positive patients. Unusual mutations were found at a higher rate in patients aged 80 years and older. Figure 4A and B shows the results for male patients. Exon 19 deletion was most common in

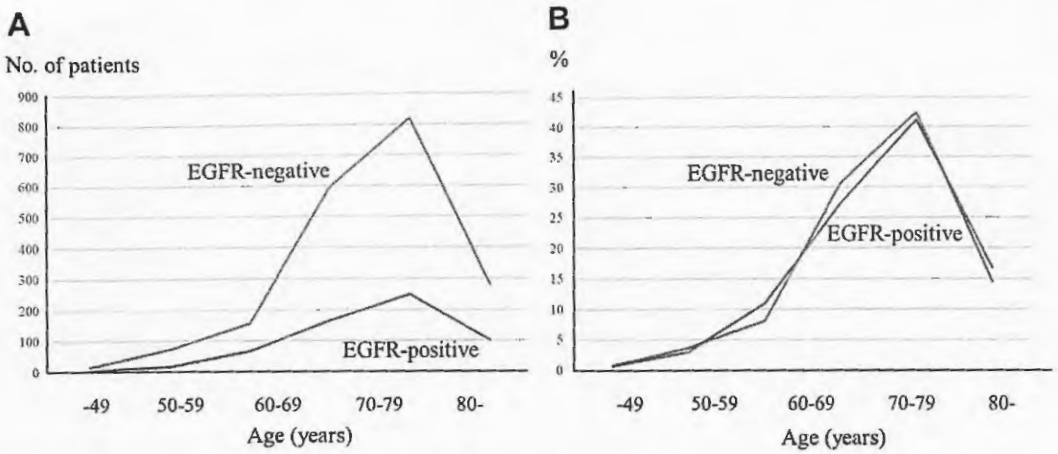


Figure 1. *Age distribution of EGFR-positive patients: number of patients (A) and percentage of patients (B).*

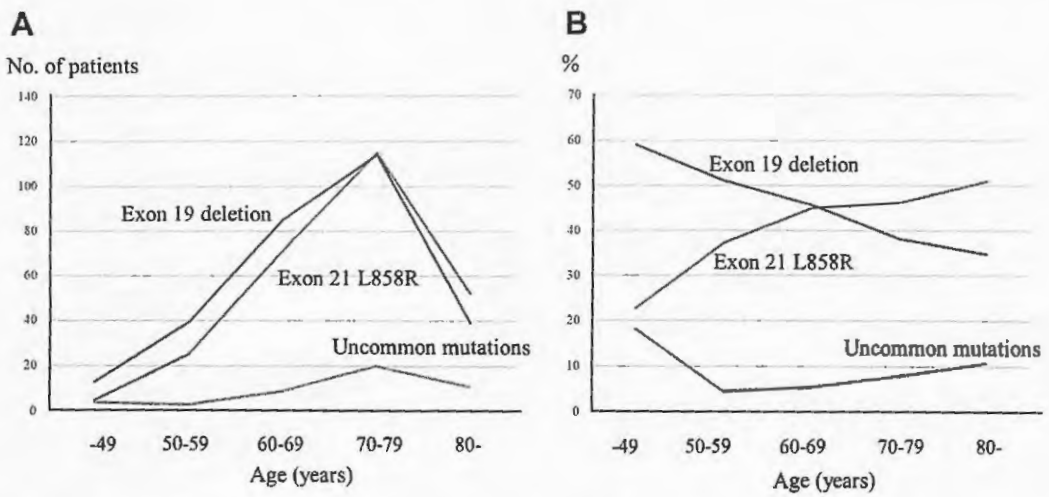


Figure 2. *Age distribution of EGFR gene subtypes in all patients: number of patients (A) and percentage of patients (B).*

people in their 70s; however, exon 21 L858R was most common in patients in their 60s. Unusual mutations were most common between 70 and 80 years. Unusual mutations were found at a higher rate in patients under 50 years of age.

Discussion

The *EGFR* gene mutation positivity rate in this study, which included 605 patients with *EGFR*-positive NSCLC, was 23.8%. The median age of *EGFR*-positive and *EGFR*-negative patients did not differ significantly. Almost 17% of *EGFR*-positive patients were over 80 years of age, constituting a noteworthy share of this patient group. Regarding the relationship between *EGFR* gene subtype and age, we revealed that the proportion of patients with exon 19 deletion decreased with age, whereas the proportion of exon 21 L858R increased. We also identified that uncommon mutations were more frequent in patients under 50 years and those over 80 years. In this study, the proportion of non-advanced patients and that of patients with favorable PS were higher in *EGFR*-positive patients than in *EGFR*-negative patients. We speculate that this may be related to the higher proportion of patients with adenocarcinoma among *EGFR*-positive patients. Adenocarcinomas have a higher incidence of onset in peripheral lung fields, and hence may be detected more frequently incidentally during screening or treatment for other diseases.

Generalizing the findings on *EGFR* gene-positive patients, it is necessary to consider the differences in *EGFR* gene positivity

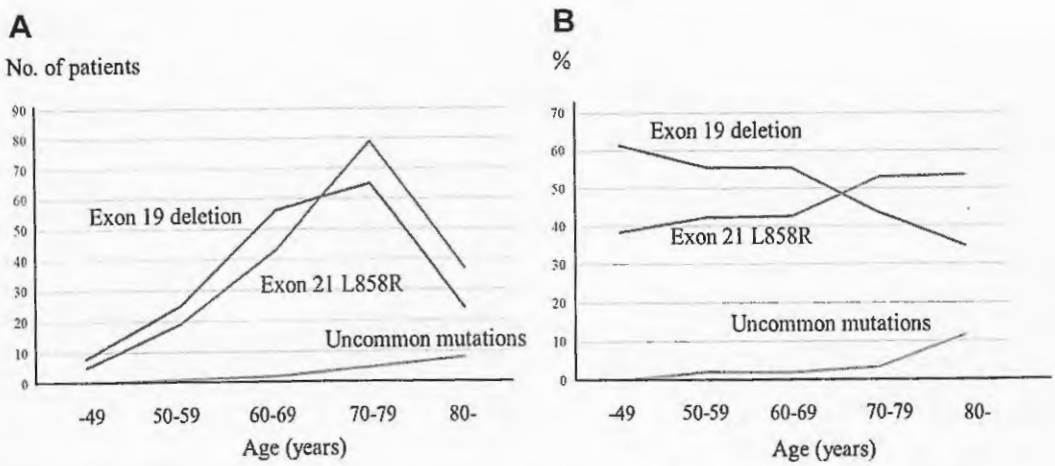


Figure 3. *Age distribution of EGFR gene subtypes in female patients: number of patients (A) and percentage of patients (B).*

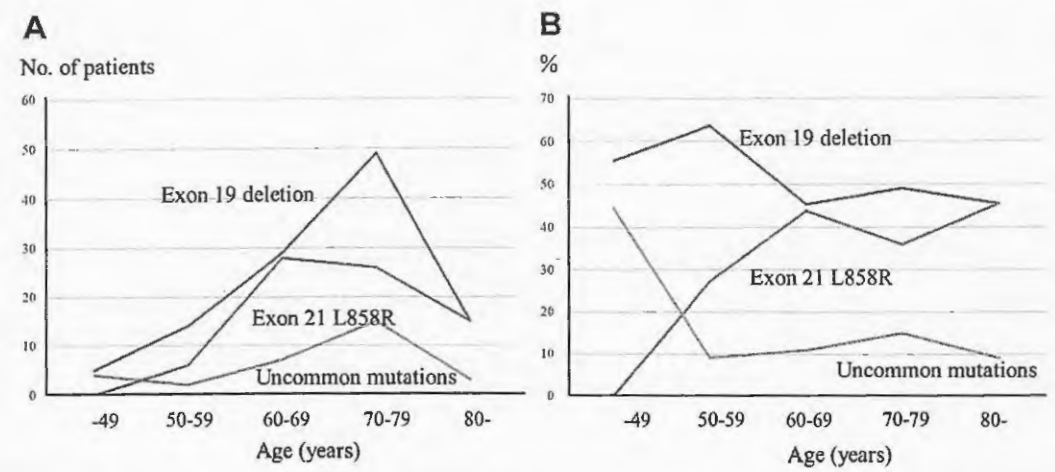


Figure 4. *Age distribution of EGFR gene subtypes in male patients: number of patients (A) and percentage of patients (B).*

between ethnicities. So far, it remains to be shown whether our results may be limited to East Asians, who have a high rate of *EGFR* gene positivity, or whether the results can be extrapolated to Caucasians as well. The *EGFR* gene positivity rate in our survey was 23.8%, which was situated between the relatively high positivity rate reported from East Asia (13, 14, 16, 21) and that of Caucasians (12, 18, 23). Therefore, we believe that our findings may be generalizable to other regions, rather than be restricted to a unique patient population. Moreover, the median age of clinical trial subjects for TKIs against *EGFR* mutations is around 60 years (4-6) as well as that of trials of adjuvant therapy in *EGFR*-positive patients (3, 8, 9), which may have led to an impression that *EGFR* gene-positive patients are younger than *EGFR*-negative patients.

In order to verify the results of clinical trials for advanced *EGFR*-positive NSCLC, retrospective studies have been performed to examine the actual situation in clinical practice. In these studies, the majority of subjects were in their mid- or late 60s (12-20), slightly older than the median age of clinical trial subjects. One study included a large number of patients with an average age of 71 years, but it was limited to patients with advanced disease (26). However, most studies included fewer than 100 *EGFR*-positive patients (21-25), and only a few large-scale research studies featured 1,000 or more *EGFR*-positive patients (10-20). Furthermore, many reports covered only advanced *EGFR*-positive patients (10, 15, 17, 22, 24-26) or only resected *EGFR*-positive patients (14, 21), and few large-scale research studies included patients at all stages (11-13). A very limited number of studies have compared the demographics of patients with EGFR positivity, both resectable and advanced, to those negative for *EGFR* treated during the same period (10-12, 14, 16, 23). Therefore, we consider it

clinically significant that in this study *EGFR*-positive patients presented relatively frequently among elderly patients.

Another noteworthy result of this study was the change in the proportion of *EGFR* gene subtypes with age. As age increased, the proportion of patients with exon 19 deletion decreased, and the proportion of patients with exon 21 L858R increased. This age-related change in common mutations was more pronounced in female patients. Similar results were obtained in the report by Evans *et al.* (18). Their report did not examine the sex of the patients, but four age groups; 40 years and younger, 41-60 years, 61-80 years, and 81 years and older (18). Our study included five age groups; 50 years and younger, 51-60 years, 61-70 years, 71-80 years, and 81 years and older, and we stratified them by sex. Interestingly, *EGFR* gene subtypes differed with age, but the reason is not clear. A relationship with smoking, especially light smoking, was suspected, as it has been reported that non-smokers or light smokers present more frequently among *EGFR* gene-positive patients (35). However, future studies are needed to confirm this. It might also be important to note that patients with rare mutations were found in all age groups, but were more common among younger and older patients. In such cases, testing methods such as upfront multiplex gene test that can examine multiple driver genes in detail will become increasingly useful (36).

Despite obtaining significant results, this study had some limitations. First of all, we must mention the limitations regarding the various testing methods for the *EGFR* gene. Because this study was a multicenter study, several *EGFR* tests were allowed for determination of *EGFR* positivity. Second, the inclusion period was relatively long. Several driver genes other than *EGFR* were discovered during this time, but patients positive for these genes were treated as *EGFR* gene-negative patients. Third, we could not collect information on smoking. Despite these limitations, we believe that the study yielded some important novel findings.

Conclusion

As revealed in this study of *EGFR*-positive patients in clinical practice, elderly NSCLC patients presented relatively frequently as positive for *EGFR* mutations. Therefore, even in elderly NSCLC patients, *EGFR* gene testing should be performed, taking into account the possibility that patients could benefit from treatment with EGFR-TKIs, which have a relatively low incidence of serious side effects that impair quality of life.

Conflicts of Interest

The Authors declare that they have no competing interests in relation to this study.

Authors' Contributions

YM, SO, KM and HS designed the study. YM, MT, TK, TI, TN, TS, KM, RN, KI, TS, and SH collected the data. YM, SO, KM and HS analyzed the data. YM, KM, HS and NH prepared the manuscript. TE, TS, HS and NH supervised the study. All Authors approved the final version for submission.

Funding

No funding was received.

References

1 Christofyllakis K, Monteiro AR, Cetin O, Kos IA, Greystoke A, Luciani A: Biomarker guided treatment in oncogene-driven advanced non-small cell lung cancer in older adults: A Young International Society of Geriatric Oncology report. J Geriatr Oncol 13(8): 1071-1083, 2022. DOI: 10.1016/j.jgo.2022.04.013

2 Minchom A, Yu KC, Bhosle J, O'Brien M: The diagnosis and treatment of brain metastases in EGFR mutant lung cancer. CNS Oncol 3(3): 209-217, 2014. DOI: 10.2217/cns.14.19

3 Zhang YL, Yuan JQ, Wang KF, Fu XH, Han XR, Threapleton D, Yang ZY, Mao C, Tang JL: The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget 7(48): 78985-78993, 2016. DOI: 10.18632/oncotarget.12587

4 Zhou C, Wu YL, Chen G, Feng J, Liu XQ, Wang C, Zhang S, Wang J, Zhou S, Ren S, Lu S, Zhang L, Hu C, Hu C, Luo Y, Chen L, Ye M, Huang J, Zhi X, Zhang Y, Xiu Q, Ma J, Zhang L, You C: Erlotinib *versus* chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. Lancet Oncol 12(8): 735-742, 2011. DOI: 10.1016/S1470-2045(11)70184-X

5 Soria JC, Felip E, Cobo M, Lu S, Syrigos K, Lee KH, Göker E, Georgoulias V, Li W, Isla D, Guclu SZ, Morabito A, Min YJ, Ardizzoni A, Gadgeel SM, Wang B, Chand VK, Goss GD, LUX-Lung 8 Investigators: Afatinib *versus* erlotinib as second-line treatment of patients with advanced squamous cell carcinoma of the lung (LUX-Lung 8): an open-label randomised controlled phase 3 trial. Lancet Oncol 16(8): 897-907, 2015. DOI: 10.1016/S1470-2045(15)00006-6

6 Soria JC, Ohe Y, Vansteenkiste J, Reungwetwattana T, Chewaskulyong B, Lee KH, Dechaphunkul A, Imamura F, Nogami N, Kurata T, Okamoto I, Zhou C, Cho BC, Cheng Y, Cho EK, Voon PJ, Planchard D, Su WC, Gray JE, Lee SM, Hodge R, Marotti M, Rukazenkov Y, Ramalingam SS, FLAURA Investigators: Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. N Engl J Med 378(2): 113-125, 2018. DOI: 10.1056/NEJMoa1713137

7 Zhong WZ, Wang Q, Mao WM, Xu ST, Wu L, Shen Y, Liu YY, Chen C, Cheng Y, Xu L, Wang J, Fei K, Li XF, Li J, Huang C, Liu ZD, Xu S, Chen KN, Xu SD, Liu LX, Yu P, Wang BH, Ma HT, Yan HH, Yang XN, Zhou Q, Wu YL; ADJUVANT investigators: Gefitinib *versus* vinorelbine plus cisplatin as adjuvant treatment for stage II-IIIA (N1-N2) EGFR-mutant NSCLC (ADJUVANT/CTONG1104): a randomised, open-label,

phase 3 study. Lancet Oncol 19(1): 139-148, 2018. DOI: 10.1016/S1470-2045(17)30729-5

8 Wu YL, Tsuboi M, He J, John T, Grohe C, Majem M, Goldman JW, Laktionov K, Kim SW, Kato T, Vu HV, Lu S, Lee KY, Akewanlop C, Yu CJ, de Marinis F, Bonanno L, Domine M, Shepherd FA, Zeng L, Hodge R, Atasoy A, Rukazenkov Y, Herbst RS, ADAURA Investigators: Osimertinib in resected EGFR-mutated non-small-cell lung cancer. N Engl J Med 383(18): 1711-1723, 2020. DOI: 10.1056/NEJMoa2027071

9 Tada H, Mitsudomi T, Misumi T, Sugio K, Tsuboi M, Okamoto I, Iwamoto Y, Sakakura N, Sugawara S, Atagi S, Takahashi T, Hayashi H, Okada M, Inokawa H, Yoshioka H, Takahashi K, Higashiyama M, Yoshino I, Nakagawa K, West Japan Oncology Group: Randomized phase III study of gefitinib *versus* cisplatin plus vinorelbine for patients with resected stage II-IIIA non-small-cell lung cancer with EGFR mutation (IMPACT). J Clin Oncol 40(3): 231-241, 2022. DOI: 10.1200/JCO.21.01729

10 Choi YL, Sun JM, Cho J, Rampal S, Han J, Parasuraman B, Guallar E, Lee G, Lee J, Shim YM: EGFR mutation testing in patients with advanced non-small cell lung cancer: a comprehensive evaluation of real-world practice in an East Asian tertiary hospital. PLoS One 8(2): e56011, 2013. DOI: 10.1371/journal.pone.0056011

11 Hong S, Fang W, Hu Z, Zhou T, Yan Y, Qin T, Tang Y, Ma Y, Zhao Y, Xue C, Huang Y, Zhao H, Zhang L: A large-scale cross-sectional study of ALK rearrangements and EGFR mutations in non-small-cell lung cancer in Chinese Han population. Sci Rep 4: 7268, 2014. DOI: 10.1038/srep07268

12 Rosell R, Moran T, Queralt C, Porta R, Cardenal F, Camps C, Majem M, Lopez-Vivanco G, Isla D, Provencio M, Insa A, Massuti B, Gonzalez-Larriba JL, Paz-Ares L, Bover I, Garcia-Campelo R, Moreno MA, Catot S, Rolfo C, Reguart N, Palmero R, Sánchez JM, Bastus R, Mayo C, Bertran-Alamillo J, Molina MA, Sanchez JJ, Taron M, Spanish Lung Cancer Group: Screening for epidermal growth factor receptor mutations in lung cancer. N Engl J Med 361(10): 958-967, 2009. DOI: 10.1056/NEJMoa0904554

13 Ueno T, Toyooka S, Suda K, Soh J, Yatabe Y, Miyoshi S, Matsuo K, Mitsudomi T: Impact of age on epidermal growth factor receptor mutation in lung cancer. Lung Cancer 78(3): 207-211, 2012. DOI: 10.1016/j.lungcan.2012.09.006

14 Nishii T, Yokose T, Miyagi Y, Daigo Y, Ito H, Isaka T, Imai K, Murakami S, Kondo T, Saito H, Oshita F, Yamada K, Matsukuma S, Tsuboi M, Nakayama H, Masuda M: Clinicopathological features and EGFR gene mutation status in elderly patients with resected non-small-cell lung cancer. BMC Cancer 14: 610, 2014. DOI: 10.1186/1471-2407-14-610

15 Inoue A, Yoshida K, Morita S, Imamura F, Seto T, Okamoto I, Nakagawa K, Yamamoto N, Muto S, Fukuoka M: Characteristics and overall survival of EGFR mutation-positive non-small cell lung cancer treated with EGFR tyrosine kinase inhibitors: a retrospective analysis for 1660 Japanese patients. Jpn J Clin Oncol 46(5): 462-467, 2016. DOI: 10.1093/jjco/hyw014

16 Wu SG, Chang YL, Yu CJ, Yang PC, Shih JY: Lung adenocarcinoma patients of young age have lower EGFR mutation rate and poorer efficacy of EGFR tyrosine kinase inhibitors. ERJ Open Res 3(3): 00092-2016, 2017. DOI: 10.1183/23120541.00092-2016

17 Okamoto I, Morita S, Tashiro N, Imamura F, Inoue A, Seto T, Yamamoto N, Ohe Y, Nakagawa K, Fukuoka M: Real world treatment and outcomes in EGFR mutation-positive non-small cell lung cancer: Long-term follow-up of a large patient cohort. Lung Cancer 117: 14-19, 2018. DOI: 10.1016/j.lungcan.2018.01.005

18 Evans M, O'Sullivan B, Smith M, Hughes F, Mullis T, Trim N, Taniere P: Large-scale EGFR mutation testing in clinical practice: Analysis of a series of 18,920 non-small cell lung cancer cases. Pathol Oncol Res 25(4): 1401-1409, 2019. DOI: 10.1007/s12253-018-0460-2

19 Palacio S, Pontes L, Prado E, Arshad J, Ali R, Piha T, Bacchi CE, Mudad R, Lopes G: EGFR mutation testing: Changing patterns of molecular testing in Brazil. Oncologist 24(4): e137-e141, 2019. DOI: 10.1634/theoncologist.2018-0254

20 Montella T, Zalis M, Zukin M, Cordeiro de Lima VC, Baldotto C, De Marchi P, Salles P, Mathias C, Barrios C, Kawamura C, Calabrich A, Araújo LH, Castro G, Bustamante C, Santa Maria A, Reis M, Ferreira CG: EGFR mutation detection in Brazilian patients with non-small-cell lung cancer: lessons from real-world data scenario of molecular testing. JCO Glob Oncol 9: e2200426, 2023. DOI: 10.1200/GO.22.00426

21 Choi YH, Lee JK, Kang HJ, Lee TS, Kim HR, Kim CH, Koh JS, Baek HJ, Lee JC, Na II: Association between age at diagnosis and the presence of EGFR mutations in female patients with resected non-small cell lung cancer. J Thorac Oncol 5(12): 1949-1952, 2010. DOI: 10.1097/jto.0b013e3181f38816

22 Feng S, Wang Y, Cai K, Wu H, Xiong G, Wang H, Zhang Z: Randomized adjuvant chemotherapy of EGFR-mutated non-small cell lung cancer patients with or without icotinib consolidation therapy. PLoS One 10(10): e0140794, 2015. DOI: 10.1371/journal.pone.0140794

23 Kelly D, Mc Sorley L, O'Shea E, Mc Carthy E, Bowe S, Brady C, Sui J, Dawod MA, O'Brien O, Graham D, McCarthy J, Burke L, Power D, O'Reilly S, Bambury RM, Mahony DO: A regional analysis of epidermal growth factor receptor (EGFR) mutated lung cancer for HSE South. Ir J Med Sci 186(4): 855-857, 2017. DOI: 10.1007/s11845-017-1579-y

24 Ding PN, Roberts TL, Chua W, Becker TM, Descallar J, Yip PY, Bray V: Clinical outcomes in patients with advanced epidermal growth factor receptor-mutated non-small-cell lung cancer in South Western Sydney Local Health District. Intern Med J 47(12): 1405-1411, 2017. DOI: 10.1111/imj.13555

25 Hirsch FR, Sequist LV, Gore I, Mooradian M, Simon G, Croft EF, DeVincenzo D, Munley J, Stein D, Freivogel K, Sifakis F, Bunn PA Jr: Long-term safety and survival with gefitinib in select patients with advanced non-small cell lung cancer: Results from the US IRESSA Clinical Access Program (ICAP). Cancer 124(11): 2407-2414, 2018. DOI: 10.1002/cncr.31313

26 Helland Å, Andersen KK, Myklebust TÅ, Johannesen TB, Aarøe J, Enerly E: EGFR-mutation testing and TKI treatment patterns in locally advanced and metastatic NSCLC in Norway – A nationwide retrospective cohort study. Cancer Treat Res Commun 33: 100636, 2022. DOI: 10.1016/j.ctarc.2022.100636

27 Tufman A, Kahnert K, Duell T, Kauffmann-Guerrero D, Milger K, Schneider C, Stump J, Syunyaeva Z, Huber RM, Reu S: Frequency and clinical relevance of EGFR mutations and EML4-ALK translocations in octogenarians with non-small cell lung cancer. Onco Targets Ther 10: 5179-5186, 2017. DOI: 10.2147/OTT.S140472

28 Corre R, Gervais R, Guisier F, Tassy L, Vinas F, Lamy R, Fraboulet G, Greillier L, Doubre H, Descourt R, Chouaid C, Auliac JB: Octogenarians with EGFR-mutated non-small cell lung cancer treated by tyrosine-kinase inhibitor: a multicentric real-world study

assessing tolerance and efficacy (OCTOMUT study). Oncotarget 9(9): 8253-8262, 2018. DOI: 10.18632/oncotarget.23836
29 Hashimoto S, Fujii Y, Ujiie M, Miyazaki K, Sato S, Kodama T, Nagatsu T, Kayauchi N, Satoh H: Refusal of subsequent treatment in patients with EGFR-mutant non-small-cell lung cancer after response to EGFR-TKIs. Cancer Diagn Progn 3(2): 251-256, 2023. DOI: 10.21873/cdp.10209
30 Agraso S, Lázaro M, Firvida XL, Santomé L, Fernández N, Azpitarte C, Leon L, Garcia C, Hudobro G, Areses MC, Campos B, Quiroga N, García J, Casal J: Real-world data with afatinib in Spanish patients with treatment-naïve non-small-cell lung cancer harboring exon 19 deletions in epidermal growth factor receptor (Del19 EGFR): Clinical experience of the Galician Lung Cancer Group. Cancer Treat Res Commun 33: 100646, 2022. DOI: 10.1016/j.ctarc.2022.100646
31 Tsubata Y, Masuda T, Hamai K, Taniwaki M, Tanino A, Hotta T, Hamaguchi M, Hamaguchi S, Yamasaki M, Ishikawa N, Fujitaka K, Sutani A, Isobe T: Efficacy of erlotinib and its effects on the quality of life of older patients with epidermal growth factor receptor-mutant non-small cell lung cancer: A prospective, multicenter, dose-modification study. Geriatr Gerontol Int 21(10): 881-886, 2021. DOI: 10.1111/ggi.14243
32 Nakao M, Muramatsu H, Sone K, Aoki S, Akiko H, Kagawa Y, Sato H, Kunieda T: Epidermal growth factor receptor-tyrosine kinase inhibitors for non-small-cell lung cancer patients aged 80 years or older: A retrospective analysis. Mol Clin Oncol 3(2): 403-407, 2015. DOI: 10.3892/mco.2014.453
33 WHO Classification of Tumours Editorial Board. World Health Organization Classification of Tumours, Thoracic Tumours 5th ed. Lyon, France, IARC Press, 2021.
34 Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, Nicholson AG, Groome P, Mitchell A, Bolejack V, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, Advisory Boards, and Participating Institutions, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee Advisory Boards and Participating Institutions: The IASLC lung cancer staging project: Proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. J Thorac Oncol 11(1): 39-51, 2016. DOI: 10.1016/j.jtho.2015.09.009
35 Kim IA, Lee JS, Kim HJ, Kim WS, Lee KY: Cumulative smoking dose affects the clinical outcomes of EGFR-mutated lung adenocarcinoma patients treated with EGFR-TKIs: a retrospective study. BMC Cancer 18(1): 768, 2018. DOI: 10.1186/s12885-018-4691-0
36 Kanasaki H, Ozawa Y, Nakamura N, Nagasaki K, Matsuyama W, Akahori D, Niwa M, Ogasawara T, Sato J: Upfront multiplex gene test helps prolong survival in advanced non-small cell lung cancer. Anticancer Res 44(2): 723-730, 2024. DOI: 10.21873/anticanres.16863

Received January 20, 2024
Revised February 10, 2024
Accepted February 12, 2024

Received: 22 August 2023 | Revised: 27 March 2024 | Accepted: 15 April 2024

DOI: 10.1002/jrsm.1720

RESEARCH ARTICLE

Research Synthesis Methods WILEY

The impact of continuity correction methods in Cochrane reviews with single-zero trials with rare events: A meta-epidemiological study

Yasushi Tsujimoto[1,2,3,4] | **Yusuke Tsutsumi**[4,5,6] | **Yuki Kataoka**[4,7,8,9] | **Akihiro Shiroshita**[4,10] | **Orestis Efthimiou**[11,12] | **Toshi A. Furukawa**[1]

[1]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan

[2]Division of Rheumatology, Department of Internal Medicine, Showa University School of Medicine, Tokyo, Japan

[3]Oku Medical Clinic, Osaka, Japan

[4]Scientific Research Works Peer Support Group (SRWS-PSG), Osaka, Japan

[5]Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[6]Department of Emergency Medicine, National Hospital Organization Mito Medical Center, Ibaraki, Japan

[7]Department of Healthcare Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto University, Kyoto, Japan

[8]Department of Internal Medicine, Kyoto Min-iren Asukai Hospital, Kyoto, Japan

[9]Section of Clinical Epidemiology, Department of Community Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

[10]Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

[11]Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland

[12]Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

Correspondence
Toshi A. Furukawa, Department of Health Promotion and Human Behavior, School of Public Health in the Graduate School of Medicine, Kyoto University, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.
Email: furukawa@kuhp.kyoto-u.ac.jp

Funding information
Japan Society for the Promotion of Science, Grant/Award Number: 22K10423

Abstract

Meta-analyses examining dichotomous outcomes often include single-zero studies, where no events occur in intervention or control groups. These pose challenges, and several methods have been proposed to address them. A fixed continuity correction method has been shown to bias estimates, but it is frequently used because sometimes software (e.g., RevMan software in Cochrane reviews) uses it as a default. We aimed to empirically compare results using the continuity correction with those using alternative models that do not require correction. To this aim, we reanalyzed the original data from 885 meta-analyses in Cochrane reviews using the following methods: (i) Mantel–Haenszel model with a fixed continuity correction, (ii) random effects inverse variance model with a fixed continuity correction, (iii) Peto method (the three models available in RevMan), (iv) random effects inverse variance model with the treatment arm continuity correction, (v) Mantel–Haenszel model without correction, (vi) logistic regression, and (vii) a Bayesian random effects model

Yasushi Tsujimoto and Yusuke Tsutsumi are co-first authors.

with binominal likelihood. For each meta-analysis we calculated ratios of odds ratios between all methods, to assess how the choice of method may impact results. Ratios of odds ratios <0.8 or <1.25 were seen in ∼30% of the existing meta-analyses when comparing results between Mantel–Haenszel model with a fixed continuity correction and either Mantel–Haenszel model without correction or logistic regression. We concluded that injudicious use of the fixed continuity correction in existing Cochrane reviews may have substantially influenced effect estimates in some cases. Future updates of RevMan should incorporate less biased statistical methods.

KEYWORDS

continuity correction, meta-analysis, single-zero studies, zero event, zero-cell correction

Highlights

What is already known

- Meta-analyses of dichotomous outcomes often use "continuity correction" to avoid computational issues in studies with zero events in one arm (single-zero studies). However, this could potentially bias study estimates.
- Despite the availability of more advanced models for handling single-zero studies, RevMan, the official Cochrane software for meta-analyses, continues to use continuity corrections.
- While simulation studies have highlighted potential issues associated with the use of continuity corrections and despite the fact that more advanced methods have been proposed, the impact of using alternative methods on effect estimates within established Cochrane reviews remains uncertain.

What is new

- We contrasted the effect estimates obtained by several alternative methods against the Mantel–Haenszel model with a continuity correction (RevMan MH), the default model in RevMan software, using data from 885 established meta-analyses within Cochrane reviews.
- Meta-analyses including single-zero studies were commonly seen in Cochrane reviews, and 64% of them used RevMan MH to deal with single-zero studies.
- In scenarios with low event rates and small sample sizes, a difference in the point estimates of odds ratios of 25% or more was observed in approximately 30% of existing meta-analyses, when comparing results between RevMan MH and either MH or logistic regression.

Potential impact for Research Synthesis Methods readers

- Our findings underscore that researchers, readers, peer-reviewers, and journal editors of meta-analyses should interpret results carefully when they appraise meta-analyses involving single-zero studies, as some commonly used methods may alter the estimates. Future meta-analyses should avoid using suboptimal methods.

1 | INTRODUCTION

Systematic reviews play an important role in decision-making.[1] A meta-analysis is the statistical combination of results from two or more separate studies. It yields an overall estimate of the effectiveness of an intervention compared with a control treatment. There are many methods that can be used for the meta-analysis of dichotomous outcomes, four of which are more widely used than others, and are also available as analysis options in RevMan web and RevMan 5, the current and prior official software used for Cochrane reviews.[1–3] These are three fixed-effect models, that is, Mantel–Haenszel (MH), Peto, and inverse variance (IV) methods, and one random-effects model, that is, the DerSimonian and Laird method.[1]

A meta-analysis of odds ratios or risk ratios sometimes includes studies in which one arm has zero events (single-zero studies). The inclusion of single-zero studies in meta-analyses can introduce computational errors; hence, several methods have been proposed to address these issues.[4] The fixed continuity correction, commonly implemented in software like RevMan, adds a specific value (usually 0.5) to all cells of the two-by-two tables in each study. Simulation studies, however, have raised concerns about its validity.[4–6] This method can artificially move the point estimate away from extremes, consequently leading to a bias toward the null effect, especially when the true effect is substantial. In contrast, alternative methods such as the MH without continuity correction, logistic regression models, and the Peto method may outperform methods utilizing the fixed continuity correction as they do not require zero-cell correction. However, RevMan automatically implements the fixed continuity correction when using the Mantel–Haenszel and IV methods and may result in biases toward no treatment effect.[1]

Although issues associated with the use of fixed continuity correction have been increasingly recognized and illustrated through simulation studies, the practice has not changed. This might be because, despite the theoretical arguments and simulations, review authors were uncertain about how various zero-cell correction methods would impact their results. No previous study has investigated whether the application of more sophisticated methods such as MH without corrections, logistic regression, or the Peto method would impact the effect estimates in existing meta-analyses within Cochrane reviews.[7] In the present study, we aim to empirically compare results after using the continuity correction with results from modes that do not require correction, in terms of the effect estimates and their interpretation in Cochrane reviews.

2 | METHODS

The protocol of the present study was posted on the OSF registry.[8]

2.1 | Eligibility criteria

We included all Cochrane reviews of interventions that included only randomized controlled trials (RCTs). We only included the most recent version of the reviews, when one or more updates are available. Reviews were eligible if a pair-wise meta-analysis using risk ratio or odds ratio was performed and included at least one single-zero study, that is, a study with zero events in one but not both treatment arms. Since this study primarily focused on methods to address single-zero studies, when reviews had included double-zero studies (studies with zero events in both treatment arms), we only included single-zero studies and excluded double-zero studies from our analyses. We selected meta-analyses with event rates of <5% and sample sizes of <1000, as a previous study showed that the method of dealing with single-zero studies only played a role in meta-analyses with small sample sizes and rare events.[9] We excluded a review if it was not possible to extract the 2×2 table from any of the meta-analyses therein, or if we could not download the data by clicking "Download statistical data" in the Cochrane library.

2.2 | Searches and study selection

We searched the Cochrane Database of Systematic Reviews by using a filter aimed at returning reviews on interventions from inception to September 30, 2022. We scraped the relevant reviews and downloaded data from the Cochrane library's website for each review using Python selenium package version 3.141.0.[14] Thereafter, we loaded the data to R statistical software and checked the eligibility according to the criteria above.

2.3 | Data extraction

All data used in the present study was obtained from publicly available dataset of Cochrane reviews. For each meta-analysis that included a single-zero study we extracted the following data: number of participants and events included in each study in the meta-analysis, types of statistical models used for the meta-analysis.

2.3.1 | Outcomes of interest of this study

The primary outcome was the difference in effect estimates between the results of each meta-analysis obtained after using Mantel–Haenszel model with a fixed continuity correction, that is, adding 0.5 to all cells of the 2 × 2 table (the model implemented in the RevMan software—"RevMan MH") with results obtained after using various models listed in the next paragraph. Difference in estimates between two models was quantified as a ratio of odds ratios (ROR).

2.4 | Statistical analysis

We reported the proportion of reviews with a meta-analysis including single-zero studies, and the proportion of such reviews with event rates of <5% and sample sizes of <1000 among all Cochrane reviews. We tabulated the following characteristics of included meta-analyses; number of participants and events included in each study in the meta-analysis, types of statistical models used for the meta-analysis.

We repeated each meta-analysis in the included reviews using the following methods (i) RevMan MH, (ii) random effects inverse variance model with a fixed continuity correction (REIV), (iii) random effects inverse variance model with the treatment arm continuity correction (REIV TACC), (iv) Peto method, (v) Mantel–Haenszel model without correction, (vi) the random effects logistic regression model, and (vii) a Bayesian random effects logistic regression model using a binomial likelihood for the outcome. We used uninformative prior distributions for the log-odds of the reference treatment and the treatment effects (N_0, $\sigma^2 = 1000$). We also used a vague half normal prior for the heterogeneity parameter τ^2. We ran 4 chains of 5000 iterations after 1000 burn in. The R script used for conducting the Bayesian analysis can be found in the Appendix S1.

Odds ratios were used for the effect estimates. For each meta-analysis, we calculated the ROR between these models. We presented the distribution of RORs using histograms and scatter plots. Further, we characterized their sizes using the following predefined categories:

1. Small, ROR in the range of 0.9–1.11.
2. Moderate, ROR in the range of 0.8–0.9 or 1.11–1.25.
3. Large, ROR in the range of ≥1.25 or ≤0.8.

For meta-analyses with extremely large differences, we further explored their characteristics in terms of the participants, intervention, comparator, outcomes, number of events in each study in the meta-analyses, and effect sizes from each statistical method. The criteria for selecting instances of exceptionally large differences were based on the highest and lowest values of the RORs observed in each comparison. All analyses were performed by meta package (ver.2.4-0) of R version 4.1.2, and RevMan 5.4.1.

3 | RESULTS

Our search identified 2300 Cochrane reviews of interventions, among which there were 1540 reviews that included 20,535 meta-analyses of dichotomous outcomes. Of those, 856 (56%) reviews included 4984 meta-analyses incorporating at least one single-zero study. Ultimately, we selected 383 (25%) reviews with 885 meta-analyses for our main analysis, each having a control event rate of less than 0.05 and a total participant count fewer than 1000 (Figure 1).

Table 1 shows characteristics of the included meta-analyses with a control event rate of <0.05 and a total participant count fewer than 1000. Median (interquartile range, IQR) number of studies included in the meta-analyses was 3 (2–4). RevMan MH model was the most frequently used model in the original Cochrane reviews.

Figure 2 visualizes the agreement between the results of meta-analyses using RevMan MH and other methods. Ten panels in the lower-left of the figure show histograms of logRORs, the denominator of which can be read in the diagonal element to the right, and the numerator to the diagonal element above. The upper-right panels present scatter plots of logORs, with the vertical axis representing the statistical methods indicated at diagonal element at the bottom and the horizontal axis representing the methods indicated at the left.

We found that the Peto method tended to be inconsistent with any other methods, indicated by the wide distribution of logROR histogram and by many outliners off the 45° diagonal in the scatterplot. RevMan MH was consistent with REIV or REIV TACC but not with MH, logistic regression, and the Bayesian model. MH was mostly consistent with the logistic regression model.

Table 2 shows the characteristics of the included meta-analyses categorized according to the RORs for each method versus RevMan MH. A total of 109 meta-analyses using MH and logistic regression did not converge due to various computational errors. In the remaining sample ($n = 776$), we observed a large

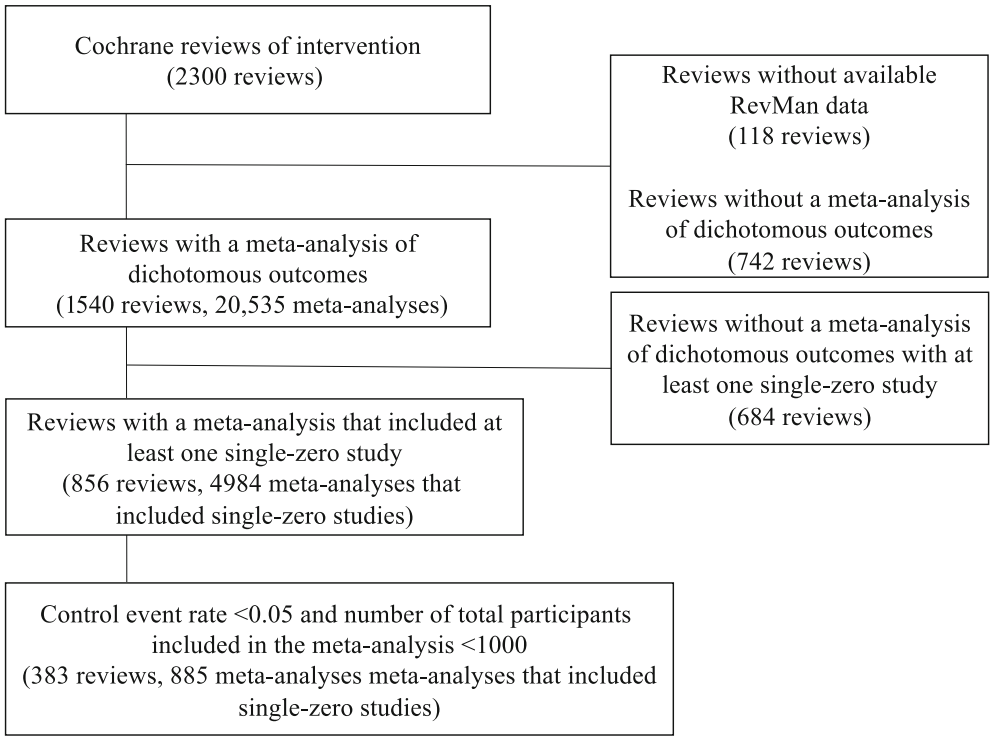


FIGURE 1 Flow diagram of the present study.

TABLE 1 Characteristics of included meta-analyses.

Characteristic	$n = 885$
Number of included studies	3 (2, 4)
Number of participants	430 (257, 690)
Number of events	10 (5, 20)
Statistical methods used in the Cochrane reviews	
Fixed IV method with a fixed correction	17 (2%)
RevMan MH model[a]	566 (64%)
REIV	302 (34%)

Note: Values in parentheses show percentages or interquartile range.
Abbreviations: MH, Mantel–Haenszel; REIV, random effects inverse variance method with a fixed correction.
[a]The Mantel–Haenszel model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table, implemented in the RevMan software.

difference between the results obtained from RevMan MH and MH, logistic regression, and Bayesian model in 27%, 32%, and 63% of cases, respectively. Meta-analyses with large RORs for methods other than the REIV or REIV TACC versus RevMan MH tended to have a small number of participants and events.

Table 3 shows details from meta-analyses that showed extremely large size of ROR. The extremely large sizes of ROR were observed when there were many events in either the intervention group or the control group, and few events in the opposite group.

4 | DISCUSSION

We found that many of existing Cochrane reviews included at least one meta-analysis including single-zero studies. Most were analyzed using RevMan MH or Random effects IV models, which involve using a fixed continuity correction. Our reanalysis showed RevMan MH gave many times substantially different results than MH without correction or the logistic regression model; both have been advocated as superior models for handling single-zero studies. MH without correction and logistic regression models showed agreement with each other, but the Peto method tended to be inconsistent with all other methods. Moreover, a substantial difference was evident in ∼30% of the existing meta-analyses when comparing the results between RevMan MH and either MH or logistic regression. Such large differences were mainly seen in meta-analyses with a smaller number of participants and events. The extremely large difference in the ORs from RevMan MH and other methods was observed when there were many events in either the intervention group or the control group, and few events in the opposite group.

We found that a substantial proportion of the ORs from RevMan MH were either 25% smaller or larger than the ORs from MH or logistic regression models. Such large difference of treatment effects may alter conclusions drawn from meta-analyses. Our results strengthen previously expressed concerns that the use of continuity

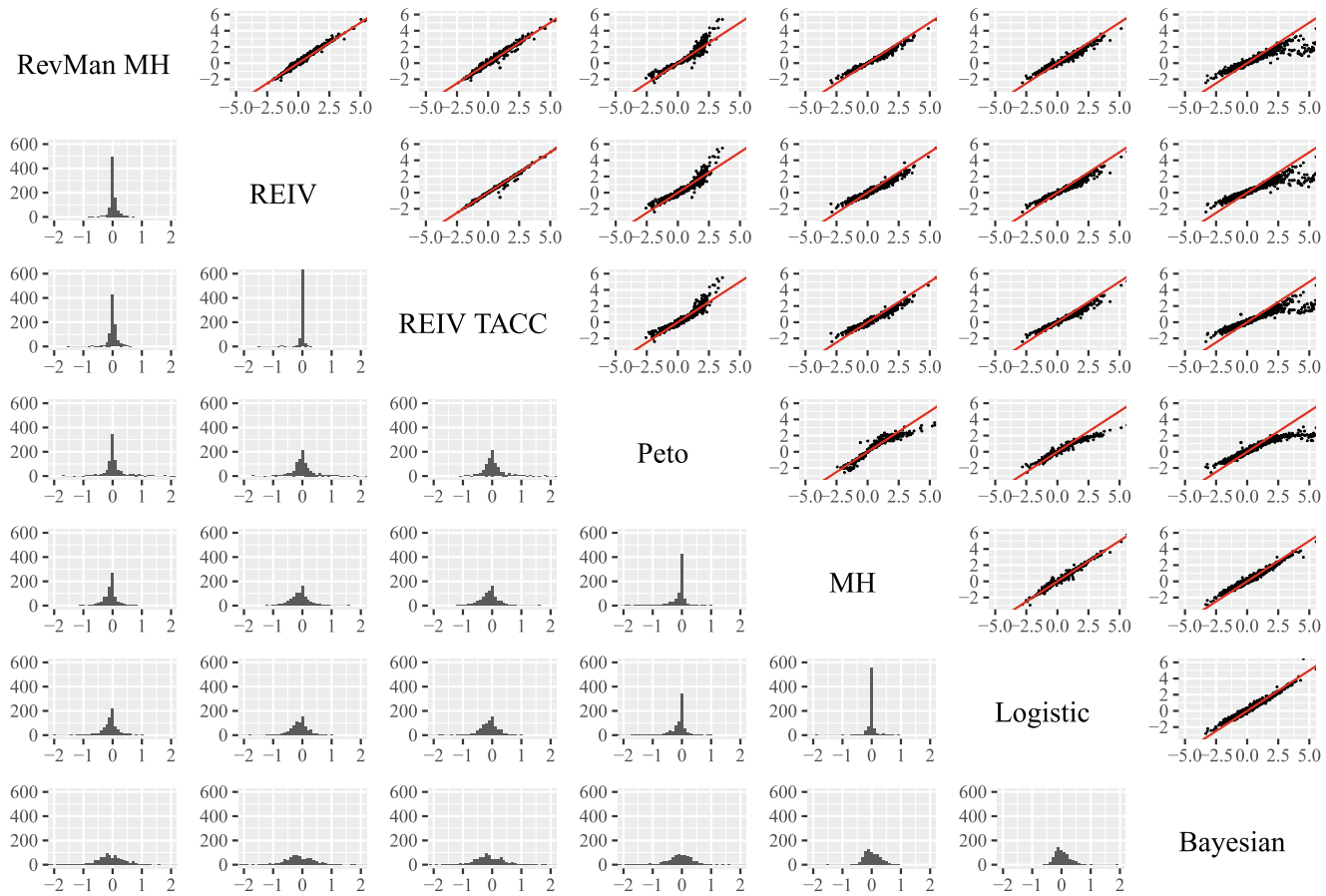


FIGURE 2 Agreement in effect estimates obtained by each statistical method to deal with single-zero studies. Panels at the lower-left present histograms of LogROR, for which the denominator of the ROR appears in the diagonal at the right and the numerator above. Upper right panels present scatter plots of Log odds ratios, with the vertical axis representing the statistical methods indicated in the diagonal at the bottom and the horizontal axis representing the methods indicated at the left. RevMan MH denotes the Mantel–Haenszel model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table, which is the model implemented in the RevMan software. REIV denotes random effects inverse variance model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table. REIV TACC denotes random effects inverse variance model with the treatment arm continuity correction. MH denotes the Mantel–Haenszel model without correction. Bayesian random effects model using a binomial likelihood for the outcome.

correction is common, even though it is not generally recommended.[5–7] This might be mainly due to the fact that RevMan, which is widely used in preparing systematic reviews and meta-analyses, has implemented the method as the default setting.[15] The Peto methods prone to break down when ORs are large. This observation is in agreement with prior studies which indicate that the approximation employed for calculating the log OR is reliable for modest effects of the intervention.[1,16]

Our results were consistent with a previous study that showed that the use of fixed continuity correction led to biased estimates, while MH without correction or logistic regression yielded the least biased estimates.[6] The agreement that we found between MH and logistic regression was also seen in a previous study.[4] Thus, our findings are not new; however, we were able to illustrate the issues related to using a fixed continuity correction utilizing real data from published meta-analyses.

Our descriptive analysis showed some situations where there is a high likelihood of large differences between comparing RevMan MH and other methods, namely (i) meta-analyses with small number of participants and events, or (ii) meta-analyses with many events in either the intervention group or the control group and few events in the opposite group (i.e., large effects). This information is crucial for informing researchers, readers, peer-reviewers, and journal editors of meta-analyses that when they encounter such situations, they need to interpret the results with caution, as the estimates might be considerably affected by zero-cell correction methods.

To our knowledge, this study is the largest to investigate the impact of different zero-cell correction methods

TABLE 2 Characteristics of the included meta-analyses categorized according to the ratio of odds ratios for each method versus RevMan Mantel–Haenszel model.[a]

	Sizes of RORs (vs. RevMan MH)[b]		
Characteristics	**Small**	**Moderate**	**Large**
Methods			
REIV[c]	658 (74%)	144 (16%)	83 (9%)
REIV TACC[d]	606 (68%)	179 (20%)	100 (11%)
Peto	492 (56%)	176 (20%)	217 (25%)
MH[e]	399 (51%)	167 (22%)	210 (27%)
Logistic regression	359 (46%)	167 (22%)	250 (32%)
Bayesian model[f]	136 (15%)	194 (22%)	555 (63%)
Median number of studies in meta-analysis			
REIV[c]	3 (2, 4)	3 (2, 5)	3 (2, 4)
REIV TACC[d]	3 (2, 4)	3 (2, 4)	2 (2, 4)
Peto	3 (2, 4)	3 (2, 4)	2 (2, 3)
MH[e]	3 (2, 4)	3 (2, 4)	2 (2, 3)
Logistic regression	3 (2, 4)	3 (2, 4)	3 (2, 4)
Bayesian model[f]	4 (3, 5)	3.5 (3, 4.75)	2 (2, 3)
Median number of participants			
REIV[c]	408 (242, 691)	507 (269, 675)	490 (262, 685)
REIV TACC[d]	430 (237, 702)	400 (270, 662)	443 (253, 683)
Peto	518 (298, 754)	359 (220, 627)	326 (191, 594)
MH[e]	552 (317, 777)	490 (273, 735)	330 (196, 551)
Logistic regression	528 (310, 761)	426 (242, 699)	398 (228, 654)
Bayesian model[f]	582 (319, 845)	558 (337, 746)	360 (206, 622)
Median number of events			
REIV[c]	9 (4, 18)	11 (7, 25)	21 (10, 36)
REIV TACC[d]	10 (5, 18)	10 (6, 21)	19 (8, 33)
Peto	13 (7, 22)	8 (4, 16)	5 (3, 19)
MH[e]	15 (8, 26)	11 (6, 24)	7 (4, 14)
Logistic regression	14 (7, 24)	10 (6, 23)	8 (5, 17)
Bayesian model[f]	15 (6, 25)	13 (8, 27)	8 (4, 17)

Note: Values in parentheses shows percentage or interquartile range. Percentage calculations were based on a denominator that excluded 109 missing values resulting from non-convergence for the MH and logistic regression models.

Abbreviations: MH, Mantel–Haenszel; REIV, random effects inverse variance; ROR, ratio of odds ratios; TACC, treatment arm continuity correction.

[a]The Mantel–Haenszel model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table, which is the model implemented in the RevMan software.

[b]Small, ROR in the range of 0.9–1.11; Moderate, ROR in the range of 0.8–0.9 or 1.11–1.25; Large, ROR in the range of ≥1.25 or ≤0.8.

[c]Random effects inverse variance model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table.

[d]Random effects inverse variance model with the treatment arm continuity correction.

[e]The Mantel–Haenszel model without correction.

[f]Bayesian random effects model using a binomial likelihood for the outcome.

on existing meta-analyses. We followed a pre-specified protocol and adhered to standard reporting guidelines.[8,17] Our study illuminates the real-world application of zero cell correction methods, aiding authors in understanding the importance of the choice between these methods. However, our study has several limitations. First, we included meta-analyses having a control event rate of <0.05 and a total participant count of fewer than 1000. We found a significant number of Cochrane reviews included such meta-analyses, enabling to test our

TABLE 3 narrative summary of the meta-analyses that had extremely large size of ROR.[a]

CDSR	Participants, intervention, control, and outcome	Events in the intervention arm for each study	Events in the control arm for each study	ORs from different models	
CD006633[10]	Participants: patients with schizophrenia	Study 1: 28/31	Study 1: 1/32	RevMan MH[b]	182.73
	Intervention: clozapine	Study 2: 23/36	Study 2: 0/36	REIV[c]	209.00
	Control: quetiapine			REIV TACC[d]	209.00
	Outcome: hypersalivation—short term			Peto	23.96
				MH[e]	530.83
				Logistic	372.02
				Bayesian model[f]	582.42
CD005067[11]	Participants: patients with Old World cutaneous leishmaniasis	Study 1: 49/60	Study 1: 0/60	RevMan MH[b]	221.89
	Intervention: oral dapsone	Study 2: 18/20	Study 2: 2/20	REIV[c]	156.47
	Control: placebo			REIV TACC[d]	156.47
	Outcome: Participants complete cure			Peto	26.89
				MH[e]	326.00
				Logistic	276.91
				Bayesian model[f]	457.41
CD005590[12]	Participants: non-HIV immunocompromised patients	Study 1: 0/30	Study 1: 1/30	RevMan MH[b]	0.12
	Intervention: TMP/SMX	Study 2: 0/80	Study 2: 18/80	REIV[c]	0.25
	Control: placebo, no treatment or non-PCP drug	Study 3: 1/22	Study 3: 0/20	REIV TACC[d]	0.25
	Outcome: documented PCP infections—hematological cancer subgroup	Study 4: 0/61	Study 4: 0/59	Peto	0.14
		Study 5: 0/74	Study 5: 0/63	MH[e]	0.05
		Study 6: 0/27	Study 6: 0/61	Logistic	0.11
		Study 7: 0/52	Study 7: 0/50	Bayesian model[f]	0.04
		Study 8: 0/74	Study 8: 0/64		
CD008370[13]	Participants: patients received pancreatic surgery	Study 1: 16/16	Study 1: 0/16	RevMan MH[b]	19.38
	Intervention: somatostatin analogues	Study 2: 2/38	Study 2: 1/37	REIV[c]	40.48
	Control: none	Study 3: 0/35	Study 3: 0/32	REIV TACC[d]	40.51
	Outcome: number with adverse effects due to treatment	Study 4: 0/107	Study 4: 0/104	Peto	21.81
				MH[e]	19.65
				Logistic	634.97
				Bayesian model[f]	92.25

TABLE 3 (Continued)

CDSR	Participants, intervention, control, and outcome	Events in the intervention arm for each study	Events in the control arm for each study	ORs from different models	
CD010328[14]	Participants: patients with symptomatic lumbar disc herniation	Study 1: 0/166	Study 1: 1/159	RevMan MH[b]	0.21
	Intervention: minimally invasive discectomy	Study 2: 0/55	Study 2: 3/57	REIV[c]	0.22
	Control: microdiscectomy/open discectomy	Study 3: 0/10	Study 3: 1/12	REIV TACC[d]	0.22
	Outcome: surgical site and other infections	Study 4: 0/100	Study 4: 4/100	Peto	0.21
		Study 5: 1/70	Study 5: 7/142	MH[e]	0.09
		Study 6: 0/30	Study 6: 0/30	Logistic	0.08
				Bayesian model[f]	0.04

Abbreviations: CDSR, Cochrane Database of Systematic Reviews; MH, Mantel–Haenszel; ORs, odds ratios; REIV, random effects inverse variance; ROR, ratio of odds ratios; TACC, treatment arm continuity correction.

[a]ROR in the range of ≥1.25 or ≤0.8.

[b]The Mantel–Haenszel model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table, which is the model implemented in the RevMan software.

[c]Random effects inverse variance model with a fixed continuity correction, adding 0.5 to all cells of the 2 × 2 table.

[d]Random effects inverse variance model with the treatment arm continuity correction.

[e]The Mantel–Haenszel model without correction.

[f]Bayesian random effects model using a binomial likelihood for the outcome.

hypothesis. However, this also means that the generalizability of our findings beyond such scenarios might be limited. Furthermore, we did not examine bias in estimates per se, only differences in effect estimate across models. Second, although we used a generic Bayesian random effects model for all meta-analyses, we acknowledge that this is not an optimal way of performing Bayesian statistics. Ideally, such analyses should have been done in separation, with a more careful selection of prior distributions (e.g., after using informative priors for heterogeneity, as in Turner et al.[18]), by carefully checking convergence, and so on. We used the generic Bayesian model here, however, as a means of exploring general differences between models. Third, because we focused on Cochrane reviews of interventions, the proportion of reviews that used RevMan MH might be smaller in non-Cochrane reviews. Fourthly, we did not directly compare the effect estimates from RevMan MH obtained by the RevMan software, but instead, we compared those obtained using R software. Nevertheless, we expect potential differences to be minor, as they are based on the same mathematical models. Finally, although a 25% difference in odds ratios was deemed large, we were unable to explore whether the interpretation of the effect estimates would change when using different methods. In Cochrane reviews, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach is used to assess the certainty of evidence. When applying GRADE to the estimates, the small number of events in our sample could result in a downgrade of imprecision. If the certainty of evidence was already very low in the original reviews, GRADE guides us to interpret the evidence as being very uncertain about the effect of the intervention on an outcome. Therefore, the difference in effect estimates might not significantly impact the interpretation of the results in such cases.

In conclusion, the influence of zero-cell correction methods on effect estimates could be significant in existing Cochrane reviews. We strongly propose that RevMan web, the software for Cochrane reviews, should incorporate more advanced statistical methods such as MH with no continuity correction and logistic regression. Even more advanced models such as the beta-binomial model, which have been shown to perform well, should be also ideally utilized.[19] Additional research is required to determine if the use of these improved methods will change interpretation of results.

AUTHOR CONTRIBUTIONS

Yasushi Tsujimoto: Conceptualization; methodology; software; data curation; investigation; validation; formal

analysis; funding acquisition; visualization; project administration; writing – original draft; writing – review and editing; resources. **Yusuke Tsutsumi:** Conceptualization; methodology; software; data curation; investigation; validation; formal analysis; writing – original draft; writing – review and editing. **Yuki Kataoka:** Conceptualization; methodology; software; data curation; investigation; writing – review and editing. **Akihiro Shiroshita:** Conceptualization; methodology; writing – review and editing; supervision. **Orestis Efthimiou:** Conceptualization; methodology; writing – review and editing; supervision; visualization. **Toshi A. Furukawa:** Conceptualization; project administration; writing – review and editing; supervision; visualization.

ACKNOWLEDGMENTS

We thank the members of the Research Group on Meta-Epidemiology at the Kyoto University School of Public Health (Drs Tomoko Fujii, Edoardo Ostinelli, Morihiro Katsura, Akira Onishi, Ethan Sahker, Yan Luo, Satoshi Funada, Kenji Omae, and Aran Tajika). They provided many valuable comments on this research.

FUNDING INFORMATION

The present study was supported by a grant from the JSPS Kakenhi Grant Number 22K10423 and 22K19688.

CONFLICT OF INTEREST STATEMENT

Yasushi Tsujimoto: YT is a board member of Cochrane Japan, and received grants from Pfizer Health Research Foundation; Akihiro Shiroshita: AS received financial support for his doctoral study from Vanderbilt University Medical Center, Center for Asthma Research and the Fulbright Association. Toshi A. Furukawa: TAF reports personal fees from DT Axis, Kyoto University Original, MSD and SONY, and a grant from Shionogi, outside the submitted work; In addition, Toshi A. Furukawa has patents 2020-548587 and 2022-082495 pending, and intellectual properties for Kokoro-app licensed to Mitsubishi-Tanabe. The rest of the author did not have any conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available at Cochrane Library (https://www.cochranelibrary.com/). Additional data that support the findings of this study are available upon reasonable request from the corresponding author.

ORCID

Yasushi Tsujimoto https://orcid.org/0000-0002-7214-5589
Yusuke Tsutsumi https://orcid.org/0000-0002-9160-0241
Yuki Kataoka https://orcid.org/0000-0001-7982-5213
Akihiro Shiroshita https://orcid.org/0000-0003-0262-459X
Orestis Efthimiou https://orcid.org/0000-0002-0955-7572
Toshi A. Furukawa https://orcid.org/0000-0003-2159-3776

REFERENCES

1. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, eds. *WV. Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022)*. Cochrane; 2022. www.training.cochrane.org/handbook
2. The Cochrane Collaboration. Review Manager Web (RevMan Web). 2023.
3. The Cochrane Collaboration. Review Manager 5 (RevMan 5). Copenhagen: The Cochrane Collaboration. 2020.
4. Efthimiou O. Practical guide to the meta-analysis of rare events. *Evid Based Ment Health*. 2018;21(2):72-76.
5. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med*. 2004;23(9):1351-1375.
6. Bradburn MJ, Deeks JJ, Berlin JA, Russell LA. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med*. 2007;26(1):53-77.
7. Weber F, Knapp G, Ickstadt K, Kundt G, Glass A. Zero-cell corrections in random-effects meta-analyses. *Res Synth Methods*. 2020;11(6):913-919.
8. Tsujimoto Y. The impact of zero-cell correction methods in Cochrane reviews with meta-analyses of single-zero studies: A meta-epidemiological study. 2022. https://osf.io/ez87c/
9. Efthimiou O, Rucker G, Schwarzer G, Higgins JPT, Egger M, Salanti G. Network meta-analysis of rare events using the Mantel-Haenszel method. *Stat Med*. 2019;38(16):2992-3012.
10. Asenjo Lobos C, Komossa K, Rummel-Kluge C, et al. Clozapine versus other atypical antipsychotics for schizophrenia. *Cochrane Database Syst Rev*. 2010;(11):CD006633.
11. Heras-Mosteiro J, Monge-Maillo B, Pinart M, et al. Interventions for old world cutaneous leishmaniasis. *Cochrane Database Syst Rev*. 2017;(12):CD005067.
12. Stern A, Green H, Paul M, Vidal L, Leibovici L. Prophylaxis for pneumocystis pneumonia (PCP) in non-HIV immunocompromised patients. *Cochrane Database Syst Rev*. 2014;(10):CD005590.
13. Gurusamy KS, Koti R, Fusai G, Davidson BR. Somatostatin analogues for pancreatic surgery. *Cochrane Database Syst Rev*. 2013;(4):CD008370.
14. Rasouli MR, Rahimi-Movaghar V, Shokraneh F, Moradi-Lakeh M, Chou R. Minimally invasive discectomy versus microdiscectomy/open discectomy for symptomatic lumbar disc herniation. *Cochrane Database Syst Rev*. 2014;(9):CD010328.
15. Buchter RB, Weise A, Pieper D. Reporting of methods to prepare, pilot and perform data extraction in systematic reviews: analysis of a sample of 152 Cochrane and non-Cochrane reviews. *BMC Med Res Methodol*. 2021;21(1):240.

16. Brockhaus AC, Bender R, Skipka G. The Peto odds ratio viewed as a new effect measure. *Stat Med*. 2014;33(28):4861-4874.
17. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med*. 2007; 147(8):W163-W194.
18. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Stat Med*. 2015;34(6):984-998.
19. Felsch M, Beckmann L, Bender R, Kuss O, Skipka G, Mathes T. Performance of several types of beta-binomial models in comparison to standard approaches for meta-analyses with very few studies. *BMC Med Res Methodol*. 2022; 22(1):319.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Tsujimoto Y, Tsutsumi Y, Kataoka Y, Shiroshita A, Efthimiou O, Furukawa TA. The impact of continuity correction methods in Cochrane reviews with single-zero trials with rare events: A meta-epidemiological study. *Res Syn Meth*. 2024;15(5):769-779. doi:10.1002/jrsm.1720

Archives of Osteoporosis (2024) 19:44
https://doi.org/10.1007/s11657-024-01393-4

REVIEW

Association between a history of major osteoporotic fractures and subsequent hip fracture: a systematic review and meta-analysis

Takashi Ariie[1,2] · Norio Yamamoto[2,3,4] · Yusuke Tsutsumi[2,5,6] · Shuri Nakao[2,7] · Akihiro Saitsu[2,8,9] · Takahiro Tsuge[2,4,10] · Haruka Tsuda[2,11] · Yuki Nakashima[2,12] · Takanori Miura[2,13] · Yousuke Bandai[2,14] · Ryota Okoba[15] · Shunsuke Taito[2,12]

Received: 24 January 2024 / Accepted: 21 April 2024 / Published online: 31 May 2024

Abstract

Purpose A history of fractures involving the distal radius, proximal humerus, spine, and hip may be associated with the incidence of subsequent hip fractures in older people. However, a comprehensive summary of this association using a rigorous methodology is lacking. Our objective was to systematically review the literature and examine the association between four major osteoporotic fractures and subsequent hip fractures in individuals aged ≥ 50 years.

Methods We searched MEDLINE, Embase, CENTRAL, ICTRP, and ClinicalTrials.gov on February 15, 2023. The search included cohort or case–control studies investigating the association between these four types of osteoporotic fractures and subsequent hip fractures. We pooled the hazard ratios (HRs) with 95% confidence intervals (CI) using the random-effects model. We used the Quality In Prognosis Studies tool to assess the risk of bias in the included studies, and the grading of recommendations assessment, development, and evaluation approach to determine the certainty of evidence.

Results The selection process identified 48 studies for qualitative synthesis and 23 studies (2,239,217 participants) for meta-analysis. The overall methodological quality had a low risk of bias in 65% of the included studies. The association between a history of major osteoporotic fractures and subsequent hip fracture varied, with a high certainty of evidence for a history of proximal humerus and hip fractures (HR 2.02, 95% CI 1.75–2.33 and 2.86, 95% CI 1.92–4.25, respectively), moderate certainty for distal radius fractures (HR 1.66, 95% CI 1.53–1.81), and low certainty for spine fractures (HR 1.53, 95% CI 1.38–1.69).

Conclusions In conclusion, a history of major osteoporotic fractures, particularly distal radius, proximal humerus, and hip fractures, is associated with subsequent hip fractures in older adults. Further research is needed to verify the association between a history of spine fracture and subsequent hip fractures.

Protocol registration: Open Science Framework (https://osf.io/7fjuc).

Keywords Osteoporosis · Subsequent fractures · Prognosis · Systematic review · Meta-analysis

Introduction

In older people, hip fracture is one of the major osteoporotic complications and can significantly increase medical costs [1] and reduce healthy life-years [2]. Observational studies have explored the link between prior major osteoporotic fractures affecting the distal radius [3], proximal humerus [4], spine [4], and hip [4]) and subsequent hip fracture. However, a comprehensive synthesis of the association between these fracture types and subsequent hip fractures using the same methodology is lacking [5–7]. Furthermore, methodological limitations, including sub-optimal search strategies [5–7] and the lack of assessments for bias risk [5, 6] and evidence certainty [5–7] in previous reviews, raise the risk of inaccurate association estimations.

Since the publication of previous reviews, several large longitudinal studies have emerged [8, 9], and updating the evidence with rigorous methodologies is important for guiding clinical decision-making and shedding light on the suboptimal adherence to secondary prevention measures [10, 11]. Moreover, based on previous systematic reviews, the association between proximal humerus fractures, which are a major osteoporotic fracture, and subsequent hip fractures is unclear.

Therefore, our objective was to conduct a systematic review of the association between a history of fractures of the distal radius, proximal humerus, spine, and hip and subsequent hip fractures in older people.

Extended author information available on the last page of the article

Methods

Study design

We conducted a systematic review and meta-analysis of prognostic factors. The review adhered to the Cochrane Handbook [12] and the meta-analysis of observational studies in epidemiology (MOOSE) guidelines [13] (Supplementary Table 1), and its protocol was registered on Open Science Framework (https://osf.io/7fjuc).

Search strategy

We conducted searches on MEDLINE (via PubMed), EMBASE (via Dialog), and CENTRAL, without limiting by publication language or year. We also searched ongoing or unpublished studies on ICTRP and ClinicalTrials.gov. The searches were conducted on February 15, 2023 (Supplementary Table 2–6). The articles cited in the included articles, as well as those citing the included studies, were also reviewed via citationchaser (https://estech.shinyapps.io/citationchaser/).

Eligibility criteria

We used the following eligibility criteria:

1. Study design: We included observational studies and secondary analyses of randomized controlled trials written in any language. We accepted both published and unpublished studies. We excluded ecological studies, case reports, and case series.
2. Participants/population: We included studies involving those aged ≥ 50 years. Studies involving participants aged < 50 years were included if the mean age was ≥ 50 years.
3. Exposure: The exposure group contained individuals with a history of major osteoporotic fractures, including those of the distal radius, proximal humerus, spine, and hip. Fractures of the radius, wrist, or forearm were considered distal radius fractures. Fractures of the humerus, shoulder, or upper arm were considered proximal humerus fractures. Vertebral deformities were considered spine fractures. Exposure was confirmed (fracture diagnosis) using imaging (e.g., X-ray), International Classification of Diseases (ICD), or medical records. We also accepted the exposure definitions used in the original studies.
4. Comparison: The comparison group contained individuals without a history of major osteoporotic fractures as defined in the exposure. If the exposure group included individuals with distal radius fractures, the comparison group contained those without distal radius fractures.
5. Outcome: The primary outcome was hip fracture incidence. Outcome confirmation was the same as exposure, with the outcome occurring after the exposure. If multiple outcome measurement time points were reported, we selected the one closest to a 2-year follow-up.

Study selection and data extraction

Five pairs of reviewers (T.A. and Y.B., T.A. and R.O., S.N. and Y.N., T.M. and A.S., and T.T. and H.T.) independently screened the titles and abstracts of the retrieved studies based on the eligibility criteria. Next, eight pairs of reviewers (T.A. and S.N., A.S., T.T., H.T., Y.N., T.M., Y.B., or R.O.) independently assessed the full texts of the included studies to determine their eligibility. If we could not assess the studies because of insufficient information, we contacted the authors to obtain the information.

The same eight pairs of full-text reviewers also independently extracted data using a predefined data extraction form. We collected information on study characteristics (e.g., publication year, study design, population, follow-up period, exposure and comparison definitions, and outcome definition) and results (e.g., crude or adjusted effect estimates). Where studies reported subgroup results (e.g., age or gender) rather than the overall results, we requested the authors to provide the overall results. Disagreements between reviewers were resolved through discussion, and if they persisted, a third reviewer (No.Y., S.T., or Y.T.) arbitrated.

Quality assessment

The same eight pairs of reviewers (T.A. and S.N., A.S., T.T., H.T., Y.N., T.M., Y.B., or R.O.) independently assessed the risk of bias in the included studies using the Quality In Prognosis Studies (QUIPS) tool [14]. If necessary, we contacted the authors of the included studies. Disagreements between reviewers were resolved through discussion, and if they persisted, a third reviewer (No.Y., S.T., or Y.T.) arbitrated. For QUIPS assessment, we defined the following variables as confounders before commencing the review: age [15], gender [15], bone mineral density [15], body mass index [16], a history of fracture [15], and the use of oral glucocorticoids [17].

Data synthesis

We performed meta-analyses using a random effects model. As primary analysis, we pooled adjusted hazard ratios (HRs). If adjusted HRs were unavailable, we used crude HRs. If the original studies reported several models using

different variables to calculate adjusted HRs, we selected the model adjusted for our predefined confounders (age, gender, bone mineral density, body mass index, a history of fracture, and the use of oral glucocorticoids) the most. If other effect estimates, e.g., odds ratios (ORs), rate ratios (RRs), or risk ratios, were available, they were separately pooled. In case subgroup results were available without overall results, their data were pooled separately. If different studies reported the same exposure using the same cohort data for the same enrollment period, we selected one for primary analysis based on the following hierarchy:

1. Studies reporting HRs
2. Studies published in full text
3. Studies that did not use the comprehensive definition of exposure, including fractures of the radius, wrist, forearm, humerus, shoulder, upper arm, or vertebral deformities
4. Studies reporting adjusted effect estimates
5. Studies with large sample sizes
6. Newly published studies

We assessed between-study heterogeneity using forest plots and I^2 statistics [12]. When heterogeneity was detected ($I^2 > 50\%$), we explored the potential reasons using predefined subgroup analyses. We used the Cochran χ^2 test (Q-test) and a P value of 0.10 indicated statistical significance.

The following subgroup analyses were planned for the primary outcome:

1. Study setting: hospital or nursing home versus community.
2. Age: < 80 years old versus ≥ 80 years old.
3. Gender: men versus women.
4. History of osteoporosis: history of osteoporosis versus no history of osteoporosis.
5. History of fracture: history of fracture before exposure versus no history of fracture.
6. Use of oral glucocorticoids: oral glucocorticoids use versus no use.

We planned sensitivity analysis for the primary outcomes, excluding studies that used crude HRs, those with a high risk of bias in the overall quality, or those that used a comprehensive definition of exposures (e.g., wrist fractures instead of distal radius fractures).

We assessed publication bias using funnel plots and Egger's test ($P < 0.05$), if > 10 studies were identified.

These analyses were performed using RevMan 5.4 (Cochrane Collaboration, London, UK) and STATA 16.0 (Stata-Corp LP, College Station, TX, US). $P < 0.05$ indicated statistical significance.

We created a summary of findings table for the outcomes using HR and used the grading of recommendations assessment, development, and evaluation (GRADE) approach [18, 19] to assess the certainty of evidence. This approach considered eight domains that can influence the certainty of the effect estimates, including factors that may reduce certainty (risk of bias, inconsistency, indirectness, imprecision, and publication bias) and those that can enhance certainty (large effect, dose–response, and plausible confounding). The level of certainty was determined based on the study design and the concerns associated with these areas and was classified as high, moderate, low, or very low. A single reviewer (T.A.) assessed the certainty and another reviewer (No.Y.) verified it. Any discrepancies were resolved through discussion and if they persisted, a third reviewer (S.T or Y.T) was involved to reach a consensus. To increase transparency, we provided reasons for downgrading in a summary of findings. We used the median event rate in the comparison group across the included studies to calculate the anticipated absolute effects.

Difference between the protocol and manuscript

During the protocol development, we had not planned the sensitivity analysis excluding studies in which the comparison group included fracture history other than exposure fracture, studies that included traumatic or pathological fractures, and studies that identified spine fracture cases based on self-reporting and ICD. However, we conducted post hoc sensitivity analyses only using these studies to examine the robustness of our findings.

Results

Study selection and characteristics

The selection process is illustrated in Fig. 1 and the list of excluded studies is shown in Supplementary Table 7. We screened 4639 records from databases and registers and 8280 records from citation searching. Finally, 48 studies were included for qualitative syntheses, and 23 studies, involving 2,239,217 participants, were included for quantitative syntheses as primary analysis. Table 1 shows a summary of study characteristics, which are detailed in Supplementary Table 8. Most of the included studies were conducted in community settings, and only four were conducted in nursing homes or hospitals. In the included studies, the median of the participants' mean age was 71.7 years old (range 58.5–86.2 years). About half of the included studies involved men and women, and < 10%

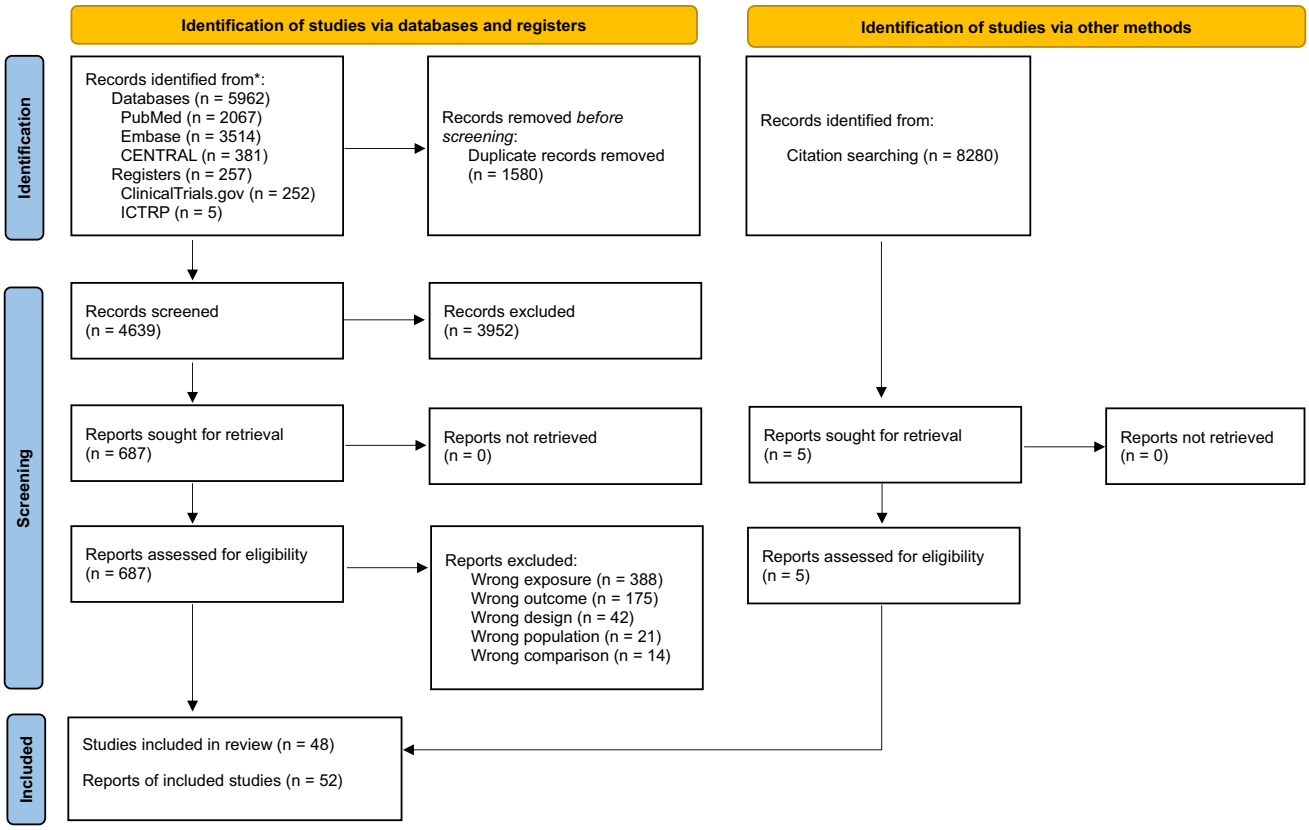


Fig. 1 Preferred reporting items for systematic reviews and meta-analysis (PRISMA) 2020 flow diagram showing the study selection

involved men only. Around 80% of the included studies reported adjusted effect estimates, and among them, 28% adjusted for a history of fractures. Several studies used comprehensive definitions of exposure, including fractures of the wrist or forearm as distal radius fractures, fractures of the humerus, upper arms, and shoulder as proximal humerus fractures, and vertebral deformity as spine fractures. A total of 19 out of 33 studies defined spine fractures using imaging without clinical symptoms (morphometric fractures). One study included in the quantitative synthesis had a comparison group that included participants with fractures. The results of the included studies are presented in Supplementary Table 9.

Quality assessment

The methodological quality of the studies included in the meta-analyses was assessed using the QUIPS tool (Supplementary Table 10). The proportion of a low risk of bias in the overall quality was 65%. Most included studies (83%) had a moderate-to-high risk of bias in the study confounding domain because of insufficient adjustment of prespecified confounders.

Primary analyses

Table 2 presents a summary of findings. The median event rate for subsequent hip fractures ranged from 0.4 to 10% in the comparison group for four different exposures among the included studies. Figure 2 shows the forest plots of the association between a history of fracture in the four sites and subsequent hip fractures.

History of distal radius fracture and subsequent hip fractures

Eleven studies [3, 4, 8, 20–27] ($n = 2{,}123{,}288$) that reported crude or adjusted HRs were included in the meta-analysis. The median of the mean follow-up period was 6.5 years (range 1.0–15.3 years). The median prevalence of prior distal radius fractures was 5.7% (range 3.4–10.9%). A history of distal radius fractures was probably associated with an increased risk of subsequent hip fractures (pooled HR 1.66, 95% CI 1.53–1.81, $I^2 = 74\%$, moderate certainty of evidence). Pooling studies that reported ORs and RRs did not change the results (Supplementary Fig. 1).

Table 1 Characteristics of the included studies

Variables	Overall $n=48$
Country	
USA	23 (48)
Australia	4 (8)
Canada	4 (8)
Others	17 (35)
Study design	
Prospective cohort	39 (81)
Retrospective cohort	6 (13)
Case–control	4 (8)
Setting	
Population-based	31 (65)
Multicenter	11 (23)
Single center	7 (15)
Number of participants, median (IQR)	8485 (2808–23,524)
Site of fracture history (exposure)	
Spine	33
Distal radius	19
Hip	12
Proximal humerus	8
Exposure confirmation	
Imaging	22 (46)
Self-report	14 (29)
International classification of disease	10 (21)
Medical record	3 (6)
Outcome confirmation (subsequent hip fracture)	
Imaging	22 (46)
International classification of disease	17 (35)
Self-report	5 (10)
Medical record	5 (10)
Gender of participants	
Both men and women	25 (51)
Women only	21 (43)
Men only	3 (6)

Data are presented as the number (percentage) of studies unless otherwise indicated

IQR interquartile range

History of proximal humerus fracture and subsequent hip fractures

Four studies [4, 21, 25, 28] ($n=1,647,988$) that reported adjusted HRs were included in the meta-analysis. The median of the follow-up period was 7.0 years (range 2.0–12.7 years). The median prevalence of prior proximal humerus fractures was 2.8% (range 1.9–4.0%). A history of proximal humerus fracture was associated with an increased risk of subsequent hip fractures (pooled HR 2.02, 95% CI 1.75–2.33, $I^2=85\%$, high certainty of evidence). Pooling studies that reported ORs and RRs did not change the results (Supplementary Fig. 2).

History of spine fracture and subsequent hip fractures

Fourteen studies [4, 21, 22, 24, 25, 29–37] ($n=1,752,959$) that reported the crude or adjusted HRs were included in the meta-analysis. One study was conducted in a hospital setting and the rest in community settings. The cases were confirmed by imaging ($n=5$), self-reporting ($n=5$), and ICD ($n=4$). The median of the mean follow-up period was 4.1 years (range 1.0–14.7 years). The median prevalence of prior spine fractures was 9.8% (range 0.4–24.2%). A history of spine fracture may be associated with an increased risk of subsequent hip fractures (pooled HR 1.53, 95% CI 1.38–1.69, $I^2=35\%$, low certainty of evidence). Pooling studies that reported ORs and RRs did not change the results (Supplementary Fig. 3).

History of hip fracture and subsequent hip fractures

Six studies [4, 21, 22, 25, 38, 39] ($n=1,656,303$) that reported adjusted HRs were included in the meta-analysis. Two studies were conducted in a nursing home setting and the rest in community settings. The median of the mean follow-up period was 3.4 years (range 1.0–12.7 years). The median prevalence of prior hip fractures was 1.7% (range 1.5–14.0%). A history of hip fracture was associated with an increased risk of subsequent hip fractures (pooled HR 2.86, 95% CI 1.92–4.25, $I^2=98\%$, high certainty of evidence). Pooling studies that reported ORs and RRs did not change the results (Supplementary Fig. 4).

Subgroup analysis

There was substantial between-study heterogeneity in the analysis of distal radius fractures ($I^2=74\%$), proximal humerus fractures ($I^2=85\%$), and considerable heterogeneity in hip fracture analysis ($I^2=98\%$). For all exposures, subgroup analyses of the histories of osteoporosis, fractures, and oral glucocorticoid use were not conducted because of insufficient data.

Subgroup analysis of the history of distal radius and proximal humerus fractures based on gender revealed significant heterogeneity between men and women ($I^2=84.4\%$ and 86.1%, respectively, Supplementary Figs. 5, 6). Subgroup analyses based on study setting and age were not conducted because of limited data.

Table 2 Summary of the association between a history of four major osteoporotic fractures and subsequent hip fracture

Fracture site	Anticipated absolute effects (95% CI)[a]				
	Assumed risk with control[b]	Corresponding risk with a history of fracture	Relative effect (95% CI)	No. of participants (studies)	Certainty of evidence (GRADE)
Distal radius	4 per 1000	6.6 per 1000 (6.1–7.2)	HR 1.66 (1.53–1.81)	2,123,288 (11 studies)	Moderate[c]
Proximal humerus	100 per 1000	192 per 1000 (168–218)	HR 2.02 (1.75–2.33)	1,647,988 (4 studies)	High
Spine	40 per 1000	61 per 1000 (55–67)	HR 1.53 (1.38–1.69)	1,752,959 (14 studies)	Low[cd]
Hip	12 per 1000	34 per 1000 (23–50)	HR 2.86 (1.92–4.25)	1,656,303 (6 studies)	High

GRADE Working Group grades of evidence: High certainty: We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty: We are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty: Our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect. Very low certainty: We have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect

CI confidence interval, *HR* hazard ratio, *GRADE* the grading of recommendations assessment, development, and evaluation

[a]The corresponding risk (95% CI) was based on the assumed risk in the nonexposed group and the relative effect of the exposure (95% CI)

[b]Median event rate of the included studies

[c]Downgraded one point because of a moderate or high risk of bias for the domain of study participation, study attrition, prognostic factor measurement, outcome measurement or statistical analysis, and reporting

[d]Downgraded one point because of publication bias

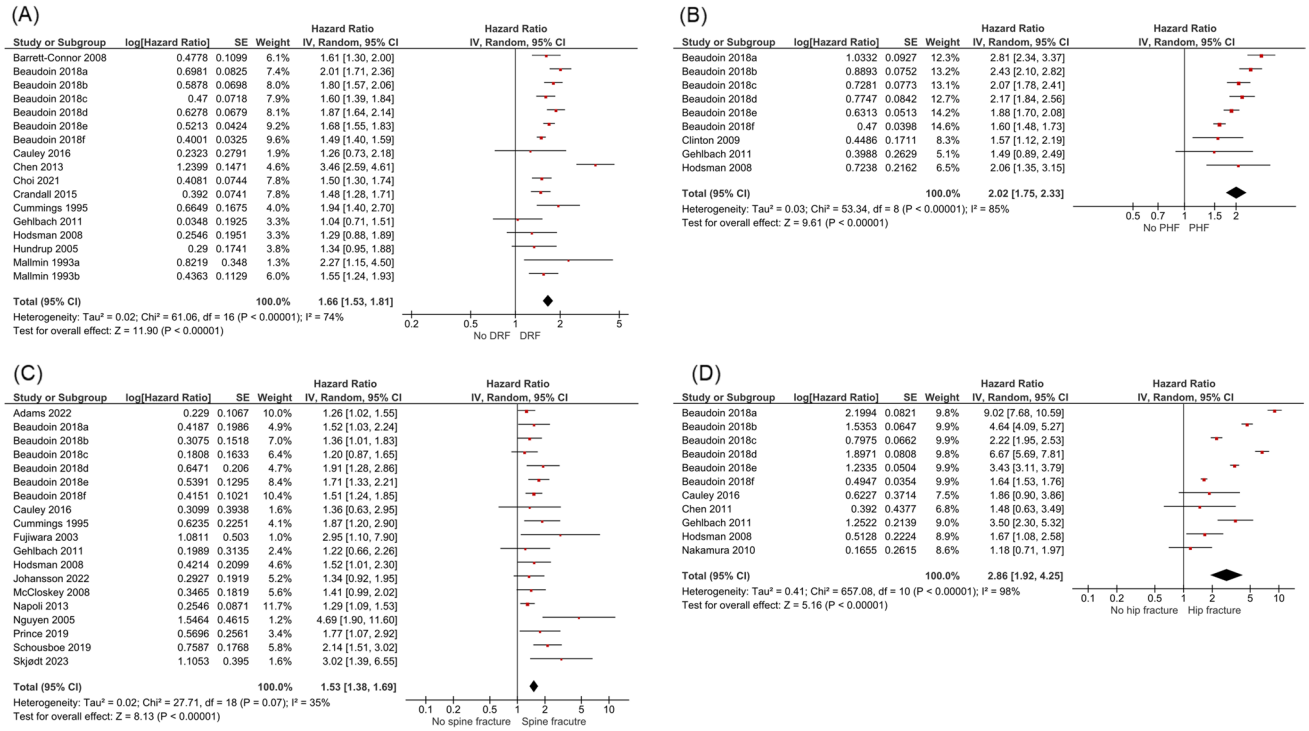


Fig. 2 The hazard ratio for subsequent hip fracture incidence after four major osteoporotic fractures. Abbreviations: SE, standard error; IV, inverse variance; CI, confidence interval; DRF, distal radius fracture; PHF, proximal humerus fracture. **A** Distal radius fracture. **B** Proximal humerus fracture. **C** Spine fracture. **D** Hip fracture. Beaudoin 2018 reported subgroup results only, based on gender and age (a, men aged 66 years; b, men aged 75 years; c, men aged 85 years; d, women aged 66 years; e, women aged 75 years; f, women aged 85 years). Mallmin 1995 reported subgroup results only, based on gender (a, men; b, women)

Subgroup analysis of the history of hip fracture based on study settings revealed significant heterogeneity between hospital or nursing home settings and community settings ($I^2 = 89.4\%$, Supplementary Fig. 7). Subgroup analysis based on age also showed significant heterogeneity between individuals aged < 80 and those aged ≥ 80 ($I^2 = 93.0\%$, Supplementary Fig. 8). The participants of studies conducted in nursing home settings were older than those of studies conducted in community settings. The effect estimates were adjusted for age in all studies, except in one community-setting study [21]. Even without this study, there was still substantial heterogeneity (66%) between settings. Subgroup analysis based on gender did not show significant heterogeneity between men and women ($I^2 = 0\%$, Supplementary Fig. 9).

Sensitivity analyses

For each association, at least one of the predefined sensitivity analyses, excluding studies that reported crude HRs, those with a high overall risk of bias, or those with a comprehensive definition of exposure, was performed (Supplementary Figs. 10–17). Post hoc analyses were also performed, excluding studies in which the comparison group included other fractures except index fractures, studies that included traumatic or pathological fractures, and studies that identified spine fracture cases based on self-reporting and ICD (Supplementary Figs. 18–20). All analyses revealed a consistent association between the history of fracture and subsequent hip fractures.

Publication bias

Publication bias assessment was performed for fractures of the distal radius, spine, and hip (Supplementary Figs. 21–23). Egger's test analysis of funnel plot asymmetry was statistically significant for spine fractures ($P = 0.007$) but not for distal radius and hip fractures ($P = 0.513$ and 0.517, respectively).

Discussion

Our analyses revealed that a history of fractures of the proximal humerus, hip, and probably distal radius is associated with subsequent hip fractures in individuals aged ≥ 50 years. A history of spine fractures may be associated with subsequent hip fractures since the certainty of evidence was low. The outcome focused on subsequent hip fractures, rather than any fracture because hip fracture is the most serious major osteoporotic fracture [40, 41]. Notably, a history of fractures of the proximal humerus and hip increases the risk of future hip fractures by twofold, based on a high certainty of evidence. Sensitivity analysis further supported the robustness of these results.

A history of all four different body parts of major osteoporotic fractures was associated with an increased risk of subsequent hip fractures. Although several systematic reviews have reported on the history of some major osteoporotic fractures as exposures, our systematic and comprehensive analysis quantified and confirmed the association between all four types of major osteoporotic fractures and subsequent hip fractures using the same methodology. Previous systematic reviews reported an association between previous fractures of the wrist [5–7], spine [6], and hip [6] and subsequent hip fractures (relative risk 1.82–3.26, 2.20–3.54, and 2.3, respectively) without assessing evidence certainty. When compared with previous reviews, our analysis, based on a larger sample size, yielded relatively smaller pooled effect estimates for distal radius and spine fractures (HR 1.66 and 1.53, respectively) and higher estimates for hip fractures (HR 2.86). We downgraded the certainty of evidence for fractures of the distal radius and spine because of concern about the risk of bias and publication bias, which were not sufficiently evaluated in previous reviews [5–7].

Clinicians and patients should be aware that the fracture site, particularly the upper extremities, is a risk factor for subsequent hip fractures. While there is evidence that secondary fracture prevention is beneficial [42], a treatment gap exists in high-risk populations worldwide [43]. A study in a healthcare center reported that only 3% of patients with proximal humerus fractures received antiresorptive therapy [44]. Moreover, surgeons performed subsequent fracture prevention on about 10% of patients who underwent surgery for distal radius fractures [11]. A fracture liaison service is one form of post-fracture care, but about 30% of patients do not attend the service [45]. These findings may be attributable to insufficient awareness among medical staff and patients about the risk of subsequent fractures. Additionally, unlike our review, previous reviews [5–7] did not report the association between proximal humerus fracture and subsequent hip fractures. In a clinical practice guideline, previous hip or spine fractures, but not distal radius or proximal humerus fractures, are specified in the pathways of management algorithm for pharmacotherapy recommendation to gain the most benefit from osteoporosis and fracture prevention, and a previous fracture that does not involve the hip or spine does not warrant a pharmacotherapy recommendation for postmenopausal women or men aged ≥ 50 years [46]. Another focus on spine fracture management, such as vertebral fracture assessment, may increase the use of pharmacologic medication for subsequent fracture prevention [47], and this might have increased opportunities for secondary prevention in people with spine fractures, resulting in the

relatively lower risk observed in our review. Our findings may provide reasonable evidence supporting the update of clinical practice guidelines and the inclusion of fractures of the distal radius and proximal humerus as independent risk factors.

Based on subgroup analyses, several effect modifiers may exist in the association between a history of fractures of the distal radius, proximal humerus or hip, and subsequent hip fractures. Our findings on gender and age subgroup analyses were consistent with previous reviews [7, 48]. Men with a history of distal radius fractures have a higher risk of subsequent hip fractures than women [7]. A previous fracture at a younger age is also associated with a higher risk of subsequent fractures than a fracture at an older age [48] but a competing event, such as higher mortality rate, may also be a reason. Subgroup analysis based on study settings revealed that community settings were associated with a higher risk than hospitals or nursing homes, which is different from previous studies showing a higher incidence of hip fractures in nursing home settings [49]. This is probably because people with a history of fractures in hospitals or nursing homes may receive more attention, such as through the implementation of fall prevention measures [50]. Future studies should assess why the risk of subsequent hip fractures differs between settings and develop strategies for reducing subsequent fractures in community settings.

This review has several strengths. First, it presents updated comprehensive evidence on the association between a history of major osteoporotic fractures and subsequent hip fractures. Specifically, recent clinical practice guidelines do not discuss the association between a history of fractures of the distal radius and proximal humerus and an increased risk of subsequent hip fractures [46, 51]. Hence, our findings may fill the gap and guide the identification of high-risk patients, who may proceed to secondary prevention management. Second, our review adhered to transparent reporting and rigorous methodology by following MOOSE guidelines, the Cochrane Handbook, and GRADE recommendations [12, 13, 18]. Reviewers independently assessed study eligibility, extracted data, and evaluated risk of the bias using the QUIPS tool.

However, our review has some limitations. First, we used research design terms (e.g., prognosis, risk factor, or hazard) in searching for practical reasons, which may have reduced the sensitivity to identify articles. Nonetheless, we identified around 5000 records from multiple databases, including trial registries, and conducted citation searches for more than 8000 records. Second, we could not perform some planned analyses, such as subgroup, sensitivity, and publication bias analyses because of limited data. In particular, the information on previous fractures before exposure is important for the estimation of a more precise association. Other participant characteristics included in studies may impact effect estimates. Non-fragility fractures, such as those caused by high-energy trauma, may be included. We did not exclude studies that might include traumatic fractures because it may underestimate the number of fragility fracture cases [52]. Moreover, the comparison group may include cases with other fracture sites or multiple previous fractures. This information is hard to measure because of recall bias or asymptomatic fractures. Specifically, studies using self-reporting or ICD to define spine fracture cases might have missed morphometric fracture cases without clinical symptoms. Although we tried to examine the impact of these characteristics in post hoc analyses, there may still be some effects. Future studies should consider the measurement method or analysis to adjust for these impacts and reveal additional effect modifiers, result robustness, and the risk of publication bias. Third, although the GRADE approach is recommended for the assessment of evidence certainty [12], measurement bias by subjectivity is inevitable. Therefore, we used multiple reviewers with experience in carrying out assessments using GRADE to increase assessment validity, and for transparency, we reported the reasons for downgrading.

In conclusion, a history of major osteoporotic fractures, particularly fractures of distal radius, proximal humerus, and hip, is associated with subsequent hip fractures in older adults. Strategies to raise awareness about these risk factors among clinicians and patients should be implemented. Well-designed studies are needed to determine if a history of spine fracture is a risk factor for subsequent hip fractures.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11657-024-01393-4.

Acknowledgements We would like to thank Dr. Tony M. Keaveny, Dr. Annette L. Adams, Prof. Eric S. Orwoll for kindly providing unpublished information and WORDVICE for their English language editing.

Author contribution Takashi Ariie: conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; visualization; writing—original draft; writing—review and editing.

Norio Yamamoto: conceptualization; methodology; project administration; supervision; writing—review and editing.

Yusuke Tsutsumi and Shunsuke Taito: conceptualization; methodology; supervision; writing—review and editing.

Shuri Nakao, Akihiro Saitsu, Takahiro Tsuge, Haruka Tsuda, Yuki Nakashima, Takanori Miura, Yousuke Bandai, and Ryota Okoba: methodology; validation; writing—review and editing.

Funding This systematic review was funded by a 2022 IUHW Research Grant. The funder was not involved in any aspect of the project, including the design of the review protocol and analysis plan.

Data availability The data used in this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest None.

References

1. Tajeu GS, Delzell E, Smith W et al (2014) Death, debility, and destitution following hip fracture. J Gerontol A Biol Sci Med Sci 69(3):346–353. https://doi.org/10.1093/gerona/glt105
2. Papadimitriou N, Tsilidis KK, Orfanos P et al (2017) Burden of hip fracture using disability-adjusted life-years: a pooled analysis of prospective cohorts in the CHANCES consortium. Lancet Public Health 2(5):e239–e246. https://doi.org/10.1016/S2468-2667(17)30046-4
3. Chen CW, Huang TL, Su LT et al (2013) Incidence of subsequent hip fractures is significantly increased within the first month after distal radius fracture in patients older than 60 years. J Trauma Acute Care Surg 74(1):317–321. https://doi.org/10.1097/ta.0b013e31824bb325
4. Hodsman AB, Leslie WD, Tsang JF, Gamble GD (2008) 10-year probability of recurrent fractures following wrist and other osteoporotic fractures in a large clinical cohort: an analysis from the Manitoba Bone Density Program. Arch Intern Med 168(20):2261–7. https://doi.org/10.1001/archinte.168.20.2261
5. Haentjens P, Autier P, Collins J, Velkeniers B, Vanderschueren D, Boonen S (2003) Colles fracture, spine fracture, and subsequent risk of hip fracture in men and women: A meta-analysis. J Bone Joint Surg Am 85(10):1936–43. https://doi.org/10.2106/00004623-200310000-00011
6. Klotzbuecher CM, Ross PD, Landsman PB, Abbott TA, Berger M (2000) Patients with prior fractures have an increased risk of future fractures: a summary of the literature and statistical synthesis. J Bone Miner Res 15(4):721–739. https://doi.org/10.1359/jbmr.2000.15.4.721
7. Johnson NA, Stirling ER, Divall P, Thompson JR, Ullah AS, Dias JJ (2017) Risk of hip fracture following a wrist fracture-a meta-analysis. Injury 48(2):399–405. https://doi.org/10.1016/j.injury.2016.11.002
8. Choi HG, Kim DS, Lee B, Youk H, Lee JW (2021) High risk of hip and spinal fractures after distal radius fracture: a longitudinal follow-up study using a national sample cohort. Int J Environ Res Public Health 18(14). https://doi.org/10.3390/ijerph18147391
9. Morin SN, Yan L, Lix LM, Leslie WD (2021) Long-term risk of subsequent major osteoporotic fracture and hip fracture in men and women: a population-based observational study with a 25-year follow-up. Osteoporos Int 32(12):2525–2532. https://doi.org/10.1007/s00198-021-06028-9
10. Kim TI, Choi JH, Kim SH, Oh JH (2016) The adequacy of diagnosis and treatment for osteoporosis in patients with proximal humeral fractures. Clin Orthop Surg 8(3):274–279. https://doi.org/10.4055/cios.2016.8.3.274
11. Baba T, Hagino H, Nonomiya H et al (2015) Inadequate management for secondary fracture prevention in patients with distal radius fracture by trauma surgeons. Osteoporos Int 26(7):1959–1963. https://doi.org/10.1007/s00198-015-3103-4
12. Higgins JPT, Thomas J, Chandler J et al (2023) Cochrane handbook for systematic reviews of interventions version 6.4 (updated August 2023). Cochrane, Available from Available from www.training.cochrane.org/handbook [Accessed October 20 2023]
13 Stroup DF, Berlin JA, Morton SC et al (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 283(15):2008–12. https://doi.org/10.1001/jama.283.15.2008
14. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C (2013) Assessing bias in studies of prognostic factors. Ann Intern Med 158(4):280–286. https://doi.org/10.7326/0003-4819-158-4-201302190-00009
15. Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E (2008) FRAX and the assessment of fracture probability in men and women from the UK. Osteoporos Int 19(4):385–397. https://doi.org/10.1007/s00198-007-0543-5
16. De Laet C, Kanis JA, Oden A et al (2005) Body mass index as a predictor of fracture risk: a meta-analysis. Osteoporos Int 16(11):1330–1338. https://doi.org/10.1007/s00198-005-1863-y
17. Kanis JA, Johansson H, Oden A et al (2004) A meta-analysis of prior corticosteroid use and fracture risk. J Bone Miner Res 19(6):893–9. https://doi.org/10.1359/JBMR.040134
18. Foroutan F, Guyatt G, Zuk V et al (2020) GRADE Guidelines 28: use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. J Clin Epidemiol 121:62–70. https://doi.org/10.1016/j.jclinepi.2019.12.023
19. Santesso N, Glenton C, Dahm P et al (2020) GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. J Clin Epidemiol 119:126–135. https://doi.org/10.1016/j.jclinepi.2019.10.014
20. Barrett-Connor E, Sajjan SG, Siris ES, Miller PD, Chen YT, Markson LE (2008) Wrist fracture as a predictor of future fractures in younger versus older postmenopausal women: results from the National Osteoporosis Risk Assessment (NORA). Osteoporos Int 19(5):607–613. https://doi.org/10.1007/s00198-007-0508-8
21. Beaudoin C, Jean S, Moore L et al (2018) Number, location, and time since prior fracture as predictors of future fracture in the elderly from the general population. J Bone Miner Res 33(11):1956–1966. https://doi.org/10.1002/jbmr.3526
22. Cauley JA, Cawthon PM, Peters KE et al (2016) Risk factors for hip fracture in older men: the Osteoporotic Fractures in Men Study (MrOS). J Bone Miner Res 31(10):1810–1819. https://doi.org/10.1002/jbmr.2836
23. Crandall CJ, Hovey KM, Cauley JA et al (2015) Wrist fracture and risk of subsequent fracture: findings from the Women's Health Initiative Study. J Bone Miner Res 30(11):2086–2095. https://doi.org/10.1002/jbmr.2559
24 Cummings SR, Nevitt MC, Browner WS et al (1995) Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group. N Engl J Med. 332(12):767–73. https://doi.org/10.1056/NEJM199503233321202
25. Gehlbach S, Saag KG, Adachi JD et al (2012) Previous fractures at multiple sites increase the risk for subsequent fractures: the Global Longitudinal Study of Osteoporosis in Women. J Bone Miner Res 27(3):645–653. https://doi.org/10.1002/jbmr.1476
26 Hundrup YA, Ekholm O, Hoidrup S, Davidsen M, Obel EB (2005) Risk factors for hip fracture and a possible effect modification by hormone replacement therapy The Danish nurse cohort study. Eur J Epidemiol. 20(10):871–7. https://doi.org/10.1007/s10654-005-2151-z
27. Mallmin H, Ljunghall S, Persson I, Naessen T, Krusemo UB, Bergstrom R (1993) Fracture of the distal forearm as a forecaster of subsequent hip fracture: a population-based cohort study with 24 years of follow-up. Calcif Tissue Int 52(4):269–272. https://doi.org/10.1007/BF00296650

28. Clinton J, Franta A, Polissar NL et al (2009) Proximal humeral fracture as a risk factor for subsequent hip fractures. J Bone Joint Surg Am 91(3):503–11. https://doi.org/10.2106/JBJS.G.01529
29. Adams AL, Ryan DS, Li BH et al (2022) Outcomes post fragility fracture among members of an integrated healthcare organization. Osteoporos Int 33(4):783–790. https://doi.org/10.1007/s00198-021-06205-w
30. Fujiwara S, Kasagi F, Masunari N, Naito K, Suzuki G, Fukunaga M (2003) Fracture prediction from bone mineral density in Japanese men and women. J Bone Miner Res 18(8):1547–1553. https://doi.org/10.1359/jbmr.2003.18.8.1547
31. Johansson L, Johansson H, Axelsson KF et al (2022) Improved fracture risk prediction by adding VFA-identified vertebral fracture data to BMD by DXA and clinical risk factors used in FRAX. Osteoporos Int 33(8):1725–1738. https://doi.org/10.1007/s00198-022-06387-x
32. McCloskey EV, Vasireddy S, Threlkeld J et al (2008) Vertebral fracture assessment (VFA) with a densitometer predicts future fractures in elderly women unselected for osteoporosis. J Bone Miner Res 23(10):1561–1568. https://doi.org/10.1359/jbmr.080515
33. Napoli N, Schwartz AV, Palermo L et al (2013) Risk factors for subtrochanteric and diaphyseal fractures: the study of osteoporotic fractures. J Clin Endocrinol Metab 98(2):659–667. https://doi.org/10.1210/jc.2012-1896
34. Nguyen ND, Pongchaiyakul C, Center JR, Eisman JA, Nguyen TV (2005) Identification of high-risk individuals for hip fracture: a 14-year prospective study. J Bone Miner Res 20(11):1921–1928. https://doi.org/10.1359/JBMR.050520
35. Prince RL, Lewis JR, Lim WH et al (2019) Adding lateral spine imaging for vertebral fractures to densitometric screening: improving ascertainment of patients at high risk of incident osteoporotic fractures. J Bone Miner Res 34(2):282–289. https://doi.org/10.1002/jbmr.3595
36. Schousboe JT, Lix LM, Morin SN et al (2019) Prevalent vertebral fracture on bone density lateral spine (VFA) images in routine clinical practice predict incident fractures. Bone 121:72–79. https://doi.org/10.1016/j.bone.2019.01.009
37. Skjodt MK, Nicolaes J, Smith CD et al (2023) Fracture risk in men and women with vertebral fractures identified opportunistically on routine computed tomography scans and not treated for osteoporosis: an observational cohort study. JBMR Plus 7(5):e10736. https://doi.org/10.1002/jbm4.10736
38. Chen JS, Cameron ID, Simpson JM et al (2011) Low-trauma fractures indicate increased risk of hip fracture in frail older people. J Bone Miner Res 26(2):428–433. https://doi.org/10.1002/jbmr.216
39. Nakamura K, Takahashi S, Oyama M et al (2010) Prior nonhip limb fracture predicts subsequent hip fracture in institutionalized elderly people. Osteoporos Int 21(8):1411–1416. https://doi.org/10.1007/s00198-009-1081-0
40. Bliuc D, Nguyen ND, Milch VE, Nguyen TV, Eisman JA, Center JR (2009) Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. JAMA 301(5):513–521. https://doi.org/10.1001/jama.2009.50
41. Alarkawi D, Bliuc D, Tran T et al (2020) Impact of osteoporotic fracture type and subsequent fracture on mortality: the Tromso Study. Osteoporos Int 31(1):119–130. https://doi.org/10.1007/s00198-019-05174-5
42. Wells GA, Hsieh SC, Zheng C, Peterson J, Tugwell P, Liu W. (2022) Risedronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. Cochrane Database Syst Rev. 5(5). https://doi.org/10.1002/14651858.CD004523.pub4
43. McCloskey E, Rathi J, Heijmans S et al (2021) The osteoporosis treatment gap in patients at risk of fracture in European primary care: a multi-country cross-sectional observational study. Osteoporos Int 32(2):251–259. https://doi.org/10.1007/s00198-020-05557-z
44. Piple A, Smith CT, Barton DW, Carmouche JJ (2020) Proximal humerus fractures in the geriatric population present an opportunity to improve recognition and treatment of osteoporosis. Geriatr Orthop Surg Rehabil 11:2151459320935103. https://doi.org/10.1177/2151459320935103
45. van den Berg P, van Haard PMM, Geusens PP, van den Bergh JP, Schweitzer DH (2019) Challenges and opportunities to improve fracture liaison service attendance: fracture registration and patient characteristics and motivations. Osteoporos Int 30(8):1597–1606. https://doi.org/10.1007/s00198-019-05016-4
46. Morin SN, Feldman S, Funnell L et al (2023) Clinical practice guideline for management of osteoporosis and fracture prevention in Canada: 2023 update. CMAJ 195(39):E1333–E1348. https://doi.org/10.1503/cmaj.221647
47. Schousboe JT, Lix LM, Morin SN et al (2019) Vertebral fracture assessment increases use of pharmacologic therapy for fracture prevention in clinical practice. J Bone Miner Res 34(12):2205–2212
48. Kanis JA, Johansson H, McCloskey EV et al (2023) Previous fracture and subsequent fracture risk: a meta-analysis to update FRAX. Osteoporos Int 34(12):2027–2045. https://doi.org/10.1007/s00198-023-06870-z
49. Berry SD, Lee Y, Zullo AR, Kiel DP, Dosa D, Mor V (2016) Incidence of hip fracture in U.S. nursing homes. J Gerontol A Biol Sci Med Sci. 71(9):1230–4. https://doi.org/10.1093/gerona/glw03450
50. Schoberer D, Breimaier HE, Zuschnegg J, Findling T, Schaffer S, Archan T (2022) Fall prevention in hospitals and nursing homes: clinical practice guideline. Worldviews Evid Based Nurs 19(2):86–93. https://doi.org/10.1111/wvn.12571
51. Gregson CL, Armstrong DJ, Bowden J et al (2022) UK clinical guideline for the prevention and treatment of osteoporosis. Arch Osteoporos 17(1):58. https://doi.org/10.1007/s11657-022-01061-5
52. Sanders KM, Pasco JA, Ugoni AM et al (1998) The exclusion of high trauma fractures may underestimate the prevalence of bone fragility fractures in the community: the Geelong Osteoporosis Study. J Bone Miner Res 13(8):1337–1342

Authors and Affiliations

Takashi Ariie[1,2] · Norio Yamamoto[2,3,4] · Yusuke Tsutsumi[2,5,6] · Shuri Nakao[2,7] · Akihiro Saitsu[2,8,9] · Takahiro Tsuge[2,4,10] · Haruka Tsuda[2,11] · Yuki Nakashima[2,12] · Takanori Miura[2,13] · Yousuke Bandai[2,14] · Ryota Okoba[15] · Shunsuke Taito[2,12]

✉ Takashi Ariie
tks.ar1212@gmail.com

1 Department of Physical Therapy, School of Health Sciences at Fukuoka, International University of Health and Welfare, 137-1 Enokizu, Okawa-Shi, Fukuoka 831-8501, Japan

2 Scientific Research WorkS Peer Support Group (SRWS-PSG), Osaka, Japan

3 Department of Orthopedic Surgery, Hashimoto Hospital, Kagawa, Japan

4 Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan

5 Department of Emergency Medicine, National Hospital Organization Mito Medical Center, Ibaraki, Japan

6 Human Health Science, Kyoto University Graduate School of Medicine, Kyoto, Japan

7 Division of Rehabilitation Medicine, Shimane University Hospital, Shimane, Japan

8 Department of Orthopedic Surgery, Jichi Medical University, Tochigi, Japan

9 Medical Education Center, Jichi Medical University, Tochigi, Japan

10 Department of Rehabilitation, Kurashiki Medical Center, Okayama, Japan

11 Akihabara Medical Clinic, Tokyo, Japan

12 Division of Rehabilitation, Department of Clinical Practice and Support, Hiroshima University Hospital, Hiroshima, Japan

13 Department of Orthopedic Surgery, Akita Rosai Hospital, Akita, Japan

14 Department of Rehabilitation Medicine, Shimada Hospital, Shimada Social Medical Corporation, Fukuoka, Japan

15 Department of Physical Therapy, Faculty of Health and Welfare, Prefectural University of Hiroshima, Hiroshima, Japan

ANTICANCER RESEARCH
International Journal of Cancer Research and Treatment
ISSN: 0250-7005

Real Clinical Practice of Combined Atezolizumab Plus Chemotherapy in Patients With Small Cell Lung Cancer

SATOSHI ANO[1,2], NORIHIRO KIKUCHI[1], SHINICHIRO OKAUCHI[3], TAKESHI NUMATA[4], RYOTA NAKAMURA[4], TOSHIHIRO SHIOZAWA[2], HIROKO WATANABE[5], TOMOHIRO TAMURA[6], KUNIHIKO MIYAZAKI[7], SHIGEN HAYASHI[8], TAKAAKI YAMASHITA[9], KOICHI KURISHIMA[10], MASAHARU INAGAKI[11], TAKAYUKI KABURAGI[6], TAKEO ENDO[4], HIROAKI SATOH[3] and NOBUYUKI HIZAWA[2]

[1]Division of Respiratory Medicine, National Hospital Organization Kasumigaura Medical Center, Tsuchiura, Japan;
[2]Division of Respiratory Medicine, Faculty of Clinical Medicine, University of Tsukuba, Tsukuba, Japan;
[3]Division of Respiratory Medicine, Mito Medical Center, University of Tsukuba, Mito, Japan;
[4]Departments of Respiratory Medicine and Surgery, National Hospital Organization Mito Medical Center, Ibarakimachi, Japan;
[5]Division of Respiratory Medicine, Tsukuba Memorial Hospital, Tsukuba, Japan;
[6]Respiratory Center, Ibaraki Prefectural Central Hospital, Kasama, Japan;
[7]Division of Respiratory Medicine, Ryugasaki Saiseikai Hospital, Ryugasaki, Japan;
[8]Division of Respiratory Medicine, Ibaraki Seinan Medical Center Hospital, Sakai, Japan;
[9]Division of Respiratory Medicine, JA Toride Medical Center Hospital, Toride, Japan;
[10]Division of Respiratory Medicine, Tsukuba Medical Center Hospital, Tsukuba, Japan;
[11]Division of Thoracic Surgery, Tsuchiura Kyodo General Hospital, Tsuchiura, Japan

Reprinted from
ANTICANCER RESEARCH *44*: 2725-2730 (2024)
doi:10.21873/anticanres.17080

ANTICANCER RESEARCH

International Journal of Cancer Research and Treatment

HighWire Press® STANFORD UNIVERSITY

ISSN (print): 0250-7005
ISSN (online): 1791-7530

Editorial Board

Editorial Office: International Institute of Anticancer Research, 1st km Kapandritiou-Kalamou Rd., Kapandriti, P.O. Box 22, Attiki 19014, Greece. Tel / Fax: +30-22950-53389.

U.S. Branch: Anticancer Research USA, Inc., 111 Bay Avenue, Highlands, NJ 07732, USA.

E-mails: Editorial Office: journals@iiar-anticancer.org
Managing Editor: editor@iiar-anticancer.org

ANTICANCER RESEARCH supports: (a) the establishment and the activities of the INTERNATIONAL INSTITUTE OF ANTICANCER RESEARCH (IIAR; Kapandriti, Attiki, Greece); and (b) the organization of the International Conferences of Anticancer Research. The IIAR is a member of UICC. For more information about ANTICANCER RESEARCH, IIAR and the Conferences, please visit the IIAR website: www.iiar-anticancer.org

Publication Data: ANTICANCER RESEARCH (AR) is published bimonthly from January 1981 to December 2008 and monthly from January 2009. Each annual volume comprises 12 issues. Annual Author and Subject Indices are included in the last issue of each volume. ANTICANCER RESEARCH Vol. 24 (2004) and onwards appears online with Stanford University HighWire Press from April 2009. All published articles are deposited in PubMed Central.

Annual Subscription Rates 2024 per volume: Institutional subscription US$ 1,898.00 (online) or US$ 2,277.00 (print & online). Personal subscription US$ 897.00 (online) or US$ 1,277.00 (print & online). Prices include rapid delivery and insurance. The complete previous volumes of Anticancer Research (Vol. 1-43, 1981-2023) are available at 50% discount on the above rates.

Subscription Orders: Orders can be placed at agencies, bookstores, or directly with the Publisher. (e-mail: subscriptions@iiar-anticancer.org)

Advertising: All correspondence and rate requests should be addressed to the Editorial Office.

Book Reviews: Recently published books and journals should be sent to the Editorial Office. Reviews will be published within 2-4 months.

Articles in ANTICANCER RESEARCH are regularly indexed in all bibliographic services, including Current Contents Life Sciences and Medical Sciences, Science Citation Index Expanded, Index Medicus, Biological Abstracts, PubMed, PubMed Central, Chemical Abstracts, BIOSIS, Previews, Essential Science Indicators, Excerpta Medica, University of Sheffield Biomedical Information Service, Current Clinical Cancer, AIDS Abstracts, Elsevier Bibliographic Database, EMBASE, Compendex, GEOBASE, EMBiology, Elsevier BIOBASE, FLUIDEX, World Textiles, Scopus, Progress in Palliative Care, Cambridge Scientific Abstracts, Cancergram (International Cancer Research Data Bank), MEDLINE, Reference Update - RIS Inc., PASCAL-CNRS, Inpharma-Reactions (Datastar, BRS), CABS, Immunology Abstracts, Telegen Abstracts, Genetics Abstracts, Nutrition Research Newsletter, Dairy Science Abstracts, Current Titles in Dentistry, Inpharma Weekly, BioBase, MedBase, CAB Abstracts/Global Health Databases, Investigational Drugs Database, VINITI Abstracts Journal, Leeds Medical Information, PubsHub, Sociedad Iberoamericana de Información Científíca (SIIC) Data Bases.

D.T.P. BY IIAR

PRINTED BY ENTYPO, ATHENS, GREECE. PRINTED ON ACID-FREE PAPER

ANTICANCER RESEARCH *44*: 2725-2730 (2024)
doi:10.21873/anticanres.17080

Real Clinical Practice of Combined Atezolizumab Plus Chemotherapy in Patients With Small Cell Lung Cancer

SATOSHI ANO[1,2], NORIHIRO KIKUCHI[1], SHINICHIRO OKAUCHI[3], TAKESHI NUMATA[4], RYOTA NAKAMURA[4], TOSHIHIRO SHIOZAWA[2], HIROKO WATANABE[5], TOMOHIRO TAMURA[6], KUNIHIKO MIYAZAKI[7], SHIGEN HAYASHI[8], TAKAAKI YAMASHITA[9], KOICHI KURISHIMA[10], MASAHARU INAGAKI[11], TAKAYUKI KABURAGI[6], TAKEO ENDO[4], HIROAKI SATOH[3] and NOBUYUKI HIZAWA[2]

[1]*Division of Respiratory Medicine, National Hospital Organization Kasumigaura Medical Center, Tsuchiura, Japan;*
[2]*Division of Respiratory Medicine, Faculty of Clinical Medicine, University of Tsukuba, Tsukuba, Japan;*
[3]*Division of Respiratory Medicine, Mito Medical Center, University of Tsukuba, Mito, Japan;*
[4]*Departments of Respiratory Medicine and Surgery, National Hospital Organization Mito Medical Center, Ibarakimachi, Japan;*
[5]*Division of Respiratory Medicine, Tsukuba Memorial Hospital, Tsukuba, Japan;*
[6]*Respiratory Center, Ibaraki Prefectural Central Hospital, Kasama, Japan;*
[7]*Division of Respiratory Medicine, Ryugasaki Saiseikai Hospital, Ryugasaki, Japan;*
[8]*Division of Respiratory Medicine, Ibaraki Seinan Medical Center Hospital, Sakai, Japan;*
[9]*Division of Respiratory Medicine, JA Toride Medical Center Hospital, Toride, Japan;*
[10]*Division of Respiratory Medicine, Tsukuba Medical Center Hospital, Tsukuba, Japan;*
[11]*Division of Thoracic Surgery, Tsuchiura Kyodo General Hospital, Tsuchiura, Japan*

Abstract. *Background/Aim: Atezolizumab, an anti-PD-L1 antibody, has been increasingly administered in combination with chemotherapy to patients with small cell lung cancer (SCLC). This study aimed to determine how patients with extensive disease (ED) -SCLC responded to atezolizumab with chemotherapy and found factors affecting long-term response and survival. Patients and Methods: This study focused on patients with SCLC who were treated with a combination of atezolizumab and chemotherapy in Japan between 2019 and 2023. Patient information and tumor response were analyzed, along with adverse events. We compared data and estimated survival probabilities. Results: In our clinical trial, 95 patients with SCLC who received this treatment had a median progression-free survival of 6.0 months and a median overall survival of 15.0 months. Immune-related adverse events were observed in 13.7% of the patients, with grade 3 or higher in 5.3%. The efficacy and immune-related adverse events associated with this treatment regimen were comparable to those reported in previous clinical trials. Progression-free survival >2 years was observed in a small number of patients (5.3%). Conclusion: Our research will offer important insights for the future care of patients with extensive-stage SCLC by utilizing atezolizumab in combination with chemotherapy. Accumulation and confirmation of clinical practice results will have important implications for the future implementation of this therapy.*

Correspondence to: Satoshi Ano, Department of Respiratory Medicine National Hospital Organization Kasumigaura Medical Center, 2-7-14 Shimotakatsu, Tsuchiura, Ibaraki 300-8585, Japan. Tel: +81 298225050, e-mail: satoshi.ano@gmail.com

Key Words: Atezolizumab, chemotherapy, progression-free survival, overall survival, immune-related adverse events.

T lymphocytes, which play a central role in antitumor immunity, are activated through interactions with antigen-presenting cells to attack tumor cells (1, 2). Anti-programmed death ligand 1 (PD-L1) antibodies, atezolizumab and durvalumab, bind to PD-L1 expressed by cancer cells and antigen-presenting cells, inhibiting their interaction with PD-1 on T cells (1, 2). As a result, T signaling inhibition is reduced, and T cell activation is maintained (1, 2). Similar to other immune checkpoint inhibitors (ICIs), such as anti-PD-1 antibodies, atezolizumab was initially administered as a single agent for non-small cell lung cancer (NSCLC) (3); however, it is now commonly used in combination with chemotherapy, especially for the treatment of SCLC (4). Furthermore, both the effects of ICIs and their add-on effects as first-line therapies for extensive disease small cell lung cancer (ED-SCLC) have been demonstrated in recent years.

The IMpower133 trial showed that adding atezolizumab to carboplatin and etoposide prolonged the overall survival (OS) (5). In the CASPIAN trial, platinum, etoposide, and durvalumab were equally effective (6). Currently, ICI therapy for SCLC is primarily administered in combination with chemotherapy, and understanding its status in clinical practice is important. However, few studies have included adequate number of patients to confirm the clinical response and duration of the response. Therefore, we accumulated as much patient information as possible to best reflect the clinical practice. In clinical practice, we conducted a study to determine the response status of patients with ED-SCLC to atezolizumab plus chemotherapy and identify factors contributing to long-term response and survival.

Patients and Methods

This retrospective study included patients with SCLC who received combination therapy of atezolizumab plus chemotherapy between September 2019 and December 2023 at 11 hospitals in Ibaraki Prefecture, Japan. Clinical information of patients with pathologically diagnosed SCLC who received atezolizumab plus chemotherapy was compiled. Pathological diagnoses were based on the WHO classification (7). All patients underwent TNM classification (8) using brain computed tomography (CT) or magnetic resonance imaging (MRI), bone scans, ultrasonography, and abdominal CT before initiating atezolizumab plus chemotherapy. Eligible patients were identified from each hospital's clinical database, and the following information was extracted: patient demographics [age, sex, Eastern Cooperative Oncology Group performance status (PS), histology, and stage] and objective tumor response at the start of combination therapy. This research was approved by the University of Tsukuba Mito Medical Center, Mito Kyodo General Hospital (NO-22-42-CHEMO), and each Institutional Review Board. Comprehensive patient consent for this clinical retrospective study and its publication was obtained from each patient.

Tumor response was assessed as complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD), or failure to respond according to the Response Evaluation Criteria in Solid Tumors (9). Adverse events were classified using the National Cancer Institute Common Terminology Criteria for Adverse Events (version 5.0) (10).

The chi-square test was used to test for differences in proportions. The Mann–Whitney *U*-test was used to compare values between two unpaired groups, such as when comparing patient age. The log-rank test and Cox proportional hazards model were used to estimate survival probabilities using the Kaplan–Meier method. *p*-Values less than 0.01 were considered to indicate a significant difference.

Results

Patient characteristics. Clinical data were collected for 95 patients during the study period. Table I shows the patient characteristics. A total of 71 (74.7%) patients were male, with a median age of 71 years (range=47-85 years); 71 (74.7%) patients had a PS of 0-1, and only one patient had mixed-type SCLC. No epidermal growth factor receptor (EGFR) genetic testing or immunostaining was performed to examine PD-L1 expression. The treatment regimens and timing of administration were as follows: atezolizumab plus etoposide plus carboplatin (IMpower133 regimen) as first-line therapy in 74 patients (77.9%), second-line therapy in 12 patients (12.6%), and third-line or later therapy in nine patients (9.5%).

Table I. *Backgrounds of clinical features in small cell lung cancer (SCLC) patients treated with atezolizumab and chemotherapy.*

	Number of patients
Total number of patients	95
Sex, male: female	71:24
Age, median (range) years	71(47-85)
Age <70 years: age ≥70 years	34:61
Performance status (ECOG), 0:1:2:3	13:58:19:5
Pathology, SCLC: combined SCLC	94:1

ECOG: Eastern Cooperative Oncology Group.

Response to treatment. Figure 1 shows the specific treatment sequence of the patient cohort. Table II shows the treatment responses [CR, PR, SD, PD, response rate (RR), and disease control rate (DCR)] for all patients and those treated with each regimen. The RR of 95 patients treated with atezolizumab plus chemotherapy was 57.9%. SD was observed in 17 patients (17.9%) with a DCR of 75.8%.

Survival analysis. Of the 95 evaluable patients, 64 died at the time of analysis. The median follow-up was 13.0 months [95% confidence interval (CI)=13.1-18.0 months]. 1- and 2-year OS was 29.0% (95%CI=22.8-31.6%) and 10.0% (95%CI=7.5-13.2%), respectively. Median progression-free survival (PFS) was 6.0 months (95%CI=5.0-7.0 months; Figure 2A), and median OS was 15.0 months (95%CI=12.4-17.6 months; Figure 2B).

To identify favorable factors affecting PFS and OS, a univariate analysis was performed using sex, PS, age, stage, immune-related adverse events (irAEs), and line of treatment as variables. In univariate analysis, none of the factors favored PFS or OS (Table III).

Of the 95 evaluable patients, 13 (13.7%) had PFS ≥1 year, and five (5.3%) had PFS >2 years. Thirteen patients with PFS ≥1 year, compared with 82 patients with PFS <1 year, showed no significant clinical differences (Table IV).

Toxicity. Table V shows irAE. Thirteen of the 95 patients (13.7%) had irAEs of any grade, of which five (5.3%) had grade 3 or higher irAEs. Twelve patients were identified as having a single irAE, and one had two types of irAEs. Pulmonary irAEs were the most frequent and serious adverse

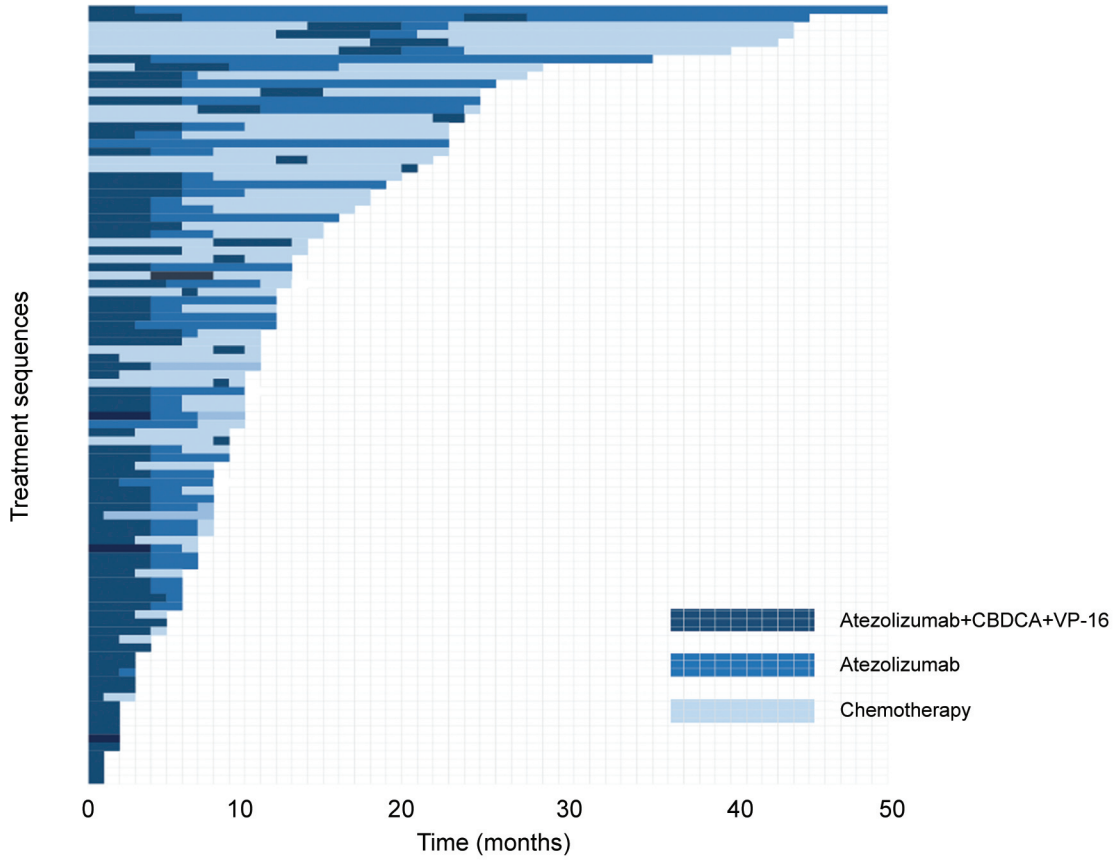


Figure 1. *The specific treatment sequences for the 95 patients with small cell lung cancer treated with atezolizumab plus chemotherapy are shown.*

events. Most irAEs were manageable and reversible. However, grade 5 irAEs occurred in one patient in the lung. The median OS of 13 patients with irAEs of any grade was 12 months.

Discussion

Background factors associated with improved PFS and OS were examined in the 95 patients evaluated in this study, but no significant factors were identified. Only five patients (5.3%) had PFS >2 years, and no characteristic clinical findings were found in these five patients compared to the other 90 patients. Among ICIs, anti-PD-1 antibodies, such as nivolumab and pembrolizumab, suppress immune checkpoints by binding to PD-1 on T-cells (1, 2). In contrast, atezolizumab and durvalumab, anti-PD-L1 antibodies, suppress immune checkpoints by binding to PD-L1 in cancer cells and antigen-presenting cells (1, 2). Carboplatin + etoposide + pembrolizumab has not been shown to prolong OS, and future studies are needed to determine what factors influence clinical outcomes (11). Platinum-based agents and the novel PD-L1 antibodies serplulimab (ASTRUM-005 trial) and debrelimab (CAPSTONE-1 trial) have also been shown to prolong OS and are promising agents for the future (12, 13). Treatment of SCLC with ICI has not shown efficacy with ICI monotherapy (14). However, ICI monotherapy has shown promising antitumor activity and a favorable safety profile in patients with SCLC after platinum-based or at least one prior therapy

Table II. *Treatment response.*

Treatment line of Atezolizumab + chemotherapy	1st: 2nd: 3rd or later
Response	
Complete response (CR)	3:0:0
Partial response (PR)	48:3:1
Stable disease (SD)	10:5:2
Progressive disease (PD)	13:4:6
Response rate (CR+PR)	55 (57.9%)
Disease control rate (CR+PR+SD)	72 (75.8%)
All patients	95

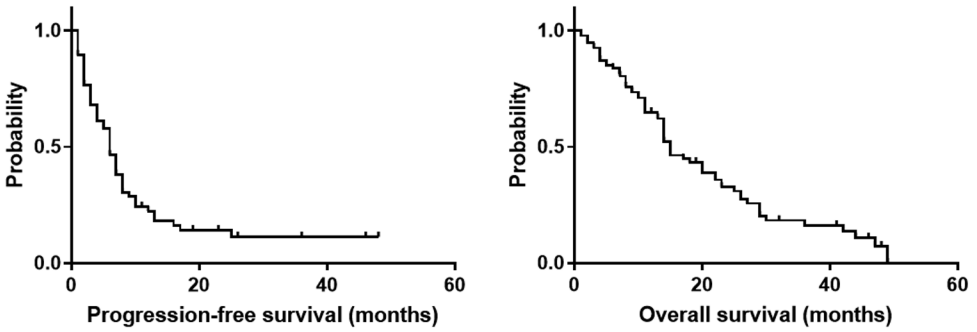


Figure 2. *In the 95 patients treated with atezolizumab plus chemotherapy, median progression-free survival was 6.0 months (95%CI=5.0-7.0 months) (A) and median overall survival was 15.0 months (95%CI=12.4-17.6 months) (B).*

Table III. *Univariate analysis of factors affecting progression-free survival (PFS) and overall survival (OS).*

	p-Value for PFS	*p*-Value for OS
Sex, male : female	0.4977	0.1429
Performance status (ECOG), 0-1 : 2-4	0.0983	0.0848
Age <70 years : Age ≥70 years	0.7295	0.9148
irAEs, absent : present	0.3115	0.8534

SCLC: Small cell lung cancer; ECOG: Eastern Cooperative Oncology Group.

Table IV. *Comparison of clinical characteristics of small cell lung cancer (SCLC) patients who had progression-free survival (PFS) more than one year and those who did not.*

	Patients with PFS ≥1 year	Patients with PFS <1 year	*p*-Value
Number of patients	13	82	
Sex, male : female	8:5	63:19	0.2385
Performance status (ECOG), 0-1 : 2-4	9:4	61:21	0.6947
Age <70 years : Age ≥70 years	5:8	29:53	0.8287
irAEs, absent : present	11:2	71:11	0.8477

ECOG: Eastern Cooperative Oncology Group.

Table V. *Immune-related adverse events (irAE).*

Number of irAE	Number of patients who had irAE		
Not developed	82		
Developed	13		
One	12		
Two	1		
Three	0		
All patients	95		
irAE	**Grade 2**	**Grade 3-4**	**Grade 5**
Pulmonary toxicity	2	1	1
Skin toxicity	1	1	0
Thyroid dysfunction	3	0	0
Diarrhea	1	0	0
Pericarditis	1	0	0
Type 1 diabetes	0	1	0
Febrile neutropenia	0	1	0

(14). Attention should be paid to adverse events, particularly in cases complicated by interstitial pneumonia, and platinum plus etoposide should be considered (14). To the best of our knowledge, there are few reports on the results of atezolizumab plus chemotherapy for SCLC in clinical practice, and the median age was similar to that of previously reported clinical trials (4); however, the population with PS 0-1 was smaller than that previously reported (15). Owing to the real-world clinical setting, many patients (25.3%) had a PS score of 2 or higher.

Despite these differences in patient background, PFS and OS were comparable to those in previous retrospective studies, confirming no differences in race or medical environment (16). Additionally, this study examined clinical factors associated with long-term responses, but no significant results were found. However, it was discovered that five patients had PFS of two years or longer. It is essential to confirm the existence of such patients, and future studies are warranted to elucidate factors associated with long-term responses. Caution should be exercised when administering ICIs in combination with chemotherapy. In the IMpower133 study of atezolizumab with chemotherapy, 39.9% of patients experienced irAEs, with serious irAEs including rash (2%), hepatitis (1.5%), infusion-related reactions (2%), and colitis (1%) (5). In the IFCT-1603 study, the most common irAEs were musculoskeletal or connective tissue disorders (12.5%) and gastrointestinal disorders (18.8%); however, no irAEs above G3 have been reported

(17). Our results showed that the incidence of all grades of irAEs was low (13.7%), and irAEs of G3 or higher were observed in 5.3% of patients, suggesting that the relatively more serious irAEs were due to the relatively high number of patients with high PS. However, the low overall incidence of irAEs could not be ruled out as a possible underestimation in this retrospective study. This study had other limitations that should be mentioned. Although the study included sufficient number of patients for medical statistics, it was a retrospective study of patients with various background characteristics. It is essential to compare the PFS and OS in patients who received atezolizumab plus chemotherapy with those who received chemotherapy alone, as in other studies. Selection bias due to background factors in choosing therapy (*e.g.*, the presence of pulmonary fibrosis and high PS) may be a major obstacle in comparing PFS and OS. Genomic profiling of SCLC revealed extensive chromosomal rearrangements and high mutational burden, likely including functional inactivation of the tumor suppressor genes *TP53* and *RB1*, and the four major subtypes of SCLC are defined based on high expression of ASCL1 (SCLC-A subtype), NEUROD1 (SCLC-N), POU2F3 (SCLC-P), and YAP1 (SCLC-Y), defined based on their high expression (18). However, this was not investigated in this study. The frequency and severity of irAEs were similar to those in studies combining ICI and chemotherapy and in previously published retrospective studies, indicating that long-term treatment is feasible for some patients (16).

Conclusion

To maximize the efficacy of combination therapies, including atezolizumab, close attention should be paid to irAEs, and actual clinical outcomes should be assembled and utilized. Moreover, the elucidation of biomarker factors that indicate the efficacy of long-term administration will have a beneficial impact on clinical practice. We believe that this information will provide useful insights for future treatments of ED-SCLC with atezolizumab plus chemotherapy.

Conflicts of Interest

The Authors declare that they have no conflicts of interest in relation to this study.

Authors' Contributions

SA, NK, and HS designed this study. NK, SO, TN, RN, TS, HW, TT, KM, SH, TY, KK, MI, HS, TK, and TE collected the data. SA, NK, and HS analyzed the data. SA, NK, HS, and NH drafted the manuscript. HS and NH supervised this study. All Authors have approved the final version of the manuscript for submission.

Acknowledgements

The Authors would like to thank Editage (https://www.editage.com) for English language editing.

Funding

No funding was received for this study.

References

1 Khadela A, Postwala H, Rana D, Dave H, Ranch K, Boddu SHS: A review of recent advances in the novel therapeutic targets and immunotherapy for lung cancer. Med Oncol 40(5): 152, 2023. DOI: 10.1007/s12032-023-02005-w

2 Lahiri A, Maji A, Potdar PD, Singh N, Parikh P, Bisht B, Mukherjee A, Paul MK: Lung cancer immunotherapy: progress, pitfalls, and promises. Mol Cancer 22(1): 40, 2023. DOI: 10.1186/s12943-023-01740-y

3 Gandara DR, von Pawel J, Mazieres J, Sullivan R, Helland Å, Han JY, Ponce Aix S, Rittmeyer A, Barlesi F, Kubo T, Park K, Goldschmidt J, Gandhi M, Yun C, Yu W, Matheny C, He P, Sandler A, Ballinger M, Fehrenbacher L: Atezolizumab treatment beyond progression in advanced NSCLC: Results from the randomized, phase III OAK study. J Thorac Oncol 13(12): 1906-1918, 2018. DOI: 10.1016/j.jtho.2018.08.2027

4 Liu X, Xing H, Liu B: Current status and future perspectives of immune checkpoint inhibitors in extensive-stage small cell lung cancer. Am J Cancer Res 12(6): 2447-2464, 2022.

5 Horn L, Mansfield AS, Szczęsna A, Havel L, Krzakowski M, Hochmair MJ, Huemer F, Losonczy G, Johnson ML, Nishio M, Reck M, Mok T, Lam S, Shames DS, Liu J, Ding B, Lopez-Chavez A, Kabbinavar F, Lin W, Sandler A, Liu SV, IMpower133 Study Group: First-line atezolizumab plus chemotherapy in extensive-stage small-cell lung cancer. N Engl J Med 379(23): 2220-2229, 2018. DOI: 10.1056/NEJMoa1809064

6 Paz-Ares L, Dvorkin M, Chen Y, Reinmuth N, Hotta K, Trukhin D, Statsenko G, Hochmair MJ, Özgüroğlu M, Ji JH, Voitko O, Poltoratskiy A, Ponce S, Verderame F, Havel L, Bondarenko I, Kazarnowicz A, Losonczy G, Conev NV, Armstrong J, Byrne N, Shire N, Jiang H, Goldman JW, CASPIAN investigators: Durvalumab plus platinum-etoposide *versus* platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer (Caspian): a randomised, controlled, open-label, phase 3 trial. Lancet 394(10212): 1929–1939, 2019. DOI: 10.1016/S0140-6736(19)32222-6

7 WHO Classification of Tumours Editorial Board: World Health Organization Classification of Tumours, Thoracic Tumours 5th edn. IARC Press, 2021.

8 Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, Nicholson AG, Groome P, Mitchell A, Bolejack V, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, Advisory Boards, and Participating Institutions and International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee Advisory Boards and Participating Institutions: The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the

TNM classification for lung cancer. J Thorac Oncol 11(1): 39-51, 2016. DOI: 10.1016/j.jtho.2015.09.009

9 Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). 1 45(2): 228-247, 2009. DOI: 10.1016/j.ejca.2008.10.026

10 National Cancer Institute Common Terminology Criteria for Adverse Events version 5.0. Available at: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcae_v5_quick_reference_5x7.pdf [Last accessed on February 27, 2024]

11 Rudin CM, Awad MM, Navarro A, Gottfried M, Peters S, Csőszi T, Cheema PK, Rodriguez-Abreu D, Wollner M, Yang JC, Mazieres J, Orlandi FJ, Luft A, Gümüş M, Kato T, Kalemkerian GP, Luo Y, Ebiana V, Pietanza MC, Kim HR, KEYNOTE-604 Investigators: Pembrolizumab or placebo plus etoposide and platinum as first-line therapy for extensive-stage small-cell lung cancer: randomized, double-blind, phase III KEYNOTE-604 study. J Clin Oncol 38(21): 2369-2379, 2020. DOI: 10.1200/JCO.20.00793

12 Cheng Y, Han L, Wu L, Chen J, Sun H, Wen G, Ji Y, Dvorkin M, Shi J, Pan Z, Shi J, Wang X, Bai Y, Melkadze T, Pan Y, Min X, Viguro M, Li X, Zhao Y, Yang J, Makharadze T, Arkania E, Kang W, Wang Q, Zhu J, ASTRUM-005 Study Group: Effect of first-line serplulimab vs placebo added to chemotherapy on survival in patients with extensive-stage small cell lung cancer: the ASTRUM-005 randomized clinical trial. JAMA 328(12): 1223-1232, 2022. DOI: 10.1001/jama.2022.16464

13 Wang J, Zhou C, Yao W, Wang Q, Min X, Chen G, Xu X, Li X, Xu F, Fang Y, Yang R, Yu G, Gong Y, Zhao J, Fan Y, Liu Q, Cao L, Yao Y, Liu Y, Li X, Wu J, He Z, Lu K, Jiang L, Hu C, Zhao W, Zhang B, Shi W, Zhang X, Cheng Y, CAPSTONE-1 Study Group: Adebrelimab or placebo plus carboplatin and etoposide as first-line treatment for extensive-stage small-cell lung cancer (CAPSTONE-1): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 23(6): 739-747, 2022. DOI: 10.1016/S1470-2045(22)00224-8

14 Fu Y, Zheng Y, Wang PP, Ding ZY: Toxicities of immunotherapy for small cell lung cancer. Front Oncol 11: 603658, 2021. DOI: 10.3389/fonc.2021.603658

15 Fujimoto D, Morimoto T, Tamiya M, Hata A, Matsumoto H, Nakamura A, Yokoyama T, Taniguchi Y, Uchida J, Sato Y, Yokoi T, Tanaka H, Furuya N, Masuda T, Sakata Y, Miyauchi E, Hara S, Saito G, Miura S, Kanazu M, Yamamoto N, Akamatsu H: Outcomes of chemoimmunotherapy among patients with extensive-stage small cell lung cancer according to potential clinical trial eligibility. JAMA Netw Open 6(2): e230698, 2023. DOI: 10.1001/jamanetworkopen.2023.0698

16 Zou Y, Ren X, Zhang H, Wang Y, Wang H, Bai R, Zhang Z, Sun G, Xu L: Efficacy and safety of durvalumab + chemotherapy vs. atezolizumab + chemotherapy in the treatment of small cell lung cancer: a retrospective comparative cohort study. J Thorac Dis 15(6): 3339-3349, 2023. DOI: 10.21037/jtd-23-588

17 Pujol JL, Greillier L, Audigier-Valette C, Moro-Sibilot D, Uwer L, Hureaux J, Guisier F, Carmier D, Madelaine J, Otto J, Gounant V, Merle P, Mourlanette P, Molinier O, Renault A, Rabeau A, Antoine M, Denis MG, Bommart S, Langlais A, Morin F, Souquet PJ: A randomized non-comparative phase II study of anti-programmed cell death-ligand 1 atezolizumab or chemotherapy as second-line therapy in patients with small cell lung cancer: results from the IFCT-1603 trial. J Thorac Oncol 14(5): 903-913, 2019. DOI: 10.1016/j.jtho.2019.01.008

18 Rudin CM, Brambilla E, Faivre-Finn C, Sage J: Small-cell lung cancer. Nat Rev Dis Primers 7(1): 3, 2021. DOI: 10.1038/s41572-020-00235-0

Received March 11, 2024
Revised March 26, 2024
Accepted March 27, 2024

INSTRUCTIONS FOR AUTHORS

General Policy

ANTICANCER RESEARCH (AR) will accept original high-quality works and reviews on all aspects of experimental and clinical cancer research. The Editorial Policy suggests that priority will be given to papers advancing the understanding of cancer causation, and to papers applying the results of basic research to cancer diagnosis, prognosis, and therapy. Each article should include a concrete conclusion constituting of a "new piece of knowledge" backed up by unambiguous and accurate scientific evidence. AR will also accept the following for publication: (a) Abstracts and Proceedings of scientific meetings on cancer, following consideration and approval by the Editorial Board; (b) Announcements of meetings related to cancer research; (c) Short reviews (of approximately 120 words) and announcements of newly received books and journals related to cancer, and (d) Announcements of awards and prizes.

The principal aim of AR is to provide prompt publication (print and online) for original works of high quality, generally within 1-2 months from final acceptance. Manuscripts will be accepted on the understanding that they report original unpublished works in the field of cancer research that are not under consideration for publication by another journal, and that they will not be published again in the same form. All authors should sign a submission letter confirming the approval of their article contents. All material submitted to AR will be subject to peer-review, when appropriate, by two to three suitable referees. All manuscripts submitted to AR are urgently treated with absolute confidence, with access restricted to the Managing Editor, the journal's secretary, the reviewers, and the printers. The Editors reserve the right to improve manuscripts on grammar and style.

AR requires that all manuscripts be prepared in accordance with the **"Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals"** (https://www.icmje.org/icmje-recommendations.pdf) as published by the International Committee of Medical Journal Editors (ICMJE). We also support and adhere to the **"Principles of Transparency and Best Practice in Scholarly Publishing"** (https://publicationethics.org/resources/guidelines/principles-transparency-and-best-practice-scholarly-publishing) (a joint statement by COPE, DOAJ, WAME, and OASPA).

The Editors and Publishers of AR accept no responsibility for the contents and opinions expressed by the contributors. Authors should warrant due diligence in the creation and issuance of their work.

AR is a monthly print and online hybrid open-access journal (a subscription journal in which some of the articles are open access). All articles that are published as open access are with gold OA, which means that the final published version is permanently and freely available to anyone. All articles of Anticancer Research in HighWire become open access two years after their publication. Our open access articles are distributed under the terms and conditions of the **Creative Commons Attribution (CC BY-NC-ND) 4.0 international license (https://creativecommons.org/licenses/by-nc-nd/4.0).**

Manuscript Format

Three types of papers may be submitted: (i) Full papers containing completed original work, (ii) review articles concerning fields of recognizable progress, and (iii) letters to the Editor. Papers should contain all essential data in order to make the presentation clear. Reasonable economy should be exercised with respect to the number of tables and illustrations used. Papers should be written in clear, concise American English.

Submitted original manuscripts exceeding 4 printed pages will be subject to excess page charges. The 4 printed pages correspond approximately to twelve (12) document pages (~250 words per double-spaced typed page in Arial 12), including abstract, text, tables, figures, and references. Excess pages are charged USD 230.00 each. Each color page is charged USD 350.00. Review articles should not exceed 35 pages (approximately 250 words per double-spaced typed page) including all tables, figures, and references.

Sections

All manuscripts should be divided into the following sections:

a. First page including (i) the title of the presented work [not exceeding fifteen (15) words], (ii) full names and affiliations of all authors (with a maximum of 20 authors), (iii) name of the corresponding author(s) (with a maximum of 2 corresponding authors) to whom proofs are to be sent (with affiliation, full postal address, telephone and e-mail), (iv) key words, (v) an abbreviated running title, (vi) an indication "review", "clinical", "epidemiological", or "experimental" study, and (vii) the date of submission. Note: The order of the authors is not necessarily indicative of their contribution to the work. Authors may note their individual contribution(s) in the appropriate section(s) of the presented work. Affiliations should be indicated with a superscript number immediately after each author's name and in front of the appropriate address. Affiliations should not include street, box number or postal (zip) code.

b. Abstract not exceeding 250 words, organized according to the following headings: Background/Aim – Materials and Methods/Patients and Methods – Results – Conclusion. For Case Reports the structure should be as follows: Background/Aim – Case Report – Conclusion.

c. Introduction;

d. Materials and Methods/Patients and Methods/Case Report;

e. Results (not needed in a Case Report);

f. Discussion;

g. Conclusion;

h. Conflicts of Interest;

i. Authors' Contributions;

j. Acknowledgements;

k. Funding;

l. References.

All pages must be numbered consecutively. Footnotes should be avoided. Review articles may follow a different style according to the subject matter and the author's opinion.

Headings and Subsections

The article should be divided into clearly defined unnumbered sections. Main headings should be typed in bold on a separate line on the left of the page. The subheadings should be typed in bold italics at the left of the page on a separate line, and only the first word should begin with a capital letter. The sub-subheadings should be typed in italics on a new line, aligned full left. The text should start on the same line with subheadings and sub-subheadings.

Figures

All figures should appear **at the end** of the submitted document file and should be numbered with Arabic numerals (1, 2, 3, etc.) according to their sequence in the text. Once a manuscript is accepted all figures and graphs should be submitted separately in either jpg, tiff, or pdf format and at a minimum resolution of 300 dpi. Graphs must be submitted as pictures made from drawings and must not require any artwork, typesetting, or size modifications. Symbols, numbering, and lettering should be clearly legible. The number and top of each figure must be indicated.

Tables

All tables should appear **at the end** of the submitted document file and should be numbered with Latin numerals (I, II, III, etc.) according to their sequence in the text. Once a manuscript is accepted, each table should be submitted separately in an editable format, typed double-spaced. Tables should include a short title. Tables should not be divided into two or more parts, should not contain vertical rules, and the main body of the table should not contain horizontal rules.

Numerals

The authors should write numbers of 10 or more as numerals except at the beginning of a sentence. Numbers one to nine should be written in words, unless they precede units of measure or are used as designators. The authors should use decimal points (not decimal commas) and a comma for thousands (1,000 and above). Decimals should not be quoted with naked points, for example the authors should quote 0.01, not .01. *p*-Values for significant outcomes can be quoted as below a threshold significance value (*e.g.*, $p<0.05$, 0.01, 0.001), but wherever possible should be quoted as an exact probability value. Departure from a significance threshold of 0.05 should be stated and justified in the Methods. Nonsignificant outcomes should be indicated with an exact probability value whenever possible, or as NS or $p>0.05$, as appropriate for the test.

Supplementary Material

The journal does not have provision for use of supplementary material (Tables, Figures, Videos, or other material). The authors may (i) include their supplementary Tables/Figures as standard material or (ii) provide their own http/ftp link and upload the material on a website maintained by the authors (in this case the links for the supplementary material are given at the end of the paper under the section "Supplementary Material") or (iii) exclude the material from publication and provide it only for Reviewers' attention.

Conflicts of Interest and Authors' Contributions

All authors will be asked to supply authors' contributions and conflicts of interest information. We encourage authors to outline their individual contributions to the paper using the relevant CRediT roles: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Roles/Writing - original draft; Writing - review & editing.

References

Authors must assume responsibility for the accuracy of the references used. Citations for the reference sections of submitted works should follow the form below and must be numbered consecutively. In the text, references should be cited by number in parenthesis, *e.g.*, (1, 2). Examples:

1 Kenyon J, Liu W, Dalgleish A: Report of objective clinical responses of cancer patients to pharmaceutical-grade synthetic cannabidiol. Anticancer Res 38(10): 5831-5835, 2018. DOI: 10.21873/anticanres.12924 (DOIs only if applicable)

2 McGuire WL and Chamnes GC: Studies on the oestrogen receptor in breast cancer. In: Receptors for Reproductive Hormones. O' Malley BW, Chamnes GC (eds.). New York City, NY, USA, Plenum Publ Corp., pp 113-136, 1973.

3 Global Health Estimates 2015: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000-2015. Geneva, Switzerland, World Health Organisation, 2016. Available at: **http://www.who.int/healthinfo/global_burden_disease/estimates/en/index2.html** [Last accessed on April 3, 2018] (The web address should link directly to the cited information and not to a generic webpage)

You may download our journal's style for Endnote at **https://iiar-anticancer.org/wp-content/uploads/2023/04/IIAR_Anticancer-Res_2023.zip**

Nomenclature and Abbreviations

Nomenclature should follow that given in "Chemical Abstracts", "Index Medicus", "Merck Index", "IUPAC -IUB", "Bergey's Manual of Determinative Bacteriology", The CBE Manual for Authors, Editors and Publishers (6th edition, 1994), and MIAME Standard for Microarray Data. Human gene symbols may be obtained from the HUGO Gene Nomenclature Committee (HGNC) **(http://www.gene.ucl.ac.uk/).** Approved mouse

nomenclature may be obtained from **http://www.informatics.jax.org/.** Standard abbreviations are preferable. The authors should define abbreviations that are not standard in this field at their first mention in the abstract, main text, Figures and Table legends, and should ensure consistency of abbreviations throughout the article.

Definitions

Sex generally refers to a set of biological attributes that are associated with physical and physiological features (*e.g.*, chromosomal genotype, hormonal levels, internal and external anatomy). In humans, a binary sex categorization (male/female) is usually designated at birth ('sex assigned at birth'), most often based solely on the visible external anatomy of a newborn. Gender generally refers to socially constructed roles, behaviours and identities of women, men, and gender-diverse people that occur in a historical and cultural context and may vary across societies and over time. The terms 'sex' and 'gender' can be ambiguous; thus, it is important for authors of studies on human subjects to define the way they are used.

Submission Process

Submission of Manuscripts

Please follow the Instructions for Authors regarding the format of your manuscript and references.
Manuscripts must be submitted only through our online submission system at: **http://www.iiar-submissions.com/login.html**
In case a submission is incomplete, the corresponding author will be notified accordingly. Questions regarding difficulties in using the online submission system should be addressed to email: **journals@iiar-anticancer.org**

Article Transfer Service

If the Editor feels that the submitted manuscript is more suitable for an alternative journal, the authors might be asked to consider transferring the manuscript to such a journal. If they agree, the manuscript will be transferred, though the authors will have the opportunity to make changes to the manuscript before the submission is complete. The manuscript will be independently reviewed by the new journal.

Revision of Manuscripts

When the authors revise their paper, they need to prepare a detailed explanation of how they have dealt with the reviewers' comments and include their response in the first page of the revised manuscript file. In addition, the authors should use the reviewers' edited manuscript file for their corrections (not the original submitted file) and submit online a highlighted version of their revised manuscript. For the highlighted version, the authors may use the Track Changes tool in MS Word or highlight their changes in yellow.

Galley Proofs

Unless otherwise indicated, galley proofs will be sent to the corresponding author of the submission. Corrections of galley proofs should be limited to typographical errors. Galley proofs should be returned corrected to the Editorial Office by email **(iiar@iiar-anticancer.org)** within 24 hours.

Specific Information and Additional Instructions for Authors

1. Anticancer Research (AR) closely follows the new developments in all fields of experimental and clinical cancer research by (a) inviting reviews on topics of immediate importance and substantial progress in the last three years, and (b) providing the highest priority for rapid publication to manuscripts presenting original results judged to be of exceptional value. Theoretical papers will only be considered and accepted if they bear a significant impact or formulate existing knowledge for the benefit of research progress.
2. AR will consider the publication of conference proceedings and/or abstracts provided that the material submitted fulfils the quality requirements and instructions of the journal, following the regular review process by two-three suitable referees.
3. An acknowledgement of receipt, including the article number, title and date of receipt is sent to the corresponding author of each manuscript upon receipt. If this receipt is not received within 20 days from submission, the author should call or write to the Editorial Office to ensure that the manuscript (or the receipt) was properly uploaded during the electronic submission.
4. Each manuscript submitted to AR is sent for peer-review (single-blind) in confidence to two-three suitable referees with the request to return the manuscript with their comments to the Editorial Office within 12 days from receipt. If reviewers need a longer time or wish to send the manuscript to another expert, the manuscript may be returned to the Editorial Office with a delay. All manuscripts submitted to AR, are treated in confidence, without access to any person other than the Managing Editor, the journal's secretary, the reviewers, and the printers.
5. All accepted manuscripts are carefully corrected in style and language, if necessary, to make presentation clear (there is no fee for this service). Every effort is made (a) to maintain the personal style of the author's writing and (b) to avoid change of meaning. Authors will be requested to examine carefully manuscripts which have undergone language correction at the pre-proof or proof stage.
6. Authors should pay attention to the following points when writing an article for AR:

- The Instructions to Authors must be followed in every detail.
- Authors submitting manuscripts for review in our journal are kindly requested to utilize their *institutional e-mail addresses* instead of personal ones. This ensures efficient communication and maintains academic integrity throughout the review process.
- The presentation of the experimental methods should be clear and complete in every detail facilitating reproducibility by other scientists.
- The presentation of results should be simple and straightforward in style. Results and discussion should not be combined into one section, unless the paper is short.
- Results given in figures should not be repeated in tables.

- Figures (graphs or photographs) should be prepared at a width of 8 or 17 cm with legible numbers and lettering.
- Photographs should be clear with high contrast, presenting the actual observation described in the legend and in the text. Each legend should provide a complete description, being self-explanatory, including technique of preparation, information about the specimen and magnification.
- Statistical analysis should be elaborated wherever it is necessary. Simplification of presentation by giving only numerical or % values should be avoided.
- Fidelity of the techniques and reproducibility of the results should be points of particular importance in the discussion section. Authors are advised to check the correctness of their methods and results carefully before writing an article. Probable or dubious explanations should be avoided.
- Authors should not cite results submitted for publication in the reference section. Such results may be described briefly in the text with a note in parenthesis (submitted for publication by… authors, year).
- References. Each article should address, list, and discuss the entire spectrum of current publications relevant to its field. All cited references must provide sufficient and valid peer-reviewed results leading to clear and reliable conclusions.
- By following these instructions, Authors will facilitate a more rapid review and processing of their manuscripts and will provide the readers with concise and useful papers.

7. Following review and acceptance, a manuscript is examined in language and style, and galley proofs are rapidly prepared. Second proofs are not sent unless required.
8. Authors should correct their galley proofs very carefully and preferably twice. An additional correction by a colleague always proves to be useful. Particular attention should be paid to chemical formulas, mathematical equations, symbols, medical nomenclature etc. Any system of correction marks can be used in a clear manner, preferably in red. Additions or clarifications are allowed provided that they improve the presentation but do not bring new results (no fee).
9. Articles submitted to AR may be rejected without review if:

- they do not fall within the journal's policy.
- they do not follow the instructions for authors.
- language is unclear.
- results are not sufficient to support a final conclusion.
- results are not objectively based on valid experiments.
- they repeat results already published by the same or other authors before the submission to AR.
- plagiarism is detected by plagiarism screening services.

[Rejection rate (2022): 71%].

10. Authors who wish to prepare a review should contact the Managing Editor of the journal in order to get confirmation of interest in the particular topic of the review. The expression of interest by the Managing Editor does not necessarily imply acceptance of the review by the journal.
11. Authors may inquire information about the status of their manuscript(s) by sending an e-mail to **journals@iiar-anticancer.org**
12. Authors who wish to edit a special issue on a particular topic should contact the Managing Editor.

This text is a combination of advice and suggestions contributed by Editors, Authors, Readers, and the Managing Editor of AR.

doi • 10.5578/tt.202402913
Tuberk Toraks 2024;72(2):107-113
Received: 20.04.2024 • Accepted: 13.05.2024

RESEARCH ARTICLE

Investigation of age and smoking in NSCLC patients with uncommon EGFR mutations

Yosuke MAEZAWA[1] (ID)
Manato TAGUCHI[2] (ID)
Takeshi KAWAKAMI[2] (ID)
Toshihide INUI[3] (ID)
Shinichiro OKAUCHI[1] (ID)
Takeshi NUMATA[5] (ID)
Toshihiro SHIOZAWA[3] (ID)
Kunihiko MIYAZAKI[6] (ID)
Ryota NAKAMURA[5] (ID)
Kesato IGUCHI[1] (ID)
Takeo ENDO[5] (ID)
Tohru SAKAMOTO[4] (ID)
Hiroaki SATOH[1] (ID)
Nobuyuki HIZAWA[3] (ID)

[1] Divisions of Respiratory Medicine and Thoracic Surgery, Mito Medical Center, University of Tsukuba-Mito Kyodo General Hospital, Mito, Japan
[2] Division of Respiratory Medicine, Kobari General Hospital, Noda, Japan
[3] Division of Respiratory Medicine, University of Tsukuba Faculty of Medicine, Tsukuba, Japan
[4] Division of Respiratory Medicine, Tsukuba Memorial Hospital, Tsukuba, Japan
[5] Departments of Respiratory Medicine and Surgery, National Hospital Organization Mito Medical Center, Ibarakimachi, Japan
[6] Division of Respiratory Medicine, Ryugasaki Saiseikai Hospital, Ryugasaki, Japan

Cite this article as: *Maezawa Y, Taguchi M, Kawakami T, İnui T, Okauchi S, Numata T, et al. Investigation of age and smoking in NSCLC patients with uncommon EGFR mutations. Tuberk Toraks 2024;72(2):107-113.*

Address for Correspondence

Dr. Hiroaki SATOH

Divisions of Respiratory Medicine and Thoracic Surgery, Mito Medical Center, University of Tsukuba-Mito Kyodo General Hospital,

MITO-JAPAN

e-mail: hirosato@md.tsukuba.ac.jp

ABSTRACT

Investigation of age and smoking in NSCLC patients with uncommon EGFR mutations

Introduction: *In addition to the two common epidermal growth factor receptor (EGFR) mutations, there are many uncommon mutations. Due to the high number of uncommon types, as well as the rarity of patients, there is lack of information regarding patient demographics, especially age distribution and smoking status. Against this background, we conducted an analysis to clarify the background of patients with uncommon EGFR mutations, especially considering their age distribution and smoking status.*

Materials and Methods: *We retrospectively reviewed the medical records of non-small cell lung cancer (NSCLC) patients diagnosed in a multicenter clinical practice from 2002 to 2023. Patients included all cases of non-advanced and advanced NSCLC with uncommon EGFR mutations.*

Results: *Information on 158 patients with uncommon EGFR mutation was collected. Median age was 72 years, with the age distribution showing that most patients were in their 70s. There was a significant difference between the proportion of patients aged up to 59 years and the proportion aged 75 years or older. In 88 patients with a smoking habit history, a significant correlation was found between smoking index and age. Among non-smokers, there was a peak between ages 70 and 74, which was older than the peak among smokers.*

Conclusion: *Even in elderly patients and NSCLC patients with a history of smoking, although it is unclear whether EGFR mutation is common or uncommon, EGFR gene testing should be performed considering the possibility of these patients being EGFR-positive.*

Key words: *Epidermal growth factor receptor; non-small cell lung cancer; uncommon mutation; age*

ÖZ

Nadir görülen EGFR mutasyonları olan KHDAK hastalarında yaş ve sigara kullanımının araştırılması

Giriş: *İki yaygın epidermal büyüme faktörü reseptörü (EGFR) mutasyonuna ek olarak, pek çok mutasyon da vardır. Nadir görülen tiplerin çokluğu ve hastaların nadirliği nedeniyle, hasta demografik özellikleri, özellikle yaş dağılımı ve sigara içme durumu hakkında bilgi eksikliği bulunmaktadır. Bu arka plana dayanarak, nadir görülen EGFR mutasyonlarına sahip hastaların geçmişini, özellikle yaş dağılımlarını ve sigara içme durumlarını göz önünde bulundurarak açıklığa kavuşturmak için bir analiz gerçekleştirilmiştir.*

Materyal ve Metod: *2002'den 2023'e kadar çok merkezli bir klinik uygulamada teşhis edilen küçük hücreli dışı akciğer kanseri (KHDAK) hastalarının tıbbi kayıtlarını retrospektif olarak inceledik. Hastalar, yaygın olmayan EGFR mutasyonları olan tüm ileri evre olmayan ve ilerlemiş KHDAK vakalarını içeriyordu.*

Bulgular: *Yaygın olmayan EGFR mutasyonuna sahip 158 hasta hakkında bilgi toplandı. Ortalama yaş 72 idi ve yaş dağılımı çoğu hastanın 70'li yaşlarda olduğunu gösteriyordu. Elli dokuz yaşına kadar olan hastaların oranı ile 75 yaş ve üzeri hastaların oranı arasında anlamlı bir fark vardı. Sigara içme öyküsü olan 88 hastada sigara içme indeksi ile yaş arasında anlamlı ilişki saptandı. Sigara içmeyenler arasında 70 ile 74 yaşları arasında bir zirve vardı ve bu, sigara içenler arasındaki zirveden daha düşüktü.*

Sonuç: *Yaşlı hastalarda ve sigara içme öyküsü olan KHDAK hastalarında bile, EGFR mutasyonunun yaygın mı yoksa nadir mi olduğu belirsiz olsa da, bu hastaların EGFR pozitif olma olasılığı göz önünde bulundurularak EGFR gen testi yapılmalıdır.*

Anahtar kelimeler: *Epidermal büyüme faktörü reseptörü; küçük hücreli dışı akciğer kanseri; yaygın olmayan mutasyon; yaş*

INTRODUCTION

The epidermal growth factor receptor (EGFR) mutation was the first driver gene discovered for non-small cell lung cancer (NSCLC) (Attili) (John). Among EGFR mutations, *Ex19* deletion and *Exon 21 L858R* are two of the most common mutations, and there are multiple uncommon mutations (1,2). Among the uncommon mutations, *G719X*, *L861Q*, and *S768I* mutations are relatively frequent (1,2). Therefore, there have been many reports that treat these mutations collectively as major uncommon mutations (1,2). It is known that patients with these gene mutations respond to second-generation EGFR-tyrosine kinase inhibitors (TKIs), but patients with Exon 20 insertions do not respond to EGFR-TKIs (3,4). At present, the existence of patients with many compound mutations with common or uncommon EGFR mutations has been recognized (4). Not only are they rare, but they are also genetically heterogeneous populations, and their responses to therapeutic drugs are not the same. As such, there are not many studies investigating patient backgrounds, such as age and smoking, in detail (5-19). In particular, only a few studies have shown information on more than 100 patients with uncommon mutations (6-8,11,13,15,18).

In view of this, we conducted this study to clarify clinical characteristics, with particular focus on age and smoking history, of NSCLC patients with uncommon EGFR mutations.

MATERIALS and METHODS

The medical records of all NSCLC patients diagnosed at 14 medical institutions in our prefecture from July 2002 to December 2023 were examined. Based on the World Health Organization classification, the pathological diagnosis of each NSCLC patient was made (20). Before starting treatment, all patients underwent TNM classification using head computed tomography or magnetic resonance imaging, bone or positron emission scan, and abdominal ultrasound and/or computed tomography (21). At the time of NSCLC diagnosis, the following patient background characteristics were investigated: Sex, age, Eastern Cooperative Oncology Group performance status (PS), clinical stage, presence of EGFR mutation and EGFR mutation subtype. The 'number of cigarettes smoked per day' and 'years of smoking' were also investigated. The product of these indices was used as the smoking index (22,23).

For statistical analyses, the Chi-squared test was used to test for differences in proportions. The Mann-Whitney U test was used to compare values between two unmatched groups, such as patient age and smoking index. Correlations were examined using the Spearman correlation coefficient. A P-value less than 0.01 was considered to indicate a significant difference.

This study was approved by the Institutional Review Board of University of Tsukuba Mito Medical Center/ Mito Kyodo General Hospital (NO-23-53) and by each institute that participated in this study.

RESULTS

Characteristics of Patients

During the study period, clinical information on 158 patients with uncommon EGFR mutations was collected from 14 institutions. Median age of these patients was 72 years (range, 35-92 years), and there were 86 male and 72 female patients. There were 153 patients with adenocarcinoma and five patients with other histological types. The clinical stage was IA-IIIC in 83 patients, and IVA-B in 75 patients. With regard to PS, 137 patients had PS 0-1, and 21 patients had PS 2-4. There were 98 patients with major uncommon mutations, *G719X*, *L861Q*, and *S768I*, 41 with compound mutations, and 19 with Exon 20 insertions. Shows the age distribution of all patients Figure 1. The highest number of patients were in their 70s. There were 25 patients aged up to 59 years, and 47 patients aged 75 years or older. There was a significant difference between the proportion of patients aged up to 59 years and that aged 75 years and older (p= 0.0046).

Comparison Among Uncommon EGFR Mutations

A comparison of patient background factors was performed with patients with three major mutations, *G719X*, *L861Q*, and *S768I*, as Group 1, patients with compound mutations as Group 2, and patients with Exon 20 insertions as Group 3. Patient background factors among the three groups are shown in Table 1. There were no significant differences in age, sex histology, clinical stage, or PS among the three groups. In addition, we focused on smoking and compared the percentage of non-smokers, the percentage of light smokers (smoking index of 100 or less), and the smoking index, but there were no significant differences among the three groups.

Correlation and Comparison between Age and Smoking Index in the Three Groups Due to Uncommon EGFR mutations

Figure 2-A shows the correlation between smoking index and age in all 158 patients. There was no significant correlation between smoking index and age in these patients (Spearman's rank correlation co-efficient p= 0.9059, p= 0.009). Next, we investigated the correlation between smoking index and age among the 88 smokers. The results are shown in Figure 2-B. For smokers only, there was a significant correlation between smoking index and age (Spearman's rank correlation co-efficient p= 0.0002, p= 0.397).

Patients were divided into non-smokers and smokers, and their age distributions are shown in Figures 3-A and B. In both groups, 70 non-smokers and 88 smokers, the most modal value for age was in the 70s. For non-smokers, the peak was at ages 70-74,

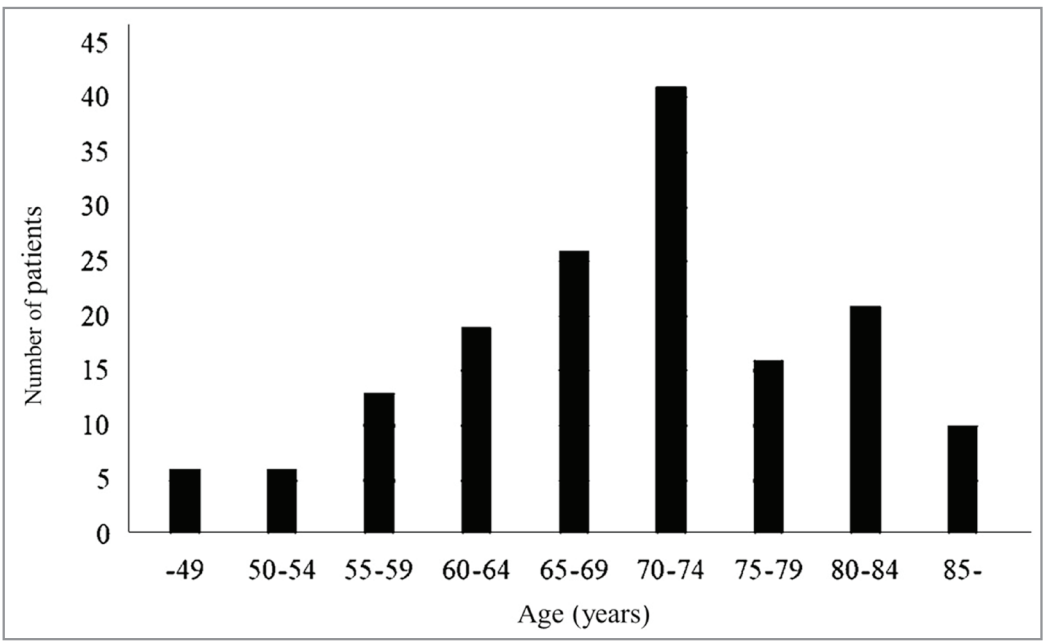


Figure 1. Age distribution of all 158 NSCLC patients with uncommon EGFR mutation (86 male patients and 72 female patients).

Table 1. Comparison of patient background factors among the three groups of patients: Group 1 (major mutations, *G719X*, *L861Q*, and *S768I*), Group 2 (compound mutations), and Group 3 (Exon 20 insertions)

	Group 1	Group 2	Group 3	p
Number of patients	98	41	19	
Age, median (range) years	72 (47-92)	68 (38-92)	73 (35-89)	0.2664
Sex				
Male	50 (51.0%)	23 (56.1%)	13 (68.4%)	0.3670
Female	48 (49.0%)	18 (43.9%)	6 (31.6%)	
Pathology				
AD	95 (96.9%)	39 (95.1%)	19 (100%)	0.6014
Others	3 (3.1%)	2 (4.9%)	0 (0%)	
Stage				
IA-IIIC	53 (54.1%)	20 (48.8%)	10 (52.6%)	0.8497
IVA-B	45 (45.9%)	21 (51.2%)	9 (47.4%)	
PS				
0-1	85 (86.7%)	36 (87.8%)	16 (84.2%)	0.9237
2-4	13 (13.3%)	5 (12.2%)	3 (15.8%)	
Non-smoker	45 (45.9%)	18 (43.9%)	7 (36.8%)	0.7653
Smoker	53 (54.1%)	23 (56.1%)	12 (63.2%)	
Non-light-smoker (SI< 100)	48 (49.0%)	19 (46.3%)	9 (47.4%)	0.8583
Smoker	50 (51.0%)	22 (53.7%)	10 (52.6%)	
SI, median (range)	130 (0-2400)	180 (0-1600)	150 (0-1500)	0.6227

AD: Adenocarcinoma, PS: Performance status, SI: Smoking index.

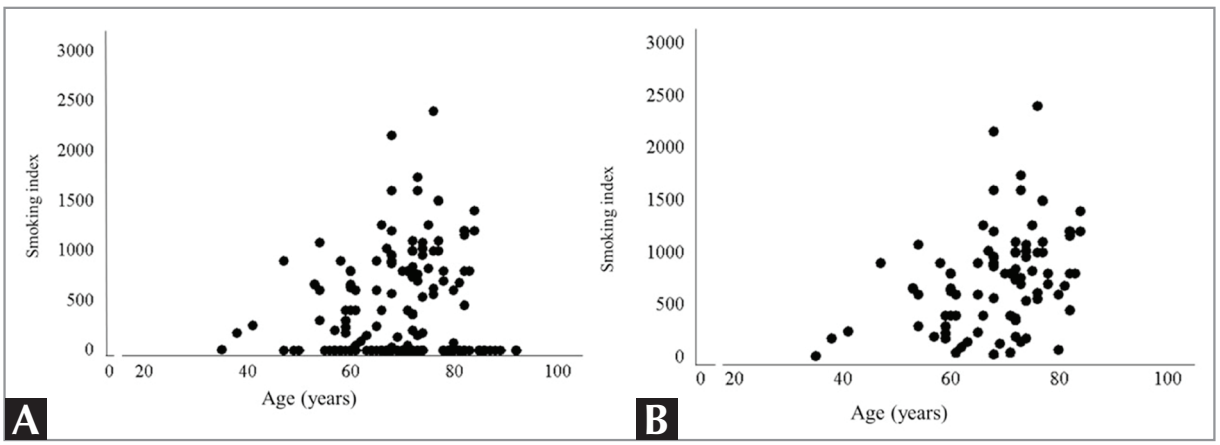


Figure 2. Correlation between smoking index and age in all 158 NSCLC patients **(A)**. There was no significant correlation between smoking index and age in these patients (Spearman's rank correlation co-efficient p= 0.9059, p= 0.009). Correlation between smoking index and age in 88 NSCLC patients with smoking habit **(B)**. There was a significant correlation between smoking index and age (Spearman's rank correlation co-efficient p= 0.0002, p= 0.397).

and for smokers, it was at ages 65-69. There was a significant difference in the age distribution of the non-smoker and smoker groups (p= 0.0192, Chi-squared test).

DISCUSSION

This study confirmed the following results: The median age of the 158 patients with EGFR uncommon mutations was 72 years. Regarding the age distribution of all patients, the most modal value for age was in the 70s. There was a significant difference between the proportion of the patients aged up to 59 years and the proportion of those aged 75 years or older. In 88 patients with a smoking habit, a significant correlation was found between the smoking index and age. Among non-smokers, there was a peak between ages 70 and 74, which was older than the peak among smokers.

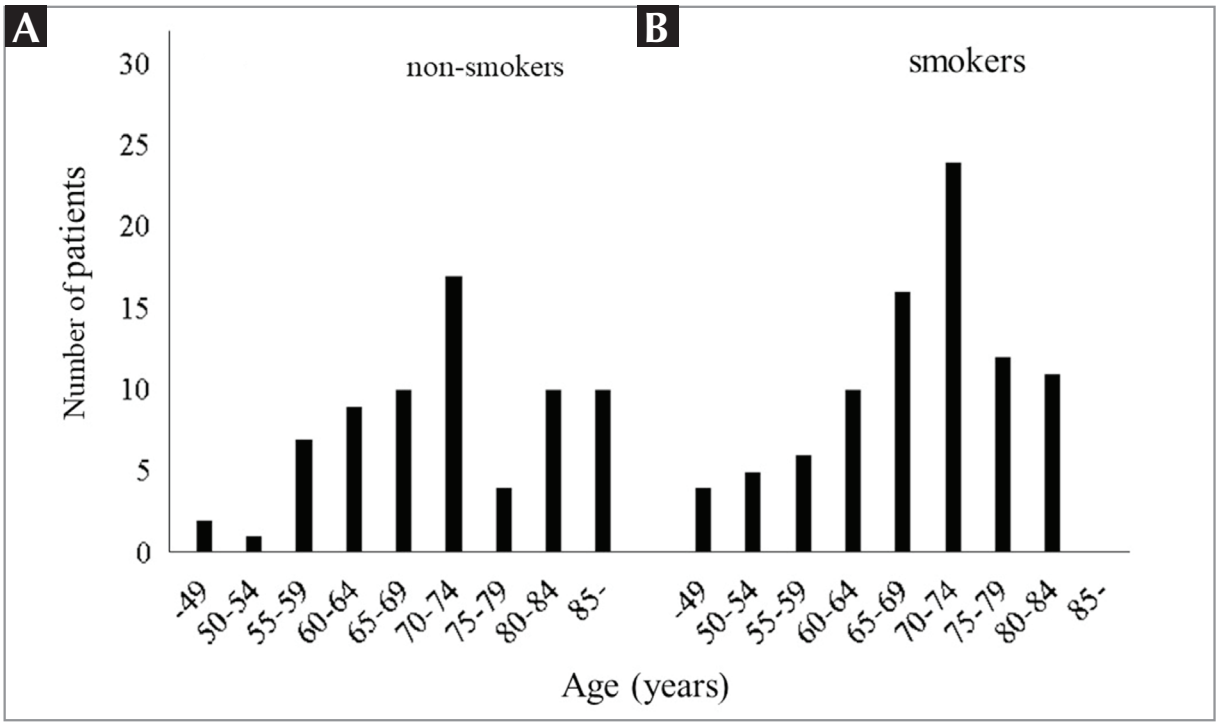


Figure 3. Age distribution of 70 NSCLC patients without smoking habit **(A)** and 88 NSCLC with smoking habit **(B)**. Significant difference in the age distribution between these two groups of patients (p= 0.0192, Chi-squared test).

Most studies to date on patients with uncommon EGFR mutations have focused on stage III-IV patients (5-7,10,12-17), and very few reports have included data on patients at all stages (9,19). Furthermore, the number of patients evaluated in these previous studies has been very small; 26 and 40 patients, respectively (9,19). In past surveys of patients with stage III-IV disease and uncommon EGFR mutations, the proportion of female patients was 37%-75% (5-7,9-19), and the proportion of patients with PS 0-1 was 60%-100% (6,9,10,12-17). The proportion of patients with adenocarcinoma was 16.7% in a study of 291 patients by Evans et al. (11), but other studies have generally reported proportions over 90% (5,7,10-17). In previous studies involving more than 100 patients, uncommon mutations have been classified into three groups: Major uncommon mutations, *G719X*, *L861Q*, and *S768I*; compound mutations; and Exon 20 insertions (6,7,11,13,15,18). Our study involved a relatively large number of patients, including patients at all stages of NSCLC from stage IA to stage IVB. In this survey, 42.4% were women, 86.7% were patients with PS 0-1, and 96.8% were adenocarcinoma patients. Focusing on uncommon mutation subtypes, 98 patients (62%) had major, 41 (25.9%) had compound, and 19 (12%) had Exon 20 insertions. It is known that the positive rate of EGFR mutations in NSCLC differs between Asians and Caucasians (24), and it was necessary to confirm these background factors. However, these results were not significantly different from previous studies (5-19).

In previously conducted EGFR-TKI clinical trials, median age of the patients with uncommon EGFR mutations was 58-64 years (6,12,14). In a recent TKI clinical trial of over 40 patients with uncommon EGFR mutations, median age has been found as 72 years (19). On the other hand, in most studies in clinical practice except one (11), median age has been found as 59-68 years (5,7-11,13,15-18). The exception is a study of 291 patients in the United Kingdom by Evans et al., in which the average age is 70.1 years (11). The results from clinical practice from Evans et al. and our study suggest that even patients older than 70 years might harbor uncommon EGFR mutations (11). Not conducting a search for driver genes in NSCLC due to advanced age should be avoided, as this might limit treatment options.

In studies conducted so far, the proportion of smokers among patients with uncommon EGFR mutations has been found as 44%-59% in clinical trials (12,14) and as 48%-69% in clinical practice, except for one study

from India involving 40 patients, which has shown a non-smoking rate of 83% (7,9,10,13,15-18). In the present study, the proportion of non-smokers was 44.3%. Although it has been reported that the proportion of smokers is higher in patients with uncommon EGFR mutations than in patients with common mutations, to the best of our knowledge, there have been no reports that have considered both smoking history and age (13). The results of this study show that in both groups, 70 smokers and 88 non-smokers, the most modal value for age was in the 70s. On the other hand, there was a significant difference in age distribution between the non-smoker and smoker groups. In other words, the number of patients increased up to the age of 74 in both non-smoking and smoking groups, and after that, the distribution showed a difference between the two groups. A significant correlation was found between age and smoking index in patients with smoking history. It has been speculated that this is not simply due to an increase in the number of years of smoking, but that there may be some other cause that remains unknown.

Although the above novel findings were obtained, this study has some limitations. We used several testing methods for EGFR mutations, but comparisons could not be made because the testing methods were not integrated due to the multicenter nature of the study, and there was no information on EGFR gene-negative patients who were treated around the same time. In addition, the study period was long because it was intended to collect a large number of patients. However, there are few reports that have investigated more than 150 patients, and we do believe that the information obtained might be useful for the future medical treatment of patients with uncommon EGFR mutations.

CONCLUSION

The implementation of driver gene testing for NSCLC is expected to provide important information for selecting treatment options tailored to the patient. Therefore, even in NSCLC patients who are elderly or who have a history of smoking, although it is unclear whether EGFR mutations are common or uncommon, EGFR gene testing should be performed in case an EGFR mutation is present.

Acknowledgment

We would like to thank the researchers at the following facilities for their cooperation in this study: Hitachi General Hospital, Hitachinaka General Hospital, Ibarakihigashi Hospital, Tsukuba Medica Center Hospital, Tsuchiura Kyodo General Hospital, Ibaraki Medical Center-Tokyo Medical University.

Ethical Committee Approval: This study was obtained from General Hospital Mito Kyodo Hospital Director of Hospital Ethics Committee (Decision no: 23-53, Date: 13.12.2023).

CONFLICT of INTEREST

The authors declare that they have no conflict of interest.

AUTHORSHIP CONTRIBUTIONS

Concept/Design: YM, SO, KM, HS

Analysis/Interpretation: YM, SO, KM, HS

Data acqusition: YM, MT, TK, TL, TN, TS, KM, RN, KI, TS

Writing: YM, KM, HS

Clinical Revision: HS

Final Approval: HS, NH

REFERENCES

1. *Attili I, Passaro A, Pisapia P, Malapelle U, de Marinis F. Uncommon EGFR compound mutations in non-small cell lung cancer (NSCLC): A systematic review of available evidence. Curr Oncol 2022; 29: 255-66. https://doi.org/10.3390/curroncol29010024*
2. *John T, Taylor A, Wang H, Eichinger C, Freeman C, Ahn MJ. Uncommon EGFR mutations in non-small-cell lung cancer: A systematic literature review of prevalence and clinical outcomes. Cancer Epidemiol 2022; 76: 102080. https://doi.org/10.1016/j.canep.2021.102080*
3. *Dorta-Suárez M, de Miguel M, Amor-Carro O, Calderón JM, González-Ortega M, Rodríguez-Abreu D. The state of the art of EGFR exon 20 insertions in non-small cell lung cancer: Diagnosis and future perspectives. Cancer Treat Rev 2023; 124: 102671. https://doi.org/10.1016/j.ctrv.2023.102671*
4. *Pretelli G, Spagnolo CC, Ciappina G, Santarpia M, Pasello G. Overview on therapeutic options in uncommon EGFR mutant non-small cell lung cancer (NSCLC): New lights for an unmet medical need. Int J Mol Sci 2023; 24: 8878. https://doi.org/10.3390/ijms24108878*
5. *Heigener DF, Schumann C, Sebastian M, Sadjadian P, Stehle I, Märten A, et al. Afatinib Compassionate Use Consortium (ACUC). Afatinib in non-small cell lung cancer harboring uncommon EGFR Mutations pretreated with reversible EGFR inhibitors. Oncologist 2015; 20: 1167-74. https://doi.org/10.1634/theoncologist.2015-0073*

6. *Yang JC, Sequist LV, Geater SL, Tsai CM, Mok TS, Schuler M, et al. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: A combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. Lancet Oncol 2015; 16: 830-8. https://doi.org/10.1016/S1470-2045(15)00026-1*

7. *Tu HY, Ke EE, Yang JJ, Sun YL, Yan HH, Zheng MY, et al. A comprehensive review of uncommon EGFR mutations in patients with non-small cell lung cancer. Lung Cancer 2017; 114: 96-102. https://doi.org/10.1016/j.lungcan.2017.11.005*

8. *Wu SG, Chang YL, Yu CJ, Yang PC, Shih JY. Lung adenocarcinoma patients of young age have lower EGFR mutation rate and poorer efficacy of EGFR tyrosine kinase inhibitors. ERJ Open Res 2017; 3: 00092-2016. https://doi.org/10.1183/23120541.00092-2016*

9. *Frega S, Lorenzi M, Fassan M, Indraccolo S, Calabrese F, Favaretto A, et al. Clinical features and treatment outcome of non-small cell lung cancer (NSCLC) patients with uncommon or complex epidermal growth factor receptor (EGFR) mutations. Oncotarget 2017; 8: 32626-38. https://doi.org/10.18632/oncotarget.15945*

10. *Chantharasamee J, Poungvarin N, Danchaivijitr P, Techawatanawanna S. Clinical outcome of treatment of metastatic non-small cell lung cancer in patients harboring uncommon EGFR mutation. BMC Cancer 2019; 19: 701. https://doi.org/10.1186/s12885-019-5913-9*

11. *Evans M, O'Sullivan B, Smith M, Hughes F, Mullis T, Trim N, et al. Large-scale EGFR mutation testing in clinical practice: Analysis of a series of 18,920 non-small cell lung cancer cases. Pathol Oncol Res 2019; 25: 1401-9. https://doi.org/10.1007/s12253-018-0460-2*

12. *Cho JH, Lim SH, An HJ, Kim KH, Park KU, Kang EJ, et al. Osimertinib for patients with non-small-cell lung cancer harboring uncommon EGFR mutations: A multicenter, open-label, phase II trial (KCSG-LU15-09). J Clin Oncol 2020; 38: 488-95. https://doi.org/10.1200/JCO.19.00931*

13. *Ko HW, Shie SS, Wang CW, Chiu CT, Wang CL, Yang TY, et al. Association of smoking status with non-small cell lung cancer patients harboring uncommon epidermal growth factor receptor mutation. Front Immunol 2022; 13: 1011092. https://doi.org/10.3389/fimmu.2022.1011092*

14. *Li HS, Yang GJ, Cai Y, Li JL, Xu HY, Zhang T, et al. Dacomitinib for advanced non-small cell lung cancer patients harboring major uncommon EGFR alterations: A dual-center, single-arm, ambispective cohort study in China. Front Pharmacol 2022; 13: 919652. https://doi.org/10.3389/fphar.2022.919652*

15. *Popat S, Hsia TC, Hung JY, Jung HA, Shih JY, Park CK, et al. Tyrosine kinase inhibitor activity in patients with NSCLC harboring uncommon EGFR mutations: A retrospective international cohort study (UpSwinG). Oncologist 2022; 27: 255-65. https://doi.org/10.1093/oncolo/oyac022*

16. *Ullas B, Shrinidhi N, Mansi S, Narayan S, Parveen J, Surender D, et al. All EGFR mutations are (not) created equal: Focus on uncommon EGFR mutations. J Cancer Res Clin Oncol 2023; 149: 1541-9. https://doi.org/10.1007/s00432-022-04033-x*

17. *Bar J, Peled N, Schokrpur S, Wolner M, Rotem O, Girard N, et al. Uncommon EGFR mutations: International case series on efficacy of osimertinib in real-life practice in first-line setting (UNICORN). J Thorac Oncol 2023; 18: 169-80. https://doi.org/10.1016/j.jtho.2022.10.004*

18. *Wang C, Zhao K, Hu S, Dong W, Gong Y, Xie C. Clinical outcomes of afatinib versus osimertinib in patients with non-small cell lung cancer with uncommon EGFR mutations: A pooled analysis. Oncologist 2023; 28: e397-e405. https://doi.org/10.1093/oncolo/oyad111*

19. *Okuma Y, Kubota K, Shimokawa M, Hashimoto K, Kawashima Y, Sakamoto T, et al. Tokyo Cooperative Oncology Group (TCOG). First-line osimertinib for previously untreated patients with NSCLC and uncommon EGFR mutations: The UNICORN phase 2 nonrandomized clinical trial. JAMA Oncol 2024; 10: 43-51. https://doi.org/10.1001/jamaoncol.2023.5013*

20. *WHO Classification of Tumours Editorial Board. World Health Organization Classification of Tumours, Thoracic Tumours 5th ed. IARC Press, 2021.*

21. *Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, et al. International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, Advisory Boards, and Participating Institutions, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee Advisory Boards and Participating Institutions The IASLC lung cancer staging project: Proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. J Thorac Oncol 2016; 11: 39-51. https://doi.org/10.1016/j.jtho.2015.09.009*

22. *Maeshima AM, Tochigi N, Tsuta K, Asamura H, Matsuno Y. Histological evaluation of the effect of smoking on peripheral small adenocarcinomas of the lung. J Thorac Oncol 2008; 3: 698-703. https://doi.org/10.1097/JTO.0b013e31817c60ae*

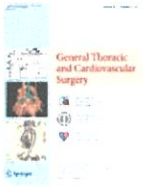
23. *Nakamura R, Inage Y, Tobita R, Yoneyama S, Numata T, Ota K, et al. Sarcopenia in resected NSCLC: Effect on postoperative outcomes. J Thorac Oncol 2018; 13: 895-903. https://doi.org/10.1016/j.jtho.2018.04.035*

24. *Batra U, Biswas B, Prabhash K, Krishna MV. Differential clinicopathological features, treatments and outcomes in patients with Exon 19 deletion and Exon 21 L858R EGFR mutation-positive adenocarcinoma non-small-cell lung cancer. BMJ Open Respir Res 2023; 10: e001492. https://doi.org/10.1136/bmjresp-2022-001492*

Serum C-reactive protein and procalcitonin levels in patients with pneumonia and anastomotic leakage in the postoperative period after esophagectomy

Original Article Published: 29 July 2024

Volume 72, pages 746–751, (2024) Cite this article



General Thoracic and Cardiovascular Surgery

Aims and scope

Submit manuscript

Hirotaka Ishida ✉, Toshiaki Fukutomi, Yusuke Taniyama, Chiaki Sato, Hiroshi Okamoto, Yohei Ozawa, Ryohei Ando, Yasuharu Shinozaki, Michiaki Unno & Takashi Kamei

Abstract

Objective

Despite being a less-invasive procedure, esophagectomy can cause severe infectious complications, such as pneumonia and anastomotic leakage. Herein, we aimed to clarify the inflammatory characteristics of pneumonia/anastomotic leakage after esophagectomy by assessing the difference between the postoperative trends of serum C-reactive protein (CRP) and procalcitonin (PCT) levels in patients with

pneumonia/anastomotic leakage using the values on the consecutive postoperative day (POD).

Methods

This study included 439 patients who underwent minimally invasive esophagectomy. Serum CRP and PCT levels were measured on PODs 1–7, 10, and 14. Pneumonia and anastomotic leakage were defined as Clavien–Dindo grades ≥ 2.

Results

Pneumonia and anastomotic leakage occurred in 96 and 51 patients, respectively. The CRP and PCT levels peaked on POD 3 (11.6 ± 6.8 mg/dL) and POD 2 (0.69 ± 2.9 ng/mL), respectively. Between PODs 3 and 14, CRP levels were significantly higher in patients with pneumonia and anastomotic leakage than in those without complications ($P < 0.001$). Between PODs 3 and 14, PCT levels were significantly higher in patients with pneumonia; however, on most PODs, there were no significant differences in PCT levels between patients with and without anastomotic leakage.

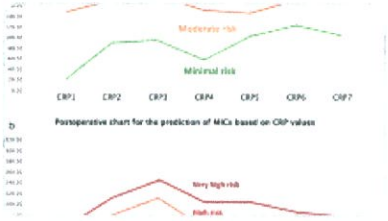
Conclusion

Inflammatory reactions caused by pneumonia may be more intense than those caused by anastomotic leakage after esophagectomy. Postoperative trends in serum CRP and PCT levels may vary depending on the complication type. Pneumonia and anastomotic leakage after esophagectomy can be potentially distinguished by the postoperative trend of PCT values before detailed examinations, such as computed tomography and endoscopy.

Similar content being viewed by others

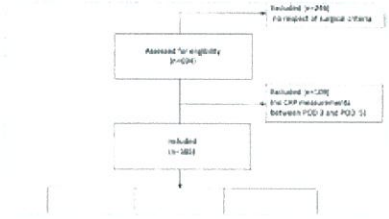
Utility of C-reactive protein as predictive biomarker of anastomotic leak afte...

Article 07 March 2018



Optimal Predictors of Postoperative Complications After Gastrectomy: Results...

Article 12 December 2022



C-reactive protein identifies patients at low risk of anastomotic leak after...

Article 08 October 2022

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Schlottmann F, Angeramo CA, Bras Harriott C, Casas MA, Herbella FAM, Patti MG. Transthoracic esophagectomy: hand-sewn versus side-to-side linear-stapled versus circular-stapled anastomosis: a systematic review and meta-analysis. Surg Laparosc Endosc Percutan Tech. 2022;32:380–92. https://doi.org/10.1097/SLE.0000000000001050.

Article PubMed Google Scholar

2. Kikuchi H, Endo H, Yamamoto H, Ozawa S, Miyata H, Kakeji Y, et al. Impact of reconstruction route on postoperative morbidity after esophagectomy: analysis of esophagectomies in the Japanese National clinical database. Ann Gastroenterol Surg. 2022;6:46–53. https://doi.org/10.1002/ags3.12501.

Article PubMed Google Scholar

3. Casas MA, Angeramo CA, Bras Harriott C, Schlottmann F. Surgical outcomes after totally minimally invasive Ivor Lewis esophagectomy. A systematic review and meta-analysis. Eur J Surg Oncol. 2022;48:473–81. https://doi.org/10.1016/j.ejso.2021.11.119.

Article PubMed Google Scholar

4. Motoyama S, Yamamoto H, Miyata H, Yano M, Yasuda T, Ohira M, et al. Impact of certification status of the institute and surgeon on short-term outcomes after surgery for thoracic esophageal cancer: evaluation using data on 16,752 patients from the National Clinical Database in Japan. Esophagus. 2020;17:41–9. https://doi.org/10.1007/s10388-019-00694-9.

Article PubMed Google Scholar

5. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. J Clin Invest. 2003;111:1805–12. https://doi.org/10.1172/JCI18921.

Article PubMed PubMed Central CAS Google Scholar

6. Lau DCW, Dhillon B, Yan H, Szmitko PE, Verma S. Adipokines: molecular links between obesity and atheroslcerosis. Am J Physiol Heart Circ Physiol. 2005;288:H2031–41. https://doi.org/10.1152/ajpheart.01058.2004.

Article PubMed CAS Google Scholar

7. Assicot M, Gendrel D, Carsin H, Raymond J, Guilbaud J, Bohuon C. High serum procalcitonin concentrations in patients with sepsis and infection. Lancet. 1993;341:515–8. https://doi.org/10.1016/0140-6736(93)90277-n.

Article PubMed PubMed Central CAS Google Scholar

8. Harada K, Matsumoto C, Toihata T, Kosumi K, Iwatsuki M, Baba Y, et al. C-reactive protein levels after esophagectomy are associated with increased surgical complications and poor prognosis in esophageal squamous cell carcinoma patients. Ann Surg Oncol. 2023;30:1554–63. https://doi.org/10.1245/s10434-022-12831-3.

Article PubMed Google Scholar

9. Richter F, Mehdorn AS, Fedders T, Reichert B, Egberts JH, Becker T, et al. C-reactive protein as predictor for infectious complications after robotic and open esophagectomies. J Clin Med. 2022;11:5654. https://doi.org/10.3390/jcm11195654.

Article PubMed PubMed Central CAS Google Scholar

10. Asti E, Bonitta G, Melloni M, Tornese S, Milito P, Sironi A, et al. Utility of C-reactive protein as predictive biomarker of anastomotic leak after minimally invasive esophagectomy. Langenbecks Arch Surg. 2018;403:235–44. https://doi.org/10.1007/s00423-018-1663-4.

Thorac Surg. 2013;95(4):1154–60 (**discussion 60–1**).

Article PubMed Google Scholar

22. Lerut T, Coosemans W, Decker G, De Leyn P, Nafteux P, van Raemdonck D. Anastomotic complications after esophagectomy. Dig Surg. 2002;19(2):92–8.

Article PubMed CAS Google Scholar

23. Bruce J, Krukowski ZH, Al-Khairy G, Russell EM, Park KG. Systematic review of the definition and measurement of anastomotic leak after gastrointestinal surgery. Br J Surg. 2001;88(9):1157–68.

Article PubMed CAS Google Scholar

24. van Heijl M, van Wijngaarden AK, Lagarde SM, Busch OR, van Lanschot JJ, van Berge Henegouwen MI. Intrathoracic manifestations of cervical anastomotic leaks after transhiatal and transthoracic oesophagectomy. Br J Surg. 2010;97:726–31. https://doi.org/10.1002/bjs.6971.

Article PubMed Google Scholar

25. Kubo N, Sakurai K, Tamura T, Toyokawa T, Tanaka H, Muguruma K, et al. The duration of systemic inflammatory response syndrome is a reliable indicator of long-term survival after curative esophagectomy for esophageal squamous cell carcinoma. Esophagus. 2021;18:548–58. https://doi.org/10.1007/s10388-021-00821-5.

Article PubMed Google Scholar

26. Gregoriano C, Heilmann E, Molitor A, Schuetz P. Role of procalcitonin use in the management of sepsis. J Thorac Dis. 2020;12(Suppl 1):S5-15. https://doi.org/10.21037/jtd.2019.11.63.

Article PubMed PubMed Central Google Scholar

27. Schuetz P, Birkhahn R, Sherwin R, Jones AE, Singer A, Kline JA, et al. Serial procalcitonin predicts mortality in severe sepsis patients: results from the multicenter procalcitonin MOnitoring SEpsis (MOSES) study. Crit Care Med. 2017;45:781–9. https://doi.org/10.1097/CCM.0000000000002321.

Article PubMed PubMed Central CAS Google Scholar

28. Wacker C, Prkno A, Brunkhorst FM, Schlattmann P. Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis. Lancet Infect Dis. 2013;13:426–35. https://doi.org/10.1016/S1473-3099(12)70323-7.

Article PubMed CAS Google Scholar

Acknowledgements

We would like to thank Editage (www.editage.jp) for English language editing.

Funding

None.

Author information

Authors and Affiliations

Department of Surgery, Tohoku University Graduate School of Medicine, 1-1, Seiryo-Machi, Aoba-ku, Sendai-Shi, Miyagi, Japan

Hirotaka Ishida, Toshiaki Fukutomi, Yusuke Taniyama, Chiaki Sato, Hiroshi Okamoto, Yohei Ozawa, Ryohei Ando, Yasuharu Shinozaki, Michiaki Unno & Takashi Kamei

Department of Surgery, National Hospital Organization Mito Medical Center, Ibaraki, Japan

Toshiaki Fukutomi

Corresponding author

Correspondence to Hirotaka Ishida.

Ethics declarations

Conflict of interest

The authors have no conflicts of interest.

Additional information

Publisher's Note

Supplementary Information

Below is the link to the electronic supplementary material.

Supplementary file1 (XLSX 11 KB)

Rights and permissions

About this article

Cite this article

Ishida, H., Fukutomi, T., Taniyama, Y. *et al.* Serum C-reactive protein and procalcitonin levels in patients with pneumonia and anastomotic leakage in the postoperative period after esophagectomy. *Gen Thorac Cardiovasc Surg* **72**, 746–751 (2024). https://doi.org/10.1007/s11748-024-02065-3

Received	Accepted	Published
18 June 2024	22 July 2024	29 July 2024

Issue Date
November 2024

DOI
https://doi.org/10.1007/s11748-024-02065-3

Keywords

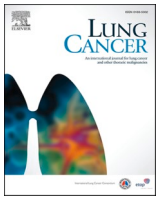
Anastomotic leakage | C-reactive protein | Esophagectomy | Pneumonia | Procalcitonin

Lung Cancer 195 (2024) 107917

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan



Mechanisms of resistance and correlation between pre-treatment co-alterations and p-prognosis to osimertinib in chemo-naïve advanced non-small cell lung cancer

Akihiro Tamiya [a], Mitsuo Osuga [b], Daijiro Harada [c], Shun-ichi Isa [d], Yoshihiko Taniguchi [a], Keiichi Nakamura [e], Yasuyuki Mizumori [f], Tsutomu Shinohara [g], Hidetoshi Yanai [h], Katsumi Nakatomi [i], Masahide Oki [j], Masahide Mori [k], Tomohito Kuwako [l], Koji Yamazaki [m], Atsuhisa Tamura [n], Masahiko Ando [o], Yasuhiro Koh [b,p,*]

[a] *Department of Internal Medicine, National Hospital Organization Kinki-Chuo Chest Medical Center, Sakai, Osaka, Japan*
[b] *Center for Biomedical Sciences, Wakayama Medical University, Wakayama, Japan*
[c] *Department of Thoracic Oncology and Medicine, National Hospital Organization Shikoku Cancer Center, Matsuyama, Ehime, Japan*
[d] *Clinical Research Center, National Hospital Organization Kinki-Chuo Chest Medical Center, Sakai, Osaka, Japan*
[e] *Department of Respiratory Medicine, National Hospital Organization Asahikawa Medical Center, Asahikawa, Hokkaido, Japan*
[f] *Department of Respiratory Medicine, National Hospital Organization Himeji Medical Center, Himeji, Hyogo, Japan*
[g] *Department of Respiratory Medicine, National Hospital Organization Kochi Hospital, Kochi, Japan*
[h] *Department of Respiratory Medicine, National Hospital Organization Mito Medical Center, Ibaraki, Japan*
[i] *Department of Respiratory Medicine, National Hospital Organization Ureshino Medical Center, Ureshino, Saga, Japan*
[j] *Department of Respiratory Medicine, National Hospital Organization Nagoya Medical Center, Nagoya, Aichi, Japan*
[k] *Department of Thoracic Oncology, National Hospital Organization Osaka Toneyama Medical Center, Toyonaka, Osaka, Japan*
[l] *Department of Respiratory Medicine, National Hospital Organization Shibukawa Medical Center, Shibukawa, Gunma, Japan*
[m] *Department of Thoracic Surgery, National Hospital Organization Kyushu Medical Center, Fukuoka, Kyushu, Japan*
[n] *Department of Respiratory Medicine, National Hospital Organization Tokyo National Hospital, Tokyo, Japan*
[o] *Department of Advanced Medicine and Clinical Research, Nagoya University Hospital, Nagoya, Japan*
[p] *Internal Medicine III, Wakayama Medical University, Wakayama, Japan*

ARTICLE INFO

Keywords:
Osimertinib
Non-small cell lung cancer
Epidermal growth factor receptor
Resistance mechanism
circulating tumour DNA

ABSTRACT

Background: Several patients treated with osimertinib experience progressive disease. The aim was to clarify the mechanisms underlying resistance to osimertinib.
Methods: ELUCIDATOR: A multi-centre, prospective, observational study involved chemotherapy-naive patients with advanced non-small cell lung cancer receiving osimertinib. Mutations in cancer-associated genes, detected via ultrasensitive next-generation sequencing of circulating tumour deoxyribonucleic acid samples, were collected at baseline and after progressive disease detection. These paired plasma samples were compared.
Results: Of 188 patients enrolled (May 2019-January 2021), 178 (119 females [67 %]) median age 74 years, were included. Patients, n = 95 (53 %) had epidermal growth factor receptor exon 19 deletion mutations. Among 115 patients with progressive disease, circulating tumour deoxyribonucleic acid levels of 85 patients were analysed. *MET* amplification (n = 4), *TP53* mutations (n = 4), *PIK3CA* mutations (n = 3), *BRINP3* mutation (n = 2), *BRAF* mutation (n = 2), *APC* mutation (n = 1), *RET* mutation (n = 1) and epidermal growth factor receptor (EGFR) resistance mutation, and *C797S* (n = 1) were detected. Patients with baseline *TP53* mutations, with *MET* or *EGFR* amplification had shorter progression-free (PFS) and overall survival. Patients with *PIK3CA* mutations tended to shorter PFS.

Abbreviations: BRINP3, bone morphogenetic protein/retinoic acid inducible neural-specific protein-3; ct, circulating tumour; CI, confidence interval; CT, computed tomography; EGFR, epidermal growth factor receptor; HR, hazard ratio; ICI, immune checkpoint inhibitor; JRCT, Japanese Register of Clinical Trials; MET, mesenchymal–epithelial transition; MRI, magnetic resonance imaging; NGS, next-generation sequencing; NR, not reached; NSCLC, non-small cell lung cancer; OS, overall survival; PET, positron emission tomography; PD, progressive disease; PD-L1, programmed death ligand-1; RFS, relapse-free survival; TKI, tyrosine kinase inhibitor; TPS, tumour proportion score. TNBC, triple-negative breast cancer; VAF, variant allele frequency.

* Corresponding author at: Center for Biomedical Sciences, Wakayama Medical University, 811-1 Kimiidera, Wakayama-shi 641-8509, Japan.
E-mail address: ykoh@wakayama-med.ac.jp (Y. Koh).

https://doi.org/10.1016/j.lungcan.2024.107917
Received 26 May 2024; Received in revised form 22 July 2024; Accepted 1 August 2024
Available online 3 August 2024

Conclusion: MET amplification and *PIK3CA* mutation mechanisms underly resistance to osimertinib in patients. Patients with coexisting mutations or amplifications at baseline had shorter PFS and overall survival.

1. Introduction

Lung cancer is the leading cause of cancer-related death worldwide. Non-small cell lung cancer (NSCLC) accounts for approximately 75 % of all cases of lung cancer [1]. Targeted therapies are being developed to improve the efficacy of treatment in selected patient populations with sensitising genetic mutations [2]. Epidermal growth factor receptor (*EGFR*) tyrosine kinase inhibitors (TKI) are the treatment for NSCLC harbouring *EGFR*-sensitising mutations owing to their superior efficacy [3,4,5,6]. Osimertinib, a third-generation, irreversible, oral EGFR-TKI, can potently and selectively inhibit *EGFR* harbouring *EGFR*-sensitising mutations and *EGFR T790M* resistance mutation [7]. Moreover, it is more effective at prolonging progression-free survival (PFS) and overall survival (OS) than first-generation EGFR-TKIs, such as gefitinib or erlotinib from the FLAURA study, a phase III trial [8,9]. Therefore, osimertinib is the standard treatment for NSCLC harbouring *EGFR*-sensitising mutations.

Previous studies have investigated the mechanisms underlying acquired resistance to osimertinib in patients with *EGFR T790M*-positive NSCLC following treatment with first- or second-generation EGFR-TKIs. The AURA3 study, a phase III trial that analysed circulating tumour DNA (ctDNA) samples and retrospective analyses, revealed that *EGFR* C797S mutation, amplification of mesenchymal–epithelial transition (*MET*), and human epidermal growth factor receptor 2 (HER2) are common resistance mechanisms [10,11,12].

Post-treatment strategies for cases after osimertinib failure remain to be established. *MET* amplification and *EGFR* C797S mutation were the most frequent resistance mechanisms detected using ctDNA and next-generation sequencing (NGS) in the FLAURA study [13]. In addition, Leonetti A, et al [14] and Choudhury NJ, et al [15] were also reported to the osimertinib resistance. While these studies are mainly tissue-based studies, our study is ct-DNA-based study. However, the number of the previous studies was small, therefore the data on the histological or molecular mechanisms underlying resistance to first-line osimertinib are still limited. Thus, the mechanisms underlying resistance to first-line osimertinib remain unknown. Further research is warranted to identify the mechanisms underlying this resistance to first-line osimertinib. Understanding the factors influencing progressive disease (PD) plays a crucial role in formulating treatment strategies for initial therapy and the subsequent strategy after osimertinib treatment. This primary report of the ELUCIDATOR study, focused on acquired resistance mechanisms and natural resistance factors against osimertinib present at baseline.

2. Methods

2.1. Study design and patients

The ELUCIDATOR study, a prospective observational study, was conducted across multiple medical centres of the National Hospital Organization Group in Japan. The details of the ELUCIDATOR study have been described in a previous report [16]. The ELUCIDATOR study primarily aimed to evaluate resistance-related mutations that affect the efficacy of osimertinib, a first-line therapy used for advanced NSCLC harbouring EGFR-sensitising mutations who had not previously received treatment. Patients who met the following criteria were eligible for inclusion in this study: a definitive diagnosis of non-squamous NSCLC confirmed via biopsy or cytology, presence of *EGFR* mutations (exon 19 deletion or exon 21-point mutation L858R), osimertinib administered as first-line therapy, and available blood specimens. Written informed consent was obtained from all patients. The study protocol was approved by the central review board and registered with the Japanese Register of Clinical Trials (JRCT; registration number: jRCTs031180051).

The ctDNA analyses presented herein were exploratory, prespecified, and prospective analyses of patients who received osimertinib as a first-line treatment. Plasma samples were collected from patients with advanced NSCLC harbouring *EGFR*-sensitising mutations at baseline and subsequent PD points and analysed to identify the acquired resistance mechanisms and potential primary resistance mechanisms to first-line osimertinib. All eligible patients who underwent NGS at baseline for the detection of plasma *EGFR* mutations were included in the analyses.

Collection of PD plasma samples included in the paired analysis was continued until December 2022. Clinical data collected until January 31, 2023, the data cut-off point, were analysed. Cases wherein events were recorded after January 31, 2023, were excluded from the analysis.

2.2. Plasma ctDNA analysis

Serial plasma samples were collected at baseline, 3 months, 12 months, and following the detection of PD. The analysis was conducted using paired plasma samples collected at baseline and after detecting PD until January 2023. The use of osimertinib after PD was permitted; however, samples from these patients were collected after the detection of PD. The patients were followed up for as long as possible without treatment if the administration of osimertinib was discontinued owing to the incidence of adverse events, and samples were collected after PD. The plasma ctDNA samples were analysed using NGS (AVENIO ctDNA Surveillance Kit; Roche Diagnostics, Indianapolis, IN, USA), which targeted somatic mutations in the whole exon and hotspot regions of 197 cancer-related genes. AVENIO ctDNA Analysis Software (Roche Diagnostics, Indianapolis, IN, USA) was used to perform variant calls. The limit of detection for the variant allelic fraction was 0.01 %, and the variant allelic fraction was set as $\geq$ 0.1 % to avoid the possibility of false positives.

2.3. Assessments

The Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 was used to assess PD [17]. Paired plasma samples of patients with and without detectable plasma *EGFR* mutations collected at baseline and PD points were compared to identify acquired resistance mechanisms. In addition, natural resistance mechanisms were identified using the plasma samples collected at baseline. The treatment period was defined as the duration from the administration of osimertinib until PD as determined by the attending physician. And PFS was defined as the duration between the administration of osimertinib and confirmation of disease progression or death. OS was defined as the duration between the administration of osimertinib and death.

2.4. Statistical analysis

The original statistical considerations of the ELUCIDATOR study have been described elsewhere [10]. Briefly, we calculated the incidence of acquired resistance in the present study. In addition, a one of secondary key analysis was performed, and the data were summarised using descriptive statistics. Briefly, we analyse the correlations between baseline resistance-related gene factors and disease progression with osimertinib.

3. Results

3.1. Patients characteristics

Among the 188 patients enrolled in the ELUCIDATOR study between May 2019 and January 2021, 10 patients were excluded (nine patients withdrew consent during the study and one patient did not meet the inclusion criteria). Consequently, plasma samples from 178 patients were analysed using NGS. Among them, 115 patients experienced PD before the data cut-off point. Thus, paired plasma samples were obtained from 85 patients (Fig. 1). The other thirty patients were unable to obtain ct-DNA from blood samples taken during PD. Analysis of resistance mechanisms during PD was performed on a total of 85 patients. And, when we conduct the prognosis analyses (PFS and OS) using baseline characteristic and genetic data, we included in all patients who could examine genetic mutations.

Of the 178 enrolled patients (median age: 74 [range 36–91] years), 119 (67 %) were female, 153 (86 %) had a performance status of 0 or 1, 95 (53 %) had an *EGFR* exon 19 deletion mutation, and 104 (58 %) were never-smokers (Table 1). Genetic analysis was conducted via NGS using the ctDNA samples of 167 patients collected at baseline. The genes of the remaining 11 patients could not be analysed as the amount of ctDNA was insufficient (Fig. 1). Among these 167 patients, *EGFR* mutations, *TP53* mutations, *EGFR* amplification, *MET* amplification, and *PIK3CA* mutations were detected in 112 (67 %), 63 (38 %), 42 (25 %), 18 (11 %), and seven (4 %) patients, respectively (Table 1). Compound mutations of *EGFR* were detected in 22 patients.

3.2. Acquired resistance mechanism of osimertinib and the incidence

Fig. 2 shows a pie chart depicting the genetic changes that induced acquired resistance. The following adaptive mutations or amplification were observed at PD: *MET* amplification and *TP53* mutations (n = 4, 4.7 %); *PIK3CA* mutations (n = 3, 3.5 %); *BRINP3* mutations (n = 2, 2.4 %); *BRAF* mutations (n = 2, 2.4 %); and *APC* mutation and *RET* mutation (n = 1, 1.2 %). Furthermore, an additional *EGFR* resistance mutation, *C797S*, was detected in one patient (1.2 %).

3.3. PFS and OS observed in the present study classified according to the EGFR mutation status

The median PFS and OS of all participants were 19.1 (95 % CI: 13.5–22.2) months and 36.0 (95 % CI: 30.4–not reached [NR]) months, respectively (Supplementary Fig. 1-a and 1-b). The PFS of the patients with *EGFR L858R* was significantly shorter than that of those with *EGFR exon 19 deletion* (median PFS: 17.2 [95 % CI: 12.0–19.5] months vs. 23.3 [95 % CI: 14.5–30.2] months, respectively; hazard ratio [HR]: 1.56, 95 % CI: 1.08–2.26; Supplementary Fig. 1-c). The OS of the patients with *EGFR L858R* tended to be shorter OS than that of those with an *EGFR exon 19 deletion* (median OS: 30.4 [95 % CI: 24.6–NR] months vs. 40.3 [95 % CI: 32.3–NR] months, respectively; HR: 1.54, 95 % CI: 0.98–2.45; Supplementary Fig. 1-d).

Table 1
Patient characteristics.

General information at baseline		Observation patients (n = 178)
Age:	Median (range)	74 (36–91) years
Sex:	Male/Female	59/119
Performance status:	0/1/2/3	68/85/21/4
Smoking status:	Current/Former/Never	8/66/104
Stage:	−IIIC/IVA/IVB	13/75/90
Activated EGFR mutation:	19del/L858R	95/83
Brain metastasis:	Positive/negative	48/130
Liver metastasis:	Positive/negative	17/161
ctDNA information at baseline		
Sensitizing EGFR mutation from ctDNA:	19del/L858R/Both/None/NE	54/57/1/55/11
TP53 mutation from ctDNA:	Positive/negative/NE	63/104/11
EGFR amplification from ctDNA:	Positive/negative/NE	42/125/11
MET amplification from ctDNA:	Positive/negative/NE	18/149/11
PIK3CA mutation from ctDNA:	Positive/negative/NE	7/160/11
EGFR compound mutations from ctDNA:	Positive/negative/NE	22/145/11

3.4. Impact of the detection of genetic mutations or amplifications at baseline on PFS and OS

Primary analysis revealed that patients with specific mutations and amplifications at baseline had shorter PFS and OS. The PFS and OS of the patients with *TP53* mutations in ctDNA were significantly shorter than those of the patients without *TP53* mutations in ctDNA (median PFS: 12.5 [95 % CI: 7.1–19.3] months and 22.3 [95 % CI: 17.3–NR] months, respectively; HR: 1.93, 95 % CI: 1.31–2.84; Figure 3-a) (median OS: 27.8 [95 % CI: 23.3–36.0] months vs. NR [95 % CI: 35.6–NR] months, respectively; HR: 2.02, 95 % CI: 1.26–3.25; Figure 3-b). The PFS and OS of the patients with *EGFR* amplification in ctDNA were significantly shorter than those of the patients without *EGFR* amplification in ctDNA (median PFS: 10.8 [95 % CI: 5.6–17.7] months vs. 20.0 [95 % CI: 17.5–26.6] months, respectively; HR: 1.84, 95 % CI: 1.20–2.85; Figure 3-c) (median OS: 25.2 [95 % CI: 17.6–40.9] months vs. NR [95 % CI: 32.2–NR] months, respectively; HR: 2.11, 95 % CI: 1.28–3.43; Figure 3-d). The PFS and OS of the patients with *MET* amplification in

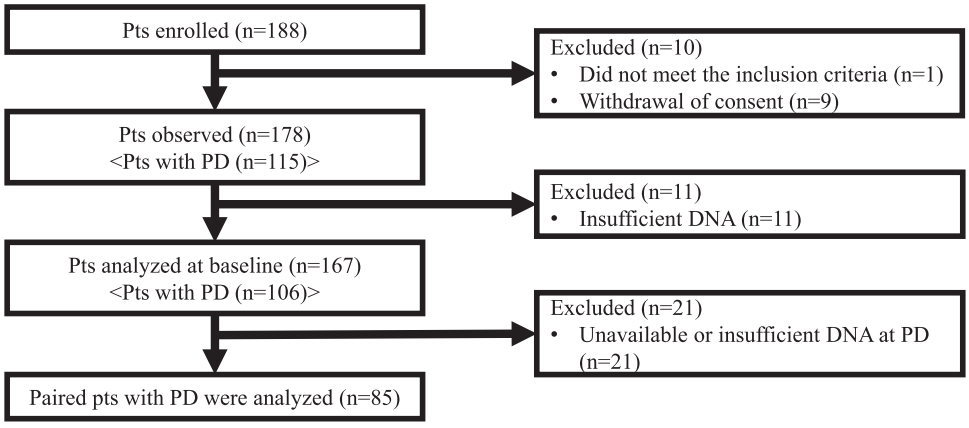


Fig. 1. Flow diagram of patient selection. In total, 188 patients were enrolled in the ELUCIDATOR study between May 2019 and January 2021. Pts: patients; PD: progressive disease.

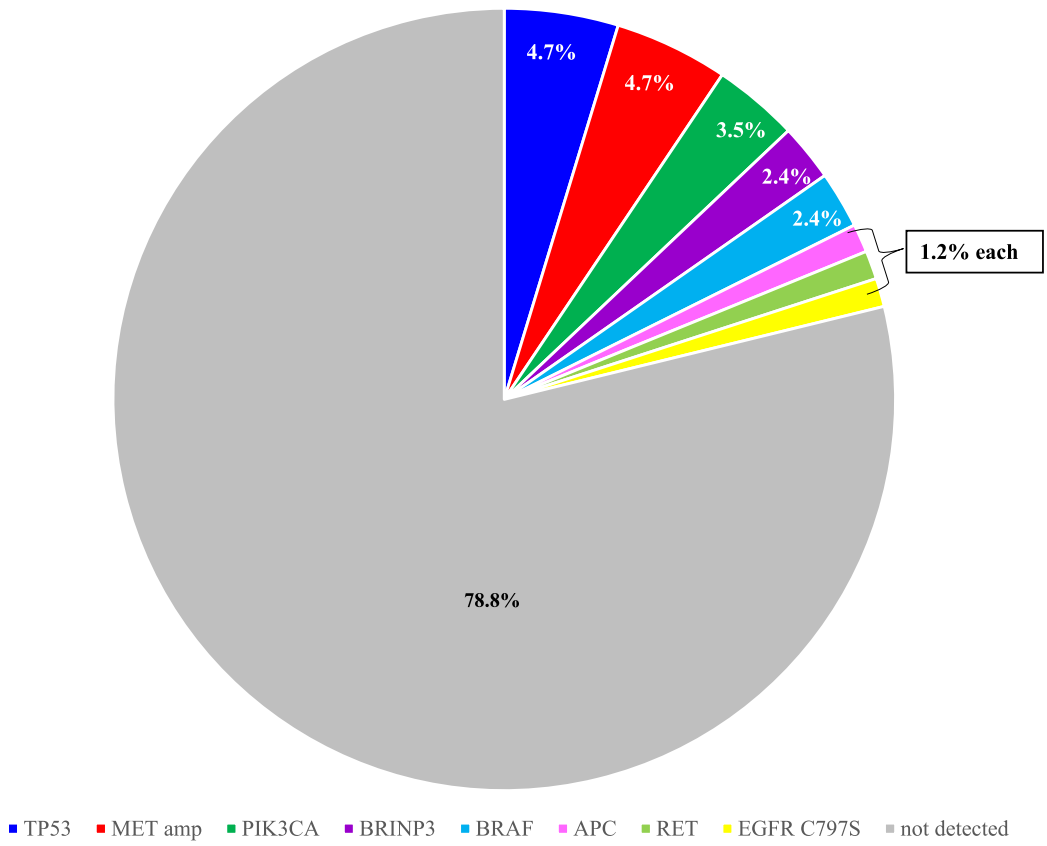


Fig. 2. Pie chart of acquired resistance mechanism (n = 85) EGFR: epidermal growth factor receptor.

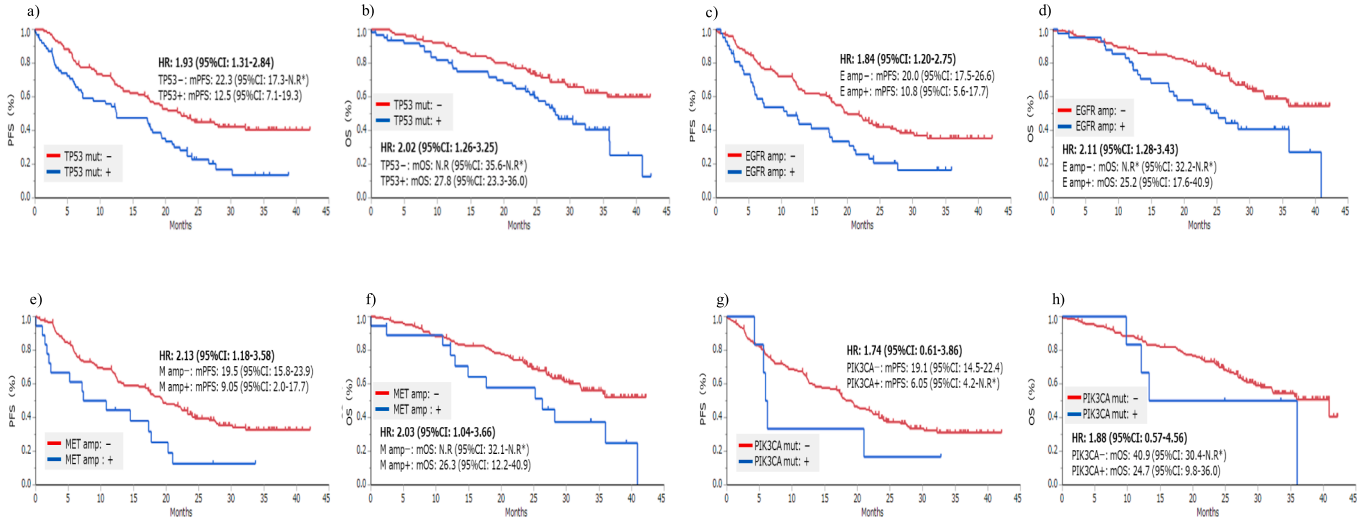


Fig. 3. Kaplan–Meier Curve of progression-free survival (PFS) and overall survival (OS) according to the gene mutations or amplifications at baseline. Comparison of the proportion of patients with *TP53* mutations and the patients without *TP53* mutations. (a) Kaplan–Meier curve of PFS. (b) Kaplan–Meier curve of overall survival (OS). Comparison of the proportion of patients with *EGFR* amplifications and the patients without *EGFR* amplifications. (c) Kaplan–Meier curve of PFS. (d) Kaplan–Meier curve of OS. Comparison of the proportion of patients with *MET* amplifications and the patients without *MET* amplifications. (e) Kaplan–Meier curve of PFS. (f) Kaplan–Meier curve of OS. Comparison of the proportion of patients with *PIK3CA* mutations and the patients without *PIK3CA* mutations. (g) Kaplan–Meier curve of PFS. (h) Kaplan–Meier curve of OS. EGFR: epidermal growth factor receptor HR: hazard ratio.

ctDNA were significantly shorter than those of the patients without *MET* amplification in ctDNA (median PFS: 9.1 [95 % CI: 2.0–17.7] months and 19.5 [95 % CI: 15.8–23.9] months, respectively; HR: 2.13, 95 % CI: 1.18–3.58; Figure 3-e) (median OS: 26.3 [95 % CI: 12.2–40.9) months and NR [95 % CI: 32.1–NR] months, respectively; HR: 2.03, 95 % CI: 1.04–3.66; Figure 3-f). The PFS of the patients with *PIK3CA* mutations tended to be shorter than that of the patients without *PIK3CA* mutations (median PFS: 6.1 [95 % CI: 4.2–NR] months vs. 19.1 [95 % CI: 14.5–22.4] months, respectively; HR: 1.74, 95 % CI: 0.61–3.86; Figure 3-g) (median OS: 24.7 [95 % CI: 9.8–36.0] months vs. 40.9 [95 %

CI: 30.4–NR] months, respectively; HR: 1.88, 95 % CI: 0.57–4.56; Figure 3-h).

4. Discussion

This primary analysis of the ELUCIDATOR study data provides insights into the mechanisms underlying acquired resistance and the natural resistance factors present at baseline among chemo-naïve patients with NSCLC who were treated with osimertinib. ctDNA analysis using NGS revealed that *TP53* mutations, *MET* amplification, and PIK3CA mutations were the most common acquired resistance mechanisms. In contrast, secondary *EGFR* point mutations, such as C797S, were less common. Furthermore, the findings of the present study suggest that the presence of natural *TP53* mutations, *EGFR* amplification, and *MET* amplification at baseline are poor prognostic factors.

The detection rate of adaptive resistance observed in the present study was lower than that reported in a previous study [13,14,15]. However, the trend of genetic variation was similar to that reported previously [13]. The incidence of acquired *EGFR* mutations and acquired amplifications reported by the FLAURA trial was higher than that reported in the present study [13]. The discrepancy between the detection rates reported by the two studies may be attributed to the limit of the variant allelic fraction. NGS (Guardant Health, Guardant 360 74 gene, or Guardant OMNI 500 gene) was performed using a limit of variant allelic fraction of 0.04–0.06 % in the FLAURA study. In contrast, NGS (AVENIO ctDNA Surveillance Kit; Roche Diagnostics, Indianapolis, IN, USA) was performed using a limit of detection for the variant allelic fraction of 0.01 %, however, the variant allelic fraction was set as $\geq$ 0.1 % to avoid false positives in the present study. In our study, a minimum of 10 ng of DNA was used for the analyses, and theoretically a detection sensitivity of 0.1 % can be maximally achievable without false positive events. Therefore, we set the cutoff at 0.1 %. Consequently, the detection rate of adaptive resistance may be lower. In addition, the number of cases wherein ctDNA could be analysed before and after the administration of osimertinib was lower than planned in the present study, leading to a reduction in the power of the analysis. This may be attributed to discontinuation of treatment owing to toxicities, such as interstitial lung disease and heart failure, which hindered post-treatment analysis. However, the present analysis of acquired resistance to osimertinib is one of the largest analyses to date. Moreover, it has contributed valuable information.

The present findings also suggest that genetic co-mutations at baseline can affect the therapeutic efficacy of osimertinib and emphasise the importance of spontaneous resistance at baseline. *MET* amplifications are associated with a poor response to EGFR-TKIs [18]. However, previous studies used *MET* fluorescence in situ hybridisation to analyse *MET* amplification, and the EGFR-TKIs did not include osimertinib. Moreover, the number of patients with *MET* amplification was only five; therefore, the results of these previous studies demonstrating a poor prognosis are underpowered. In contrast, NGS was performed using ctDNA to analyse the *MET* amplification in the present study, and the EGFR-TKI used was osimertinib, the recent first-line standard therapy. To the best of our knowledge, this is the first report to demonstrate that *MET* amplification at baseline is a poor prognostic factor for response to osimertinib in treatment-naïve patients with metastatic sensitising EGFR mutation-positive NSCLC. However, it should be considered that *MET* amplification and *EGFR* amplification are poor prognostic factors for osimertinib treatment while selecting the first-line strategy for EGFR mutation-positive NSCLC. Thus, other treatment strategies to adjust *MET* amplification and *EGFR* amplification may be more suitable for these patients. Tepotinib + gefitinib induced a moderate response in patients with *EGFR* mutation-positive NSCLC with MET overexpression or *MET* amplification and acquired resistance to previous *EGFR* inhibitors. The median PFS was 4.9 (90 %CI: 3.9–6.9) months and median OS was 17.3 (12.1–37.3) months [19].

Amivantamab is a human bispecific antibody that binds to the *EGFR* and *MET* receptor to inhibit ligand binding, promote downregulation of cell surface receptors, induce Fc-dependent trogocytosis, and promote antibody-dependent cellular cytotoxicity [20,21,22,23]. Amivantamab has shown anti-tumour activity across diverse EGFR-driven and MET-driven NSCLC [24,25]. A phase 1 trial of amivantamab + lazertinib (a third-generation *EGFR* inhibitor) revealed that amivantamab + lazertinib showed an overall response rate of 36 % (95 %CI: 22–51 %) in patients with osimertinib-relapsed *EGFR* mutation-positive NSCLC, with a median PFS of 4.9 months [26]. In addition, the amivantanab + lazertinib (third-generation *EGFR* inhibitors) combination in chemotherapy (MARIPOSA-2 study) yielded positive outcomes in a population with limited options after disease progression following treatment with osimertinib. The median PFS was 8.3 (95 %CI: 6.8–9.1) months and objective response rate was 63 % (95 %CI: 57–69 %) [27]. Moreover, amivantamab + lazertinib yielded more clinically meaningful improvements than osimeritnib as first-line drugs for the treatment of sensitising EGFR mutation-positive NSCLC (MARIPOSA study), with a median PFS of 23.7 (95 %CI: 19.1–27.7) months vs. 16.6 (14.8–18.5) months [28]. Furthermore, the HR of PFS was statistically larger (HR: 0.70, 95 %CI: 0.58–0.85).

Missense *TP53* mutations are associated with poor response to EGFR-TKIs [29,30]. *TP53* mutations, the most prevalent concurrent mutations, have been detected in 50 % of patients with *EGFR* mutation-positive NSCLC [30]. EGFR-TKIs used in previous studies did not include osimertinib. In contrast, osimertinib was used as the first-line treatment in the present study. To the best of our knowledge, this is one of the first study to demonstrate that *TP53* mutation is a prognostic factor indicating poor response to osimertinib in treatment-naïve patients with metastatic sensitising *EGFR* mutation-positive NSCLC. This factor should also be considered before selecting the first-line treatment strategy for *EGFR* mutation-positive NSCLC.

The addition of a VEGF inhibitor improves outcomes in patients with *TP53* mutations [31]. In the study, erlotinib + ramucirumab is an effective first-line treatment option for *EGFR* mutation-positive NSCLC with *TP53* mutations. Although the PFS of patients with a concurrent *TP53* mutation was shorter than that of the patients with TP53 wild-type tumours (12.25 vs. 19.35 months, respectively; HR 1.867; 95 % CI, 1.448–2.407), treatment with erlotinib + ramucirumab yielded superior PFS than treatment with erlotinib + placebo in patients with *TP53* mutations, with a median PFS of 15.2 months and 10.6 months, respectively (HR 0.54; 95 % CI, 0.37–0.79). On the other hands, the median PFS of patients with *TP53* wild-type was 20.8 months and 15.7 months for erlotinib + ramucirumab and erlotinib + placebo, respectively (HR 0.79; 95 % CI 0.55–1.12). Thus, erlotinib + ramucirumab was more effective in the treatment of *EGFR* mutation-positive NSCLC with *TP53* mutation than erlotinib monotherapy. However, treatment with erlotinib + ramucirumab also resulted in poorer prognosis compared with that observed in patients with *TP53* wild-type. Therefore, establishing a treatment strategy for EGFR mutations with TP53 co-mutations is necessary, as the present study demonstrated that the single use of osimertinib is not a satisfied solution.

This study has some limitations. First, plasma NGS analysis only focused on genomic alterations that can be detected in ctDNA. Consequently, non-genetic mechanisms of resistance, such as histological transformation, co-existence of different histological tissues, and alterations in protein expression (such as PD-L1 expression), could not be evaluated. For instance, the transformation to small cell lung cancer could not be pathologically confirmed [32,33]. Furthermore, changes between the mutation observed in tissue and plasma could not be compared. Liquid biopsy provides valuable information that aids in monitoring and identifying emerging resistance mechanisms; however, complementary tissue testing must be conducted in future studies to facilitate a complete histological diagnosis. Second, although ultrasensitive NGS with a detection sensitivity of $\geq$ 0.1 % was used to analyse ctDNA, it was not detectable in the plasma of some patients. Thus, this analysis is descriptive and there may be additional resistance

mechanisms that remain to be identified. Third, this study only included patients from Japan; thus, the results may not be generalisable to individuals of other ethnicities due to different detection rate of EGFR mutations and different proportion of EGFR L858R. However, the *TP53* mutation and MET amplification rates observed in this study did not differ from those observed in the FLAURA study [13]. Finally, this is a genetic mutation with a possible resistance mechanism that requires further basic investigation to confirm the resistance mechanism, such as TP53.

5. Conclusion

ctDNA analysis revealed acquired resistance to osimertinib in approximately 20 % of cases. *MET* amplification and *PIK3CA* mutation were the most common mechanisms; however, EGFR C797S was detected in one case. In addition, *TP53* mutations, *EGFR* amplification, and *MET* amplification detected at baseline (before osimertinib treatment) were identified as factors predicting poor response to osimertinib. The median PFS of these patients was only half of that of the patients without mutations or amplifications.

Tamiya A received honoraria from Eli Lilly, Ono Pharmaceutical, Chugai Pharmaceutical, Boehringer Ingelheim, AstraZeneca, Bristol-Myers Squibb, Amgen, Taiho Pharmaceutical, Kyowa Kirin, MSD, Takeda Pharmaceutical, Nihon-Kayaku, Novartis, Thermo Fischer, Amgen, Tsumura, Daiich-Sankyo and Merck BioFarma, and research funding from Daiichi-Sankyo, Beigene and AstraZeneca. Taniguchi Y received honoraria from Chugai Pharmaceutical, Ono Pharmaceutical, AstraZeneca, and MSD. Harada has received honoraria from Takeda Pharmaceutical, Eli Lilly, Chugai Pharmaceutical, AstraZeneca, Taiho Pharmaceutical, Ono Pharmaceutical, Bristol-Myers Squibb, Towa Pharmaceutical, and Boehringer Ingelheim. Oki M received honoraria from AMCO, AstraZeneca, Canon Medical Systems, Chugai Pharmaceutical, Fujifilm Toyama Chemical, Kaneka Medix, Merit Medical Japan, Novartis Pharma, Olympus and Sanofi, and research funding from AbbVie, AstraZeneca, Chugai Pharmaceutical, Fujifilm Toyama Chemical, GlaxoSmithKline, Janssen Pharmaceutical, MSD, Ono Pharmaceutical, Parxel International, Pfizer, Sanofi. Mori M received honoraria from AstraZeneca, Boehringer Ingelheim, MSD, Eli Lilly, Novartis, Chugai Pharmaceutical, Taiho Pharmaceutical, Kyowa-kirin, Ono Pharmaceutical, Otsuka, Nihon-kayaku, Pfizer, Daiici-Sankyo, Takeda Pharmaceutical, and Shionogi and research funding from Chugai Pharmaceutical, Ono Pharmaceutical, MSD, and Delt-fly. Koh Y received honoraria from Chugai Pharmaceutical, Guardant Health, Amgen, Takeda Pharmaceutical, and Tosoh Corporation and received consulting or advisory roles from Tosoh Corporation and research funding from Boehringer Ingelheim, AstraZeneca, Chugai Pharmaceutical, Tosoh Corporation, Daiichi Sankyo, Zeon Corporation, Amgen, and Takeda Pharmaceutical. The other co-authors received no honoraria or research funding.

6. Financial supports

This study was supported by AstraZeneca and the National Hospital Organization.

CRediT authorship contribution statement

Akihiro Tamiya: Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mitsuo Osuga:** Visualization, Validation, Software, Resources, Formal analysis, Data curation. **Daijiro Harada:** Writing – review & editing, Supervision, Project administration, Investigation. **Shun-ichi Isa:** Writing – review & editing, Supervision, Resources, Data curation, Conceptualization. **Yoshihiko Taniguchi:** Project administration, Investigation, Conceptualization. **Keiichi Nakamura:** Resources, Investigation. **Yasuyuki Mizumori:** Writing – review & editing, Supervision, Project administration, Investigation. **Tsutomu Shinohara:** Writing – review & editing, Project administration, Investigation. **Hidetoshi Yanai:** Writing – review & editing, Project administration, Investigation. **Katsumi Nakatomi:** Writing – review & editing, Project administration, Investigation. **Masahide Oki:** Writing – review & editing, Supervision, Project administration, Investigation. **Masahide Mori:** Writing – review & editing, Supervision, Project administration, Investigation. **Tomohito Kuwako:** Writing – review & editing, Project administration, Investigation. **Koji Yamazaki:** Writing – review & editing, Project administration, Investigation. **Atsuhisa Tamura:** Writing – review & editing, Project administration, Investigation. **Masahiko Ando:** Writing – review & editing, Supervision, Software, Methodology, Formal analysis, Data curation. **Yasuhiro Koh:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I want to thank all the participating patients, families, and care givers. In addition, I appreciated the writing editing check by Editage.

Trial Registration: Japanese Register of Clinical Trials registration number: jRCTs031180051. This study was supported by AstraZeneca (ESR-17-13245) and National Hospital Organization (H28-EBM-01).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lungcan.2024.107917.

References

[1] P.C. Hoffman, A.M. Mauer, E.E. Vokes, Lung cancer, Lancet 355 (2000) 479–485, https://doi.org/10.1016/S0140-6736(00)82038-3.
[2] R.S. Herbst, J.V. Heymach, S.M. Lippman, Lung cancer, N. Engl. J. Med. 359 (2008) 1367–1380, https://doi.org/10.1056/NEJMra0802714.
[3] M. Maemondo, A. Inoue, K. Kobayashi, et al., Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR, N Engl. J. Med. 362 (2010) 2380–2388.
[4] T. Mitsudomi, S. Morita, Y. Yatabe, et al., Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial, Lancet Oncol. 11 (2010) 121–128.
[5] C. Zhou, Y.L. Wu, G. Chen, et al., Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study, Lancet Oncol. 12 (2011) 735–742.
[6] R. Rosell, E. Carcereny, R. Gervais, et al., Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomized phase 3 trial, Lancet Oncol. 13 (2012) 239–246.
[7] D.A. Cross, S.E. Ashton, S. Ghiorghiu, C. Eberlein, C.A. Nebhan, P.J. Spitzler, et al., AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer, Cancer Discov. 4 (2014) 1046–1061, https://doi.org/10.1158/2159-8290.CD-14-0337.
[8] S.S. Ramalingam, J. Vansteenkiste, D. Planchard, B.C. Cho, J.E. Gray, et al., Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC, N. Engl. J. Med. 382 (2020) 41–50, https://doi.org/10.1056/NEJMoa1913662.
[9] J.C. Soria, Y. Ohe, J. Vansteenkiste, T. Reungwetwattana, B. Chewaskulyong, K. H. Lee, et al., Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer, N. Engl. J. Med. 378 (2018) 113–125, https://doi.org/10.1056/NEJMoa1713137.
[10] K.S. Thress, C.P. Paweletz, E. Felip, B.C. Cho, D. Stetson, B. Dougherty, et al., Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M, Nat. Med. 21 (2015) 560–562, https://doi.org/10.1038/nm.3854.
[11] X. Le, S. Puri, M.V. Negrao, M.B. Nilsson, J. Robichaux, T. Boyle, et al., Landscape of EGFR-dependent and -independent resistance mechanisms to osimertinib and

continuation therapy beyond progression in EGFR-mutant NSCLC, Clin. Cancer Res. 24 (2018) 6195–6203, https://doi.org/10.1158/1078-0432.CCR-18-1542.
[12] J. Chmielecki, T. Mok, Y.L. Wu, J.Y. Han, M.J. Ahn, S.S. Ramalingam, et al., Analysis of acquired resistance mechanisms to osimertinib in patients with EGFR-mutated advanced non-small cell lung cancer from the AURA3 trial, Nat. Commun. 14 (2023) 1071, https://doi.org/10.1038/s41467-023-35962-x.
[13] J. Chmielecki, J.E. Gray, Y. Cheng, Y. Ohe, F. Imamura, B.C. Cho, et al., Candidate mechanisms of acquired resistance to first-line osimertinib in EGFR-mutated advanced non-small cell lung cancer, Nat. Commun. 14 (2023) 1070, https://doi.org/10.1038/s41467-023-35961-y.
[14] A. Leonetti, M. Verze, R. Minari, F. Perrone, L. Gnetti, P. Bordi, et al., Resistance to Osimertinib in advanced EGFR-mutated NSCLC: a prospective study of molecular genotyping on tissue and liquid biopsy, Br J Cancer. 130 (2024) 135–142, https://doi.org/10.1038/s41416-023-02475-9.
[15] N.J. Choudhury, A. Marra, J.S.Y. Sui, J. Flynn, S.R. Yang, C.J. Falcon, et al., Molecular biomarkers of disease outcomes and mechanisms of acquired resistance to first-line Osimertinib in advanced EGFR-mutant lung cancers, J Thorac Oncol. 18 (2023) 463–475, https://doi.org/10.1016/j.jtho.2022.11.022.
[16] A. Tamiya, S.I. Isa, Y. Taniguchi, H. Nakagawa, S. Atagi, M. Ando, et al., Prospective observational study of treatment resistance-related gene screening using plasma circulating tumor DNA in third-generation EGFR-TKI osimertinib therapy (Elucidator), Clin Lung Cancer 22 (2021) e336–e341, https://doi.org/10.1016/j.cllc.2020.05.023.
[17] E.A. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, et al., New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (version 1.1), Eur. J. Cancer 45 (2009) 228–247, https://doi.org/10.1016/j.ejca.2008.10.026.
[18] G.G.Y. Lai, T.H. Lim, J. Lim, P.J. Liew, X.L. Kwang, R. Nahar, et al., Clonal MET amplification as a determinant of tyrosine kinase inhibitor resistance in epidermal growth factor receptor-mutant non-small-cell lung cancer, J. Clin. Oncol. 37 (2019) 876–884, https://doi.org/10.1200/JCO.18.00177.
[19] Y.L. Wu, Y. Cheng, J. Zhou, S. Lu, Y. Zhang, J. Zhao, et al., Tepotinib plus gefitinib in patients with EGFR-mutant non-small-cell lung cancer with MET overexpression or MET amplification and acquired resistance to previous EGFR inhibitor (INSIGHT study): an open-label, phase 1b/2, multicentre, randomised trial, Lancet Respir. Med. 8 (2020) 1132–1143, https://doi.org/10.1016/S2213-2600(20)30154-5.
[20] S.L. Moores, M.L. Chiu, B.S. Bushey, K. Chevalier, L. Luistro, K. Dorn, et al., A novel bispecific antibody targeting EGFR and cMet is effective against EGFR inhibitor-resistant lung tumors, Cancer Res. 76 (2016) 3942–3953, https://doi.org/10.1158/0008-5472.CAN-15-2833.
[21] J. Neijssen, R.M.F. Cardoso, K.M. Chevalier, L. Wiegman, T. Valerius, G. M. Anderson, et al., Discovery of amivantamab (JNJ-61186372), a bispecific antibody targeting EGFR and MET, J. Biol. Chem. 296 (2021) 100641, https://doi.org/10.1016/j.jbc.2021.100641.
[22] S. Vijayaraghavan, L. Lipfert, K. Chevalier, B.S. Bushey, B. Henley, R. Lenhart, et al., Amivantamab (JNJ-61186372), an Fc enhanced EGFR/cMet bispecific antibody, induces receptor downmodulation and antitumor activity by monocyte/macrophage trogocytosis, Mol. Cancer Ther. 19 (2020) 2044–2056, https://doi.org/10.1158/1535-7163.MCT-20-0071.
[23] J. Yun, S.H. Lee, S.Y. Kim, J.H. Kim, K.H. Pyo, C.W. Park, et al., Antitumor activity of amivantamab (JNJ-61186372), an EGFR-MET bispecific antibody, in diverse models of *EGFR* exon 20 insertion-driven NSCLC, Cancer Discov. 10 (2020) 1194–1209, https://doi.org/10.1158/2159-8290.CD-20-0116.
[24] E.B. Haura, B.C. Cho, J.S. Lee, J.Y. Han, K.H. Lee, R.E. Sanborn, et al., JNJ-61186372 (JNJ-372), an EGFR-cMet bispecific antibody, in EGFR-driven advanced non-small cell lung cancer (NSCLC), J. Clin. Oncol. 37 (2019) 9009, https://doi.org/10.1200/JCO.2019.37.15_suppl.9009.
[25] M. Krebs, A.I. Spira, B.C. Cho, B. Besse, J.W. Goldman, P.A. Janne, et al., Amivantamab in patients with NSCLC with *MET* exon 14 skipping mutation: updated results from the CHRYSALIS study, J. Clin. Oncol. 40 (2022) 9008, https://doi.org/10.1200/JCO.2022.40.16_suppl.9008.
[26] B.C. Cho, D.W. Kim, A.I. Spira, J.E. Gomez, E.B. Haura, S.W. Kim, et al., Amivantamab plus lazertinib in osimertinib-relapsed EGFR-mutant advanced non-small cell lung cancer: a phase 1 trial, Nat. Med. 29 (2023) 2577–2585, https://doi.org/10.1038/s41591-023-02554-7.
[27] A. Passaro, J. Wang, Y. Wang, S.H. Lee, B. Melosky, J.Y. Shih, et al., Amivantamab plus chemotherapy with and without lazertinib in EGFR-mutant advanced NSCLC after disease progression on osimertinib: primary results from the phase III MARIPOSA-2 study, pp. S0923–7534(23), pp. 04281–3, Ann. Oncol. 35 (77–90) (2024), https://doi.org/10.1016/j.annonc.2023.10.117.
[28] Cho, B.C., Felip, E., Spira, A.I., Girard, N., Lee, J.S., Lee, S.H., et al. Amivantamab plus lazertinib vs osimertinib as first-line treatment in patients with EGFR-mutated, advanced non-small cell lung cancer (NSCLC): primary results from MARIPOSA, a phase III, global, randomized, controlled trial. *ESMO*: LBA13; 2023.
[29] M. Canale, E. Petracci, A. Delmonte, E. Chiadini, C. Dazzi, M. Papi, et al., Impact of TP53 mutations on outcome in EGFR-mutated patients treated with first-line tyrosine kinase inhibitors, Clin. Cancer Res. 23 (2017) 2195–2202, https://doi.org/10.1158/1078-0432.CCR-16-0966.
[30] H. Hou, K. Qin, Y. Liang, C. Zhang, D. Liu, H. Jiang, et al., Concurrent TP53 mutations predict poor outcomes of EGFR-TKI treatments in Chinese patients with advanced NSCLC, Cancer Manag. Res. 11 (2019) 5665–5675, https://doi.org/10.2147/CMAR.S201513.
[31] M. Nishio, L. Paz-Ares, M. Reck, K. Nakagawa, E.B. Garon, S. Popat, et al., RELAY, Ramucirumab plus erlotinib (RAM+ERL) in untreated metastatic EGFR-mutant NSCLC (EGFR+ NSCLC): association between TP53 status and clinical outcome, Clin. Lung Cancer 24 (2023) 415–428, https://doi.org/10.1016/j.cllc.2023.02.010.
[32] X. Chu, Y. Li, Z. Zhu, A case of small cell lung cancer transformation from EGFR-mutant lung adenocarcinoma with primary resistance to gefitinib, J. Thorac. Oncol. 13 (2018) e211–e214, https://doi.org/10.1016/j.jtho.2018.05.022.
[33] J.S. Ham, S. Kim, H.K. Kim, S. Byeon, J.M. Sun, S.H. Lee, et al., Two cases of small cell lung cancer transformation from EGFR mutant adenocarcinoma during AZD9291 treatment, J. Thorac. Oncol. 11 (2016) e1–e4, https://doi.org/10.1016/j.jtho.2015.09.013.

CellPress

Med

Viewpoint

Asciminib: the next-generation bullet for first-line treatment of chronic myeloid leukemia

Chikashi Yoshida[1,*] and Tomoiku Takaku[2]

The standard of care for chronic myeloid leukemia (CML) involves tyrosine kinase inhibitors (TKIs), which suppress tyrosine kinase activity of BCR::ABL1. Hochhaus et al. reported that asciminib, a BCR::ABL1 inhibitor specifically targeting the ABL myristoyl pocket, showed superior efficacy and favorable safety compared with TKIs in the phase 3 ASC4FIRST trial in patients with newly diagnosed chronic-phase CML.[1]

Chronic myeloid leukemia (CML) is a hematological malignancy with an annual incidence of 1 to 2 cases per 100,000 persons. The disease is typically diagnosed in the chronic phase, with only an increase in peripheral blood leukocytes and platelets and few symptoms; in untreated cases, the disease progresses to an accelerated phase within 3 to 5 years, followed by a treatment-resistant, life-threatening blast phase. The main pathogenesis of CML involves the constitutive activation of tyrosine kinase caused by the BCR::ABL1 oncoprotein, translated from the *BCR::ABL1* fusion gene arising from chromosome 9 and 22 translocations (Philadelphia chromosome).[2] The prognosis associated with CML has improved markedly over the past 20 years, with a 10-year overall survival rate of 85%–90%, and life expectancy for most patients is comparable with that of healthy controls.[3,4] This favorable prognosis is due to the development of tyrosine kinase inhibitors (TKIs) that suppress tyrosine kinase activity of BCR::ABL1.[4,5] Imatinib (a first-generation TKI) and dasatinib, nilotinib, and bosutinib (second-generation TKIs) have been approved for first-line treatment, and ponatinib (a third-generation TKI) is used for second-line or subsequent treatment. In addition, the establishment of a treatment evaluation method based on minimal residual disease (MRD) also plays an important role in the prognosis. MRD in CML is assessed by measuring *BCR::ABL1* mRNA in peripheral blood by reverse-transcription quantitative polymerase chain reaction and is expressed on a standardized International Scale (IS).[5] However, several concerning issues still remained in this successful model of cancer treatment. The first is adverse events (AEs) associated with long-term TKI treatment, such as cardiovascular events.[3,5] Even minor AEs may persist for a long period of time, reducing a patient's quality of life. Second, there are economic issues for patients and society due to high drug costs.[3] Third, women should avoid pregnancy while taking TKIs because they are teratogenic.[5] To address these issues, discontinuation of TKIs has been attempted in patients showing a sustained deep molecular response (DMR), and approximately half of these patients were able to achieve so-called treatment-free remission (TFR).[5] However, only about a quarter of CML patients achieve TFR,[3] and most patients require long-term treatment. In addition, some patients are resistant or intolerant to all TKIs.[3] The development of new treatments with superior efficacy and long-term safety is therefore necessary.

Asciminib is a first-in-class, specific allosteric inhibitor specifically targeting the ABL myristoyl pocket and causes conformational changes induced by myristate binding to the N terminus of BCR::ABL1. A phase 1 trial of asciminib was conducted involving CML patients who were refractory or intolerant to at least two or more TKIs.[6] The maximum tolerated dose was not reached. Of the chronic-phase CML patients who showed hematologic relapse, 92% exhibited a complete hematological response, and 54% who did not show a complete cytogenetic response (CCyR) at the baseline achieved CCyR. A major molecular response (MMR; *BCR::ABL1* transcript level $\leq$ 0.1% on IS), an important surrogate for progression-free survival, was achieved by 12 months in 48% of evaluable patients. Responses were also observed in patients with the T315I mutation who were highly resistant to TKIs other than ponatinib. Asymptomatic elevated lipase levels and clinical pancreatitis were dose-limiting toxic effects. Subsequently, the phase 3 ASCEMBL trial was conducted to compare asciminib with bosutinib in patients with chronic-phase CML who were refractory or intolerant to two or more TKIs.[7] The MMR rate at week 24 was 25.5% with asciminib, versus 13.2% with bosutinib, meeting the primary objective of the study. Grade 3 or higher AEs occurred in 50.6% of asciminib patients and 60.5% of bosutinib patients, and AEs resulting in discontinuation occurred in 5.8% and 21.1% of patients,

[1]Department of Hematology, NHO Mito Medical Center, Ibaraki, Japan

[2]Department of Hematology, Saitama Medical University, Saitama, Japan

*Correspondence: c.yoshida@mitomedical.org

https://doi.org/10.1016/j.medj.2024.07.001

respectively. These results led to the approval of asciminib for patients with chronic-phase CML previously treated with two or more TKIs (also for patients with T315I mutation in the United States).

The results of the phase 3 pivotal ASC4-FIRST trial to investigate the efficacy and safety of asciminib compared with TKIs (imatinib, dasatinib, nilotinib, and bosutinib) in patients with newly diagnosed chronic-phase CML were recently published.[1] The trial was well designed, and before patients were randomized, investigators selected TKIs that they would take if they were assigned to the TKI treatment arm. The randomization was stratified by the prognostic score category and TKIs selected by investigators. This randomization method was used because each TKI has its own characteristic AEs, and TKI selection depends on patient background factors: age, cardiovascular risk, and comorbidities. A total of 201 patients (101 prerandomization-selected imatinib stratum and 100 prerandomization-selected second-generation TKI stratum) were assigned to receive asciminib, and 204 were assigned to receive investigator-selected TKIs (102: imatinib; 102: second-generation TKIs). The asciminib-treated patients received 80 mg once daily, which differs from the ASCEMBL study, in which 40 mg was administered twice daily.[7] The primary endpoints were MMR at week 48 for comparisons between asciminib and investigator-selected TKIs and between asciminib and imatinib in the prerandomization-selected imatinib stratum. The MMR rate at week 48 was superior in patients treated with asciminib versus those treated with investigator-selected TKIs (67.7% versus 49.0%, respectively; $p < 0.001$). In the prerandomization-selected imatinib stratum, 69.3% of patients treated with asciminib achieved MMR compared to 40.2% of patients treated with imatinib ($p < 0.001$). In the prerandomization-selected second-generation TKI stratum, the MMR rate was 66.0% in those receiving asciminib compared with 57.8% in those receiving second-generation TKIs, although the difference was not significant. The cumulative incidence of *BCR::ABL1* transcript level of $\leq$0.01% on IS (MR4) by week 48 was 34.0% with asciminib and 15.9% with investigator-selected TKIs. In the subset analysis, the cumulative incidence of MR4 was 37.0% with asciminib and 10.0% with imatinib in the prerandomization-selected imatinib stratum and 31.0% of patients treated with asciminib and 21.8% treated with second-generation TKIs in the prerandomization-selected second-generation TKI stratum. Focusing on the newly emerging *BCR::ABL1* mutations causing treatment resistance, eight asciminib-treated patients (4.0%) acquired mutations predominantly in or near the myristoyl pocket, and 4 investigator-selected TKI-treated patients (2.0%) acquired mutations predominantly in the phosphate-binding loop. Grade 3 or higher hematological AEs with asciminib were thrombocytopenia (13.0%), neutropenia (10.0%), and anemia (1.5%), which were fewer than with the investigator-selected TKIs. Nonhematological AEs of any grade observed in more than 10% of patients treated with asciminib were coronavirus disease 2019 (COVID-19), diarrhea, fatigue, headache, myalgia, rash, and increased lipase, but the incidence of grade 3 or higher was less than 3% for any AE. The incidence of arterial occlusive events was similar in both groups. This favorable efficacy and safety profile of asciminib translated into a lower permanent discontinuation rate (13.5%) compared with investigator-selected TKIs (30.3%).

Based on these results, asciminib is expected to be approved for the treatment of newly diagnosed chronic-phase CML patients. Will asciminib replace TKIs as a first-line treatment for CML patients? Imatinib has been in use for more than 20 years, and dasatinib and nilotinib for more than 10 years, and patients receiving them have demonstrated higher long-term survival rates. Their respective characteristic AE profiles are widely recognized, and most clinicians have extensive clinical experience to manage them. Furthermore, since generic imatinib is widely available and generic dasatinib has been approved in some countries,[4] asciminib use is expected to be disadvantaged because of the economic burden. Therefore, it is important to determine whether it exhibits superior long-term therapeutic efficacy and safety compared with TKIs. It is encouraging that no new AEs were reported in the long-term follow-up of the phase 1 study with asciminib.[8] Of concern is the occurrence of *BCR::ABL1* mutations around the myristoyl pocket in patients treated with asciminib, but second-generation TKIs are effective against these mutations, as used in some patients in ASC4FIRST.[1] Hypothetically, the combination or sequential administration of asciminib and TKIs may prevent the acquisition of mutations and further improve the prognosis. This combination therapy is being investigated in clinical trials. Another point of interest is whether the superior treatment efficacy of asciminib increases the TFR rate. There are currently several issues regarding TFR. First, only a small proportion of patients ultimately achieve TFR with TKI treatment.[3] Second, even patients who do achieve TFR require prolonged treatment before discontinuing TKIs. The European LeukemiaNet recommends an optimal treatment duration of 5 years before TKI discontinuation, based mainly on data with imatinib.[5] Second-generation TKIs have been shown to yield TFR rates comparable with imatinib involving a shorter treatment duration, but the optimal duration of treatment has not been established.[9] A third issue is the treatment strategy for second discontinuation in patients who have failed the first TFR. Because asciminib has been reported to promote an earlier and a deeper MR than

TKIs,[1] it is expected to induce TFR in more patients with a shorter treatment duration and be effective in those who have failed TFR with TKIs. Future studies with this promising drug may further improve the treatment strategy for CML, a successful model of cancer treatment.

DECLARATION OF INTERESTS

C.Y. received research funding from Bristol Myers Squibb. T.T. received honoraria from Novartis.

REFERENCES

1. Hochhaus, A., Wang, J., Kim, D.W., Kim, D.D.H., Mayer, J., Goh, Y.T., le Coutre, P., Takahashi, N., Kim, I., Etienne, G., et al. (2024). Asciminib in Newly Diagnosed Chronic Myeloid Leukemia. N. Engl. J. Med. *31*. Published online May 31, 2024. https://doi.org/10.1056/NEJMoa2400858.

2. Mughal, T.I., Radich, J.P., Deininger, M.W., Apperley, J.F., Hughes, T.P., Harrison, C.J., Gambacorti-Passerini, C., Saglio, G., Cortes, J., and Daley, G.Q. (2016). Chronic myeloid leukemia: reminiscences and dreams. Haematologica *101*, 541–558.

3. Senapati, J., Sasaki, K., Issa, G.C., Lipton, J.H., Radich, J.P., Jabbour, E., and Kantarjian, H.M. (2023). Management of chronic myeloid leukemia in 2023 - common ground and common sense. Blood Cancer J. *13*, 58.

4. Kantarjian, H., Branford, S., Breccia, M., Cortes, J., Haddad, F.G., Hochhaus, A., Hughes, T., Issa, G.C., Jabbour, E., Nicolini, F.E., et al. (2024). Are their new relevant therapeutic endpoints in the modern era of the BCR::ABL1 tyrosine kinase inhibitors in chronic myeloid leukemia? Leukemia *38*, 947–950.

5. Hochhaus, A., Baccarani, M., Silver, R.T., Schiffer, C., Apperley, J.F., Cervantes, F., Clark, R.E., Cortes, J.E., Deininger, M.W., Guilhot, F., et al. (2020). European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. Leukemia *34*, 966–984.

6. Hughes, T.P., Mauro, M.J., Cortes, J.E., Minami, H., Rea, D., DeAngelo, D.J., Breccia, M., Goh, Y.T., Talpaz, M., Hochhaus, A., et al. (2019). Asciminib in Chronic Myeloid Leukemia after ABL Kinase Inhibitor Failure. N. Engl. J. Med. *381*, 2315–2326.

7. Réa, D., Mauro, M.J., Boquimpani, C., Minami, Y., Lomaia, E., Voloshin, S., Turkina, A., Kim, D.W., Apperley, J.F., Abdo, A., et al. (2021). A phase 3, open-label, randomized study of asciminib, a STAMP inhibitor, vs bosutinib in CML after 2 or more prior TKIs. Blood *138*, 2031–2041.

8. Mauro, M.J., Hughes, T.P., Kim, D.W., Rea, D., Cortes, J.E., Hochhaus, A., Sasaki, K., Breccia, M., Talpaz, M., Ottmann, O., et al. (2023). Asciminib monotherapy in patients with CML-CP without BCR::ABL1 T315I mutations treated with at least two prior TKIs: 4-year phase 1 safety and efficacy results. Leukemia *37*, 1048–1059.

9. Yoshida, C., Yamaguchi, H., Doki, N., Murai, K., Iino, M., Hatta, Y., Onizuka, M., Yokose, N., Fujimaki, K., Hagihara, M., et al. (2023). Importance of TKI treatment duration in treatment-free remission of chronic myeloid leukemia: results of the D-FREE study. Int. J. Hematol. *117*, 694–705.

Vaccine 42 (2024) 126221

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Delayed peak antibody titers after the second dose of SARS-CoV-2 vaccine in solid organ transplant recipients: Prospective cohort study

Kohei Unagami [a,b,h], Mikiko Yoshikawa [o], Hiroto Egawa [c,*], Satoko Ohfuji [n], Yoichiro Natori [u], Rikako Oki [a,b], Tomomi Mori [d], Hidetoshi Hattori [e], Ayumi Ishiwatari [h], Taichi Kanzawa [f,h], Tomokazu Shimizu [a,f,h], Kazuya Omoto [f,h], Masashi Inui [f,h], Yuuki Masano [p], Takashi Ito [p], Daisuke Nakajima [q], Tetsuya Babazono [d], Toshio Takagi [f], Shinichi Nunoda [g], Yoshito Tomimaru [i], Ryoichi Imamura [j], Shigeru Miyagawa [k], Koichi Toda [k], Etsuro Hatano [p], Hiroshi Date [q], Miyaji Kyakuno [l], Shiro Takahara [m], Kenji Yuzawa [r], Naoki Tanimine [s], Hideki Ohdan [s], Hideki Ishida [a], Yoshio Hirota [t], Japan Solid Organ Transplantation COVID-19 Countermeasure Group

[a] *Department of Organ Transplant Medicine, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[b] *Department of Nephrology, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[c] *Department of Surgery, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[d] *Department of Diabetology and Metabolism, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[e] *Department of Cardiology, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[f] *Department of Urology, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[g] *Department of Therapeutic Strategy for Severe Heart Failure, Graduate School of Medicine, Tokyo Women's Medical University, Japan*
[h] *Department of Urology, Yochomachi Clinic, Tokyo, Japan*
[i] *Department of Gastroenterological Surgery, Osaka University, Graduate School of Medicine, Japan*
[j] *Department of Urology, Osaka University, Graduate School of Medicine, Japan*
[k] *Department of Cardiovascular Surgery, Osaka University, Graduate School of Medicine, Japan*
[l] *Department of Renal Transplantation, Takatsuki General Hospital, Osaka Metropolitan University, Osaka, Japan*
[m] *Department of Renal Transplantation, Kansai Medical Clinic, Osaka Metropolitan University, Osaka, Japan*
[n] *Department of Public Health, Graduate School of Medicine, Osaka Metropolitan University, Osaka, Japan*
[o] *Department of Organ Transplantation and General Surgery, Kyoto Prefectural University of Medicine, Graduate School of Medical Science, Kyoto, Japan*
[p] *Department of Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan*
[q] *Department of Thoracic Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan*
[r] *Department of Transplantation Surgery, National Hospital Organization Mito Medical Center, Ibaraki, Japan*
[s] *Department of Gastroenterological and Transplant Surgery, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan*
[t] *Clinical Epidemiology Research Center, Medical Co. LTA (SOUSEIKAI), Fukuoka, Japan*
[u] *Division of Infectious Disease, Department of Medicine, University of Miami Miller School of Medicine, Miami, Florida, United States.*

ARTICLE INFO

Keywords:
SARS-CoV-2 vaccine
Post-vaccination antibody
Solid organ transplantation
Immunosuppressive therapy
Prospective cohort study

ABSTRACT

Poor post-vaccination production of antibody against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a concern among solid organ transplant (SOT) recipients. Furthermore, the timing and kinetics of antibody titers after the second vaccine dose are unknown. We conducted a multicenter prospective observational study that included 614 SOT recipients: 460 kidney, 53 heart, 50 liver, 20 lung, and 31 simultaneous pancreas–kidney (SPK). The participants received two doses of the mRNA vaccine (Pfizer BNT162b2 or Moderna mRNA-1273), as indicated. Serum samples were collected before the first and second vaccinations and at 1, 3, and 6 months after the second vaccine dose, which were then assessed for SARS-CoV-2 antibodies. The overall seropositivity rate was 43% at 1 month after administration of the second vaccine dose; it gradually increased to 68% at 3 months after second dose administration and to 70% at 6 months. In addition, recipient of kidney, lung

Abbreviations: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SOT, solid organ transplant; SPK, simultaneous pancreas–kidney; S-IgG, anti-spike protein immunoglobulin G; N-IgG, anti-nucleocapsid protein immunoglobulin G; GMT, geometric mean antibody titer; MMF, mycophenolate mofetil.

* Corresponding author at: Department of Surgery, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan.
E-mail address: egawa@kuhp.kyoto-u.ac.jp (H. Egawa).

https://doi.org/10.1016/j.vaccine.2024.126221
Received 1 March 2024; Received in revised form 5 August 2024; Accepted 9 August 2024
Available online 24 August 2024

or SPK transplants had lower antibody titers at the 3- and 6-month time points than did the other recipients. SOT recipients acquired SARS-CoV-2 S-IgG antibodies slowly, and the peak titer differed significantly from that of the general population.

1. Introduction

Infections caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have resulted in a global pandemic. Although one of the most effective prevention strategies against the SARS-CoV-2 outbreak has been vaccination [1,2], post-vaccination antibody production in solid organ transplant (SOT) recipients often shows low reactivity [3–7]. In the general population, antibody titers peak at approximately 1 month after the second vaccination dose and gradually decline within 6 months [8]. In contrast, the antibody titer peak after vaccination is thought to be delayed in SOT recipients because of immunosuppressant-mediated inhibition of the B-cell lineage responsible for producing antibodies. Statistical evidence regarding post-vaccination antibody kinetics in SOT recipients has been unavailable, although it is essential for optimal timing of booster injections in this population.

The Japanese Society for Transplantation organized a multicenter study (across Tokyo, Osaka, and Kyoto) that included recipients of kidney, liver, heart, lung, or simultaneous pancreas–kidney (SPK) transplants to reveal long-term immunogenicity after SARS-CoV-2 vaccination in SOT patients. Here, we report the positivity and titer kinetics of SARS-CoV-2 spike antibody in SOT recipients in the 6 months after administration of the second dose of the mRNA vaccine.

2. Material and methods

This clinical trial was conducted from March 2021 through March 2022 and involved 631 patients who underwent kidney, heart, liver, lung, or SPK transplantation. Patients who had received their organ transplants <6 months previously were excluded. The study protocol was approved by the Institutional Research Ethics Committee (approval number 2021–0024) and was performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all patients.

This multicenter study included six hospitals in Tokyo, Osaka, and Kyoto as data collection centers. After enrollment, the participants received two doses of the mRNA vaccine (Pfizer BNT162b2 or Moderna mRNA-1273), as indicated by the respective manufacturer. Serum samples from each patient were collected before the first (pre1) and second (pre2) vaccinations and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccine dose.

2.1. Assessment of SARS-CoV-2 antibodies

To assess SARS-CoV-2 antibody levels, SARS-CoV-2 anti-spike protein immunoglobulin G (S-IgG) and anti-nucleocapsid protein immunoglobulin G (N-IgG) antibodies were measured at each visit (Elecsys, Roche, Basel, Switzerland). We defined positive antibody responses according to the manufacturer-suggested thresholds (>0.8 U/mL for S-IgG, >1.0 U/mL for N-IgG); high seropositivity was defined as an S-IgG antibody titer of >210 U/mL, according to a Food and Drug Administration statement [9].

2.2. Statistical analysis

The following outcomes were calculated for assessing the immunogenicity of SARS-CoV2 vaccine: geometric mean titer; mean fold rise; and the proportion of recipients who were seropositive according to manufacturer-suggested thresholds. For data processing, seronegative results were regarded as titers of 0.4 U/mL or less, and reciprocal antibody titers were handled after logarithmic transformation. Calculated values were converted back to the original scale by exponential transformation and shown as results. To account for potential confounders, the following stratified analyses were conducted: age (quartile: <45, 45–54, 55–64, or ≥ 65 years); gender (male or female); body mass index (<18.5, 18.5–24.9, or ≥ 25.0 kg/m^2); transplanted organ; vaccine type (Pfizer or Moderna); length of time since transplantation (<2, 2–5, 6–10, or ≥ 11 years), donor type (living or deceased); splenectomy (absent or present); induction immunosuppression (no use or use); creatinine level (normal, increased, or increased at least 2-fold above normal upper limits); and maintenance immunosuppression treatments (no use or use). The significance of fold rise within a category was assessed by using the Wilcoxon signed rank-sum test, whereas inter-category comparisons were made by using either the Wilcoxon rank-sum test or the Kruskal–Wallis test. In addition, chi-square, Fisher's exact, and Mantel extension tests for trend were performed when appropriate.

The independent effects of potential confounders on antibody induction at each time point were evaluated by using logistic regression. The models were constructed with seropositivity (≥0.8 U/mL) as the dependent variable and variables that were statistically significantly associated with seropositivity as explanatory variables. Odds ratios (OR) and 95% confidence intervals (95% CI) were calculated. All tests were two-sided, and a P value of <0.05 was considered statistically significant. All analyses were performed by using SAS software (version 9.4; SAS Institute, Cary, NC, USA).

2.3. Data collection

The collected variables included routine demographic variables such as age, sex, comorbidities, and transplant information. Comorbidities included, among others, chronic obstructive pulmonary disease, hypertension, diabetes mellitus, asthma, and cancer. Transplant information included organ type, the date, living or deceased donor, induction immunosuppressive therapy, and maintenance immunosuppressive therapy. All the data obtained were documented by the healthcare professionals who provided those services to patients.

3. Results

3.1. Baseline characteristics

In total, 631 patients were enrolled. Of these, 28 who had infection or positive for N-IgG antibody (or both) within the test period, or did not provide data of antibody titer at any points, were excluded. Although a total of 614 recipients were analyzed (460 kidney, 53 heart, 50 liver, 20 lung, and 31 SPK), the final number of recipients that met the inclusion criteria was 603 (Fig. 1). The patient characteristics are given in Table S1. The mean age of the whole group was 53.1 years; heart transplant recipients were young compared with the other groups (54.5 kidney, 44.8 heart, 50.0 liver, 49.9 lung, and 52.9 years SPK, $P < 0.01$). There were 366 males (60%) and 246 females (40%) across all the groups, while females were more common in the SPK group (39% male, 61% female) ($P < 0.01$). For body mass index, only the lung transplant group was small (21.9 total, 22.2 kidney, 20.6 heart, 21.5 liver, 18.6 lung, and 21.5 kg/m^2 SPK, $P < 0.01$). The duration after organ transplantation was short for the lung group but long for the heart group (8.2 total, 8.4 kidney, 10.3 heart, 7.1 liver, 3.6 lung, and 6.0 years SPK, $P < 0.01$). In terms of induction immunosuppression, anti-thymocyte globulin use was prevalent in the SPK group (3% total, 1% kidney, 4% heart, 0% liver, 0% lung, and 45% SPK, $P < 0.01$). Basiliximab was used in the kidney, liver, and lung groups, especially in the kidney group (72% total,

91% kidney, 19% heart, and 55% SPK, $P < 0.01$). Rituximab was only used in the kidney and liver groups (27% total, 36% kidney, and 6% liver, $P < 0.01$). For maintenance immunosuppression, calcineurin inhibitors were common in all SOT groups (97% total, 97% kidney, 96% heart, 98% liver, 95% lung, and 100% SPK, $P = 0.33$). Steroid use was less common in the heart and liver groups (81% total, 90% kidney, 34% heart, 30% liver, 95% lung, and 94% SPK, $P < 0.01$), MMF use was less common in the heart group (86% total, 86% kidney, 55% heart, 93% liver, 100% lung, and 97% SPK, $P < 0.01$), while everolimus use was less common in kidney, lung, and SPK transplant recipients (29% total, 25% kidney, 72% heart, 45% liver, 5% lung, and 6% SPK, $P < 0.01$).

3.2. Antibody responses after the second vaccine dose

Among all SOT recipients, the S-IgG antibody titer gradually increased after the two vaccine doses and peaked at six months after the second dose (Table S2, Fig. 2; geometric mean antibody titer (GMT): 0.40 at pre1; 0.54 at pre2; 2.91 at 1 M; 11.71 at 3 M; and 12.73 U/mL at 6 M). This same trend—i.e., peak titer at 3 to 6 months after the second of vaccination—was present within each organ transplant group. However, recipients of kidney, lung, or SPK transplants had lower peak titers than those in other organ transplant groups.

The SARS-CoV-2 antibody seropositivity rate was 43% at 1 month; it gradually increased to 68% and 70% at 3 and 6 months, respectively, after the second vaccine dose (Table S3). This trend was paralleled in the organ transplant subgroups (Table S3. Fig. 3A). Furthermore, the percentage of recipients with high antibody seropositivity (S-IgG antibody titer >210 U/mL) demonstrated a trend similar to the antibody titer peak at 3 to 6 months after the second vaccination (Fig. 3B). However, recipients of kidney, lung, or SPK transplants had poorer rates of antibody seropositivity than other organ transplant recipients.

3.3. Associations between antibody responses and risk factors

The associations between SARS-CoV-2 antibody responses and other factors are shown in Table S2.

First, classification of SOT recipients according to age (<45, 45–54, 55–64, and ≥ 65 years) revealed lower antibody titers in older age groups than in younger ones (Table S2, Fig. 4A). The antibody seropositivity rate showed a similar trend (Table S3, Fig. 4B, C).

Secondly, when SOT recipients were subdivided into groups according to the period between transplantation and vaccination (<2, 2–5, 6–10, and ≥ 11 years), patients with shorter periods between the two events had significantly poorer antibody titers than those with longer time periods (Table S2, Fig. 5A). Again, similar trends emerged for antibody seropositivity rate (Fig. 5B, C). In addition, high serum creatinine levels were associated with poorer antibody titers and seropositivity rates (Table S2 and S3, Fig. 6A, B, C).

We then performed a stratified analysis according to the induction immunosuppression regimen at the time of organ transplantation—such as anti-thymocyte globulin, basiliximab, rituximab, or anti-lymphocyte globulin—and the use of maintenance immunosuppression—such as calcineurin inhibitors, steroid, mycophenolate mofetil (MMF) and everolimus. These analyses revealed that the use of basiliximab or rituximab as induction immunosuppression was associated with lower antibody titers and seropositivity rates than when these drugs were not provided as part of the induction regimen (Table S2 and S3, Fig. 7A, B, and 8A, B). Finally, patients whose maintenance immune suppression included steroid or MMF had poorer antibody titer elevation and seropositivity than SOT recipients who received other regimens (Table S2 and S3, Fig. 9A, B, and 10A, B). Furthermore, the steroid- and MMF-induced decreases in antibody responses were dose dependent. In contrast, using of calcineurin inhibitors or everolimus was not associated with poor antibody responses. Logistic regression analysis revealed significant associations between older age, shorter period between transplantation and vaccination, elevated creatinine elevation, and steroid or MMF with lower odds ratios for seropositivity rate after the second vaccine dose (Table S4).

3.4. Seropositivity rate of each transplanted organ adjusted for risk factors

Regression analyses of the seropositivity rate adjusted for risk characteristics are shown in Table 1. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. Model 1 included age categories, sex, duration from transplantation, and serum creatinine levels at the time of transplantation. Model 2 included basiliximab and rituximab usage at the time of transplantation, and steroid,

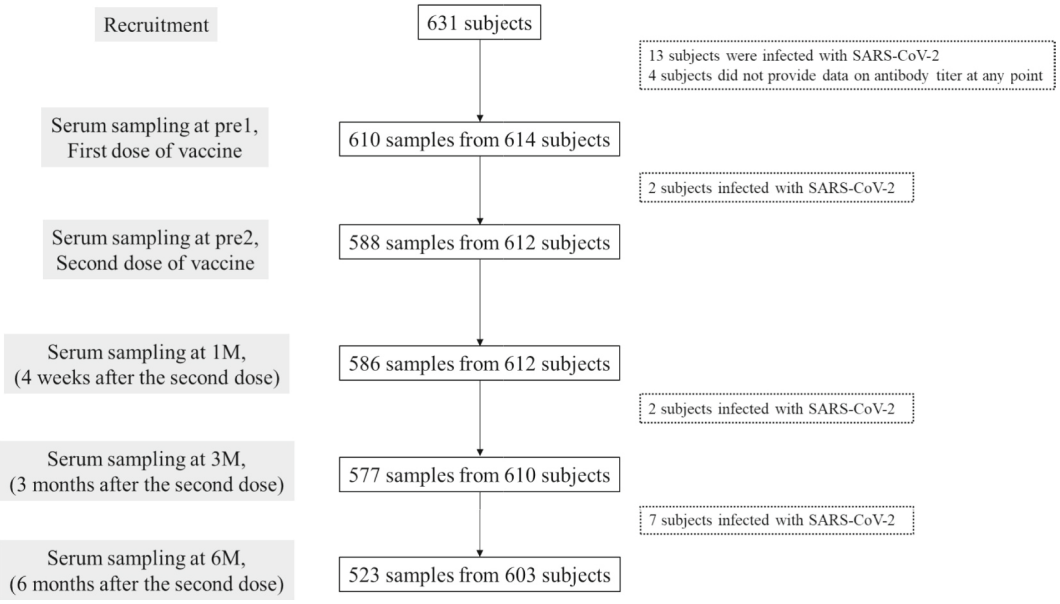


Fig. 1. Flow diagram showing the number of participants and samples included throughout the stages of the study. A total of 631 patients were enrolled in this study. Excluded patients included those who tested positive for SARS-CoV-2 infection or N-IgG antibody (or both) within the test period, as well as those who were missing data on antibody titer at any points. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

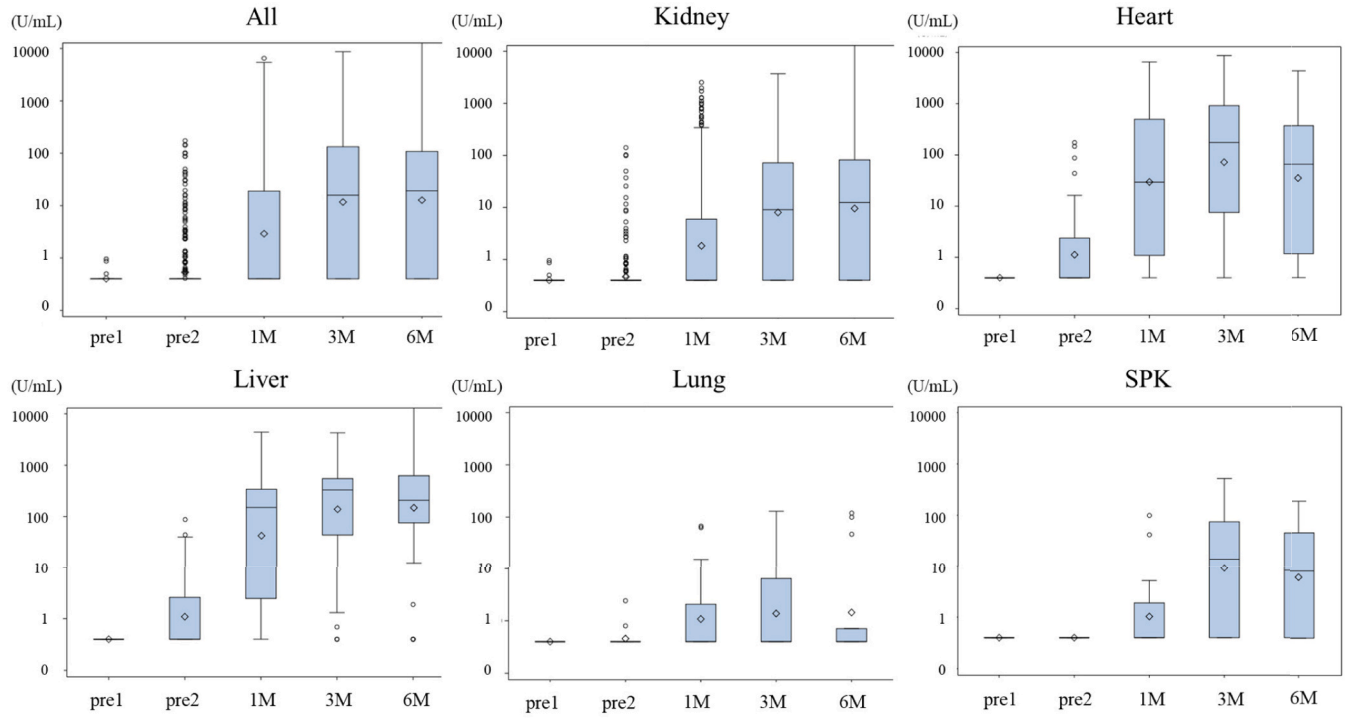


Fig. 2. SARS-CoV-2 antibody kinetics according to the transplanted organ. The test periods were as follows: before the first (pre1) and the second (pre2) dose of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; All, all types of organ transplantation; SPK, simultaneous pancreas–kidney.

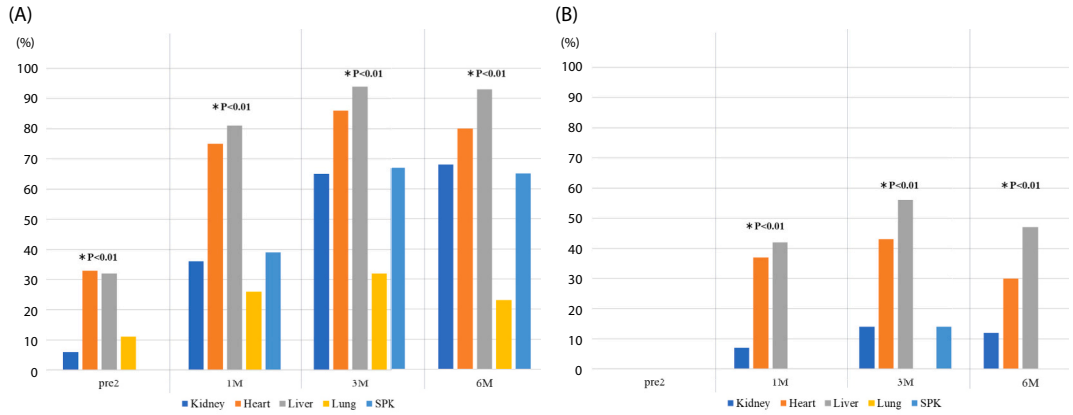


Fig. 3. SARS-CoV-2 antibody acquisition rate according to the transplanted organ. A) Antibody seropositivity rate, B) High antibody seropositivity rate. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SPK, simultaneous pancreas–kidney.

mycophenolate mofetil, and everolimus treatment for the maintenance of immunosuppression. The final model included all the variables in Models 1 and 2. The kidney and SPK transplant groups originally had low antibody titers, but in Model 2, the antibody responses were no longer low compared to the liver transplant group. The heart transplantation group originally showed a high antibody titer; however, the antibody response was low in Model 1. In lung transplantation recipients, the antibody titer remained low, regardless of which factors were adjusted for.

3.5. Subgroup analysis of kidney transplant recipients

Kidney transplant recipients constituted the majority of the study cohort and were analyzed as a subgroup. Among these patients, the SARS-CoV-2 antibody acquisition rate was 36% at 1 month, and 68% at 6 months, after the second vaccination (Table S5). As stated previously for the overall SOT recipient population, logistic regression analysis of kidney recipients showed that older age, shorter period between transplantation and vaccination, elevated creatinine levels, and steroid or MMF use were significantly associated with failed seropositivity after the second vaccination (Table S6). Furthermore, among kidney transplant recipients who were at <2 years after transplantation, the antibody acquisition rate was only 13% at 1 month and 47% at 6 months after the second vaccination. In addition, these patients who had received rituximab as induction therapy at the time of transplantation had poorer antibody responses at 6 months than the no-rituximab group (Table S7, 34% vs. 58%, $P = 0.04$).

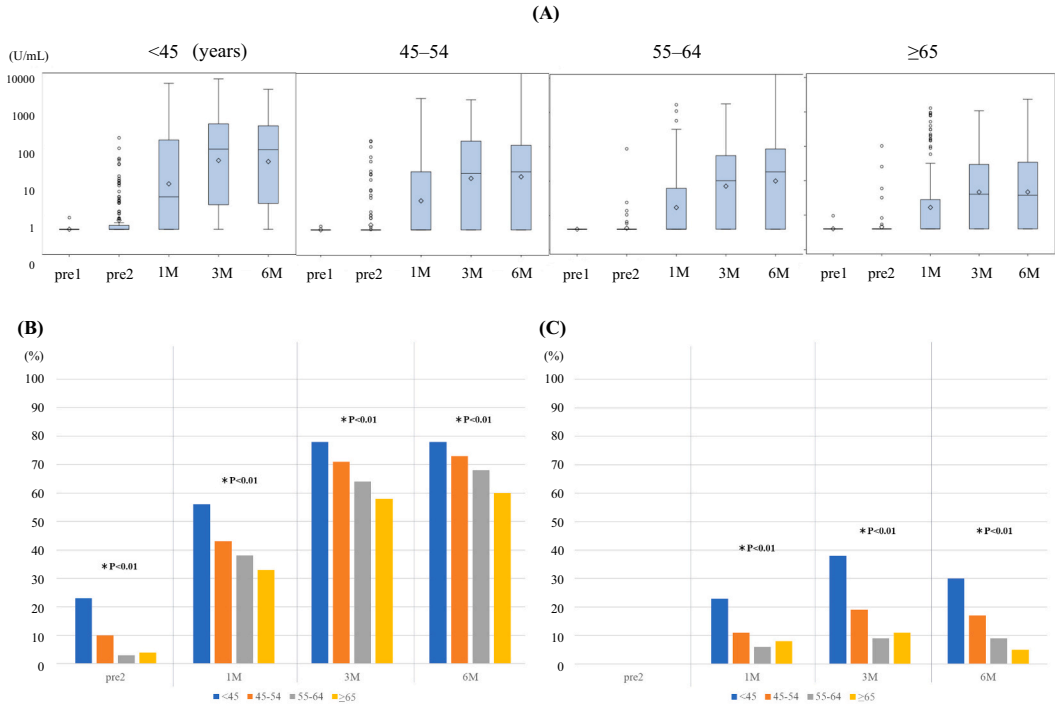


Fig. 4. SARS-CoV-2 antibody kinetics and acquisition rate according to age (years). A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

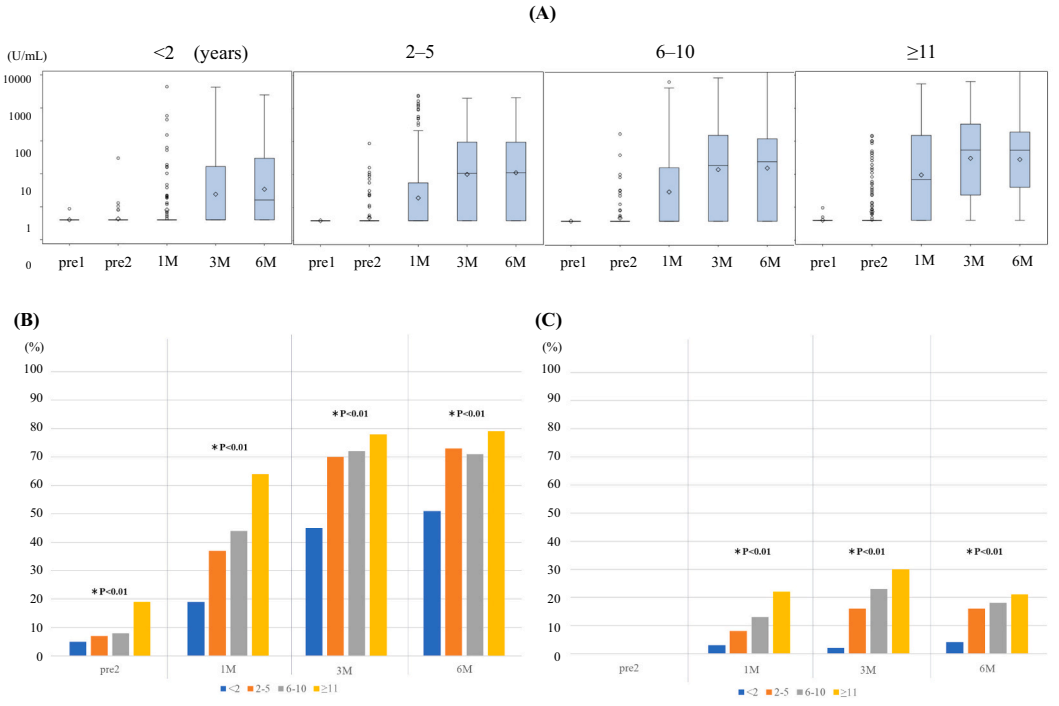


Fig. 5. SARS-CoV-2 antibody kinetics and acquisition rate according to the period between transplantation and vaccination. A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

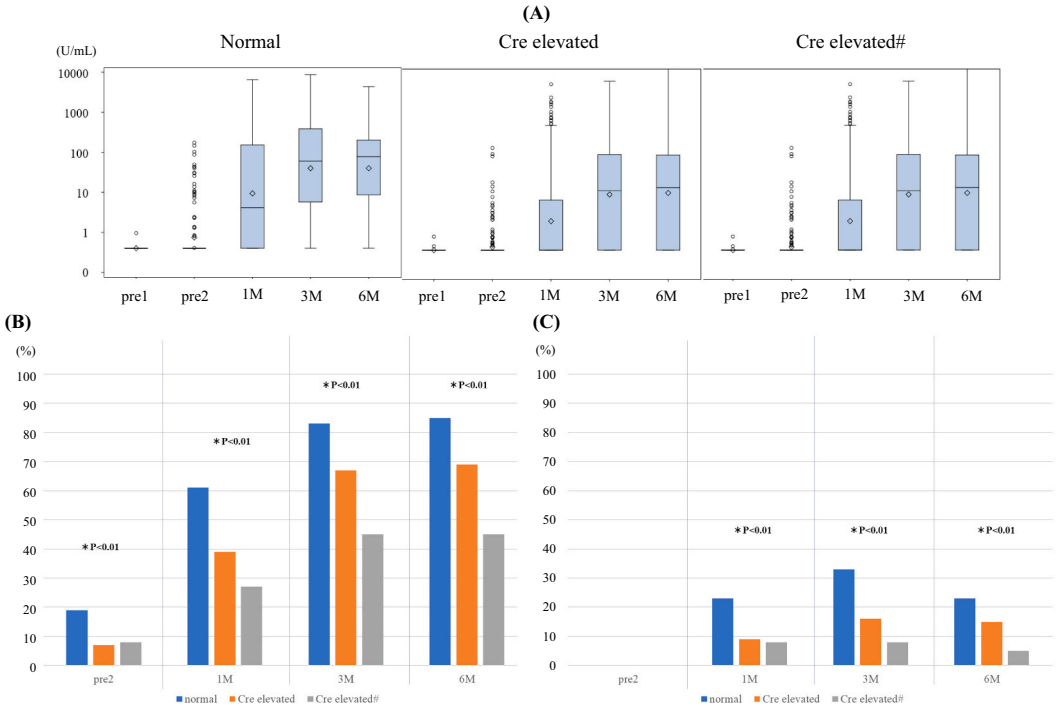


Fig. 6. SARS-CoV-2 antibody kinetics and acquisition rate according to serum creatinine level. A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Patients were classified according to serum creatinine level as follows: normal,
normal, <1.08 mg/dL for men, <0.80 mg/dL for women; Cre elevated, 1.08–2.15 mg/dL for men, 0.80–1.59 mg/dL for women; and Cre elevated#, ≥2.16 mg/dL for men, ≥1.60 mg/dL for women. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1M), 3 (3M), and 6 (6M) months after the second vaccination.
SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

4. Discussion

Poor antibody production after vaccination is a concern among SOT recipients [3–5]. Furthermore, precisely when SOT recipients demonstrate the peak antibody response after vaccination had not been investigated previously. The general population typically acquires a peak SARS-CoV-2 antibody titer approximately 1 month after the second vaccine dose, and the titer gradually decays thereafter [10,11]. In contrast, SOT recipients rarely exhibit an antibody response this early after vaccination [12], owing to the slow growth of immunosuppressed antibody-producing cells. Given these known and unknown features of antibody kinetics, a third round of vaccination is recommended as a booster vaccination for the entire population, including SOT recipients [13,14]. Characterizing the onset, peak, and decay of antibody titers is essential for optimizing the timing of the third vaccine dose, but statistical data regarding transition in antibody titers after vaccination in SOT patients are unavailable.

According to previous reports, only 25%–55% of SOT recipients acquire SARS-CoV-2 S-IgG antibodies after the second vaccination dose [4–6,13,15–21]. Even after a third dose of vaccine (booster injection), poor antibody responses were supposed [3], and the seropositivity rate remained at only 60%–70% [13,15,17,19]. In contract, approximately 40% of SOT recipients had acquired SARS-CoV-2 antibodies by 1 month after vaccination in our current study. This acquisition rate gradually increased and reached approximately 70% by 6 months after the second vaccination dose; this rate is much higher than that stated in previous reports. Furthermore, the S-IgG antibody titer gradually increased and peaked at 6 months after the second vaccination (GMT: 0.40 at pre1; 0.54 at pre2; 2.91 at 1 M; 11.71 at 3 M; and 12.73 U/mL at 6 M). These results showed that SARS-CoV-2 S-IgG antibody titers and kinetics among SOT recipients were slower than those in the general Japanese population, peaking 1 month after vaccination and diminishing thereafter [10]. In general, antibody-producing cells are short-lived, reaching peak productivity approximately 1 month after a second dose of vaccine. However, the maturation of antibody-producing cells in SOT recipients may be delayed by the use of immunosuppressants, leading to differences in timing, rates, and levels of antibody acquisition and decay from those of the general population. In SOT patients, the third dose of vaccine should be administered at approximately the same time as the antibody peak to the second vaccination, that is, at 3 to 6 months after the second vaccine. This practice is expected to increase the effectivity of the booster injection in the SOT recipients, thus increasing the proportion of seropositive patients.

Furthermore, results from other Japanese general population studies on antibody titers using the same antibody kit obtained 928 U/mL GMT 1 month after the second vaccination doses [10]. The antibody titers in SOT recipients were extremely low compared with those in those studies, such as 2.91 U/mL GMT 1 month after the second vaccine dose. Even liver transplant patients with the highest antibody response among SOT recipients yielded 42.27 U/mL GMT (17.57–101.6) 1 month after vaccination, which is significantly lower than the lower limit of the 95% confidence interval for GMT in the general population, which is 857 U/mL, and the upper limit of the 95% confidence interval for liver transplant patients is 101.6 U/mL. Various reports suggest that increased antibody or neutralizing antibody titer from vaccination [22–24] results in greater protection against infection. However, on the other hand, the antibody titer in SOT recipients is extremely low and may be inadequate to prevent infection. In this study, the infection prevention effect was not investigated, so the adequate level of antibody titer is unknown; however, Yamamoto et al. found that the antibody titer required to obtain a 50% protection rate against infection in the general population was 27,000 AU/mL (using the AdviseDx SARS-CoV-2 IgG II assay, Abbott)

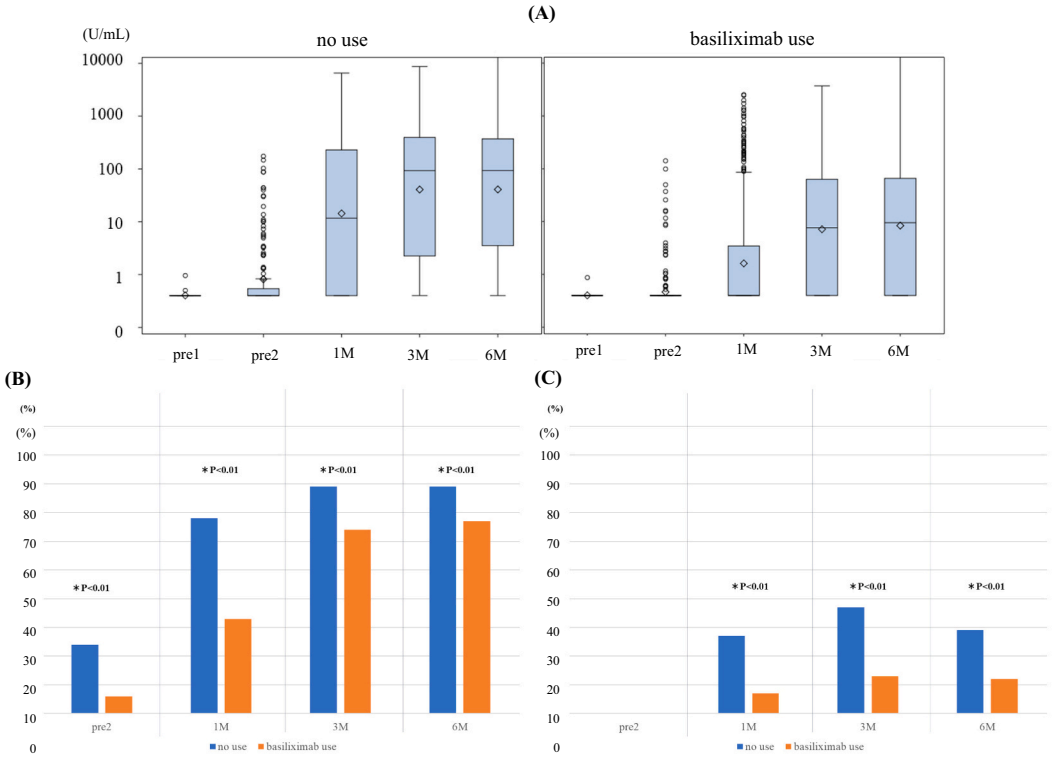


Fig. 7. SARS-CoV-2 antibody kinetics and acquisition rate according to basiliximab use. A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

[22], approximately 3000 U/mL when converted to Roche's antibody reagent. This was investigated by Matsumura [10]; however, most cases did not reach the antibody titer found in our current study. Therefore, it is highly likely that SOT recipients will not acquire sufficient antibody titers after two doses of the vaccine and a booster vaccination will be required.

As seen in our current study, many previous reports have inferred that the peak titers of SOT recipients likely would be lower than those of the general population because of a reduced humoral response to vaccination owing to immunosuppressive therapy [4,5,13,17–19]. Immunosuppression induction therapy, including basiliximab and rituximab use, was a risk factor for low reactivity in our patients, consistent with previous reports [25]. Likewise, steroid and MMF use for maintenance immunosuppressive therapy resulted in poor production reactivity, as seen previously [4,5,18,20,21,26]. In this regard, a new finding of our current study was the dose-dependent nature of the steroid- and MMF-associated poor responses.

The antibody seropositivity rate and titer kinetics differed significantly according to the transplant organ type. In particular, recipients of kidney, lung, or SPK transplant recipients had lower antibody titers than other transplant patients. In addition, antibody production levels differed depending on the immunosuppressant type or dose. Furthermore, our study showed that older age, vaccination administration relatively soon after transplantation, and decreased kidney function were risk factors for poor antibody production; these findings were similar to those in previous reports [4–6,13,18,20,21]. These factors may also influence the humoral response to vaccination overall.

Table S1 suggests the influence of the characteristics of each transplanted organ. For example, heart transplant recipients are young and have a long duration since organ transplantation. Basiliximab use was more common in kidney and SPK transplant recipients. Rituximab use was more common in kidney transplant recipients. Steroid use was less common in heart or liver transplant recipients. MMF use was less common in heart transplant recipients. Everolimus use was less common in kidney, lung, or SPK transplant recipients. Furthermore, in the subanalyses adjusted for these risk characteristics (Table 1), kidney and SPK transplant recipients originally had low antibody titers, but in Model 2, which was adjusted for immunosuppressive drugs, the antibody responses were no longer low compared to the liver transplant group, suggesting that the immunosuppressants had an adverse effect on the antibody response. In contrast, the heart transplantation group originally had a high antibody titer; however, the antibody response was low in Model 1, which was adjusted for age categories and years since transplantation, suggesting that these factors had a positive effect on the antibody response. However, in lung transplantation recipients, the antibody titer remained low regardless of which factors were adjusted for. The reason for this is unknown, but it may be due to the characteristics of the transplanted organ. For this reason, if there is concern about poor antibody response to vaccination in kidney or SPK transplant recipients, attenuating the use of immunosuppressants may be effective in acquiring antibodies, considering the risk of rejection. Furthermore, it is necessary to treat lung transplant recipients with special attention due to the low acuity of the antibodies after vaccination.

The current study had several limitations. First of all, this study did not include the general population as a control group, so it was not possible to compare actual differences in response. However, it was suggested that peak antibody titer production was delayed compared to studies in similar healthy Japanese populations [10]. Furthermore, approximately 70% of SOT recipients acquired S-IgG antibodies, but the titer was often significantly lower than in the general population [10] and would be insufficient for infection control. Second, the actual infection prevention effect was not investigated in this study; therefore, the actual adequate antibody titer is unknown. Furthermore, the T-cell response is thought to have a strong influence on vaccine antibody

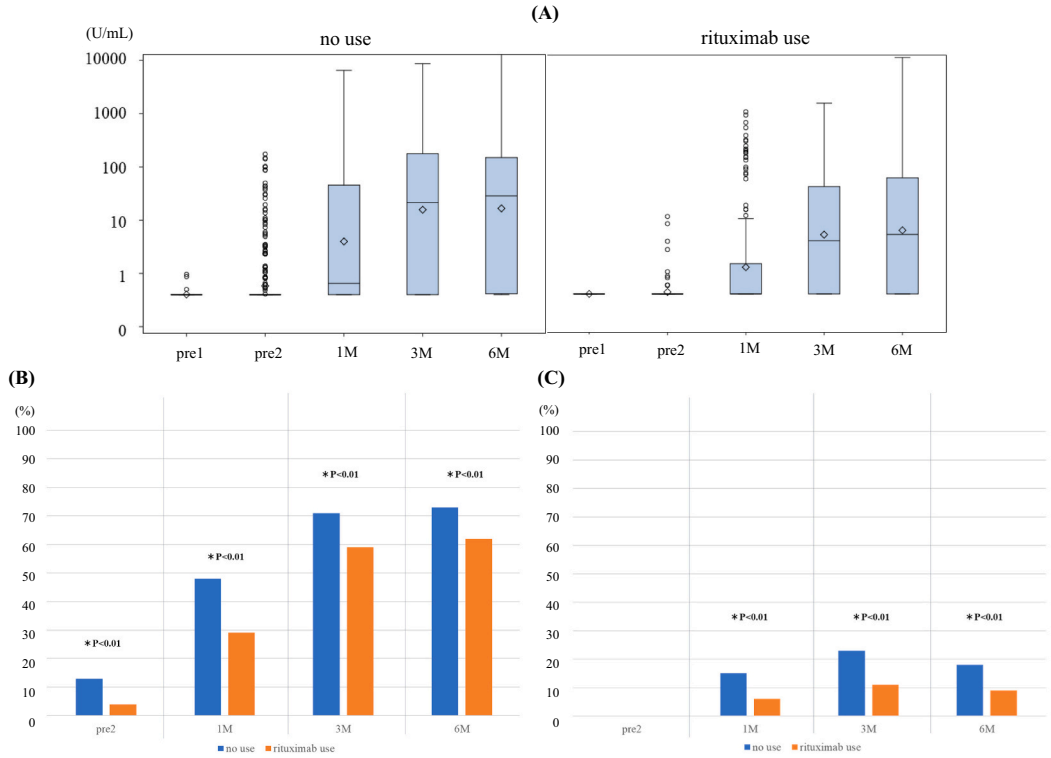


Fig. 8. SARS-CoV-2 antibody kinetics and acquisition rate according to rituximab use. A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

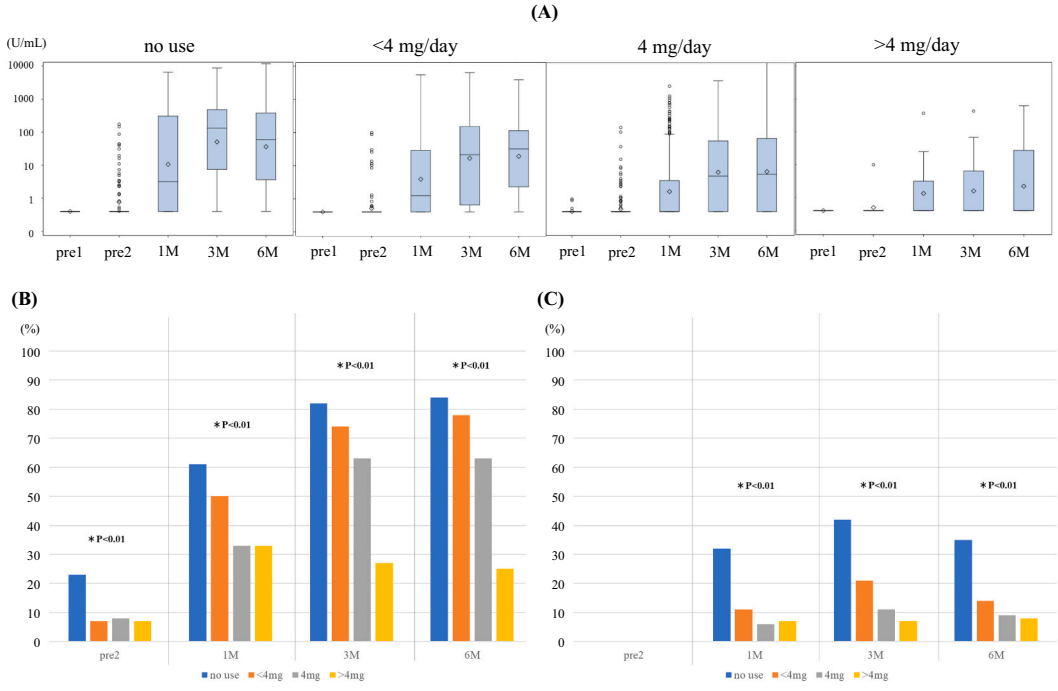


Fig. 9. SARS-CoV-2 antibody kinetics and acquisition rate according to steroid dose. A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Patients were classified according to steroid dose as follows: no use, no steroids include in maintenance immunosuppressive regimen; < 4 mg, <4 mg daily; 4 mg, 4 mg daily; and > 4 mg, >4 mg daily. Steroid dose was converted to methylprednisolone dose. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and the second (pre2) doses of vaccine and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

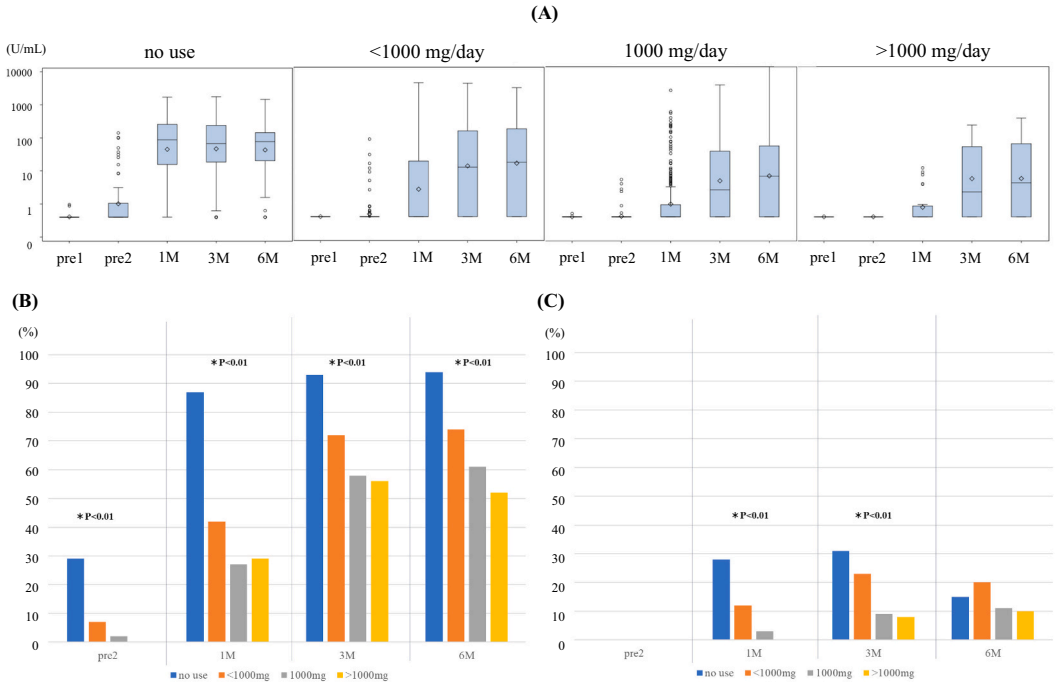


Fig. 10. SARS-CoV-2 antibody kinetics and acquisition rate according to MMF dose. A) Antibody kinetics, B) Antibody seropositivity rate, C) High antibody seropositivity rate. Patients were classified according to the MMF dose as follows: no use, no inclusion of MMF in maintenance immunosuppressive regimen; <1000 mg, <1000 mg daily; 1000 mg, 1000 mg daily; and > 1000 mg, >1000 mg daily. Antibody seropositivity was defined as ≥0.8 U/mL, according to the test usage guidelines. High antibody seropositivity was defined as ≥210 U/mL, according to the FDA statement[8]. The test periods were as follows: before the first (pre1) and second (pre2) doses of vaccination and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination. SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; MMF, mycophenolate mofetil.

production [17,27–29]. In this study, antibody production was weak in the basiliximab group, so it is important to determine the T-cell activation response, which was not assessed in the present study. Further research is required on the association between antibody acquisition and infection. In addition, we focused on S-IgG and did not assess neutralizing antibody kinetics. However, S-IgG levels are well correlated with neutralizing activity kinetics [26], and the S-IgG measurement method is widely available; therefore, S-IgG kinetics may be more useful for clinical management. Furthermore, neutralizing antibody increases more slowly than S-IgG [30], potentially leading to an even wider divergence between SOT recipients and the general population. Finally, the administration of third and later doses of SARS-CoV-2 vaccines has begun [15,31–33]; further evaluation, including studies of antibody kinetics and rates of acquisition and of seropositivity, is required. To determine the appropriate timing of the third SARS-CoV-2 vaccine dose for SOT recipients, assessing the onset of the S-IgG peak after the second dose and determining its wanning kinetics thereafter are important.

5. Conclusion

SARS-CoV-2 S-IgG antibody titers and kinetics among SOT recipients differed from those of the general population and had delayed and dampened peak. In the context of the prolonged SARS-CoV-2 epidemic, it is considered important not only to acquire a high antibody titer quickly but also to maintain it for a long time for infection control. Therefore, to maximize effectiveness, a third vaccine dose should be administered to SOT recipients at 3 to 6 months after the second dose, i.e. at the time when antibody titers peaked in our patients.

Funding

This study was supported by a Grant-in-Aid for Investigation of Promotion of Health Labor Administration (Research Project for Promotion of Policies for Emerging and *Re*-emerging Infectious Diseases and Immunization) [Principal Investigator: Yoshio Hirota; Grant Number: 20HA2001].

Supplementary data to this article can be found online at https://doi.org/10.1016/j.vaccine.2024.126221.

CRediT authorship contribution statement

Kohei Unagami: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Investigation, Data curation, Conceptualization. **Mikiko Yoshikawa:** Visualization, Resources, Project administration, Investigation, Data curation, Conceptualization. **Hiroto Egawa:** Writing – review & editing, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Satoko Ohfuji:** Writing – review & editing, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Yoichiro Natori:** Writing – review & editing, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Conceptualization. **Rikako Oki:** Writing – review & editing, Visualization, Validation, Resources, Methodology, Investigation, Conceptualization. **Tomomi Mori:** Visualization, Resources, Investigation, Data curation. **Hidetoshi Hattori:** Visualization, Resources, Investigation, Data curation. **Ayumi Ishiwatari:** Visualization, Resources, Investigation, Data curation. **Taichi Kanzawa:** Visualization, Resources, Investigation, Data curation. **Tomokazu Shimizu:** Visualization, Resources, Investigation, Data curation. **Kazuya Omoto:** Visualization, Resources, Investigation, Data curation. **Masashi Inui:** Visualization, Resources, Investigation, Data curation. **Yuuki Masano:** Visualization, Resources, Investigation, Data curation. **Takashi Ito:** Visualization, Resources, Investigation, Data curation. **Daisuke Nakajima:** Visualization, Resources, Investigation, Data curation. **Tetsuya Babazono:** Visualization, Resources,

Table 1
Regression analyses for seropositivity rate of each transplanted organ adjusted for risk factors.

Model	Transplanted organ	pre2				1 M				3 M				6 M			
		SPR	OR	95% CI	*P*-value	SPR	OR	95% CI	P-value	SPR	OR	95% CI	P-value	SPR	OR	95% CI	P-value
Crude	kidney	**6%**	**0.14**	**0.07–0.29**	**<0.01**	**36%**	**0.13**	**0.06–0.28**	**<0.01**	**65%**	**0.12**	**0.04–0.41**	**<0.01**	**68%**	**0.05**	**0.05–0.51**	**<0.01**
	heart	33%	1.07	0.46–2.49	0.88	75%	0.69	0.27–1.81	0.45	86%	0.42	0.10–1.73	0.23	80%	0.28	0.07–1.11	0.07
	liver	32%	ref			81%	ref			94%	ref			93%	ref		
	lung	11%	0.25	0.05–1.23	0.09	**26%**	**0.08**	**0.02–0.29**	**<0.01**	**32%**	**0.03**	**0.01–0.14**	**<0.01**	**23%**	**0.02**	**0.00–0.12**	**<0.01**
	SPK	0%	NA			**39%**	**0.15**	**0.05–0.45**	**<0.01**	**67%**	**0.13**	**0.03–0.59**	**<0.01**	**65%**	**0.13**	**0.03–0.61**	**<0.01**
Adjusted model 1																	
	kidney		**0.11**	**0.04–0.28**	**<0.01**		**0.10**	**0.04–0.23**	**<0.01**		**0.12**	**0.03–0.41**	**<0.01**		**0.18**	**0.05–0.63**	**<0.01**
	heart		0.43	0.15–1.23	0.12		**0.27**	**0.09–0.81**	**0.02**		**0.18**	**0.04–0.87**	**0.03**		**0.17**	**0.04–0.77**	**0.02**
	liver		ref				ref				ref				ref		
	lung		0.32	0.06–1.77	0.19		**0.07**	**0.02–0.27**	**<0.01**		**0.02**	**0.00–0.09**	**<0.01**		**0.01**	**0.00–0.08**	**<0.01**
	SPK		NA				**0.11**	**0.03–0.39**	**<0.01**		**0.07**	**0.01–0.35**	**<0.01**		**0.09**	**0.02–0.50**	**<0.01**
Adjusted model 2																	
	kidney		**0.12**	**0.02–0.85**	**0.03**		0.43	0.12–1.59	0.21		0.19	0.03–1.25	0.08		0.35	0.05–2.65	0.31
	heart		0.37	0.06–2.46	0.31		0.50	0.13–1.97	0.32		0.25	0.03–1.79	0.17		0.13	0.02–1.02	0.05
	liver		ref				ref				ref				ref		
	lung		0.38	0.03–5.24	0.47		**0.18**	**0.03–0.89**	**0.04**		**0.05**	**0.01–0.41**	**<0.01**		**0.03**	**0.00–0.38**	**<0.01**
	SPK		NA				0.44	0.10–2.00	0.29		0.17	0.02–1.40	0.10		0.13	0.01–1.22	0.07
Final model																	
	kidney		**0.09**	**0.01–0.88**	**0.04**		0.28	0.06–1.39	0.12		0.14	0.02–1.21	0.07		0.24	0.03–2.30	0.22
	heart		0.41	0.05–3.49	0.41		**0.16**	**0.03–0.90**	**0.04**		**0.08**	**0.01–0.80**	**0.03**		**0.05**	**0.01–0.51**	**0.01**
	liver		ref				ref				ref				ref		
	lung		0.63	0.03–12.5	0.76		0.22	0.03–1.40	0.11		**0.03**	**0.00–0.27**	**<0.01**		**0.03**	**0.00–0.35**	**<0.01**
	SPK		NA				0.34	0.06–2.08	0.24		**0.08**	**0.01–0.85**	**0.04**		**0.08**	**0.01–1.00**	**0.05**

Logistic regression model. Model 1 included age categories, sex, duration from transplantation and serum creatinine levels at the time of transplantation. Model 2 included basiliximab and rituximab usage at the time of transplantation, and steroid, mycophenolate mofetil, and everolimus treatment for the maintenance of immunosuppression. The final model included all the variables in Models 1 and 2.
Antibody seropositivity was defined as $\geq$0.8 U/mL, according to the test usage guidelines. The test periods were as follows: before the first (pre1) and second (pre2) doses of vaccination and at 1 (1 M), 3 (3 M), and 6 (6 M) months after the second vaccination.
SPK, simultaneous pancreas–kidney.

Investigation, Data curation. **Toshio Takagi:** Visualization, Resources, Investigation, Data curation. **Shinichi Nunoda:** Visualization, Resources, Investigation, Data curation. **Yoshito Tomimaru:** Visualization, Resources, Investigation, Data curation. **Ryoichi Imamura:** Conceptualization. **Shigeru Miyagawa:** Visualization, Resources, Investigation, Data curation. **Koichi Toda:** Visualization, Resources, Investigation, Data curation. **Etsuro Hatano:** Visualization, Resources, Investigation, Data curation. **Hiroshi Date:** Visualization, Resources, Investigation, Data curation. **Miyaji Kyakuno:** Visualization, Resources, Investigation, Data curation. **Shiro Takahara:** Visualization, Resources, Investigation, Data curation. **Kenji Yuzawa:** Visualization, Validation, Resources, Project administration, Methodology, Investigation, Conceptualization. **Naoki Tanimine:** Validation, Project administration, Conceptualization. **Hideki Ohdan:** Visualization, Supervision, Resources, Investigation, Data curation, Conceptualization. **Hideki Ishida:** Visualization, Supervision, Resources, Investigation, Data curation, Conceptualization. **Yoshio Hirota:** Visualization, Resources, Investigation, Data curation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Yoshio Hirota reports financial support was provided by Health Labor Administration. Yoshio Hirota reports a relationship with Japan Health Labor Administration that includes: funding grants. The authors of this manuscript have no conflicts of interest to disclose as described by the Vaccine. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available owing to privacy or ethical restrictions.

References

[1] Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020;383:2603–15. https://doi.org/10.1056/NEJMoa2034577.

[2] Barouch DH. Covid-19 vaccines - immunity, variants. Boosters N Engl J Med 2022;387:1011–20. https://doi.org/10.1056/NEJMra2206573.

[3] Chen X, Luo D, Mei B, et al. Immunogenicity of COVID-19 vaccines in solid organ transplant recipients: a systematic review and meta-analysis. Clin Microbiol Infect 2023;29:441–56. https://doi.org/10.1016/j.cmi.2022.12.004.

[4] Boyarsky BJ, Werbel WA, Avery RK, et al. Antibody response to 2-dose SARS-CoV-2 mRNA vaccine series in solid organ transplant recipients. JAMA 2021;325:2204–6. https://doi.org/10.1001/jama.2021.7489.

[5] Grupper A, Rabinowich L, Schwartz D, et al. Reduced humoral response to mRNA SARS-CoV-2 BNT162b2 vaccine in kidney transplant recipients without prior exposure to the virus. Am J Transplant 2021;21:2719–26. https://doi.org/10.1111/ajt.16615.

[6] Mazzola A, Todesco E, Drouin S, et al. Poor antibody response after two doses of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) vaccine in transplant recipients. Clin Infect Dis 2022;74:1093–6. https://doi.org/10.1093/cid/ciab580.

[7] Cucchiari D, Egri N, Bodro M, et al. Cellular and humoral response after MRNA-1273 SARS-CoV-2 vaccine in kidney transplant recipients. Am J Transplant 2021;21:2727–39. https://doi.org/10.1111/ajt.16701.

[8] Doria-Rose N, Suthar MS, Makowski M, et al. Antibody persistence through 6 months after the second dose of mRNA-1273 vaccine for Covid-19. N Engl J Med 2021;384:2259–61. https://doi.org/10.1056/nejmc2103916.

[9] Food and Drug Administration home page. https://www.fda.gov/media/141477/download. [Accessed 29 February 2024].

[10] Matsuura T, Fukushima W, Nakagama Y, et al. Kinetics of anti-SARS-CoV-2 antibody titer in healthy adults up to 6 months after BNT162b2 vaccination measured by two immunoassays: a prospective cohort study in Japan. Vaccine 2022;40:5631–40. https://doi.org/10.1016/j.vaccine.2022.08.018.

[11] Grupel D, Gazit S, Schreiber L, et al. Kinetics of SARS-CoV-2 anti-S IgG after BNT162b2 vaccination. Vaccine 2021;39:5337–40. https://doi.org/10.1016/j.vaccine.2021.08.025.

[12] Yi SG, Knight RJ, Graviss EA, et al. Kidney transplant recipients rarely show an early antibody response following the first COVID-19 vaccine administration. Transplantation 2021;105:e72–3. https://doi.org/10.1097/tp.0000000000003764.

[13] Kamar N, Abravanel F, Marion O, et al. Three doses of an mRNA Covid-19 vaccine in solid-organ transplant recipients. N Engl J Med 2021;385:661–2. https://doi.org/10.1056/nejmc2108861.

[14] Hall VG, Ferreira VH, Ku T, et al. Randomized trial of a third dose of mRNA-1273 vaccine in transplant recipients. N Engl J Med 2021;385:1244–6. https://doi.org/10.1056/nejmc2111462.

[15] Karaba AH, Zhu X, Liang T, et al. A third dose of SARS-CoV-2 vaccine increases neutralizing antibodies against variants of concern in solid organ transplant recipients. Am J Transplant 2022;22:1253–60. https://doi.org/10.1111/ajt.16933.

[16] Haller MC, Kaiser RA, Langthaler S, et al. Comparison of mRNA-1273 and BNT162b2 SARS-CoV-2 mRNA vaccine immunogenicity in kidney transplant recipients. Transpl Int 2022;35:10026. https://doi.org/10.3389/ti.2021.10026.

[17] Bertrand D, Hamzaoui M, Lemée V, et al. Antibody and T-cell response to a third dose of SARS-CoV-2 mRNA BNT162b2 vaccine in kidney transplant recipients. Kidney Int 2021;100:1337–40. https://doi.org/10.1016/j.kint.2021.09.014.

[18] Russo G, Lai Q, Poli L, et al. SARS-COV-2 vaccination with BNT162B2 in renal transplant patients: risk factors for impaired response and immunological implications. Clin Transpl 2022;36:e14495. https://doi.org/10.1111/ctr.14495.

[19] Tylicki L, Dębska-Ślizień A, Muchlado M, et al. Boosting humoral immunity from mRNA COVID-19 vaccines in kidney transplant recipients. Vaccines 2021;10:56. https://doi.org/10.3390/vaccines10010056.

[20] Dębska-Ślizień A, Ślizień Z, Muchlado M, et al. Predictors of humoral response to mRNA COVID19 vaccines in kidney transplant recipients: a longitudinal study-the COViNEPH project. Vaccines 2021;9:1165. https://doi.org/10.3390/vaccines9101165.

[21] Rozen-Zvi B, Yahav D, Agur T, et al. Antibody response to SARS-CoV-2 mRNA vaccine among kidney transplant recipients: a prospective cohort study. Clin Microbiol Infect 2021;27(1173):1173.e1–4. https://doi.org/10.1016/j.cmi.2021.04.028.

[22] Yamamoto S, Mizoue T, Ohmagari N. Analysis of previous infection, vaccinations, and anti-SARS-CoV-2 antibody titers and protection against infection with the SARS-CoV-2 omicron BA.5 variant. JAMA Netw Open 2023;6:e233370. https://doi.org/10.1001/jamanetworkopen.2023.3370.

[23] Bergwerk M, Gonen T, Lustig Y, et al. Covid-19 breakthrough infections in vaccinated health care workers. N Engl J Med 2021;385:1474–84. https://doi.org/10.1056/NEJMoa2109072.

[24] Khoury DS, Cromer D, Reynaldi A, et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. Nat Med 2021;27:1205–11. https://doi.org/10.1038/s41591-021-01377-8.

[25] Mohamadou I, Nkok J, Galichon P, et al. Immediate impact of induction treatment on Postvaccination SARS-CoV-2 serology in kidney transplant recipients. Transplantation 2021;105:e135–6. https://doi.org/10.1097/tp.0000000000003862.

[26] Bae S, Alejo JL, Chiang TPY, Werbel WA, et al. mTOR inhibitors, mycophenolates, and other immunosuppression regimens on antibody response to SARS-CoV-2 mRNA vaccines in solid organ transplant recipients. Am J Transplant 2022;22:3137–42. https://doi.org/10.1111/ajt.17158.

[27] Fabris M, De Marchi G, Domenis R, et al. High T-cell response rate after COVID-19 vaccination in belimumab and rituximab recipients. J Autoimmun 2022;129:102827. https://doi.org/10.1016/j.jaut.2022.102827.

[28] Schiavoni I, Palmieri A, Olivetta E, et al. T-cell mediated response after primary and booster SARS-CoV-2 messenger RNA vaccination in nursing home residents. J Am Med Dir Assoc 2023;24:140–147.e2. https://doi.org/10.1016/j.jamda.2022.11.024.

[29] Sieiro Santos C, Calleja Antolin S, Moriano Morales C, et al. Immune responses to mRNA vaccines against SARS-CoV-2 in patients with immune-mediated inflammatory rheumatic diseases. RMD Open 2022;8:e001898. https://doi.org/10.1136/rmdopen-2021-001898.

[30] Takahashi M, Ai T, Sinozuka K, et al. Activation of SARS-CoV-2 neutralizing antibody is slower than elevation of spike-specific IgG, IgM, and nucleocapsid-specific IgG antibodies. Sci Rep 2022;12. https://doi.org/10.1038/s41598-022-19073-z. 14909.7.

[31] Bar-On YM, Goldberg Y, Mandel M, et al. Protection by a fourth dose of BNT162b2 against omicron in Israel. N Engl J Med 2022;386:1712–20. https://doi.org/10.1056/nejmoa2201570.

[32] Karaba AH, Johnston TS, Aytenfisu TY, et al. A fourth dose of COVID-19 vaccine does not induce neutralization of the omicron variant among solid organ transplant recipients with suboptimal vaccine response. Transplantation 2022;106:1440–4. https://doi.org/10.1097/tp.0000000000004140.

[33] Agrawal U, Bedston S, McCowan C, et al. Severe COVID-19 outcomes after full vaccination of primary schedule and initial boosters: pooled analysis of national prospective cohort studies of 30 million individuals in England, Northern Ireland, Scotland, and Wales. Lancet 2022;400:1305–20. https://doi.org/10.1016/S0140-6736(22)01656-7.

BMC Pulmonary Medicine

RESEARCH

Open Access

Baseline genetic abnormalities and effectiveness of osimertinib treatment in patients with chemotherapy-naïve EGFR-mutated NSCLC based on performance status

Yoshihiko Taniguchi[1*], Akihiro Tamiya[1], Mitsuo Osuga[2], Daijiro Harada[3], Shun-ichi Isa[4], Keiichi Nakamura[5], Yasuyuki Mizumori[6], Tsutomu Shinohara[7], Hidetoshi Yanai[8], Katsumi Nakatomi[9], Masahide Oki[10], Masahide Mori[11], Tomohito Kuwako[12], Koji Yamazaki[13], Atsuhisa Tamura[14], Masahiko Ando[15] and Yasuhiro Koh[3,16]

Abstract

Background/Aim For patients treated with osimertinib as first-line therapy, there have been no studies comparing both progression-free survival (PFS) and overall survival (OS) according to performance status (PS). Furthermore, no studies have examined differences in baseline genetic abnormalities between patients with poor and good PS. Therefore, we aimed to investigate differences in baseline genetic abnormalities and treatment effects between patients with poor and good PS who received osimertinib as the primary treatment.

Patients and methods This is a secondary analysis of the ELUCIDATOR study, which is a multi-center prospective observational study in Japan that assessed mechanisms underlying resistance to osimertinib as first-line treatment for advanced non-small cell lung cancer with epidermal growth factor receptor mutations.

Results There were 153 and 25 patients in the good and poor PS groups, respectively. Multivariate analysis revealed no significant between-group differences in PFS (hazards ratio [HR]: 0.98, 95% confidence interval [CI]: 0.52–1.72, $p=0.946$). Multivariate analysis of OS revealed that poor PS was a poor prognostic factor (HR: 2.67, 95% CI: 1.43–4.73, $p=0.003$). Regarding baseline genetic abnormalities, there was a significant increase in APC-positive cases (20.0% vs. 2.2%, $p=0.009$) and a trend toward more CTNNB1-positive cases in the poor PS group than in the good PS group (14.3% vs. 2.9%, $p=0.062$).

Conclusion There was no between-group difference in PFS, although OS was significantly inferior in the poor PS group. Additionally, there was a significant increase in APC-positive cases and a trend toward more CTNNB1-positive cases in the poor PS group.

Keywords EGFR mutations, Osimertinib, Overall survival, Performance status, Progression-free survival

*Correspondence:
Yoshihiko Taniguchi
taniguchi.yoshihiko.ny@mail.hosp.go.jp

Full list of author information is available at the end of the article

Introduction

The prevalence of epidermal growth factor receptor (EGFR) mutations in patients with non-small cell lung cancer (NSCLC) ranges from approximately 10–30% [1, 2]. Osimertinib is a third-generation, irreversible, oral EGFR tyrosine kinase inhibitor (TKI) that selectively inhibits both EGFR-TKI–sensitizing and EGFR p.Thr790Met (T790M) resistance mutations; further, it has shown efficacy in patients with NSCLC [3]. Among patients with treatment-naïve advanced EGFR-mutated NSCLC, osimertinib significantly prolonged overall survival (OS) compared with a first-generation EGFR-TKI [4, 5].

However, since existing phase III trials for third-generation EGFR-TKIs excluded patients with poor performance status (PS) and specific comorbidities, their results are not directly applicable to real-world settings. Several studies have indicated that osimertinib could be beneficial in patients with poor PS and EGFR T790M mutation-positive NSCLC following the progression of first- and second-generation EGFR-TKI treatments [6–8]. Moreover, several studies have evaluated the effectiveness of first-line osimertinib treatment in patients with poor PS [9–11], with inconsistent findings being reported. Additionally, no studies have compared both progression-free survival (PFS) and OS according to PS.

Co-occurring alterations in tumor suppressor genes (TSGs) affect outcomes among patients with EGFR-mutated NSCLC. These alterations, including mutations in tumor protein (TP) 53 and other TSGs, have been associated with worse outcomes in patients treated with EGFR-TKIs [12]. Furthermore, no studies have examined differences in baseline genetic abnormalities between patients with poor and good PS.

Therefore, this prospective study aimed to assess differences in baseline genetic abnormalities and treatment effects between patients with poor and good PS who received osimertinib as the primary treatment.

Table 1 Patient characteristics

*$P<0.05$	Good PS ($n=153$)	Poor PS ($n=25$)	*P*-value
Sex: Male – no. (%)	52 (34.0)	7 (28.0)	0.651
Age (range)	74 (36–91)	76 (55–86)	0.435
PS (0/1/2/3)	68/85/0/0	0/0/21/4	
Smoking: Yes – no. (%)	65 (42.5)	9 (36.0)	0.663
EGFR (L858R/Del-19)	76/77	7/18	0.053
Brain metastasis: Yes – no. (%)	40 (26.1)	8 (32.0)	0.627
Pleural effusion: Yes – no. (%)	63 (41.2)	9 (36.0)	0.667
Liver metastasis: Yes – no. (%)	13 (8.5)	4 (16.0)	0.266
Bone metastasis: Yes – no. (%)	42 (27.5)	13 (52.0)	0.019*

*$P<0.05$

PS, performance status; EGFR, epidermal growth factor receptor

Patients and methods

This study is a secondary analysis of the ELUCIDATOR study, which was a prospective observational study conducted at multiple centers of the National Hospital Organization Group in Japan. The ELUCIDATOR study evaluated resistance-related mutations and alternations of osimertinib as first-line chemotherapy for advanced NSCLC harboring EGFR-sensitizing mutations. The eligibility criteria included (i) a definitive diagnosis of non-squamous NSCLC confirmed through biopsy or cytology, (ii) the presence of EGFR mutations (exon 19 deletion or exon 21-point mutation L858R), (iii) osimertinib planned as first-line chemotherapy, and (iv) ability to provide blood specimens. Serial plasma samples were collected at baseline. Progressive disease was assessed according to the Response Evaluation Criteria in Solid Tumors version 1.1. In the present analysis, we used data regarding patient background, treatment efficacy, survival, baseline genetic abnormalities, and post-treatment. This study follows the ethical principles of the Declaration of Helsinki. Written informed consent was obtained from all patients before the study was performed; in addition, and the National Hospital Organization Review Board for Clinical Trial approved this protocol before the start of the study. This prospective observational study was registered on December 10, 2018, in the Japanese Register of Clinical Trials (clinical Trial Number: jRCTs031180051).

Statistical analyses

Survival curves were estimated using the Kaplan–Meier method, with between-group comparisons using the log-rank test. Multivariate analyses were performed using the Cox proportional hazards model. Between-group differences in continuous and categorical variables were evaluated using the Wilcoxon rank-sum and Fisher's exact tests, respectively. A p-value <0.05 was considered statistically significant. Statistical analyses were performed using the JMP statistical software program (14th version, SAS Institute Inc., Cary, NC) to compare clinical outcomes according to patient characteristics.

Results

There were 153 and 25 patients in the good (PS 0 and 1) and poor (PS 2 to 4) PS groups, respectively. There were no significant between-group differences in sex; age; smoking history; type of EGFR gene mutation; and proportion of patients with brain metastases, pleural effusion, or liver metastases. Bone metastases were more common in the poor PS group than in the good PS group (Table 1).

There was no significant difference in PFS between the good and poor PS groups (19.5, 95% confidence interval [CI] 17.2 to 23.3 months vs. 13.5, 95% CI 5.97 to NE months, hazards ratio [HR]: 1.09, 95% CI: 0.61–1.82,

p=0.758). The reasons for discontinuation of treatment were disease progression (n=11), treatment toxicity (n=4), and other reasons (n=3) in the poor PS group; and disease progression (n=74), treatment toxicity (n=26), and other reasons (n=6) in the good PS group. However, OS was significantly longer in the good PS group than in the poor PS group (40.9, 95% CI 32.3 to NE vs. 14.4, 95% CI 9.2 to NE months; HR: 2.51, 95% CI: 1.42–4.22, p<0.001) (Fig. 1). Multivariate analysis of PFS revealed that L858R was a poor prognostic factor (HR: 1.59, 95% confidence interval CI: 1.08–2.35, p=0.019); further, PFS tended to be shorter in patients with brain (HR: 1.48, 95% CI: 0.96–2.23, p=0.076), liver (HR: 1.88, 95% CI: 0.97–3.42, p=0.061), and bone metastases (HR: 1.45, 95% CI: 0.95–2.17, p=0.081); however, there was no significant between-group difference (HR: 0.98, 95% CI: 0.52–1.72, p=0.946) (Table 2). Multivariate analysis of OS revealed that poor PS was a poor prognostic factor (HR: 2.67, 95% CI: 1.43–4.73, p=0.003); further, OS tended to be shorter in patients with smoking history (HR: 1.75, 95% CI: 0.96–3.17, p=0.067) and L858R (HR:1.59, 95% CI: 0.98–2.57, p=0.060) (Table 3). During the observation period, 76 deaths were identified. Four deaths were treatment-related: 2 pulmonary embolisms, 1 drug-induced pneumonia, and 1 heart failure. The remaining deaths were tumor related.

The overall response rate was higher in the poor PS group than in the good PS group (76.0% vs. 53.6%, p=0.049), with the disease control rates being similar (88.0% vs. 82.4%, p=0.773).

Regarding genetic abnormalities before osimertinib treatment, there was a significant increase in APC-positive cases (20.0% vs. 2.2%, p=0.009) and a trend toward more CTNNB1-positive cases (14.3% vs. 2.9%, p=0.062) in the poor PS group than in the good PS group (Table 4).

Regarding post-treatment with osimertinib therapy, the proportions of patients who received post-treatment (28.0% vs. 40.5%), two or more post-treatments (16.0% vs. 27.5%), platinum-combination chemotherapy (20.0% vs. 31.4%), and immunotherapy (8.0% vs. 14.4%) were all lower in the poor PS group than in the good PS group (Table S1). Comparing the response to post-treatment in the poor PS and good PS groups, the response rate to platinum-combination chemotherapy was 29.17% vs. 0%, and the response rate to ICI was 31.82% vs. 0%.

Discussion

This is the first study to compare both PFS and OS among patients receiving osimertinib as first-line therapy between the good and poor PS groups. There was no significant between-group difference in PFS; however, OS was significantly inferior in the poor PS group. Notably, the response rate was rather favorable in the poor PS group. Additionally, this was the first study to examine

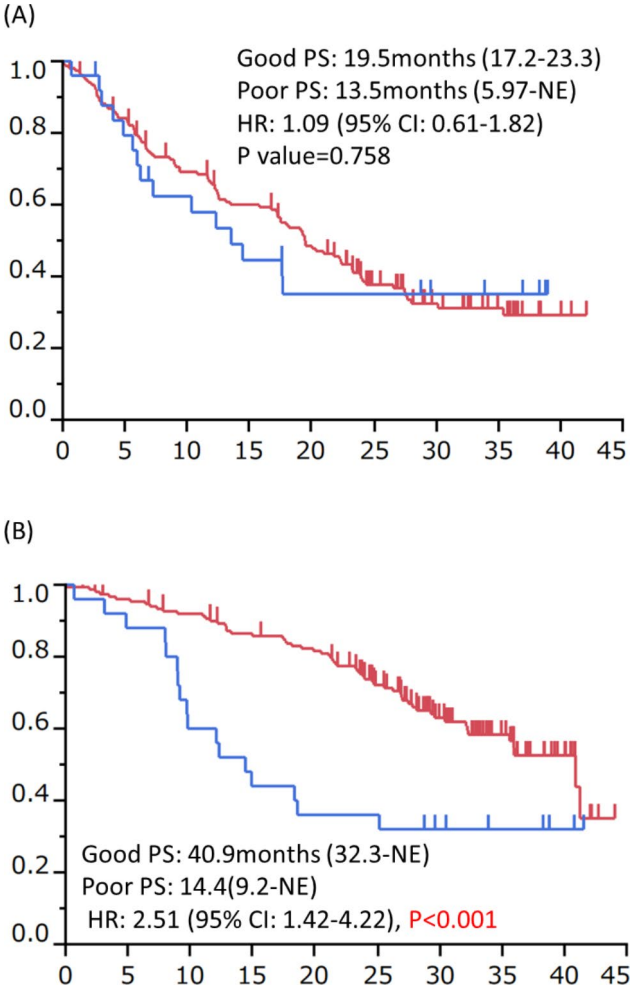


Fig. 1 Survival analysis of 178 patients with non-small cell lung carcinoma treated with osimertinib. (**a**) Progression-free survival curves of patients treated with osimertinib stratified according to the PS score. (**b**) Overall survival curves of patients treated with osimertinib stratified according to the PS score. PFS, progression-free survival; OS, overall survival; PS, performance status; HR, hazards ratio; CI, confidence interval

Table 2 Multivariate Cox proportional hazards model analysis of factors associated with progression-free survival

	HR	95% CI	*P*-value
Sex (Male)	1.20	0.71–2.03	0.498
Age (≥ 75)	1.19	0.81–1.77	0.375
PS (≥ 2)	0.98	0.52–1.72	0.946
Smoking (Yes)	1.12	0.67–1.83	0.671
EGFR (L858R)	1.59	1.08–2.35	0.019*
Brain metastasis	1.48	0.96–2.23	0.076**
Pleural effusion	1.29	0.87–1.91	0.198
Liver metastasis	1.88	0.97–3.42	0.061**
Bone metastasis	1.45	0.95–2.17	0.081**

*P<0.05, **P<0.10

HR, hazard ratio; CI, confidence interval; PS, performance status; EGFR, epidermal growth factor receptor

Table 3 Multivariate Cox proportional hazards model analysis of factors associated with overall survival

	HR	95% CI	*P*-value
Sex (Male)	1.31	0.72–2.40	0.382
Age (≥ 75)	1.42	0.89–2.27	0.138
PS (≥ 2)	2.67	1.43–4.73	0.003*
Smoking (Yes)	1.75	0.96–3.17	0.067**
EGFR (L858R)	1.59	0.98–2.57	0.060**
Brain metastasis	1.56	0.92–2.58	0.100
Pleural effusion	1.08	0.66–1.76	0.747
Liver metastasis	1.42	0.66–2.84	0.354
Bone metastasis	1.40	0.84–2.28	0.195

*$P<0.05$, **$P<0.10$

HR, hazard ratio; CI, confidence interval; PS, performance status; EGFR, epidermal growth factor receptor

Table 4 Genetic abnormality before osimertinib administration

	Good PS (*n* = 153)	Poor PS (*n* = 25)	*P*-value
TP53 (Yes/No)	56/87	11/13	0.653
EGFR amplification (Yes/No)	56/86	12/12	0.373
MET amplification (Yes/No)	21/122	4/20	0.762
ERBB2 amplification (Yes/No)	1/142	0/24	1.000
KIT (Yes/No)	2/141	0/24	1.000
MET (Yes/No)	7/136	2/22	0.619
PIK3CA (Yes/No)	6/137	2/22	0.323
CTNNB1 (Yes/No)	4/139	3/21	0.062**
APC (Yes/No)	3/139	4/20	0.009*
BRCA (Yes/No)	6/137	0/24	0.595
EGFR Compound mutation (Yes/No)	21/119	4/20	0.765

*$P<0.05$, **$P<0.10$

PS, performance status; TP, tumor protein; EGFR, epidermal growth factor receptor; MET, mesenchymal-epithelial transition; APC, adenomatous polyposis coli

differences in baseline genetic abnormalities before osimertinib treatment between patients with good and poor PS. Furthermore, there was a significant increase in APC-positive cases and a trend toward more CTNNB1-positive cases in the poor PS group.

A retrospective study of 61 patients (16 patients with poor PS) who received first-line osimertinib treatment reported that poor PS (2-4) negatively impacted PFS; however, OS was not verified [9]. Moreover, a prospective observational study on first-line osimertinib therapy in 16 patients with EGFR-mutated NSCLC who had poor PS reported that the overall objective response rate and median PFS were 56.3% and 10.5 months, respectively; further, the PS score improved in 8 of the 16 patients [10]. Furthermore, a multicenter retrospective study evaluated patients treated with osimertinib with a PS score of 2–4. In our study, we observed between-group differences in OS but not PFS. This suggests that osimertinib can be used to treat patients with poor PS and that differences in OS may have resulted from differences in post-treatment characteristics.

Among 36 patients with a PS score of 2, the median PFS, 1-year PFS, median OS, and 1-year OS were 14.5 months, 65.4%, 18.1 months, and 72.7%, respectively. Among 20 patients with a PS score of 3–4, the median PFS, 1-year PFS, median OS, and 1-year OS were 3.0 months, 27.1%, 5.0 months, and 46.1%, respectively [11]. In our study, the PFS and OS in the poor PS group were 13.5 and 14.4 months, respectively. The lack of significant between-group differences in our study could be attributed to the small sample size; however, this may have been influenced by the lower post-treatment rate in the poor PS group.

In patients with EGFR mutation-positive NSCLC, several genetic abnormalities have been shown to affect the efficacy of EGFR-TKIs. TP53 mutations, along with mutations in other TSGs such as RB1, NF1, ARID1A, BRCA1, and PTEN, have been shown to drive poor outcomes in patients with EGFR/TP53-mutated NSCLC [12]. Garon et al. identified baseline alterations co-occurring with activating EGFR in 69 genes, most commonly TP53 (43%), EGFR (other than activating EGFR; 25%), and PIK3CA (10%). Other genetic alterations included NF1 ($n=30$, 7.8%), APC ($n=27$, 7.0%), BRAF ($n=24$, 6.2%), CDK6 ($n=20$, 5.2%), and MET ($n=20$, 5.2%) [13]. In our study, there was a significant increase in APC-positive cases and a trend toward more CTNNB1-positive cases in the poor PS group. APC mutations, which result in increased intestinal epithelial cell proliferation and loss of differentiation, are crucially involved in the oncogenesis and progression of colon cancer [14, 15]. Furthermore, methylation of the APC gene promoter has been found to be higher in lung cancer tissues than in autologous controls, suggesting its importance in NSCLC carcinogenesis [16]. Moreover, APC mutations lead to the deregulation of the Wnt signaling pathway, promoting increased cell proliferation and decreased cell differentiation, which can contribute to tumor progression and metastasis. This disruption can create a more aggressive tumor phenotype, potentially diminishing the efficacy of EGFR-TKIs by fostering resistance mechanisms or by altering the tumor microenvironment [17–19]. However, there have been no previous reports regarding the role of APC in EGFR-mutated lung cancer. Aberrant activation of CTNNB1 contributes to carcinogenesis and tumor progression. Increased invasive potential of NSCLC cells in vitro has been reported with CTNNB1 co-mutation in EGFR-mutated lung cancer. This alteration is often detected at advanced stages and is involved in processes that begin late in tumor progression and dissemination [20, 21]. In addition, CTNNB1, encoding beta-catenin, is a key component of the Wnt signaling pathway. Mutations in CTNNB1 lead to the stabilization

and accumulation of beta-catenin in the cytoplasm and its subsequent translocation to the nucleus, where it activates target genes involved in cell proliferation and survival. This aberrant activation can enhance the invasive and metastatic potential of cancer cells, reducing the effectiveness of EGFR-TKIs by promoting pathways that circumvent EGFR inhibition [22–24]. Moreover, CTNNB1 mutations have been associated with poor recurrence-free postoperative survival in patients with EGFR-mutated lung adenocarcinomas [25]. Furthermore, mutations in PIK3CA and CTNNB1 are more common in advanced-stage tumors than in early-stage tumors, indicating their functional roles in malignant progression and metastasis; contrastingly, TP53, RB1, and NKX2–1 alterations appear to occur with comparable frequencies in early- and advanced-stage tumors [20, 26–29], which is consistent with our findings.

This study has several limitations. First, this study is a secondary analysis of a multicenter prospective observational study that was not designed for this study; therefore, it may have had insufficient statistical power. Second, the sample size of the poor PS group was small and did not allow sufficient comparison with the good PS group. Third, baseline genetic abnormalities were not tested in all cases due to an insufficient amount of ct-DNA.

Conclusion

We observed no significant between-group difference in PFS, although OS was significantly inferior in the poor PS group. Additionally, there was a significant increase in APC-positive cases and a trend toward more CTNNB1-positive cases in the poor PS group.

Abbreviations

EGFR	Epidermal growth factor receptor
NSCLC	Non-small cell lung cancer
OS	Overall survival
PS	Performance status
PFS	Progression-free survival
TSGs	Tumor suppressor genes
TP	Tumor protein
HR	Hazard ratio
CI	Confidence interval

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12890-024-03212-5.

Supplementary Material 1

Acknowledgements

We wish to thank Editage (www.Editage.jp) for the English language editing.

Author contributions

Yoshihiko Taniguchi, Akihiro Tamiya, Mitsuo Osuga, Shun-ichi Isa, Masahiko Ando and Yasuhiro Koh developed the study concept initiated the project. Yoshihiko Taniguchi, Akihiro Tamiya, Shun-ichi Isa, Masahiko Ando and Yasuhiro Koh coordinated protocol design. Masahiko Ando was responsible for the statistical analysis. Yoshihiko Taniguchi wrote the original draft. All authors have read and approved the final article.

Funding

This study was supported by AstraZeneca and the National Hospital Organization.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethical approval and consent to participate

This prospective observational study was registered on December 10, 2018, in the Japanese Register of Clinical Trials (JRCT; Clinical Trial Number: jRCTs031180051). Written informed consent was obtained from all patients before the study began, and the National Hospital Organization Review Board for Clinical Trials approved this protocol prior to the start of the study.

Consent for publication

Not Applicable.

Competing interests

Taniguchi Y has received honoraria from Chugai Pharmaceutical, Ono Pharmaceutical, AstraZeneca, and MSD. Tamiya A has received honoraria from Eli Lilly, Ono Pharmaceutical, Chugai Pharmaceutical, Boehringer Ingelheim, AstraZeneca, Bristol-Myers Squibb, Amgen, Taiho Pharmaceutical, Kyowa Kirin, MSD, Takeda Pharmaceutical, Nihon-Kayaku, Novartis, Thermo Fischer, Amgen, Tsumura, Daiich-Sankyo and Merck BioFarma, and research funding from Daiichi-Sankyo, Beigene, and AstraZeneca. Harada D has received honoraria from Takeda Pharmaceutical, Eli Lilly, Chugai Pharmaceutical, AstraZeneca, Taiho Pharmaceutical, Ono Pharmaceutical, Bristol-Myers Squibb, Towa Pharmaceutical, and Boehringer Ingelheim. Oki M has received Honoraria from AMCO, AstraZeneca, Canon Medical Systems, Chugai Pharmaceutical, Fujifilm Toyama Chemical, Kaneka Medix, Merit Medical Japan, Novartis Pharma, Olympus and Sanofi, and research funding from AbbVie, AstraZeneca, Chugai Pharmaceutical, Fujifilm Toyama Chemical, GlaxoSmithKline, Janssen Pharmaceutical, MSD, Ono Pharmaceutical, Parxel International, Pfizer, Sanofi. Mori M received honoraria from AstraZeneca, Boehringer Ingelheim, MSD, Eli Lilly, Novartis, Chugai Pharmaceutical, Taiho Pharmaceutical, Kyowa-kirin, Ono Pharmaceutical, Otsuka, Nihon-kayaku, Pfizer, Daiici-Sankyo, Takeda Pharmaceutical, and Shionogi and research funding from Chugai Pharmaceutical, Ono Pharmaceutical, MSD, and Delt-fly. Koh Y has received honoraria from Chugai Pharmaceutical, Guardant Health, Amgen, Takeda Pharmaceutical, and Tosoh Corporation and received consulting or advisory roles from Tosoh Corporation and research funding from Boehringer Ingelheim, AstraZeneca, Chugai Pharmaceutical, Tosoh Corporation, Daiichi Sankyo, Zeon Corporation, Amgen, and Takeda Pharmaceutical. The other co-authors received no honoraria or research funding.

Author details

[1]Department of Internal Medicine, NHO Kinki Chuo Chest Medical Center, 1180 Nagasone-cho, Kita-ku, Sakai City 591-8555, Osaka, Japan
[2]Center for Biomedical Sciences, Wakayama Medical University, Wakayama, Japan
[3]Department of Thoracic Oncology and Medicine, NHO Shikoku Cancer Center, Matsuyama, Ehime, Japan
[4]Clinical Research Center, NHO Kinki Chuo Chest Medical Center, Sakai, Osaka, Japan
[5]Department of Respiratory Medicine, NHO Asahikawa Medical Center, Asahikawa, Hokkaido, Japan
[6]Department of Respiratory Medicine, NHO Himeji Medical Center, Himeji, Hyogo, Japan
[7]Department of Respiratory Medicine, NHO Kochi Hospital, Kochi, Japan
[8]Department of Respiratory Medicine, NHO Mito Medical Center, Ibaraki, Japan
[9]Department of Respiratory Medicine, NHO Ureshino Medical Center, Ureshino, Saga, Japan

[10]Department of Respiratory Medicine, NHO Nagoya Medical Center, Nagoya, Aichi, Japan
[11]Department of Thoracic Oncology, NHO Osaka Toneyama Medical Center, Toyonaka, Osaka, Japan
[12]Department of Respiratory Medicine, NHO Shibukawa Medical Center, Shibukawa, Gunma, Japan
[13]Department of Thoracic Surgery, NHO Kyushu Medical Center, Fukuoka, Kyushu, Japan
[14]Department of Respiratory Medicine, NHO Tokyo National Hospital, Tokyo, Japan
[15]Department of Advanced Medicine, Nagoya University Hospital, Aichi, Japan
[16]Internal Medicine III, Wakayama Medical University, Wakayama, Japan

Received: 27 June 2024 / Accepted: 9 August 2024
Published online: 24 August 2024

References

1. Zhang YL, Yuan JQ, Wang KF, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget. 2016;7:78985–93. https://doi.org/10.18632/oncotarget.12587.
2. Greenhalgh J, Boland A, Bates V, et al. First-line treatment of advanced epidermal growth factor receptor (EGFR) mutation positive non-squamous non-small cell lung cancer. Cochrane Database Syst Rev. 2021;3:CD010383. https://doi.org/10.1002/14651858.CD010383.pub2.
3. Mok TS, Wu YL, Ahn MJ, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. N Engl J Med. 2016;376:629–40. https://doi.org/10.1016/j.lungcan.2018.10.027.
4. Soria JC, Ohe Y, Vansteenkiste J, et al. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. N Engl J Med. 2017;378:113–25. https://doi.org/10.1056/NEJMoa1713137.
5. Ramalingam SS, Vansteenkiste J, Planchard D, et al. Overall survival with osimertinib in untreated, *EGFR*-mutated advanced NSCLC. N Engl J Med. 2020;382:41–50. https://doi.org/10.1056/NEJMoa1913662.
6. Kato Y, Hosomi Y, Watanabe K, et al. Impact of clinical features on the efficacy of osimertinib therapy in patients with T790M-positive non-small cell lung cancer and acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors. J Thorac Dis. 2019;11:2350–60. https://doi.org/10.21037/jtd.2019.06.03.
7. Nakashima K, Ozawa Y, Daga H, et al. Osimertinib for patients with poor performance status and EGFR T790M mutation-positive advanced non-small cell lung cancer: a phase II clinical trial. Invest New Drugs. 2020;38:1854–61. https://doi.org/10.1007/s10637-020-00943-0.
8. Tsubata Y, Watanabe K, Saito R, et al. Osimertinib in poor performance status patients with T790M-positive advanced non-small-cell lung cancer after progression of first- and second-generation EGFR-TKI treatments (NEJ032B). Int J Clin Oncol. 2022;27:112–20. https://doi.org/10.1007/s10147-021-02043-2.
9. Teranishi S, Sugimoto C, Nagaoka S, et al. Retrospective analysis of independent predictors of progression-free survival in patients with EGFR mutation-positive advanced non-small cell lung cancer receiving first-line osimertinib. Thorac Cancer. 2022;13:2741–50. https://doi.org/10.1111/1759-7714.14608.
10. Igawa S, Fukui T, Kasajima M, et al. First-line osimertinib for poor performance status patients with EGFR mutation-positive non-small cell lung cancer: a prospective observational study. Invest New Drugs. 2022;40:430–7. https://doi.org/10.1007/s10637-021-01195-2.
11. Takamizawa S, Okuma Y, Kato Y, Hakozaki T, Kitagawa S, Zenke Y. First-line osimertinib in EGFR mutation-positive non-small cell lung cancer patients with poor performance status. Future Oncol. 2022;18:291–300. https://doi.org/10.2217/fon-2021-0947.
12. Stockhammer P, Grant M, Wurtz A, et al. Co-occurring alterations in multiple tumor suppressor genes are associated with worse outcomes in patients with EGFR-mutant lung cancer. J Thorac Oncol. 2024;19:240–51. https://doi.org/10.1016/j.jtho.2023.10.001.
13. Garon EB, Reck M, Nishio K, et al. Ramucirumab plus Erlotinib versus placebo plus erlotinib in previously untreated EGFR-mutated metastatic non-small-cell lung cancer (RELAY): exploratory analysis of next-generation sequencing results. ESMO Open. 2023;8:101580. https://doi.org/10.1016/j.esmoop.2023.101580.
14. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990:61:759–767. https://doi.org/10.1016/0092-8674(90)90186-i
15. Szvicsek Z, Oszvald A, Szabo L, et al. Extracellular vesicle release from intestinal organoids is modulated by apc mutation and other colorectal cancer progression factors. Cell Mol Life Sci. 2019;76:2463–76. https://doi.org/10.1007/s00018-019-03052-1.
16. Hu B, Zhang H, Wei H, et al. Does adenomatous polyposis coli gene promoter 1A methylation increase non-small cell lung cancer risk? A meta-analysis. Thorac Cancer. 2017;8:410–6. https://doi.org/10.1111/1759-7714.12450.
17. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61:759–67. https://doi.org/10.1016/0092-8674(90)90186-I.
18. Kinzler KW, Vogelstein B, Herrup K, et al. Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science. 1991;251:1366–70. https://doi.org/10.1126/science.1848370.
19. Esteller M, Fraga MF, Guo M, et al. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. Hum Mol Genet. 2001;10:3001–7. https://doi.org/10.1093/hmg/10.26.3001.
20. Skoulidis F, Heymach JV. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. Nat Rev Cancer. 2019;19:495–509. https://doi.org/10.1038/s41568-019-0179-8.
21. Pezzuto F, Hofman V, Bontoux C, et al. The significance of co-mutations in EGFR-mutated non-small cell lung cancer: optimizing the efficacy of targeted therapies? Lung Cancer. 2023;181:107249. https://doi.org/10.1016/j.lungcan.2023.107249.
22. Herbst RS, Prager D, Hermann R, et al. TRIBUTE: a phase III trial of erlotinib hydrochloride and carboplatin/paclitaxel in advanced non–small-cell lung cancer. J Clin Oncol. 2008;23:5892–9. https://doi.org/10.1200/JCO.2005.02.7347.
23. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from never smokers and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci USA. 2005;101:13306–11. https://doi.org/10.1073/pnas.0405220101.
24. Dong ZY, Zhong WZ, Zhang XC, et al. Potential biomarker for checkpoint blockade immunotherapy and treatment strategy. J Thorac Oncol. 2019;14:519–27. https://doi.org/10.1016/j.jtho.2018.12.013.
25. Kim Y, Ahn B, Yoon S, et al. An oncogenic CTNNB1 mutation is predictive of post-operative recurrence-free survival in an EGFR-mutant lung adenocarcinoma. PLoS ONE. 2023;18:e0287256. https://doi.org/10.1371/journal.pone.0287256.
26. Frampton GM, Ali SM, Rosenzweig M, et al. Activation of MET via diverse exon 14 splicing alterations occurs in multiple tumor types and confers clinical sensitivity to MET inhibitors. Cancer Discov. 2015;5:850–9. https://doi.org/10.1158/2159-8290.CD-15-0285.
27. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511:543–50. https://doi.org/10.1038/nature13385.
28. Yu HA, Suzawa K, Jordan E, et al. Concurrent alterations in EGFR-mutant lung cancers associated with resistance to EGFR kinase inhibitors and characterization of MTOR as a mediator of resistance. Clin Cancer Res. 2018;24:3108–18. https://doi.org/10.1158/1078-0432.CCR-17-2961.
29. Blakely CM, Watkins TBK, Wu W, et al. Evolution and clinical impact of co-occurring genetic alterations in advanced-stage EGFR-mutant lung cancers. Nat Genet. 2017;49:1693–704. https://doi.org/10.1038/ng.3990.

Publisher's Note

diagnostics

MDPI

Article

Relationship between Physical Characteristics and Morphological Features of the Articular Radius Surface: A Retrospective Single-Center Study

Reo Asai [1,2], Akira Ikumi [2], Yusuke Eda [3], Sho Kohyama [4], Takeshi Ogawa [5] and Yuichi Yoshii [1,*]

[1] Department of Orthopedic Surgery, Tokyo Medical University Ibaraki Medical Center, Ami 300-0395, Japan
[2] Department of Orthopedic Surgery, Faculty of Medicine, University of Tsukuba, Tsukuba 305-8577, Japan
[3] Department of Orthopaedic Surgery, Tsukuba Medical Center Hospital, 1-3-1, Amakubo, Tsukuba 305-8576, Japan
[4] Department of Orthopedic Surgery, Kikkoman General Hospital, Noda 278-0005, Japan
[5] Department of Orthopaedic Surgery, National Hospital Organization Mito Medical Center, Ibaraki 311-3193, Japan
* Correspondence: yyoshii@tokyo-med.ac.jp; Tel.: +81-298871161

Abstract: Preoperative planning is important for the osteosynthesis of distal radius fractures. Challenges arise for patients presenting with bilateral wrist injuries or a history of contralateral wrist injuries. In such cases, the estimation of the distal radius morphology and the determination of the plate size from the preoperative physical characteristics could prove beneficial. The objective of this study was to investigate the correlation between the physical characteristics and the morphology of the distal radius articular surface. A total of 79 wrist computed tomography (CT) images (41 women and 38 men) were evaluated. Physical characteristics, such as height, weight, and body mass index (BMI), were recorded. Three-dimensional CT analysis was performed to investigate the transverse and anteroposterior diameters of the distal radius. Pearson's correlation coefficient was used to assess the relationships between height, weight, and BMI and the transverse and anteroposterior diameters of the distal radius. A moderate to strong correlation was found in the overall analysis between body height and transverse diameter (r = 0.66). There were also moderate correlations between body height and anteroposterior diameter (r = 0.45) as well as weight and transverse diameter (r = 0.41), both of which were statistically significant ($p < 0.001$). Our findings indicate a statistically significant correlation between height, weight, and morphology of the distal radius. When analyzed by sex, the correlation between body height and the transverse diameter of the distal radius was found to be relatively strong in women (r = 0.47, $p = 0.002$), suggesting that it could be a useful indicator for preoperative planning, such as estimating plate size.

Keywords: preoperative planning; distal radius; volar locking plates; physical characteristics; height; weight; transverse diameter; anteroposterior diameter; three-dimensional analysis; sex differences

check for updates

Citation: Asai, R.; Ikumi, A.; Eda, Y.; Kohyama, S.; Ogawa, T.; Yoshii, Y. Relationship between Physical Characteristics and Morphological Features of the Articular Radius Surface: A Retrospective Single-Center Study. *Diagnostics* **2024**, *14*, 2005. https://doi.org/10.3390/diagnostics14182005

Academic Editors: Michał Strzelecki, Adam Piórkowski and Rafał Obuchowicz

Received: 6 August 2024
Revised: 6 September 2024
Accepted: 7 September 2024
Published: 10 September 2024

1. Introduction

Volar locking plates (VLPs) have become the preferred choice for the osteosynthesis of distal radius fractures, a common orthopedic procedure [1]. These plates provide stable fixation, allowing for the early mobilization and optimal healing of the fractured bone. The utilization of X-ray or computed tomography (CT) scans of the unaffected wrist is often incorporated into the preoperative planning phase for osteosynthesis and corrective osteotomy [2,3]. This imaging serves as a crucial reference, providing detailed anatomical information that guides surgeons in the reduction and the accurate placement of the VLP. Strict and meticulous preoperative planning is of paramount importance to ensure that the surgery is executed safely and precisely, thereby minimizing potential complications such as malalignment, hardware failure, or tendon/nerve damage [3–5].

When performing preoperative planning, orthopedic surgeons often use the unaffected side as a template for the reduction position [2,6,7]. This is also commonly performed in 3D preoperative planning. However, using the unaffected wrist as an indicator of reduction presents challenges in patients with bilateral wrist injuries or a past medical history affecting the contralateral wrist [8,9]. In such cases, preoperative planning becomes difficult because the contralateral side cannot serve as an indicator of the ideal shape of the reduction, complicating the preoperative selection of the appropriate plate size. Moreover, not all facilities are equipped for CT-based 3D preoperative planning. When preoperative planning cannot be performed, implant size must be determined intraoperatively. However, the distal radius cannot be fully exposed during surgery, and the articular surface may be displaced, making intraoperative plate selection challenging in some cases. It would be advantageous if preoperative planning could be easily performed using other factors. Estimating the morphology of the distal radius and determining the plate size from the patient's preoperative physical characteristics would reduce the operation time and minimize the risk of size mismatch between the plate and the distal radius.

Significant individual variations in the bone morphology of the distal radius articular surface have been well documented [10]. The radial inclination and palmar tilt in 3D models are both larger in women than in men [11]. The area of the distal radius articular surface is significantly larger in men than in women [11]. In addition, the width of the anterior surface in the coronal view is larger in men than in women, and the curved part of the anterior surface in men is longer and more concave than that in women [12]. The literature discusses the relationship between physical characteristics and radial morphology. Park et al. assessed the two-dimensional morphology of the distal radius on magnetic resonance imaging (MRI), revealing positive correlations between height and both transverse and anteroposterior diameters [13]. Zenke et al. investigated the distance from the extensor pollicis longus (EPL) groove to the volar cortical line of the distal radius and found a positive correlation with height [14]. Sex-based differences in the morphology of the radius have also been reported [12,15–17]. However, no study has investigated the association between physical characteristics and the morphology of the distal radius in men and women separately.

In this study, we hypothesized that there is a correlation between physical characteristics (such as height, weight, and body mass index [BMI]) and the morphology of the distal radius articular surface (such as transverse and anteroposterior diameters). If the size of the distal radius can be estimated from preoperative physical characteristics, it would allow for more convenient preoperative planning, reduce the necessity of additional imaging tests, and potentially shorten operation times. The aim of this study was to investigate the relationship between patients' physical characteristics and the morphological characteristics of the distal radius joint surface using CT images of healthy wrist joints.

2. Methods

The study protocol was approved by an institutional review board (approval No. T2022-0041). This was a retrospective case–control study (level of evidence: III). A radiographic database was accessed to identify patients who underwent CT scans of the unaffected wrist for comparison with those of the affected side. From the database, we evaluated CT images of the unaffected wrist between January 2016 and August 2022. The absence of previous history or complaints in the unaffected wrist was confirmed through interviews and medical records. Patients with a history of traumatic arm injuries were excluded from this study. Patients younger than 18 years were also excluded from the study. A total of 79 wrist CT images, including those of 41 women and 38 men (age range: 20–95 years, mean age: 58.4 years for men; age range: 26–91 years, mean age: 62.5 years for women), were evaluated. Physical characteristics such as height, weight, and BMI were recorded for each patient at the initial hospital visit, and these data were retrieved from medical records.

2.1. Three-Dimensional Bone Morphology and Analysis

CT imaging and analysis of the 3D bone model of the distal radius were performed as previously described [18]. Using computer analysis software (BoneSimulater, Orthree, Osaka, Japan), we defined the long axis of the radius. The plane containing the long axis and the radial styloid process was defined as the coronal plane. The plane containing the long axis and perpendicular to the coronal plane was defined as the sagittal plane. The plane perpendicular to the long axis was defined as the axial plane. Subsequently, three reference points, (1) the radial styloid process, (2) the volar edge of the sigmoid notch, and (3) the dorsal edge of the sigmoid notch, were marked (Figure 1a). The 3D coordinates of each reference point were assessed using the 3D images. These reference points were utilized because of their high inter-rater reliability and reproducibility in our preliminary study [18]. The vertical distance between reference points (1) and (2) on the coronal plane was defined as the radius transverse diameter (Figure 1b), and the vertical distance between reference points (2) and (3) on the sagittal plane was defined as the radius anteroposterior diameter (Figure 1c). The 3D analysis was performed by an experienced hand surgeon.

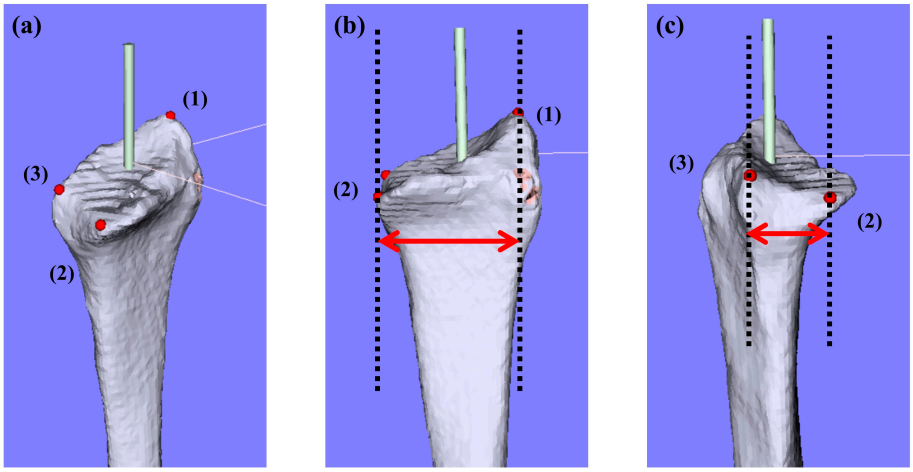


Figure 1. Analysis of the 3D bone model of the distal radius. Three reference points were identified: (1) the radial styloid process, (2) the volar edge of the sigmoid notch, and (3) the dorsal edge of the sigmoid notch (**a**). The radius transverse diameter was measured as the vertical distance between points (1) and (2) on the coronal plane (**b**), while the radius anteroposterior diameter was measured as the vertical distance between points (2) and (3) on the sagittal plane (**c**).

2.2. Statistical Analysis

The results are presented as mean ± standard deviation (SD). The mean values were compared between the sexes using Welch's *t*-test. The coefficient of variation (CV) was calculated as SD/mean. Pearson's correlation coefficient (r value) was used to assess correlations between height, weight, BMI, and transverse and anteroposterior diameters of the distal radius. Additionally, correlation coefficients were calculated separately for men and women. Correlations were interpreted as weak ($0.1 \leq |r| \leq 0.3$), moderate ($0.4 \leq |r| \leq 0.6$), or strong ($0.7 \leq |r|$) [19]. A p-value <0.05 was considered statistically significant. Statistical tests were performed using R software version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

Table 1 summarizes the measurements of the physical characteristics of the patients and the morphology of the distal radius. The mean values for height and weight were larger in men than in women ($p < 0.001$). The mean value for BMI did not differ between the sexes ($p = 0.258$). In both men and women, the CVs for weight and BMI were greater than those for height. The transverse and anteroposterior diameters of the distal radius were larger in men than in women ($p < 0.001$), but there were no differences in variability.

Table 1. Physical characteristics and distal radius morphology of the patients.

	Overall (n = 79)			Men (n = 38)			Women (n = 41)		
	Mean ± SD		CV	Mean ± SD		CV	Mean ± SD		CV
Height	161.3	± 8.8	0.05	167.4	± 7.5	0.04	155.7	± 5.7	0.04
Weight	59.6	± 14.2	0.24	65.9	± 15.8	0.24	53.7	± 9.5	0.18
BMI	22.8	± 4.3	0.19	23.3	± 4.4	0.19	22.2	± 4.2	0.19
Transverse diameter	25.9	± 2.7	0.10	27.8	± 2.3	0.08	24.2	± 1.7	0.07
Anteroposterior diameter	13.7	± 1.4	0.10	14.5	± 1.3	0.09	12.9	± 1.0	0.08

The correlation coefficients are listed in Table 2. Scatter plots of height, weight, and BMI versus transverse and anteroposterior diameters are shown in Figure 2. In the overall analysis, a moderate to strong correlation was observed between height (x) and transverse diameter (y) (r = 0.66, y = 0.20x − 6.67), and moderate correlations were observed between height (x) and anteroposterior diameter (y) (r = 0.45, y = 0.07x + 2.05) and between weight (x) and transverse diameter (y) (r = 0.41, y = 0.08x + 21.3), all of which were statistically significant ($p < 0.001$). In the analysis by sex, a moderate correlation was found between height (x) and transverse diameter (y) in women, which was also statistically significant (r = 0.47, y = 0.14x + 2.41, $p = 0.002$).

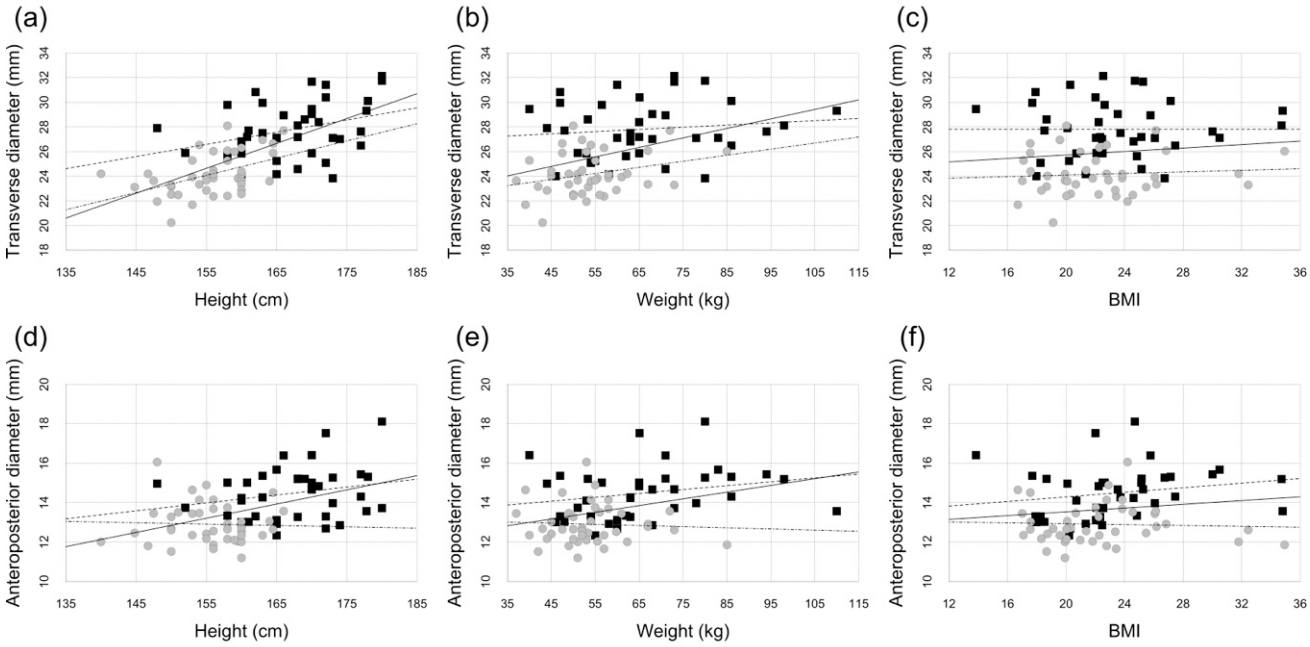


Figure 2. Scatter plots of physical characteristics versus distal radius morphology. Scatter plots show the relationship between transverse diameter and height (**a**), weight (**b**), and BMI (**c**), as well as between anteroposterior diameter and height (**d**), weight (**e**), and BMI (**f**). The male patients are represented by black squares, while the female patients are represented by grey circles. The regression line representing the overall correlation is depicted as a solid line, the regression line for men as a dashed line, and the regression line for women as a dot-dashed line.

Table 2. Correlation coefficients between physical characteristics and distal radius morphology.

		Height		Weight		BMI	
		r	***p*-Value**	**r**	***p*-Value**	**r**	***p*-Value**
Transverse diameter	Overall	**0.66**	<0.001	**0.41**	<0.001	0.11	0.325
	Men	0.33	0.046	0.12	0.459	0.00	1.000
	Women	**0.47**	0.002	0.27	0.084	0.08	0.628
Anteroposterior diameter	Overall	**0.45**	<0.001	0.35	0.002	0.15	0.197
	Men	0.23	0.168	0.24	0.154	0.20	0.238
	Women	−0.04	0.808	−0.06	0.731	−0.05	0.777

Numbers in bold indicate significant correlations.

4. Discussion

We investigated the relationship between physical characteristics and the transverse and anteroposterior diameters of the articular radial surface in a single institution. Statistically significant correlations were observed overall between transverse diameter and height, transverse diameter and weight, and anteroposterior diameter and height. In the analysis by sex, a moderate correlation was found only between transverse diameter and height in women.

Several studies have thoroughly investigated and reported the intricate relationship between the physical characteristics and the morphological features of the distal radius, as examined in our current study. These studies have aimed to understand how various physical attributes, such as height, influence the anatomy and dimensions of the distal radius [13,14]. For instance, sex identification using radial length has been reported to have a sensitivity of 83% and a specificity of 96% [15]. Additionally, one study indicated that the size of the radial head is useful for sex identification [16], although these studies were mostly conducted within the field of archaeology. From a more clinical perspective, reports highlight the differences in distal radius morphology between sexes [12] and suggest that it would be beneficial to create separate VLPs for men and women [1]. In our study, we observed that the patients' overall height showed a statistically significant correlation with both the transverse and anteroposterior diameters of the distal radius, with correlation coefficients of 0.66 and 0.45, respectively. These findings suggest a meaningful association between a person's height and the size of the distal radius. Previous studies have also obtained similar findings, indicating a connection between height and radial dimensions [13,14]; however, it is worth noting that their reference points and methodologies were different from ours, leading to variations in the results. Further investigation into the relationship between height and the morphology of the radius revealed a positive association between height and radial length [20], highlighting the potential influence of overall stature on radial dimensions. Additionally, there is a well-documented strong correlation between radial length and the transverse diameter of the distal radius, with a particularly high correlation coefficient of 0.753 reported in a study [21]. This underscores the potential interconnectedness of these anatomical features. Given these findings, it is plausible to hypothesize that height is indeed correlated with the transverse diameter of the distal radius. However, it appears that radial length may have an even stronger correlation with the transverse diameter than height alone. This suggests that while height is an important factor, radial length could serve as a more precise predictor of the transverse diameter of the distal radius. Consequently, understanding these relationships can be crucial for clinical practices, such as preoperative planning and surgical interventions, where precise anatomical knowledge is essential for optimal outcomes.

There was a moderate correlation between weight and transverse diameter (r = 0.41) and a weak to moderate correlation between weight and anteroposterior diameter (r = 0.35), both of which were statistically significant. These correlations between weight and both the transverse and anteroposterior radial diameters have not been previously reported. Weight can fluctuate throughout the day, and it is conceivable that individuals of the same

height can have significantly different weights. Comparing the CV results, it is evident that weight exhibited greater variability than height. Considering that height correlates with both transverse and anteroposterior radial diameters and can influence weight, it likely acts as a confounding factor in these results. Further investigation, such as multivariate analysis, is needed to explore this relationship in more detail. The correlations between BMI and both transverse and anteroposterior diameters were weak and not statistically significant, suggesting that body shape is not a reliable indicator for predicting the morphology of the radius.

Focusing on sex differences, this study revealed interesting distinctions in the relationship between physical characteristics and distal radius morphology. It was found that only height and the transverse radial diameter in women exhibited a moderate correlation. While both men and women showed similar coefficients of variation (CVs) for height and transverse diameter, the correlation coefficient was notably higher for women. In other words, although a weak to moderate correlation between height and transverse diameter was observed in men, a stronger and more pronounced association was demonstrated in women. This suggests that in women, height may be a more reliable predictor of transverse diameter, indicating a higher accuracy in predicting the transverse diameter from height in the female population. In contrast, when examining the relationship between height and the anteroposterior diameter, the findings were somewhat different. Although a moderate correlation between height and anteroposterior diameter was observed overall in both sexes, the correlation was only weak in men and, intriguingly, no significant correlation was observed in women. This discrepancy highlights the complex interplay between height and the anteroposterior diameter, which seems to vary significantly between the sexes.

Previous studies have reported a correlation between height and anteroposterior diameter without differentiating by sex [13,14], suggesting a generalized relationship between these parameters. However, the lack of differentiation by sex in these studies may have masked important nuances. The current findings indicate that there may be differences in the growth patterns of the anteroposterior diameter between men and women, underscoring the importance of sex-specific analyses. This necessity of sex-specific analyses has been emphasized in previous research, as certain anatomical features and growth patterns may differ significantly between men and women due to genetic, hormonal, or developmental factors [1,12]. Understanding these differences is crucial for improving clinical assessments, surgical planning, and personalized treatment strategies, ultimately enhancing patient care outcomes.

In this study, the correlation coefficient between height and transverse diameter in women was 0.47, and the coefficient of determination calculated from this was 0.22. These values are still low for predicting transverse diameter from height, suggesting that while multivariate analysis combining various factors is one approach, it is also necessary to explore factors more strongly correlated with the morphology of the distal radius. The value of this study lies in the ability to predict the morphology of the radius using relatively simple anthropometric indicators such as height and weight. Although it was shown that radial length correlates more strongly with transverse diameter than height [21], this might not be useful in cases of bilateral injury. Focusing on the lower limbs, several studies in Asian populations have found a correlation between foot length and height [22,23]. Foot length can be easily measured, and feet are unlikely to be injured simultaneously in distal radius fractures, indicating that foot length may be a potentially useful indicator for estimating the shape of the distal radius in the future studies. To enhance the accuracy of estimating distal radius morphology for preoperative planning and to scale this approach to other clinical cases, several new methods could be explored. Incorporating multiple physical characteristics, such as height, weight, radial length, and other anthropometric data, into predictive models could improve preoperative planning accuracy, particularly in complex cases like bilateral wrist injuries. Advanced image analysis, incorporating machine learning algorithms, could be used to analyze larger datasets of three-dimensional CT scans to detect patterns in bone morphology. Additionally, similar approaches for

correlating physical characteristics with bone morphology could be applied to other bones, such as the humerus or femur, to improve surgical planning for fractures in different areas of the body. These techniques should be considered in future studies.

It should be noted that this study has several limitations. First, this study was conducted at a single institution with a relatively small sample size of approximately 40 cases for each sex. This limits the ability to demonstrate significantly weak correlations. It would be beneficial to increase the sample size across multiple institutions in future studies. Furthermore, the correlation coefficient analysis was univariate, rendering it vulnerable to the impact of outliers. Therefore, it is necessary to develop and validate prediction models for transverse and anteroposterior diameters using multivariate regression analysis. Moreover, 3D analysis was conducted by a single examiner, and the reliability of the assessment could not be evaluated. Nevertheless, our preliminary study, which employed the same 3D analysis method, indicated minimal interexaminer error using intraclass correlation coefficients [18], suggesting that the impact on this study's findings is likely limited.

The findings of this study indicate a correlation between height and transverse diameter of the distal radius, with a particularly moderate correlation observed in women. Further research will be conducted with a larger sample size and with the application of multivariate analyses to predict plate size based on physical characteristics.

In conclusion, this study examined the relationship between the physical characteristics of patients and the transverse and anteroposterior diameters of the radius, as observed through analysis of 3D CT images of the wrist joint. A correlation was observed between the height and the transverse and anteroposterior diameters of the distal radius. Subsequent analysis by sex revealed that this correlation was present between the height and transverse diameter in women.

Author Contributions: R.A. acquired and analyzed the data and wrote the manuscript. A.I. acquired and analyzed the data and revised the manuscript. Y.E. acquired and analyzed the data and revised the manuscript. S.K. acquired and analyzed the data and revised the manuscript. T.O. acquired and analyzed the data, interpreted the data, and revised the manuscript. Y.Y. designed and supervised this study, acquired and analyzed the data, conducted 3D analysis, and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a Grant-in-Aid for Scientific Research (23K08618), the General Insurance Association of Japan, and Terumo Life Science Foundation. These funds were not involved in data collection, data analysis, or the preparation or editing of the manuscript.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Tokyo Medical University (protocol code T2022-0041 and date of approval, 14 December 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The datasets analyzed during the present study are available from the corresponding author upon reasonable request.

Conflicts of Interest: No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References

1. Perrin, M.; Badre, A.; Suh, N.; Lalone, E.A. Analysis of three-dimensional anatomical variance and fit of the distal radius to current volar locking plate designs. *J. Hand Surg. Glob. Online* **2020**, *2*, 277–285. [CrossRef] [PubMed]
2. Schindele, S.; Oyewale, M.; Marks, M.; Brodbeck, M.; Herren, D.B. Three-Dimensionally Planned and Printed Patient-Tailored Plates for Corrective Osteotomies of the Distal Radius and Forearm. *J. Hand Surg. Am.* **2024**, *49*, 277.e1–277.e8. [CrossRef] [PubMed]
3. Sheth, B.; Lavin, A.C.; Martinez, C.; Sabesan, V.J. The use of preoperative planning to decrease costs and increase efficiency in the OR. *JSES Int.* **2022**, *6*, 454–458. [CrossRef]
4. Prijs, J.; Schoolmeesters, B.; Eygendaal, D.; de Vries, J.P.M.; Jutte, P.C.; Doornberg, J.N.; Jaarsma, R.L.; IJpma, F.F.A.; Traumaplatform 3D Consortium. 3D virtual pre-operative planning may reduce the incidence of dorsal screw penetration in volar plating of intra-articular distal radius fractures. *Eur. J. Trauma Emerg. Surg.* **2022**, *48*, 3911–3921. [CrossRef]

5. Totoki, Y.; Yoshii, Y.; Kusakabe, T.; Akita, K.; Ishii, T. Screw Length Optimization of a Volar Locking Plate Using Three Dimensional Preoperative Planning in Distal Radius Fractures. *J. Hand Surg. Asian Pac. Vol.* **2018**, *23*, 520–527. [CrossRef]
6. Moolenaar, J.Z.; Tümer, N.; Checa, S. Computer-assisted preoperative planning of bone fracture fixation surgery: A state-of-the-art review. *Front. Bioeng. Biotechnol.* **2022**, *10*, 1037048. [CrossRef]
7. Atesok, K.; Galos, D.; Jazrawi, L.M.; Egol, K.A. Preoperative Planning in Orthopaedic Surgery. Current Practice and Evolving Applications. *Bull. Hosp. Jt. Dis. (2013)* **2015**, *73*, 257–268.
8. Ehsan, A.; Stevanovic, M. Skeletally mature patients with bilateral distal radius fractures have more associated injuries. *Clin. Orthop. Relat. Res.* **2010**, *468*, 238–242. [CrossRef] [PubMed]
9. de Alencar Neto, J.B.; Jales, C.D.S.; Coelho, J.V.V.; de Souza, C.J.D.; Cavalcante, M.L.C. Epidemiology, classification, and treatment of bilateral fractures of the distal radius. *Acta Ortop. Bras.* **2022**, *30*, e245185. [CrossRef]
10. Smilovic, J.; Bilic, R. Conservative treatment of extra-articular Colles' type fractures of the distal radius: Prospective study. *Croat. Med. J.* **2003**, *44*, 740–745.
11. Ikumi, A.; Yoshii, Y.; Eda, Y.; Ishii, T. Computer-Aided Assessment of Three-Dimensional Standard Bone Morphology of the Distal Radius. *Diagnostics* **2022**, *12*, 3212. [CrossRef] [PubMed]
12. Cho, H.J.; Kim, S.; Kwak, D.S. Morphological Study of the Anterior Surface of the Distal Radius. *Biomed. Res. Int.* **2017**, *2017*, 8963768. [CrossRef] [PubMed]
13. Park, H.Y.; Roh, Y.T.; Min, D.U.; Song, S.W.; Sur, Y.J. Two-dimensional Morphological Characteristics of the Distal Radius on Axial Magnetic Resonance Image and the Effects on Distal Screw Length. *J. Hand Surg. Asian Pac. Vol.* **2017**, *22*, 167–173. [CrossRef]
14. Zenke, Y.; Sakai, A.; Oshige, T.; Moritani, S.; Menuki, K.; Yamanaka, Y.; Furukawa, K.; Nakamura, T. Extensor pollicis longus tendon ruptures after the use of volar locking plates for distal radius fractures. *Hand Surg.* **2013**, *18*, 169–173. [CrossRef] [PubMed]
15. Akhlaghi, M.; Sheikhazadi, A.; Ebrahimnia, A.; Hedayati, M.; Nazparvar, B.; Saberi Anary, S.H. The value of radius bone in prediction of sex and height in the Iranian population. *J. Forensic Leg. Med.* **2012**, *19*, 219–222. [CrossRef]
16. Sakaue, K. Sexual determination of long bones in recent Japanese. *Anthr. Anthropol. Sci.* **2004**, *112*, 75–81. [CrossRef]
17. Thom, M.L.; Willmore, K.; Surugiu, A.; Lalone, E.; Burkhart, T.A. Females Are Not Proportionally Smaller Males: Relationships Between Radius Anthropometrics and Their Sex Differences. *Hand* **2020**, *15*, 850–857. [CrossRef]
18. Yoshii, Y.; Totoki, Y.; Shigi, A.; Oka, K.; Ogawa, T.; Murase, T.; Ishii, T. Computer-Aided Assessment of Displacement and Reduction of Distal Radius Fractures. *Diagnostics* **2021**, *11*, 719. [CrossRef]
19. Dancey, C.; Reidy, J. *Statistics without Maths for Psychology: Using SPSS for Windows*; Prentice Hall: Harlow, UK, 2004.
20. Torimitsu, S.; Makino, Y.; Saitoh, H.; Sakuma, A.; Ishii, N.; Hayakawa, M.; Yajima, D.; Inokuchi, G.; Motomura, A.; Chiba, F.; et al. Stature estimation based on radial and ulnar lengths using three-dimensional images from multidetector computed tomography in a Japanese population. *Leg. Med.* **2014**, *16*, 181–186. [CrossRef]
21. Baumbach, S.F.; Krusche-Mandl, I.; Huf, W.; Mall, G.; Fialka, C. Linear intra-bone geometry dependencies of the radius: Radius length determination by maximum distal width. *Eur. J. Radiol.* **2012**, *81*, 947–950. [CrossRef]
22. Zhang, X.; Wei, Y.; Zheng, L.; Yu, K.; Zhao, D.; Bao, J.; Li, Y.; Lu, S.; Xi, H.; Xu, G.; et al. Estimation of stature by using the dimensions of the right hand and right foot in Han Chinese adults. *Sci. China Life Sci.* **2017**, *60*, 81–90. [CrossRef] [PubMed]
23. Kim, W.; Kim, Y.M.; Yun, M.H. Estimation of stature from hand and foot dimensions in a Korean population. *J. Forensic Leg. Med.* **2018**, *55*, 87–92. [CrossRef] [PubMed]

American Journal of Emergency Medicine 85 (2024) 196–201

Contents lists available at ScienceDirect

American Journal of Emergency Medicine

journal homepage: www.elsevier.com/locate/ajem

Association between helicopter medical services for pediatric trauma patients and mortality: Systematic review and meta-analysis

Yuki Enomoto, MD, PhD [a,b], Yusuke Tsutsumi, MD, MPH, DrPH [c,d,*], Takahiro Kido, MD, PhD [b], Kazuki Nagatomo, MD [a], Asuka Tsuchiya, MD, MPH, PhD [e,f], Yoshiaki Inoue, MD, PhD [a]

[a] Department of Emergency and Critical Care Medicine, University of Tsukuba, Ibaraki, Japan
[b] Department of Pediatrics, University of Tsukuba hospital, Ibaraki, Japan
[c] Department of Emergency and Critical Care Medicine, National Hospital Organization Mito Medical Center, Ibaraki, Japan
[d] Human Health Science, Kyoto University Graduate School of Medicine, Kyoto, Japan
[e] Department of Emergency and Critical Care Medicine, Tokai University School of Medicine, Kanagawa, Japan
[f] Department of Clinical Epidemiology and Health Economics, School of Public Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Article history:
Received 2 July 2024
Received in revised form 3 September 2024
Accepted 8 September 2024

Keywords:
Helicopter emergency medical service
Child
Pediatric
Hems

ABSTRACT

Background: Helicopter emergency medical services (HEMS) have become widespread around the world. However, previous studies of the influence of HEMS on mortality were limited to adult patients only and showed inconsistent and heterogeneous results. This study aimed to examine the association between HEMS and mortality among pediatric emergencies compared to ground emergency medical service (GEMS).
Methods: We searched relevant databases (MEDLINE, EMBASE, The Cochrane Central Register of Controlled Trials) and included articles in any language. The most recent search was on January 4th, 2024. We included prospective observational cohort studies or clinical trials that compared HEMS with GEMS in pediatric patients. We excluded any study that did not compare two or more groups of participants. Two pairs of researchers blindly screened studies and evaluated risk of bias using the Risk of Bias in Nonrandomized Studies of Interventions tool. We conducted this systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. Data were extracted by four independent reviewers. We calculated the odds ratio using the random-effects model. The primary outcome was mortality.
Results: Our search strategy yielded 1454 results. Of these, seven observational studies met our eligibility criteria; no RCT met the criteria. All studies targeted trauma patients only. HEMS was associated with lower mortality (Odds ratio 0.66, 95 % CI 0.59 to 0.74). Inconsistency between trials was determined to be low due to low heterogeneity ($I^2 = 0$ %). In a subgroup analysis conducted with and without physicians on the HEMS staff, we found no significant differences ($I^2 = 0$ %, $p = 0.71$).
Conclusion: Our systematic review and meta-analysis, which was limited to trauma pediatric trauma patients, revealed that HEMS deployment correlated with decreased mortality. Further research is necessary to more effectively measure the potential influence and applicability of HEMS for pediatric emergencies.

1. Introduction

Helicopter emergency medical services (HEMS) have become widespread around the world [1,2]. HEMS may enable rapid intervention for severe trauma and acute illness in pre-hospital settings [3]. The operation of HEMS involves significant costs [4,5]. Furthermore, although HEMS incidents are rare, HEMS can cause serious harm not only to the patient but also to the operating crew and medical staff [6]. Therefore, a clear understanding of the benefits of HEMS is important.

Although several studies examined have the association of HEMS and mortality among trauma [7-10] and non-trauma conditions, including stroke [11] and acute myocardial infarction [12], these studies included adult patients only and their results were inconsistent and heterogeneous [13].

Although HEMS is now frequently also used for pediatric cases [14-21], its usefulness for these patients is not understood. Here, to evaluate evidence on pediatric HEMS, we conducted a systematic review and meta-analysis of existing studies that examined the association

* Corresponding author at: Department of Emergency and Critical Care Medicine, National Hospital Organization Mito Medical Center, 280 Sakuranosato Ibaraki-Machi Higashi-Ibaraki-Gun, Ibaraki 311-3117, Japan.
E-mail address: patachan03@yahoo.co.jp (Y. Tsutsumi).

https://doi.org/10.1016/j.ajem.2024.09.015
0735-6757/

between HEMS and mortality by comparing pediatric patients transported by HEMS with those transported by ground emergency medical services (GEMS).

2. Methods

2.1. Protocol and registration

We conducted a systematic review of published scientific literature following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [22]. The protocol for this study has been registered in the PROSPERO database of systematic reviews (registration number: CRD42021250995).

2.2. Eligibility criteria

We included prospective observational cohort studies or clinical trials that compared HEMS with GEMS in pediatric patients. The definition of "pediatric" was as used from each study.

We included studies that controlled for confounders, such as regression modeling or stratification, and excluded any study that did not compare two or more groups of participants. We included both prospective and retrospective observational studies.

2.3. Information sources and search strategy

We performed a comprehensive search of relevant databases (MEDLINE, EMBASE, The Cochrane Central Register of Controlled Trials) and included articles in any language. Details of the search strategy can be found in our online supplementary materials (Supplemental text). The most recent search was on January 4th, 2024.

2.4. Selection process

Two pairs of researchers (YE and TK; and YT and KN) independently screened both titles and abstracts to exclude irrelevant studies. Discrepancies within a pair were resolved by one of the researchers in the uninvolved pair (i.e. YE for YT and KN; and YT for YE and TK). The same two pairs of researchers screened relevant full-text articles according to the predefined inclusion criteria.

2.5. Data collection process and data items

Two pairs of researchers (YE and TK, YT and KN) independently extracted information from selected studies using a standardized data collection form. Disagreements were resolved by discussion with the authors, with one researcher in the other pair (YE or YT) acting as arbiter if necessary. We extracted the following information: study design, participant demographics, type of emergency, types of HEMS staff, adjusted confounders if the study was observational, outcomes, and the results of adjusted analyses (if applicable). In the case of missing data, we contacted the study's corresponding author.

2.6. Primary and secondary outcomes

The primary outcome was mortality. Secondary outcomes were length of hospital stay and neurological prognosis.

2.7. Study risk of bias and assessment

Two pairs of researchers (YE and TK, YT and KN) independently assessed the risk of bias in the included studies using the Cochrane risk of bias tool 2.0 (RoB 2.0) [23] for randomized-controlled trials (RCTs) and the Risk of Bias In Non-randomized Studies (ROBINS-I) [24] for observational studies. Disagreements were resolved by discussion with the authors, with one researcher in the other pair (YE or YT) acting as arbiter if necessary.

2.8. Effect measures and synthesis methods

We used RevMan software (Review Manager, version 5.4, The Nordic Cochrane Centre, The Cochrane Collaboration, 2008) for data analysis. We calculated odds ratios (ORs) and 95 % confidence interval (CIs) using the random-effects model [25]. To assess heterogeneity,

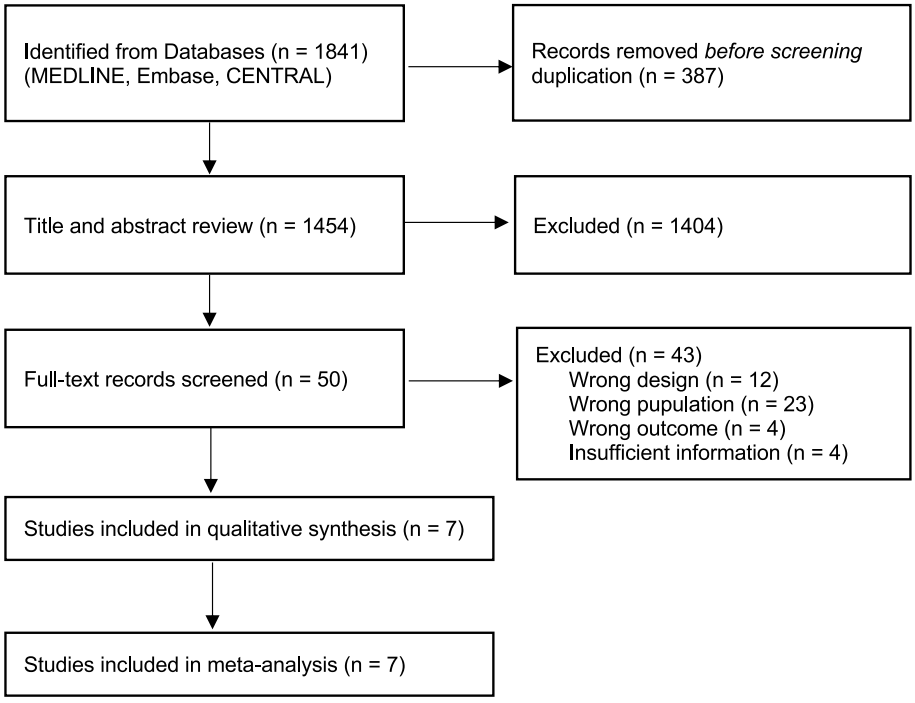


Fig. 1. Study flow diagram.

Table 1
Characteristics of included studies.

Author, year	Country	Physician-staffed	Inclusion criteria	Number of patients	Age, mean (SD)	Mortality follow-up period	Crude Mortality (%)	Mortality-adjusted OR (IQR)	Remarks
Blåsius[14], 2021	Germany, multi-center	yes	Patients ≤15 years and max AIS ≥3	H: 826 G: 1929	H: 9.0 (4.7) G: 9.1 (4.8)	In-hospital	H: 7.9/G: 7.8 ($p = 0.93$)	0.60 (0.33–1.09)	Control group was physician-staffed G
Brown[15], 2016	United States, multi-center	no	Patients ≤15 years old transported from the scene of injury to a level I/ II trauma center	H: 25837 G: 140757	H: 10 (6.7) G: 10 (6.7)	In-hospital	H: 4.3/G: 2.2 ($p < 0.01$)	0.58 (0.42–0.79)	
Englum[16], 2017	United States, multi-center	no	All trauma patients ≤18 years old transferred to a level I/II pediatric center and ISS ≥9	H: 10565 G 26808	N/A	In-hospital	H: 7.4/G: 4.8 ($p < 0.001$)	0.7 (0.6–0.8)	
Enomoto[17], 2020	Japan, multi-center	yes	Participants aged <18 years, and transported to a hospital. Registration criteria for this database were basically patients with max AIS ≥3.	H: 453 G: 5494	H: 10.5 (4.7) G: 10.7 (4.5)	In-hospital	H: 3.8/G: 1.3 ($p < 0.001$)	0.82 (0.42–1.58)	
Missios[18], 2014	United States, multi-center	no	Patients ≤15 years old transported to a level I/II trauma center and hospitalized for traumatic brain injury. Transferred to a level I/II trauma center,	H: 2445 G: 8864	H: 9.2 (4.9) G: 8.4 (5.4)	In-hospital	H: 7.5/G: 3.8 (<0.0001)	0.56 (0.40–0.80)	Group transferred to level I trauma center was included in the analysis
Moors[19], 2019	Netherlands, single center	yes	All consecutive severely injured pediatric patients (age < 18 years and ISS ≥15)	H: 196 G: 112	H: 12 (6.7) G: 13 (5.9)	30 days	H: 26.5 / G: 10.7 ($p < 0.001$)	0.83 (0.31–2.15)	
Polites[20], 2017	United States, multi-center	no	Children ≤8 years of age. Only those transported to an adult or pediatric level I/II trauma center were included.	H: 8218 G: 33305	H: 13.1 (5.3) G:13 (5.4)	In-hospital	H: 9.7 / G: 11.9 ($p < 0.003$)	0.66(0.47–0.93)	ISS ≥15 was included into the analysis.

ISS, injury severity score; AIS, abbreviated injury score, H, HEMS; G, GEMS; SD, standard deviation; OR, odds ratio; IQR, interquartile range.

we performed visual inspection of the forest plots. We also calculated the Chi^2 p-value and I^2 statistic for statistical heterogeneity. We regarded an I^2 statistic of 60 % or above as indicating substantial heterogeneity. To evaluate publication bias, we visually examined funnel plots and then conducted Egger's test for outcomes when ten or more trials were included [26]. We used odds ratio and confidence interval as the effect measure for binary outcomes. We used mean difference for continuous outcomes if the outcomes were measured by the same scale; otherwise, we used standardized mean difference. For studies that reported only the first decimal place of an effect measure, the 95 % confidence interval and point estimate were calculated with the second decimal place set to 0. The 95th percentile value on the near-zero side was set to the value inferred from the point estimate and the 95th percentile value on the contralateral side.

We planned to conduct subgroup analyses based on the type of emergency. We also conducted subgroup analyses based on the type of HEMS staff.

3. Results

3.1. Study characteristics

Our search strategy yielded 1454 results. Of these, 7 observational studies met our eligibility criteria [15-21], whereas no RCT met the criteria (Fig. 1). Table 1 shows the characteristics of all included studies.

All studies were observational studies and all but one [20] used a multi-center database. All studies targeted trauma patients only. One study was limited to children who had traumatic brain injury [19]. In another study, the control group consisted of physician-staffed ground emergency medical services [15]. Four studies were conducted in the United States [16,17,19,21], and one each was conducted in Germany [15], the Netherlands [20], and Japan [18]. All countries except the U.S. used HEMS staffed by physicians. The definition of pediatric patients varied across studies, with three studies using the definition of under 18 years of age, three studies using the definition of under 15 years of age, and one study using the definition of under 8 years of age. The mean age of study patients was generally around school-age. In those studies which allowed for crude comparison, there were no significant differences in age between the two groups. All studies were focused on patients with a moderate to severe condition, although the definition of severe varied (ISS 9 to 15, max AIS ≥3 or transferred to level I or II trauma center in the U.S.).

3.2. Risk of bias in studies

The overall risk of bias for the included studies was moderate or greater (Fig. 2, eFigure1). Two studies were found to have serious bias due to confounding.

3.3. Result of syntheses

For mortality, one study used 30-day mortality [20] while the others used in-hospital mortality. HEMS was associated with lower mortality (OR 0.66, 95 % CI 0.59 to 0.74), although more than 50 % weight was attached to one study (Fig. 3). Inconsistency between trials was determined to be low due to low heterogeneity ($I^2 = 0$ %). Indirectness was determined to be less severe due to the populations and measured outcomes. Only one study examined adjusted hospital length of stay, with a mean difference of −1.49 (95 %CI -4.94 to −1.97) [18], and no study examined neurological prognosis. We were therefore unable to analyze secondary outcomes.

3.4. Subgroup analysis and sensitivity analysis

In the subgroup analysis conducted with and without physicians on the HEMS staff, results showed no significant difference ($I^2 = 0$ %, $p =$

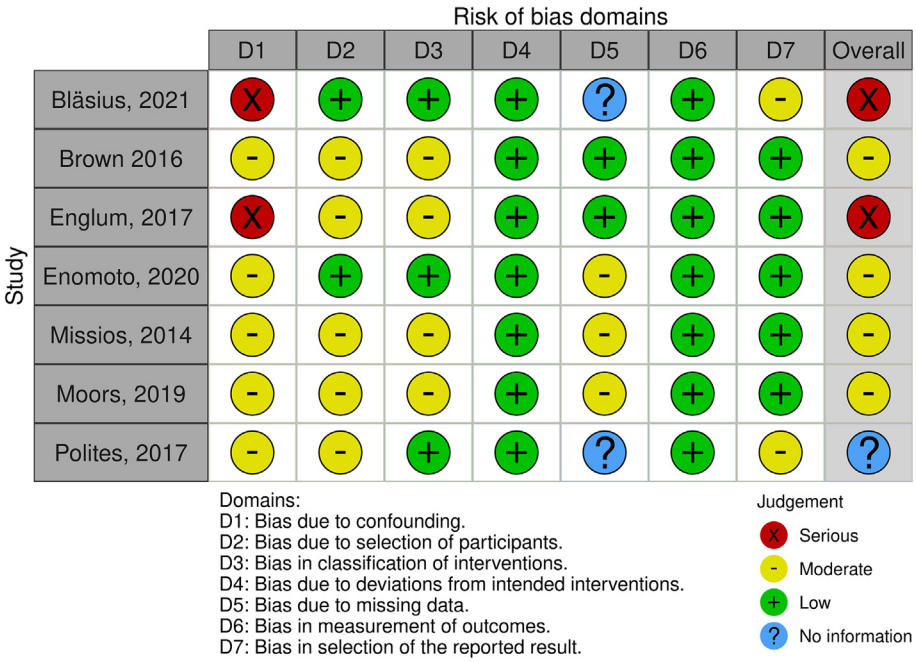


Fig. 2. Risk of bias.

0.71) (eFigure 2). Although not originally planned, we also performed two additional sensitivity analyses. First, we excluded a study that included only TBI [19]. Second, we reduced the number of studies that used data from the NTDB (National Trauma Data Bank): four studies were based on data from the NTDB, with some overlap in study period. We excluded three studies from the analysis to account for potential overrepresentation of the NTDB database [17,19,21], leaving the largest study [16]. Both results were consistent with the main analysis. (OR 0.68, 95 % CI 0.60 to 0.76 and OR 0.62, 95 % CI 0.48 to 0.80) (eFigure 3,4).

3.5. Eporting biases

We did not perform the Egger's test because the number of studies included in the study was less than 10. However, no publication bias was suspected from the funnel plots. (eFigure 5).

4. Discussion

We conducted the first systematic review and meta-analysis to assess the potential influence of HEMS on mortality among severely ill pediatric patients. The results revealed that HEMS utilization in pediatric trauma cases was clearly associated with lower mortality than GEMS. The use of HEMS could be useful in improving the survival outcomes of pediatric emergency patients.

Our systematic review on children included only observational studies for trauma. The small number of pediatric cases and the variety of disease types in pediatric disease may make it impossible to verify effectiveness for a single disease. In fact, HEMS responses for non-traumatic diseases are reported to account for only 20–32 % of all dispatch requests [27,28]. Furthermore, non-trauma cases are characteristically diverse in many regions and vary from drowning, seizures and respiratory failure to cardiac arrest [27-29]. Because it is difficult to compare the effects of different endogenous diseases on a disease-by-disease basis, HEMS studies in non-traumatic pediatric patients have been mainly descriptive studies.

In contrast to the findings of the several systematic reviews conducted in adults, which were strongly heterogeneous and inconclusive [13,30,31], our study found that the use of HEMS for children with trauma was associated with lower mortality. Possible reasons why HEMS may be beneficial for children include the ability to provide

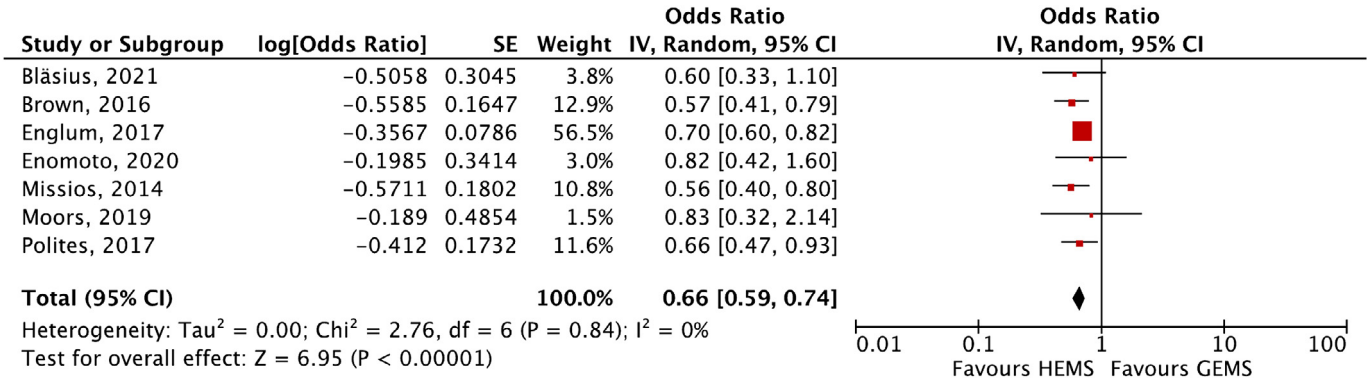


Fig. 3. Effect of helicopter emergency medical service on mortality.
HEMS, helicopter emergency medical service; GEMS, ground emergency medical service; df, degree of freedom.

adequate airway management [32] and triage for appropriate transport hospitals [10,15]. Among children, the most common severe trauma is head trauma [15,27,33]. Consequently, ensuring adequate airway management is often one of the essential procedures for stabilizing a critically ill child with TBI in the field [34]. In the prehospital setting, the number of pediatric critical care cases is generally lower than that of adults, which can make it difficult to decide when to transition to bag-valve-mask ventilation [35]. Additionally, there is insufficient experience with both noninvasive/invasive airway management techniques [35,36]. Given these factors, it is possible that HEMS staff, who more often respond to more critically ill patients, are generally more proficient in airway management techniques [37]. Furthermore, there may be differences in the equipment provided by GEMS and HEMS, as children require appropriately sized items for each age group, which can vary widely. Although there is no certainty as to whether the prehospital intubation procedure itself improves prognosis [37-39], it is possible that more experienced staff can provide appropriate airway management when dealing with patients with airway problems. A second possible reason is that trauma patients have better outcomes when transferred directly to an appropriate trauma center [10,15]. In most countries, fewer trauma centers are certified for children than for adults. On this basis, the effectiveness of HEMS in suitably triaging hospitals and quickly transporting patients to them at distant locations may be more pronounced in the pediatric population [40].

Furthermore, several studies have suggested that physician presence at the scene may have a positive association with patient prognosis [41]. In contrast, our subgroup analysis found no effect of physician staffing on mortality. Regional differences in the healthcare system, procedures that paramedics can perform in the field, and distance to nearby trauma centers may be more important than physician staffing. As mentioned above, tracheal intubation is the key technique in field procedures for pediatric trauma patients, and mortality may not differ as much as in adults when compared against paramedics with adequate tracheal intubation skills [42,43].

Our systematic review and meta-analysis had several strengths. The comprehensive search strategy enabled us to include studies published in languages other than English, broadening the scope of our findings. Integrating studies from different regions allowed us to show the association between HEMS and pediatric patient mortality in regions with different health care systems. By incorporating existing studies, we gained valuable insights in a field in which RCTs are challenging to perform [44,45].

4.1. Limitation

Our study also has several limitations. First, the lack of studies on non-trauma patients precluded examining outcomes of non-trauma cases, and the analysis was accordingly limited to trauma patients. Second, the number of studies that met the eligibility criteria was small; all were observational studies, and no RCTs were included. Acknowledging the difficulty of conducting RCTs in this field, summarizing available results from non-RCTs still holds value. Third, heterogeneity may have arisen due to variations in emergency medical systems across different countries. Fourth, it is possible that the effect size of a single study may be substantial, potentially influencing the overall results. However, given the similarity in point estimates across our included studies, any impact of this bias would have been minimal. Finally, some studies employed data from the NTDB database, yet the outcomes were comparable when the number of studies utilizing the NTDB was reduced for analysis.

The authors consider that although pediatric patients account for a relatively small percentage of prehospital emergencies, the use of HEMS for children is beneficial and important, and that further clarification of the actual situation is warranted. Larger, well-designed studies will provide deeper insights and more extensive data on effectiveness. Additionally, investigating the utility of HEMS in pediatric non-traumatic conditions, such as respiratory or cardiac emergencies, would be valuable. Further research and data collection will facilitate a better understanding of the potential benefits and limitations of HEMS, enabling us to evaluate its benefits for patients and the healthcare system. Moreover, future studies should aim to address potential confounders in order to enhance our understanding of the impact of HEMS on pediatric patients.

5. Conclusion

This systematic review and meta-analysis of pediatric trauma patients revealed that HEMS deployment correlated with decreased mortality. Further research to more effectively measure the impact and applicability of HEMS for pediatric emergencies is necessary.

CRediT authorship contribution statement

Yuki Enomoto: Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Yusuke Tsutsumi:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Takahiro Kido:** Writing – review & editing, Formal analysis. **Kazuki Nagatomo:** Writing – review & editing, Formal analysis. **Asuka Tsuchiya:** Writing – review & editing, Supervision, Resources. **Yoshiaki Inoue:** Writing – review & editing, Supervision.

Declaration of competing interest

Yusuke Tsutsumi reports financial support was provided by Japan Society for the Promotion of Science. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to express our gratitude to Dr. Yasushi Tsujimoto for his valuable advice on the methodology of our research. His expertise and insights helped us to improve the quality and rigor of our work. The authors declare that they have no competing interests. This work was supported by JSPS Grant-in-Aid for Scientific Research(C) Grant Number 21 K10386.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajem.2024.09.015.

References

[1] Masterson S, Deasy C, Doyle M, Hennelly D, Knox S, Sorensen J. What clinical crew competencies and qualifications are required for helicopter emergency medical services? A review of the literature. Scand J Trauma Resusc Emerg Med. 2020;28:28. https://doi.org/10.1186/s13049-020-00722-z.
[2] Tjelmeland IBM, Masterson S, Herlitz J, Wnent J, Bossaert L, Rosell-Ortiz F, et al. Description of emergency medical services, treatment of cardiac arrest patients and cardiac arrest registries in Europe. Scand J Trauma Resusc Emerg Med. 2020;28:103. https://doi.org/10.1186/s13049-020-00798-7.
[3] Mommsen P, Bradt N, Zeckey C, Andruszkow H, Petri M, Frink M, et al. Comparison of helicopter and ground emergency medical service: a retrospective analysis of a German rescue helicopter base. Technol Health Care. 2012;20:49–56. https://doi.org/10.3233/THC-2011-0655.
[4] Meyer MT, Gourlay DM, Weitze KC, Ship MD, Drayna PC, Werner C, et al. Helicopter interfacility transport of pediatric trauma patients: are we overusing a costly resource? J Trauma Acute Care Surg. 2016;80:313–7. https://doi.org/10.1097/TA.0000000000000904.
[5] Madiraju SK, Catino J, Kokaram C, Genuit T, Bukur M. In by helicopter out by cab: the financial cost of aeromedical overtriage of trauma patients. J Surg Res. 2017;218:261–70. https://doi.org/10.1016/j.jss.2017.05.102.
[6] Chesters A, Grieve PH, Hodgetts TJ. A 26-year comparative review of United Kingdom helicopter emergency medical services crashes and serious incidents. J

Trauma Acute Care Surg. 2014;76:1055–60. https://doi.org/10.1097/TA.0000000000000170.
[7] Tsuchiya A, Tsutsumi Y, Yasunaga H. Outcomes after helicopter versus ground emergency medical services for major trauma–propensity score and instrumental variable analyses: a retrospective nationwide cohort study. Scand J Trauma Resusc Emerg Med. 2016;24:140. https://doi.org/10.1186/s13049-016-0335-z.
[8] Brown JB, Gestring ML, Guyette FX, Rosengart MR, Stassen NA, Forsythe RM, et al. Helicopter transport improves survival following injury in the absence of a time-saving advantage. Surgery. 2016;159:947–59. https://doi.org/10.1016/j.surg.2015.09.015.
[9] Bekelis K, Missios S, Mackenzie TA. Prehospital helicopter transport and survival of patients with traumatic brain injury. Ann Surg. 2015;261:579–85. https://doi.org/10.1097/SLA.0000000000000672.
[10] Deeb A-P, Teng CY, Peitzman AB, Billiar TR, Sperry JL, Lu L, et al. Direct trauma center access by helicopter emergency medical services is associated with improved survival after severe injury. Ann Surg. 2023;278:e840–7. https://doi.org/10.1097/SLA.0000000000005812.
[11] Lukovits TG, Von Iderstine SL, Brozen R, Pippy M, Goddeau RP, McDermott ML. Interhospital helicopter transport for stroke. Air Med J. 2013;32:36–9. https://doi.org/10.1016/j.amj.2012.04.002.
[12] Nishigoori S, Kobayashi N, Shibata Y, Shirakabe A, Yagi T, Takano M, et al. Helicopter emergency medical service for patients with acute coronary syndrome: selection validity and impact on clinical outcomes. Heart Vessels. 2022;37:1125–35. https://doi.org/10.1007/s00380-022-02022-1.
[13] Risgaard B, Draegert C, Baekgaard JS, Steinmetz J, Rasmussen LS. Impact of physician-staffed helicopters on pre-hospital patient outcomes: a systematic review. Acta Anaesthesiol Scand. 2020;64:691–704. https://doi.org/10.1111/aas.13547.
[14] Nielsen VML, Bruun NH, Søvsø MB, Kløjgård TA, Lossius HM, Bender L, et al. Pediatric emergencies in helicopter emergency medical services: a National Population-Based Cohort Study from Denmark. Ann Emerg Med. 2022;80:143–53. https://doi.org/10.1016/j.annemergmed.2022.03.024.
[15] Bläsius FM, Horst K, Brokmann JC, Lefering R, Andruszkow H, Hildebrand F, et al. Helicopter emergency medical service and hospital treatment levels affect survival in pediatric trauma patients. J Clin Med. 2021;10:837. https://doi.org/10.3390/jcm10040837.
[16] Brown JB, Leeper CM, Sperry JL, Peitzman AB, Billiar TR, Gaines BA, et al. Helicopters and injured kids: improved survival with scene air medical transport in the pediatric trauma population. J Trauma Acute Care Surg. 2016;80:702–10. https://doi.org/10.1097/TA.0000000000000971.
[17] Englum BR, Rialon KL, Kim J, Shapiro ML, Scarborough JE, Rice HE, et al. Current use and outcomes of helicopter transport in pediatric trauma: a review of 18,291 transports. J Pediatr Surg. 2017;52:140–4. https://doi.org/10.1016/j.jpedsurg.2016.10.030.
[18] Enomoto Y, Tsuchiya A, Tsutsumi Y, Ishigami K, Osone J, Togo M, et al. Association between physician-staffed helicopter versus ground emergency medical services and mortality for pediatric trauma patients: a retrospective nationwide cohort study. PloS One. 2020;15:e0237192. https://doi.org/10.1371/journal.pone.0237192.
[19] Missios S, Bekelis K. Transport mode to level I and II trauma centers and survival of pediatric patients with traumatic brain injury. J Neurotrauma. 2014;31:1321–8. https://doi.org/10.1089/neu.2014.3325.
[20] Moors XRJ, Van Lieshout EMM, Verhofstad MHJ, Stolker RJ, Den Hartog D. A physician-based helicopter emergency medical services was associated with an additional 2.5 lives saved per 100 dispatches of severely injured pediatric patients. Air Med J. 2019;38:289–93. https://doi.org/10.1016/j.amj.2019.04.003.
[21] Polites SF, Zielinski MD, Fahy AS, Wagie AE, Moir CR, Jenkins DH, et al. Mortality following helicopter versus ground transport of injured children. Injury. 2017;48:1000–5. https://doi.org/10.1016/j.injury.2016.12.010.
[22] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al, The PRISMA. Statement: an updated guideline for reporting systematic reviews. BMJ. 2020;2021:n71. https://doi.org/10.1136/bmj.n71.
[23] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019:l4898. https://doi.org/10.1136/bmj.l4898.
[24] Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016:i4919. https://doi.org/10.1136/bmj.i4919.
[25] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177–88. https://doi.org/10.1016/0197-2456(86)90046-2.
[26] Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629–34. https://doi.org/10.1136/bmj.315.7109.629.
[27] Enomoto Y, Tsuchiya A, Tsutsumi Y, Kikuchi H, Ishigami K, Osone J, et al. Characteristics of children cared for by a physician-staffed helicopter emergency medical service. Pediatr Emerg Care. 2018;37:365–70. https://doi.org/10.1097/PEC.0000000000001608.
[28] Barker CL, Weatherall AD. Prehospital paediatric emergencies treated by an Australian helicopter emergency medical service. Eur J Emerg Med Off J Eur Soc Emerg Med. 2014;21:130–5. https://doi.org/10.1097/MEJ.0b013e328362dffa.
[29] Oude Alink MB, Moors XRJ, Karrar S, Houmes RJ, Hartog DD, Stolker RJ. Characteristics, management and outcome of prehospital pediatric emergencies by a Dutch HEMS. Eur J Trauma Emerg Surg Off Publ Eur Trauma Soc. 2022;48:989–98. https://doi.org/10.1007/s00068-020-01579-8.
[30] Galvagno Jr SM, Sikorski R, Hirshon JM, Floccare D, Stephens C, Beecher D, et al. Helicopter emergency medical services for adults with major trauma. Cochrane Database Syst Rev. 2015. https://doi.org/10.1002/14651858.CD009228.pub3.
[31] Tal S, Mor S. The impact of helicopter emergency medical service on acute ischemic stroke patients: a systematic review. Am J Emerg Med. 2021;42:178–87. https://doi.org/10.1016/j.ajem.2020.02.021.
[32] Knapp J, Häske D, Böttiger BW, Limacher A, Stalder O, Schmid A, et al. Influence of prehospital physician presence on survival after severe trauma: systematic review and meta-analysis. J Trauma Acute Care Surg. 2019;87:978–89. https://doi.org/10.1097/TA.0000000000002444.
[33] Osmond MH, Brennan-Barnes M, Shephard AL. A 4-year review of severe pediatric trauma in eastern Ontario: a descriptive analysis. J Trauma Acute Care Surg. 2002;52:8–12.
[34] Badjatia N, Carney N, Crocco TJ, Fallat ME, Hennes HMA, Jagoda AS, et al. Guidelines for prehospital management of traumatic brain injury 2nd edition. Prehosp Emerg Care. 2008;12(Suppl. 1):S1–52. https://doi.org/10.1080/10903120701732052.
[35] Hansen M, Meckler G, O'Brien K, Engle P, Dickinson C, Dickinson K, et al. Pediatric airway management and prehospital patient safety: results of a National Delphi Survey by the Children's safety initiative-emergency medical Services for Children. Pediatr Emerg Care. 2016;32:603–7. https://doi.org/10.1097/PEC.0000000000000742.
[36] Lyng J, Harris M, Mandt M, Moore B, Gross T, Gausche-Hill M, et al. Prehospital pediatric respiratory distress and airway management training and education: an NAEMSP position statement and resource document. Prehosp Emerg Care. 2022;26:102–10. https://doi.org/10.1080/10903127.2021.1992551.
[37] Ohashi-Fukuda N, Fukuda T, Doi K, Morimura N. Effect of prehospital advanced airway management for pediatric out-of-hospital cardiac arrest. Resuscitation. 2017;114:66–72. https://doi.org/10.1016/j.resuscitation.2017.03.002.
[38] Bossers SM, Schwarte LA, Loer SA, Twisk JWR, Boer C, Schober P. Experience in prehospital endotracheal intubation significantly influences mortality of patients with severe traumatic brain injury: a systematic review and Meta-analysis. PloS One. 2015;10:e0141034. https://doi.org/10.1371/journal.pone.0141034.
[39] Schauer SG, Naylor JF, Hill GJ, Arana AA, Roper JL, April MD. Association of prehospital intubation with decreased survival among pediatric trauma patients in Iraq and Afghanistan. Am J Emerg Med. 2018;36:657–9. https://doi.org/10.1016/j.ajem.2017.11.066.
[40] Galvagno SM. Comparative effectiveness of helicopter emergency medical services compared to ground emergency medical services. Crit Care. 2013;17:169. https://doi.org/10.1186/cc12779.
[41] Laverty C, Tien H, Beckett A, Nathens A, Rivest-Caissy JP, da Luz LT. Primary aeromedical retrieval crew composition: do different teams impact clinical outcomes? A descriptive systematic review. CJEM. 2020;22:S89–103. https://doi.org/10.1017/cem.2020.404.
[42] Burns BJ, Watterson JB, Ware S, Regan L, Reid C. Analysis of out-of-hospital pediatric intubation by an Australian helicopter emergency medical service. Ann Emerg Med. 2017;70:773–782.e4. https://doi.org/10.1016/j.annemergmed.2017.03.020.
[43] Vilke GM, Steen PJ, Smith AM, Chan TC. Out-of-hospital pediatric intubation by paramedics: the San Diego experience. J Emerg Med. 2002;22:71–4. https://doi.org/10.1016/S0736-4679(01)00439-5.
[44] Hilton Boon M, Burns J, Craig P, Griebler U, Heise TL, Vittal Katikireddi S, et al. Value and challenges of using observational studies in systematic reviews of public health interventions. Am J Public Health. 2022;112:548–52. https://doi.org/10.2105/AJPH.2021.306658.
[45] Boyko EJ. Observational research — opportunities and limitations. J Diabetes Complications. 2013;27:642–8. https://doi.org/10.1016/j.jdiacomp.2013.07.007.

Journal of Clinical Epidemiology 176 (2024) 111538

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Machine learning approaches to evaluate heterogeneous treatment effects in randomized controlled trials: a scoping review

Kosuke Inoue[a,b,*,1], Motohiko Adomi[c,1], Orestis Efthimiou[d,e], Toshiaki Komura[f], Kenji Omae[g,h], Akira Onishi[i], Yusuke Tsutsumi[j,k], Tomoko Fujii[l,m], Naoki Kondo[a], Toshi A. Furukawa[m]

[a]*Department of Social Epidemiology, Graduate School of Medicine, Kyoto University, Kyoto, Japan*
[b]*Hakubi Center, Kyoto University, Kyoto, Japan*
[c]*Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA*
[d]*Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland*
[e]*Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland*
[f]*Department of Epidemiology, School of Public Health, Boston University, Boston, MA, USA*
[g]*Department of Innovative Research and Education for Clinicians and Trainees, Fukushima Medical University Hospital, Fukushima, Japan*
[h]*Center for Innovative Research for Communities and Clinical Excellence, Fukushima Medical University, Fukushima, Japan*
[i]*Department of Advanced Medicine for Rheumatic Diseases, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[j]*Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[k]*Department of Emergency Medicine, National Hospital Organization Mito Medical Center, Ibaraki, Japan*
[l]*Intensive Care Unit, Jikei University Hospital, Tokyo, Japan*
[m]*Departments of Health Promotion and Human Behavior and of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan*

Accepted 16 September 2024; Published online 19 September 2024

Abstract

Background and Objectives: Estimating heterogeneous treatment effects (HTEs) in randomized controlled trials (RCTs) has received substantial attention recently. This has led to the development of several statistical and machine learning (ML) algorithms to assess HTEs through identifying individualized treatment effects. However, a comprehensive review of these algorithms is lacking. We thus aimed to catalog and outline currently available statistical and ML methods for identifying HTEs via effect modeling using clinical RCT data and summarize how they have been applied in practice.

Study Design and Setting: We performed a scoping review using prespecified search terms in MEDLINE and Embase, aiming to identify studies that assessed HTEs using advanced statistical and ML methods in RCT data published from 2010 to 2022.

Results: Among a total of 32 studies identified in the review, 17 studies applied existing algorithms to RCT data, and 15 extended existing algorithms or proposed new algorithms. Applied algorithms included penalized regression, causal forest, Bayesian causal forest, and other metalearner frameworks. Of these methods, causal forest was the most frequently used (7 studies) followed by Bayesian causal forest (4 studies). Most applications were in cardiology (6 studies), followed by psychiatry (4 studies). We provide example R codes in simulated data to illustrate how to implement these algorithms.

Conclusion: This review identified and outlined various algorithms currently used to identify HTEs and individualized treatment effects in RCT data. Given the increasing availability of new algorithms, analysts should carefully select them after examining model performance and considering how the models will be used in practice.

Keywords: Heterogeneous treatment effect; Individualized treatment effect; Machine learning; Randomized controlled trial; Personalized medicine; Scoping review

Funding: Kosuke Inoue (KI) was supported by grant from the Japan Society for the Promotion of Science (22K17392 and 23KK0240), the Japan Agency for Medical Research and Development (AMED; JP22rea522107), the Japan Science and Technology (JST PRESTO; JPMJPR23R2), the Japan Health Insurance Association, and the Program for the Development of Next-generation Leading Scientists with Global Insight (L-INSIGHT) sponsored by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. Study sponsors were not involved in study design, data interpretation, writing, or the decision to submit the article for publication.

[1] equally contributed.

* Corresponding author. Department of Social Epidemiology, Graduate School of Medicine, Kyoto University, Floor 2, Science Frontier Laboratory, Yoshida-konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.

E-mail address: inoue.kosuke.2j@kyoto-u.ac.jp (K. Inoue).

https://doi.org/10.1016/j.jclinepi.2024.111538

What is new?

Key findings

- In this scoping review, we identified 32 studies focusing on statistical and machine learning (ML) algorithms for assessing heterogeneous treatment effects (HTEs) in randomized controlled trial (RCT) data.
- By the end of 2022, this review categorized 17 studies as application papers and 15 studies as methodology papers.
- Although topics and algorithms varied across the studies, cardiology was the most popular field of application, and the causal forest was the most frequently applied model in healthcare literature.

What this adds to what was known?

- Despite the rapid development of statistical and ML methods to assess HTEs, evidence is limited regarding the commonality of each method's application in clinical research.
- This scoping review extends existing literature by detailing the practical application of various ML methods for HTE assessment in RCTs, offering guidance and example R codes in simulated data for implementation.
- We also described the strengths and limitations of each method, which will help researchers choose appropriate algorithms for investigating HTEs based on their research design and research purposes.

What is the implication and what should change now?

- When investigating HTEs in clinical epidemiology, researchers should carefully select algorithms based on the causal estimands of interest, the performance of the algorithms, and the practical application perspectives.

1. Introduction

Average treatment effect (ATE) is the primary focus of randomized controlled trials (RCTs), because establishing ATE is often required to obtain regulatory approval or change clinical guidelines through informing regulatory agencies and health-care practitioners about the expected treatment effects of interventions in the target population. When applying estimated ATEs to make treatment decisions, we implicitly assume that these estimated effects are applicable to all individuals in the population. This, however, may not be a plausible assumption to make in some cases, for example, when the magnitude and the direction of treatment effect vary substantially according to individual's baseline characteristics [1]. Even when RCTs report no evidence of a treatment effect on average, there may still be some individuals who benefit from treatment. For example, recent post hoc analyses have reported that a certain subpopulation may have a decreased risk of cardiovascular diseases through policy intervention [2], lifestyle intervention [3], intensive glucose control [4], and pharmacological therapy [5], while the original RCT reported a null ATE. Moreover, some participants could be harmed even when ATE indicates beneficial effects. If we only focus on ATE, such patients will miss the opportunity to receive benefit or avoid harm of the treatment. As the concept of personalized medicine has emerged over the years, the importance of assessing heterogeneous treatment effects (HTEs) has been widely recognized [6–8]. HE refers to the situation when the effect of treatment at individual levels, known as conditional average treatment effect (CATE) [9], is different across individuals or across patient subgroups. Estimating CATE allows us to prioritize individuals with high expected benefits from the intervention under the strong assumption that the results are unlikely to be false positive [10,11]. This is implemented in the "*high-benefit approach*" [12] and "*optimal treatment regimes*" [13] in the prior literature.

Over the last decades, a range of statistical and machine learning (ML) methods have been developed for assessing HTE and CATE [14], and have been implemented in open-source software packages such as R and Python [15,16]. Compared to modern methods, traditional approaches like 'one-variable-at-a-time' subgroup analysis have several limitations. The latter considers only one variable to create patient subgroups, and estimate treatment effects therein. Although straightforward and easy to use, the method has several limitations, such as an increased risk of false positives when considering many variables and the loss of statistical power of finding true effects [7]. Moreover, results from the analysis cannot be easily used to guide personalized choice of treatment; for example, if the analysis shows that the treatment is beneficial for males and older patients and harmful for females and younger patients, what about young males or older females? This issue could be addressed by dividing patients into additional groups based on more than one covariate, but such approach would intensify the multiple-testing problem and raise concerns about cherry-picking. Therefore, a principled approach is needed to determine which groups should be considered. To date, a wide range of statistical and ML methods have been proposed, which could be useful in modeling not only linear but nonlinear relationships and high-order interactions between covariates. However, the variety of methods may create confusion among epidemiologists regarding the optimal choice for practical applications. One caveat for the assessment of HTEs is that

the high predictive performance of the model does not necessarily correspond to accurate effect estimation. Although some reviews have summarized the characteristics of currently available methods for HTEs assessment [17−19], the evidence as to how common these methods are in epidemiologic research is limited.

In the Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement, two primary methods are described for assessing HTEs: risk modeling and effect modeling [8]. Risk modeling employs a multivariable approach to predict the outcome risk, followed by stratification of individuals based on the predicted risk. Effect modeling, alternatively, involves the development of models that incorporate interaction terms between treatment and baseline patient characteristics. While both approaches offer unique advantages and should not be exclusively favored, this review emphasizes effect modeling given the rapid advancement in data-driven methods for this approach. Specifically, we aimed to (1) summarize currently available statistical and ML methods for assessing HTEs via effect modeling that have been applied to RCT data and (2) provide a summary of how each algorithm works along with code for implementing it in R, exemplifying it with the use of simulated data. Our overall objective is to provide readers with guidance on how to apply methods to assess HTEs in large clinical RCTs.

2. Methods

This scoping review (ScR) was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Extension for ScRs [20] and the protocol was registered in Open Science Framework on April 2, 2023 [21]. The terminology we used is based on the PATH Statement [8].

2.1. Inclusion and exclusion criteria

Inclusion criteria are as follows: 1) no language restriction; 2) studies published from 2010 until 2022; 3) studies that developed, evaluated, or applied statistical or ML algorithms to predict HTE or CATE; 4) studies that applied an existing method to RCT datasets (ie, data involving random allocation of treatment strategies at individual levels); and 5) studies that conceptualized or modeled treatment strategies based on predicted CATE. In this review, we defined that CATE refers to the CATE, which is the treatment effect conditional on an individual's characteristics. More formally, within the counterfactual framework, CATE can be written as

$$\mathrm{E}[\mathrm{Y}_{t=1} - \mathrm{Y}_{t=0} \mid \mathrm{Z} = \mathrm{z}]$$

where Y_t denotes potential outcome Y under treatment $T = t$, and Z denotes a set of baseline characteristics.

Exclusion criteria are as follows: 1) conference abstract; 2) studies that only used summary data from RCTs; 3) studies that used datasets from cluster-randomized trials, cross-over trials, single-arm trials, or observational studies; 4) studies that only used simple regression-based methods (eg, linear or logistic regression without penalization) even with effect modeling approach such as the metalearner framework; 5) studies that conducted standard subgroup analysis (ie, stratified analysis by a single or some variables such as age, sex, and race); 6) studies that developed risk modeling and stratified patients solely based on prognostic models (ie, assess HTEs based on the predicted risk of outcome); and 7) studies that conceptualized treatment prioritization without specifying HTEs or CATE.

2.2. Search strategy

The search was conducted on two databases: MEDLINE and Embase via OVID. The search terms were determined through the meetings among all authors (Supplementary Tables 1 and 2). The search was performed on March 8, 2023 (MEDLINE), and on April 28, 2023 (Embase).

2.3. Selection of study and data extraction

After all duplicates were removed in the identified studies, six independent reviewers (KI, MA, KO, AO, YT, and TF) screened titles and abstracts. Full texts of the candidate studies were retrieved and underwent full-text screening. We also retrieved citations suggested by the authors of this review. The full-text screening was performed by the two reviewers (KI and MA). Disagreements between reviewers were resolved through consensus-driven discussion. After selecting studies, we categorized them into application studies (which aimed to assess HTEs by applying statistical and ML methods to RCT datasets) and methodological studies (which aimed to propose new methods and used RCT data as an example illustration).

The following basic information was extracted from the included studies after full-text screening and summarized in tables:

- Authors, year of publication (if the study was published online first, the corresponding year was regarded as the year of publication)
- Name, medical area, and sample size of RCT.
- Treatments randomized and outcome examined in RCT.
- Targeted measure of CATE (eg, risk difference, risk ratio, etc.)
- Analysis method used to assess HTEs (eg, decision tree, regularization, targeted learning, etc.)

For application papers, we additionally extracted the information on (i) whether ATE was significant or not in the original RCT and (ii) whether HTEs were identified or not. We also assessed whether each application paper

assessed the calibration and discrimination performance of the model on the treatment effect scale and took some approaches to avoid overfitting.

3. Results

After removing duplicates, a total of 3969 citations were identified in MEDLINE and Embase. After screening titles and abstracts, 3864 citations were excluded as they did not meet eligibility criteria. After seven citations hand-searched by the authors were added, the remaining 112 citations were reviewed in full text, 79 of which were excluded. The main reasons for exclusion were 1) did not use RCT datasets ($N = 29$); 2) did not predict CATE ($N = 13$); and 3) used an outcome prediction approach which was clarified during the full-text review ($N = 9$). Among the 33 citations included in this review, one citation was retracted in September 2023 [22], therefore, we reviewed the remaining 32 articles. The PRISMA flow diagram is shown in Figure 1.

3.1. Study characteristics

A total of 32 studies were included in the review. Of these, 17 (53%) focused on the application of existing algorithms to RCT datasets [3,23–38], and 15 (47%) focused on developing methods for HTE assessment [10,39–52]. Hereafter, we summarize the characteristics of studies separately by study type (ie, application paper or methodology paper).

3.1.1. Application papers

The characteristics of studies that used statistical and ML methods for an applied project ($N = 17$) are shown in Table 1. The most frequently used algorithms were causal forest ($N = 7$) and Bayesian Additive Regression Trees (BART) ($N = 4$). Additional algorithms included XGBoost, penalized regression, SuperLearner, support vector machines, and random forest in metalearner frameworks. The causal forest was applied multiple times to specific RCT datasets such as the Systolic Blood Pressure Intervention Trial [53] or the Action to Control Cardiovascular Risk in Diabetes study [54], and, as result, a total of six studies were in cardiovascular medicine. The remaining studies were in a variety of medical fields: geriatrics, intensive care, neurology, nutrition, psychiatry, respiratory medicine, and sociology. Regarding targeted measures of CATE, risk difference was specified in six studies, odds ratio in two studies, and risk ratio in one study. Other measures included hazard ratio ($N = 3$), difference in survival time ($N = 2$), and difference in the score of a continuous outcome ($N = 6$). The calibration performance of the model was assessed in only four studies, while discrimination performance was not formally assessed in any studies. Most studies (15 in total) employed cross-validation or similar approaches to avoid overfitting.

3.1.2. Methodology papers

The characteristics of studies that developed new methods for CATE ($N = 15$) are shown in Table 2. In most of the included methodological studies, the authors proposed the extension of a pre-existing method and compared the performance of multiple algorithms in simulations, using mean squared error of predicted CATE or population average outcome under the derived individualized treatment rule as performance measures. For example, Spanbauer et al proposed the extension of BART (mixedBART) to incorporate random effects and clustering of outcomes, and

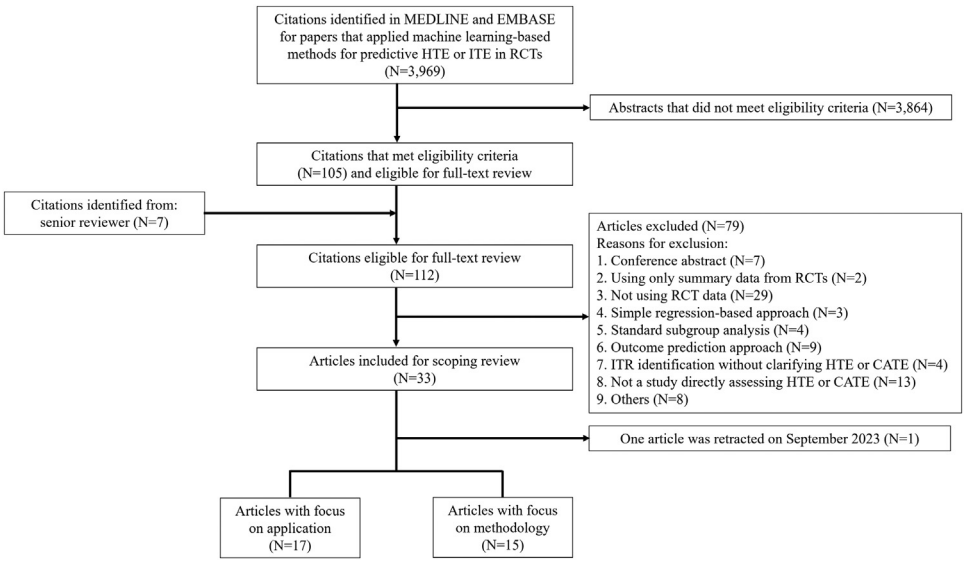


Figure 1. Study flow chart Footnote: HTE, heterogeneous treatment effect; CATE, conditional average treatment effect; RCT, randomized controlled trial; ITR, individualized treatment rule.

Table 1. Characteristics of studies that applied ML algorithms in RCTs

Author	Year	Field	Trial (sample size)	Treatment	Outcome	Causal estimand[a]	Method/Base-learner	Significant ATE reported in the original RCT?	HTE identified?
Edward	2022	Cardiovascular	1. ACCORD (N = 10,251) 2. VADT (N = 1,791)	Intensive glycemic control	MACE	RD	Causal forest	1. No 2. No	Yes
Falet	2022	Neurology	1. OPERA I (N = 821) 2. OPERA II (N = 835) 3. BRAVO (N = 1,331) 4. ORATORIO (N = 661) 5. OLYMPUS (N = 331) 6. ARPEGGIO (N = 318)	Anti-CD20 antibody	Disability progression	Survival time, HR	Deep learning	1. Yes 2. Yes 3. No 4. Yes 5. No 6. No	Yes
Kianmehr	2022	Cardiovascular	1. ACCORD (N = 10,251) 2. ACCORD-BP (N = 4,733)	1. Intensive glycemic control 2. Intensive BP control	Incident heart failure	RD, RR	Causal forest	1. No 2. No	1. Yes 2. Yes
Oikonomou	2022	Cardiovascular	1. SPRINT (N = 9,361) 2. ACCORD-BP (N = 4,733)	Intensive BP control	MACE	HR	XGBoost (along with the Gower method to define phenotypical neighborhood)	1. Yes 2. No	Yes
Sadique	2022	Intensive care	The 65 Trial (N = 2,463)	Permissive hypotension strategy	90-day mortality	RD	Causal forest	No	No
Hu	2021	Respiratory	The National Lung Screening Trial (N = 53,454)	Screening with low-dose CT vs CXR	1. Lung cancer mortality 2. Overall survival	The ratio of survival time	Accelerated failure time model with BART	Yes	Yes
Jiang	2021	Nutrition	IDEA trial (N = 343)	Exercise/diet	1. PCS 2. Weight loss 3. WOMAC scores 4. Compressive force 5. Plasma IL-6 level	RD	1. Penalized regression 2. Kernel ridge regression 3. Random forests 4. Reinforcement learning trees 5. List-based dynamic treatment regime 6. Residual weighted learning 7. BART	Yes	Yes
Kessler	2021	Psychiatry	SUN☺D (N = 1,549)	1. Sertraline only 2. Mirtazapine only 3. Sertraline + Mirtazapine	Depression remission at week 9	Difference in outcomes	SuperLearner	Yes	Yes
Raghavan	2021	Cardiovascular	1. ACCORD (N = 10,251) 2. VADT (N = 1,791)	Intensive glycemic control	All-cause mortality	RD	Causal forest	1. No 2. No	Yes
Sinha	2021	Intensive care	1. ALVEOLI (N = 549) 2. FACTT (N = 1000) 3. SAILS (N = 745)	1. PEEP management 2. Fluid management 3. Rosuvastatin	90-day mortality	OR	Unsupervised learning: 1) K-means clustering 2. Partitioning around medoids 3. Hierarchical clustering 4. Spectral clustering 5. Latent class analysis Supervised learning: 1. Model-based recursive partitioning 2. Causal forest 3. X-learner with Random Forest 4. X-learner with BART	1. No 2. No 3. No	Yes
Furukawa	2020	Psychiatry	SUN☺D (N = 1,549)	1. Sertraline only 2. Mirtazapine only 3. Sertraline + Mirtazapine	Depression remission at week 9	Difference in outcomes	1. Penalized regression (LASSO, ridge) 2. SVM 3. Neural network	Yes	Yes
Shepherd-Banigan	2020	Geriatrics	HI-FIVES (N = 241)	Caregiver education intervention	1. Number of days the veteran was not at home due to medical reason 2.Caregiver depressive symptoms at 12 month	Difference in outcomes	1. Model-based recursive partitioning 2. mCART 3. Random forest	Yes	Yes
Solnick	2020	Sociology	Original RCT (N = 3,277)	A clinical vignette was presented to the participant with a picture of the emergency department physician.	A composite of participant's confidence and satisfaction with the physician	Difference in outcomes	BART	No	No
Foster	2019	Psychiatry	TADS trial (N = 439)	1. Placebo 2. Cognitive-behavioral therapy (CBT) 3. Fluoxetine (FLX) 4. CBT and FLX	CDRS-R	Difference in outcomes	Model-based random forest	Yes	Yes

(*Continued*)

Table 1. Continued

Author	Year	Field	Trial (sample size)	Treatment	Outcome	Causal estimand[a]	Method/Base-learner	Significant ATE reported in the original RCT?	HTE identified?
Scarpa	2019	Cardiovascular	SPRINT (N = 9,361)	Intensive BP control	CV event	HR	Causal forest	Yes	Yes
Furukawa	2018	Psychiatry	1. Trial by Keller et al (N = 681) 2. REVAMP Trial (N = 296) 3. Trial by Schramm et al (N = 59)	1. Cognitive-behavioral analysis system of psychotherapy 2. Antidepressants 3. Combination	1. Depression severity 2. Dropout for any reason	Difference in outcomes, OR	Penalized regression	1. Yes 2. No 3. No	Yes
Baum	2017	Cardiovascular	Look Ahead (N = 5,145)	Weight loss intervention	CV event	RD	Causal forest	No	Yes

BART, Bayesian additive regression trees; BP, blood pressure; CDRS-R, Children's Depression Rating Scale-Revised; CT, computed tomography; CV event, cardiovascular event; CXR, chest X-rays; HR, hazard ratio; IL-6, interleukin-6; LASSO, least absolute shrinkage and selection operator; MACE, major adverse cardiovascular event; mCART, multivariate classification and regression tree; ML, machine learning; MMSE, mini-mental state examination; PCS, physical component score; PEEP, positive end-expiratory pressure; RD, risk difference; RR, risk ratio; SVM, support vector machine; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index; XGBoost, eXtreme Gradient Boosting.

[a] ''Difference in outcomes'' means that the authors investigated the difference in continuous outcomes (specified in ''outcome'' column) between treatment and control groups.

compared the performance of mixedBART and BART using root mean square error of the outcome as a metric in simulations [46]. In another example, Conzuelo Rodriguez et al compared the magnitude of bias of predicted CATE when using doubly robust learners compared to generalized linear models or flexible parametric models with spline functions [41]. In most of the 14 included studies, simulation data was first used to compare the performance of the newly proposed algorithm with pre-existing algorithms, and then an RCT dataset was used to show how the new algorithm could be applied to the real data. A total of 18 RCT datasets were used in 15 studies, and these datasets were from various fields including cardiology, endocrinology, and psychiatry.

3.2. Overview of methods for HTEs assessment

In Tables 3 and 4, we provided a summary of the identified methods for assessing HTE, outlining the way models work, and highlighting their respective strengths and limitations. In the Supplementary Method, we described two algorithms that are most commonly used in the studies identified: penalized regression and causal forest. Additionally, we discussed the Bayesian causal forest algorithm to highlight that, despite being a tree-based method for estimating CATE, it is not the Bayesian counterpart of the causal forest algorithm. We then introduced the general metalearner framework. Subsequently, we explained how to evaluate the calibration of these algorithms.

3.3. Implementation

To demonstrate how each algorithm works to identify HTEs, we provide R code in simulated data. In this implementation, we simulated a hypothetical RCT with 10,000 participants to investigate the effect of intensive blood pressure management on cardiovascular outcomes. Each individual has been attributed a potential outcome, either from the intervention (Y_1) or the placebo (Y_0); that is, Y_1 equals to the observed Y and Y_0 are not observed for the intervention group whereas Y_0 equals to the observed Y and Y_1 are not observed for the placebo group. Outcomes were labeled as 0 in the absence of events and one when events occurred. The treatment effect was calculated by contrasting these potential outcomes, where $\tau = Y_1 - Y_0$. Our data incorporated four baseline covariates, including age (continuous), systolic blood pressure (continuous), estimated glomerular filtration rate (eGFR; continuous), and statin use (binary). We simulated two scenarios of HTEs by setting different treatment effects based on eGFR values and statin use: (i) linear interaction between eGFR and treatment and (ii) nonlinear interaction between eGFR and treatment (ie, eGFR interacted with treatment only between 45 and 90 mL/min/1.73 m^2, and no interaction for eGFR <45 and ≥ 90 mL/min/1.73 m^2). Our code implements penalized regression, causal forest, Bayesian causal forest, and metalearners. The code is available online (https://github.com/Koinoue/HTE_review), and can be used for future implementation of the algorithms.

4. Discussion

In this ScR, we searched for published studies that applied existing or developed new methods for assessing HTEs in RCT data. Although topics and algorithms varied across studies, cardiology was the most popular field of application, and the causal forest was the most frequently applied model in health-care literature. We then outlined the identified algorithms, elucidating their architecture and highlighting their advantages and limitations. For example, penalized regression efficiently selects features and is less computationally intense while causal forest and Bayesian causal forest are less prone to misspecification of the nonlinear complex interaction. Regarding the metalearner framework, S-learner and T-learner are simple approaches while X-learner, DR-learner, and R-learner are particularly effective in scenarios where the sample size of

Table 2. Characteristics of studies that developed ML methods for the HTE assessment

Author	Year	Field	Trial used e.g. illustration (sample size)	Treatment	Outcome	Causal estimand	Methodological contribution
Doubleday	2022	Diabetes	DURABLE trial (N = 1,498)	Twice-daily insulin vs once-daily basal insulin	Change in HbA1c from baseline to week 24	Difference in outcomes	Proposed risk-controlled individual treatment rule (rcITR) estimation using decision tree/random forest
Montoya	2022	Psychiatry	Correctional Intervention for People with Mental Illness "Intervetions" trial (N = 441)	CBT	Recidivism	RD	Provided optimal dynamic treatment rule framework using SuperLearner
Conzuelo Rodriguez	2021	Pregnancy	EAGeR Trial (N = 1,228)	Low-dose aspirin	Live birth	RD	Compared performance between generalized linear models and DR-learner (using SuperLearner)
Du	2021	Cardiovascular	SOLVD-T (N = 2,569)	Enalapril	Time to hospitalization/ death	Difference in survival time	Proposed constrained Lasso approach
Fazzari	2021	Neurology	AADDOPT-2 (N = 569)	Acupuncture	12-week chronic pain	RD	Proposed virtual twin method
Guo	2021	Nutrition	1. Almond consumption trial (N = 68) 2. Avocado consumption trial (N = 108)	Almond/avocado consumption	Composition of GI microbiota and host characteristics	RD	Proposed Multiple Outcome Treatment Effect Forests (MOTEF)
Li	2021	Sociology	Jobs dataset (N = 2,915)	Employment program	Trainee earnings	RD	Proposed causal optimal transport model
Spanbauer	2021	Diabetes Infection	1. TBSI trial (N = 255) 2. ACTG175 Study (N = 1,762)	1. Knowledge/ motivation intervention 2. Didanosine/ azidothymidine treatment	1. HbA1c change from baseline 2. Relative change of CD4 T-cell count	RD	Extended BART by incorporating random effects and clustering for repeated measures (mixedBART)
Chen	2020	Psychiatry	STAR*D (N = 2,555)	SSRIs	Depressive symptoms (HAM-D sum score, QIDS, WSAS, and CGI)	Difference in outcomes	Proposed integrated learning framework using multi-layer neural network
Henderson	2018	Cardiovascular	1. SOLVD-T (N = 2,569) 2. SOLVD-P (N = 4,228)	ACE inhibitor	Time until death/ hospitalization	Difference in expected log-survival	Proposed Bayesian accelerated failure time models
Seibold	2018	Neurology	PRO-ACT (N = 3,306)	Riluzole	Survival time and the ALSFRS at 6 months	Difference in survival time	Proposed personalized models using model-based random forest in time-to-event data
Zhu	2017	Neurology	CATIE-AD (N = 213)	Atypical antipsychotics	Minimal improvement on the CGI scale at 12 weeks	OR	Proposed weighted random forests
Lipkovich	2016	Infection Hematology	Two RCT-datasets, name not specified (N = 470, N = 599)	1. Novel treatment for sepsis 2. Experimental therapy for hematological malignancy	1. All-cause survival at 28 days 2. Overall survival	RD, HR	Illustrated subgroup identification based on 1) Differential effect search (SIDES) 2) Virtual twins' method (VT) 3) Outcome-weighted learning (OWL)
Shen	2016	Cardiovascular	AVID (N = 1,016)	Defibrillator	Two-year Mortality	RD	Proposed Bayesian tree based latent variable model
Weiss	2015	Gastrointestinal	Primary biliary cirrhosis dataset (N = 288)	D-penicillamine	Three-year survival	RD	Compared performance between logistic regression models and AdaBoost

ALSFRS, ALS functional rating scale; BART, Bayesian additive regression trees; CBT, cognitive behavioral therapy; CGI, clinical global impression scale; DR-learner, doubly robust learner; HAM-D, Hamilton depression rating scale; HR, hazard ratio; ML, machine learning; QIDS, quick inventory of depressive symptomatology; OR, odds ratio; RD, risk difference; RR, risk ratio; SSRI, selective serotonin reuptake inhibitors; WSAS, work and social functioning.

Table 3. Summary concept of commonly used algorithms to assess heterogeneous treatment effect

Algorithm	Method paper	Brief description	Strengths	Limitations
Penalized regression (LASSO, ridge, elastic net)	Imai and Ratkovic. 2013 [55]	LASSO is a penalized regression model that shrinks regression coefficients, aiming to maximize predictive performance in new samples. It also performs variable selection, by completely removing some predictors from the model. Ridge is similar to LASSO but does not perform variable selection. Elastic net combines LASSO and ridge. Treatment-covariate interactions can be included in all penalized regression models, to model CATE.	-Feature selection -Simplicity	-Cannot account for interaction by covariates if not prespecified -Difficulty of detecting interactions across a high-dimensional set of covariates.
Causal tree/ causal forest[a]	Wager and Athey. 2018 [56]	Causal forest, a forest-based algorithm, splits samples to maximize the variance in treatment effect estimates across leaves (defined by individual characteristics), employing the R-learner framework to minimize loss function. It adopts an 'honest' estimation approach by using different subsamples for growing trees and estimating CATE, ensuring independence between tree structure and effect estimation.	-Nonparametric (less prone to misspecification of the nonlinear complex interaction) -Embedded estimation of uncertainty	-Computational intensity
Bayesian additive regression trees/ Bayesian causal forest[a]	Hahn et al 2020 [57]	Bayesian causal forest applies two Bayesian additive regression trees functions to evaluate the HTEs. This algorithm calculates the sum of base trees to predict outcomes, and updates the trees to minimize the residual iteratively (MCMC). The framework avoids overfitting and reduces ''regularization-induced confounding'' (which occurs particularly in observational studies).	-Nonparametric (less prone to misspecification of the nonlinear complex interaction) -Embedded estimation of uncertainty	-Computational intensity

CATE, conditional average treatment effect; LASSO, least absolute shrinkage and selection operator; ML, machine learning; MCMC, Markov chain Monte Carlo.

[a] Strengths and limitations of R-learner can also be applied in these methods while Bayesian Causal Forest takes some approaches to consider nonoverlap regions of covariates.

the treatment group is much larger than the other or the covariate distribution is imbalanced which often occurs in observational studies. We provided R code using simulated data; the code can be used to implement the various algorithms, exemplify their use, and facilitate the uptake of these methods in future epidemiological research.

Traditionally, the focus of RCT designs has been on the estimation of ATE in the target population. In some clinical specialties, such as cardiology, multiple studies applied ML-based methods to RCT datasets, indicating the increased interest in HTEs assessment using ML-based methods. Such HTEs assessment via CATE estimation also allows us to create treatment strategies that prioritize individuals who are expected to receive benefit from the treatment (''high-benefit approach'') rather than treating individuals at high risk of developing the outcome (''high-risk approach'') [12,63]. Thus, when designing future RCTs, researchers may prespecify the HTEs assessment methods in the protocol, and include a sufficient set of known or suspected effect modifiers in the study to enrich HTEs assessment. Meanwhile, assessing HTEs via effect modeling typically requires large samples. Given that interaction effects equal to ATE require a sample size four times larger [64], ability to assess HTEs would be limited if we use a single RCT with a small sample, as observed in several studies in our review. In such scenarios, individual participant data metaanalysis could be a viable option to overcome this limitation if data are available [65,66].

There are three important points to note. First, some algorithms assess HTEs on an absolute scale (eg, risk difference, change in score, etc.), but not on a relative scale (eg, risk ratio, odds ratio, hazard ratio). Although assessing HTEs on an absolute scale may be more relevant from a public health perspective—partially because the estimated effects need to be weighed against the harms and costs of the treatment [67,68]—assessing HTEs on a relative scale is also important, particularly when distinguishing between prognostic factors and effect measure modifiers. Indeed, the PATH statement recommends reporting treatment effects in both absolute and relative scale [8]. When researchers want

Table 4. Metalearner framework to assess heterogeneous treatment effect

Algorithm	Method paper	Brief description	Strengths	Limitations
S-learner	Hill et al 2011 [58]	It predicts outcome under treatment and control using base learners that model the interaction between treatment and covariates.	-Simple -Perform better than T-learner when the treatment effect is simple or even zero	-It is not directly optimized to estimate the treatment effect -Risks exclusion of treatment variable from the model when using methods with variable selection as base learners. -Unstable when treatment and control groups are imbalanced.
T-learner	Athey and Imbens. 2016 [59]	It predicts outcomes separately under treatment and control using base learners, and then subsequently calculates the difference in their expected outcome values.	-Simple -Explicit modeling in treatment and control groups, separately - Perform better than S-learner when the treatment effect is strongly heterogeneous	-It is not directly optimized to estimate the treatment effects -Unstable when treatment and control groups are imbalanced.
X-learner, DR-learner	Künzel et al 2019, Kennedy. 2023 [60,61]	It estimates treatment effects on treated patients and on untreated patients using the difference between observed outcomes and estimated counterfactuals for each group. It incorporates propensity score weights to address scenarios with imbalances in covariate distribution (which often occur in observational studies). DR-learner is a similar form of X-learner using a doubly robust estimator instead of propensity score weights.	-Directly estimates heterogenous treatment effects. -Particularly effective in scenarios with imbalanced designs (which often occur in observational studies).	-Complex -Unstable in the presence of extreme propensity scores. -Vulnerable to model misspecification of the propensity scores.
R-learner	Nie and Wager. 2021 [62]	It calculates propensity scores of the exposure and marginal outcomes, calculates the residuals of treatment and outcome, and then minimizes the loss function defined by these residuals. R-Learner requires ML that incorporate some form of regularization for minimizing the loss function.	-Uses different subsamples to estimate the nuisance parameters and to predict CATE by constructing a direct loss function on it.	-Complex -Unstable in the presence of extreme propensity scores. -Vulnerable to model misspecification of the propensity scores.

CATE, conditional average treatment effect; ML, machine learning.

to assess heterogeneity on a relative scale, they may want to use approaches involving the calculation of potential outcomes under treatment and control to obtain such estimands for each individual. Second, once these models are built, it is crucial to check model performance. One thing that complicates this assessment is that a model may be good at predicting absolute outcomes but may nevertheless fail in predicting treatment benefit [69,70]. Thus, assessing model calibration and discrimination, as in simple prediction models, is not enough. Moreover, while S-learner and T-learner predict outcomes among treated and untreated, some models such as causal forest and Bayesian causal forest directly predict CATE which further complicates this assessment. Recently, a range of methods and measures was developed specifically for assessing performance of models for predicting CATE [71–74]. It is also crucial that such an evaluation avoids issues related to overfitting. One way to do this is via using resampling methods (eg, bootstrapping) or data splitting methods such as k-fold cross-validation [73]. While we need more comprehensive discussion on how to evaluate the comparative performance of each method, a standard checklist would be helpful for authors to report these analyses in future applications. Lastly, our review only covers effect modeling approach to assess HTEs. In general, the effect modeling approach is prone to several pitfalls such as overfitting, low statistical power, and multiplicity owing to using multiple treatment interactions [75]. Furthermore, it sometimes lacks sufficient prior knowledge of critical effect modifiers. When these issues cannot be avoided despite employing some statistical approaches such as penalization, consideration of the risk modeling approach—another useful approach for analyzing HTEs in RCT data—is recommended.

Our study has several limitations. First, our review focused on the applications to RCTs and did not consider observational studies. Focusing on randomized datasets helped simplify and clarify the differences among the existing approaches. Given the recent development of approaches

to identify HTEs in observational studies [76,77], future work is needed targeting observational studies in addition to RCTs. Second, our review has covered the literature up to January 2023 only, while the number of publications in ML-based assessments of HTEs in health-care literature has been increasing steadily and some methods (eg, model-based forests) have been newly developed and extended after 2023. However, such is inevitable for any review of a hot and rapidly developing topic, and we believe we were able to cover essential methods. Third, several studies were excluded from this review because they focused on the identification of subgroups rather than CATE estimation. One example is the paper that applied interaction trees to identify qualitative interactions (a type of interaction where, one treatment is better than the other for some subgroups of patients, whereas the reverse is true for other subgroups) in the study population [78]. Such methods are more suitable when the primary objective is to identify specific subgroups with large (or small) treatment effects, rather than CATE. Lastly, our study did not aim to compare model performance across algorithms. While the causal forest algorithm was most frequently applied, our results do not necessarily suggest that the algorithm performs better than the others. Further simulation studies and prospective studies are required to externally validate each algorithm's performance (including the performance of algorithm-based prioritization of treatment) and assess their comparative strengths and limitations.

Due to the increasing availability of statistical and ML methods for assessing treatment effects at the individual level, epidemiologists should carefully select algorithms based on the causal estimands of interest, the performance of the algorithms, and the practical application perspectives.

Ethics statement

Not required.
Consent to participate: Not applicable.
Consent to publish: Not applicable.

Transparency statement

The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

CRediT authorship contribution statement

Kosuke Inoue: Writing − review & editing, Writing − original draft, Supervision, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Motohiko Adomi:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Orestis Efthimiou:** Writing − review & editing, Writing − original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Toshiaki Komura:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kenji Omae:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Akira Onishi:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yusuke Tsutsumi:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tomoko Fujii:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Naoki Kondo:** Writing − review & editing, Writing − original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Toshi A. Furukawa:** Writing − review & editing, Writing − original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Data availability

All data, protocols, and statistical codes are available either through the manuscript and supplementary materials, the Open Science Framework (https://osf.io/3fqgh/), or GitHub (https://github.com/Koinoue/HTE_review).

Declaration of competing interest

All authors have completed the ICMJE uniform disclosure format (available on request from the corresponding author). M.A. received financial supports from Kyoto University related to this work. There are no competing interests for any other author.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2024.111538.

References

[1] Varadhan R, Segal JB, Boyd CM, Wu AW, Weiss CO. A framework for the analysis of heterogeneity of treatment effect in patient-centered outcomes research. J Clin Epidemiol 2013;66(8):818−25. https://doi.org/10.1016/j.jclinepi.2013.02.009.

[2] Inoue K, Athey S, Baicker K, Tsugawa Y. Heterogeneous effects of Medicaid coverage on cardiovascular risk factors: secondary analysis of randomized controlled trial. BMJ 2024;386:e079377. https://doi.org/10.1136/bmj-2024-079377.

[3] Baum A, Scarpa J, Bruzelius E, Tamler R, Basu S, Faghmous J. Targeting weight loss interventions to reduce cardiovascular complications of type 2 diabetes: a machine learning-based post-hoc analysis of heterogeneous treatment effects in the Look AHEAD trial. Lancet Diabetes Endocrinol 2017;5(10):808−15. https://doi.org/10.1016/S2213-8587(17)30176-6.

[4] Kiyohara K, Kondo N, Iwami T, Yano Y, Nishiyama A, Node K, et al. Heterogeneous effects of intensive glycemic and blood pressure on cardiovascular events among diabetes by living arrangements. J Am Heart Assoc 2024;13(13):e033860. https://doi.org/10.1161/JAHA.123.033860.

[5] Desai RJ, Glynn RJ, Solomon SD, Claggett B, Wang SV, Vaduganathan M. Individualized treatment effect prediction with machine learning — salient considerations. NEJM Evid 2024;3(4):EVIDoa2300041. https://doi.org/10.1056/EVIDoa2300041.

[6] Angus DC, Chang CCH. Heterogeneity of treatment effect: estimating how the effects of interventions vary across individuals. JAMA 2021;326(22):2312−3. https://doi.org/10.1001/jama.2021.20552.

[7] Dahabreh IJ, Hayward R, Kent DM. Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. Int J Epidemiol 2016;45(6):2184−93. https://doi.org/10.1093/ije/dyw125.

[8] Kent DM, Paulus JK, van Klaveren D, D'Agostino R, Goodman S, Hayward R, et al. The predictive approaches to treatment effect heterogeneity (PATH) statement. Ann Intern Med 2020;172(1):35−45. https://doi.org/10.7326/M18-3667.

[9] Tipton E. Beyond generalization of the ATE: designing randomized trials to understand treatment effect heterogeneity. J R Stat Soc Ser A Stat Soc 2021;184(2):504−21. https://doi.org/10.1111/rssa.12629.

[10] Lipkovich I, Dmitrienko A, D'Agostino BR Sr. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. Stat Med 2017;36(1):136−96. https://doi.org/10.1002/sim.7064.

[11] van Klaveren D, Balan TA, Steyerberg EW, Kent DM. Models with interactions overestimated heterogeneity of treatment effects and were prone to treatment mistargeting. J Clin Epidemiol 2019;114:72−83. https://doi.org/10.1016/j.jclinepi.2019.05.029.

[12] Inoue K, Athey S, Tsugawa Y. Machine-learning-based high-benefit approach versus conventional high-risk approach in blood pressure management. Int J Epidemiol 2023;52(4):1243−56. https://doi.org/10.1093/ije/dyad037.

[13] Li Z, Chen J, Laber E, Liu F, Baumgartner R. Optimal treatment regimes: a review and empirical comparison. Int Stat Rev 2023;91(3):427−63. https://doi.org/10.1111/insr.12536.

[14] Powers S, Qian J, Jung K, Schuler A, Shah NH, Hastie T, et al. Some methods for heterogeneous treatment effect estimation in high dimensions. Stat Med 2018;37(11):1767−87. https://doi.org/10.1002/sim.7623.

[15] Tibshirani J, Athey S, Friedberg R, Hadad V, Hirshberg D, Miner L, et al. Grf: generalized random forests. Available at: https://cran.r-project.org/web/packages/grf/. Accessed October 22, 2023.

[16] McCulloch R, Sparapani R, Gramacy R, Pratola M, Spanbauer C, Plummer M, et al. BART: bayesian additive regression trees. Available at: https://cran.r-project.org/web/packages/BART/. Accessed October 22, 2023.

[17] Hu A. Heterogeneous treatment effects analysis for social scientists: a review. Soc Sci Res 2023;109:102810. https://doi.org/10.1016/j.ssresearch.2022.102810.

[18] Ling Y, Upadhyaya P, Chen L, Jiang X, Kim Y. Emulate randomized clinical trials using heterogeneous treatment effect estimation for personalized treatments: methodology review and benchmark. J Biomed Inform 2023;137:104256. https://doi.org/10.1016/j.jbi.2022.104256.

[19] Lipkovich I, Svensson D, Ratitch B, Dmitrienko A. Modern approaches for evaluating treatment effect heterogeneity from clinical trials and observational data. Stat Med 2024;43:10167. https://doi.org/10.1002/sim.10167.

[20] Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467−73. https://doi.org/10.7326/M18-0850.

[21] Inoue K, Adomi M, Efthimiou O, Komura T, Omae K, Onishi A, et al. Machine learning approaches to identify individualized treatment effect in randomized controlled trial: a scoping review. Available at: https://osf.io/3fqgh/. Accessed October 22, 2023.

[22] Duan T, Rajpurkar P, Laird D, Ng AY, Basu S. Clinical value of predicting individual treatment effects for intensive blood pressure therapy: a machine learning experiment to estimate treatment effects from randomized trial data. Circ Cardiovasc Qual Outcomes 2019;12(3):e005010. https://doi.org/10.1161/CIRCOUTCOMES.118.005010.

[23] Edward JA, Josey K, Bahn G, Caplan L, Reusch JEB, Reaven P, et al. Heterogeneous treatment effects of intensive glycemic control on major adverse cardiovascular events in the ACCORD and VADT trials: a machine-learning analysis. Cardiovasc Diabetol 2022;21(1):58. https://doi.org/10.1186/s12933-022-01496-7.

[24] Falet JPR, Durso-Finley J, Nichyporuk B, Schroeter J, Bovis F, Sormani MP, et al. Estimating individual treatment effect on disability progression in multiple sclerosis using deep learning. Nat Commun 2022;13(1):5645. https://doi.org/10.1038/s41467-022-33269-x.

[25] Kianmehr H, Guo J, Lin Y, Luo J, Cushman W, Shi L, et al. A machine learning approach identifies modulators of heart failure hospitalization prevention among patients with type 2 diabetes: a revisit to the ACCORD trial. J Diabetes Complications 2022;36(9):108287. https://doi.org/10.1016/j.jdiacomp.2022.108287.

[26] Oikonomou EK, Spatz ES, Suchard MA, Khera R. Individualising intensive systolic blood pressure reduction in hypertension using computational trial phenomaps and machine learning: a post-hoc analysis of randomised clinical trials. Lancet Digit Health 2022;4(11):e796−805. https://doi.org/10.1016/S2589-7500(22)00170-4.

[27] Sadique Z, Grieve R, Diaz-Ordaz K, Mouncey P, Lamontagne F, O'Neill S. A machine-learning approach for estimating subgroup- and individual-level treatment effects: an illustration using the 65 trial. Med Decis Making 2022;42(7):923−36. https://doi.org/10.1177/0272989X221100717.

[28] Hu L, Lin JY, Sigel K, Kale M. Estimating heterogeneous survival treatment effects of lung cancer screening approaches: a causal machine learning analysis. Ann Epidemiol 2021;62:36−42. https://doi.org/10.1016/j.annepidem.2021.06.008.

[29] Jiang X, Nelson AE, Cleveland RJ, Beavers DP, Schwartz TA, Arbeeva L, et al. Precision medicine approach to develop and internally validate optimal exercise and weight-loss treatments for overweight and obese adults with knee osteoarthritis: data from a single-center randomized trial. Arthritis Care Res 2021;73(5):693−701. https://doi.org/10.1002/acr.24179.

[30] Kessler RC, Furukawa TA, Kato T, Luedtke A, Petukhova M, Sadikova E, et al. An individualized treatment rule to optimize probability of remission by continuation, switching, or combining antidepressant medications after failing a first-line antidepressant in a two-stage randomized trial. Psychol Med 2022;52(15):3371−80. https://doi.org/10.1017/S0033291721000027.

[31] Raghavan S, Josey K, Bahn G, Reda D, Basu S, Berkowitz SA, et al. Generalizability of heterogeneous treatment effects based on causal forests applied to two randomized clinical trials of intensive glycemic control. Ann Epidemiol 2022;65:101−8. https://doi.org/10.1016/j.annepidem.2021.07.003.

[32] Sinha P, Spicer A, Delucchi KL, McAuley DF, Calfee CS, Churpek MM. Comparison of machine learning clustering algorithms for detecting heterogeneity of treatment effect in acute respiratory distress syndrome: a secondary analysis of three randomised controlled trials. EBioMedicine 2021;74:103697. https://doi.org/10.1016/j.ebiom.2021.103697.

[33] Furukawa TA, Debray TPA, Akechi T, Yamada M, Kato T, Seo M, et al. Can personalized treatment prediction improve the outcomes, compared with the group average approach, in a randomized trial? Developing and validating a multivariable prediction model in a pragmatic megatrial of acute treatment for major depression. J Affect Disord 2020;274:690−7. https://doi.org/10.1016/j.jad.2020.05.141.

[34] Shepherd-Banigan M, Smith VA, Lindquist JH, Cary MP, Miller KEM, Chapman JG, et al. Identifying treatment effects of an informal

caregiver education intervention to increase days in the community and decrease caregiver distress: a machine-learning secondary analysis of subgroup effects in the HI-FIVES randomized clinical trial. Trials 2020;21(1):189. https://doi.org/10.1186/s13063-020-4113-x.

[35] Solnick RE, Peyton K, Kraft-Todd G, Safdar B. Effect of physician gender and race on simulated patients' ratings and confidence in their physicians: a randomized trial. JAMA Netw Open 2020;3(2): e1920511. https://doi.org/10.1001/jamanetworkopen.2019.20511.

[36] Foster S, Mohler-Kuo M, Tay L, Hothorn T, Seibold H. Estimating patient-specific treatment advantages in the 'treatment for adolescents with depression study.'. J Psychiatr Res 2019;112:61−70. https://doi.org/10.1016/j.jpsychires.2019.02.021.

[37] Scarpa J, Bruzelius E, Doupe P, Le M, Faghmous J, Baum A. Assessment of risk of harm associated with intensive blood pressure management among patients with hypertension who smoke: a secondary analysis of the systolic blood pressure intervention trial. JAMA Netw Open 2019; 2(3):e190005. https://doi.org/10.1001/jamanetworkopen.2019.0005.

[38] Furukawa TA, Efthimiou O, Weitz ES, Cipriani A, Keller MB, Kocsis JH, et al. Cognitive-behavioral analysis system of psychotherapy, drug, or their combination for persistent depressive disorder: personalizing the treatment choice using individual participant data network metaregression. Psychother Psychosom 2018;87(3): 140−53. https://doi.org/10.1159/000489227.

[39] Doubleday K, Zhou J, Zhou H, Fu H. Risk controlled decision trees and random forests for precision Medicine. Stat Med 2022;41(4): 719−35. https://doi.org/10.1002/sim.9253.

[40] Montoya LM, Van Der Laan MJ, Luedtke AR, Skeem JL, Coyle JR, Petersen ML. The optimal dynamic treatment rule superlearner: considerations, performance, and application to criminal justice interventions. Int J Biostat 2023;19(1):217−38. https://doi.org/10.1515/ijb-2020-0127.

[41] Conzuelo Rodriguez G, Bodnar LM, Brooks MM, Wahed A, Kennedy EH, Schisterman E, et al. Performance evaluation of parametric and nonparametric methods when assessing effect measure modification. Am J Epidemiol 2022;191(1):198−207. https://doi.org/10.1093/aje/kwab220.

[42] Du Y, Chen H, Varadhan R. Lasso estimation of hierarchical interactions for analyzing heterogeneity of treatment effect. Stat Med 2021; 40(25):5417−33. https://doi.org/10.1002/sim.9132.

[43] Fazzari MJ, Kim MY. Subgroup discovery in non-inferiority trials. Stat Med 2021;40(24):5174−87. https://doi.org/10.1002/sim.9118.

[44] Guo B, Holscher HD, Auvil LS, Welge ME, Bushell CB, Novotny JA, et al. Estimating heterogeneous treatment effect on multivariate responses using random forests. Stat Biosci 2021;15: 545−61. https://doi.org/10.1007/s12561-021-09310-w.

[45] Li Q, Wang Z, Liu S, Li G, Xu G. Causal optimal transport for treatment effect estimation. IEEE Trans Neural Netw Learn Syst 2023; 34(8):4083−95. https://doi.org/10.1109/TNNLS.2021.3118542.

[46] Spanbauer C, Sparapani R. Nonparametric machine learning for precision medicine with longitudinal clinical trials and Bayesian additive regression trees with mixed models. Stat Med 2021;40(11):2665−91. https://doi.org/10.1002/sim.8924.

[47] Chen Y, Zeng D, Xu T, Wang Y. Representation learning for integrating multi-domain outcomes to optimize individualized treatments. Adv Neural Inf Process Syst 2020;33:17976−86.

[48] Henderson NC, Louis TA, Rosner GL, Varadhan R. Individualized treatment effects with censored data via fully nonparametric Bayesian accelerated failure time models. Biostatistics 2020;21(1): 50−68. https://doi.org/10.1093/biostatistics/kxy028.

[49] Seibold H, Zeileis A, Hothorn T. Individual treatment effect prediction for amyotrophic lateral sclerosis patients. Stat Methods Med Res 2018; 27(10):3104−25. https://doi.org/10.1177/0962280217693034.

[50] Zhu K, Huang Y, Zhou XH. Tree-based ensemble methods for individualized treatment rules. Biostat Epidemiol 2018;2(1):61−83. https://doi.org/10.1080/24709360.2018.1435608.

[51] Shen C, Hu Y, Li X, Wang Y, Chen P, Buxton AE. Identification of subpopulations with distinct treatment benefit rate using the Bayesian tree. Biom J 2016;58(6):1357−75. https://doi.org/10.1002/bimj.201500180.

[52] Weiss J, Kuusisto F, Boyd K, Liu J, Page D. Machine learning for treatment assignment: improving individualized risk attribution. AMIA Annu Symp Proc 2015;2015:1306−15.

[53] SPRINT Research Group, Wright JT Jr, Williamson JD, Whelton PK, Snyder JK, Sink KM. A randomized trial of intensive versus standard blood-pressure control. N Engl J Med 2015;373(22):2103−16. https://doi.org/10.1056/NEJMoa1511939.

[54] ACCORD Study Group, Cushman WC, Evans GW, Byington RP, Goff DC Jr, Grimm RH Jr. Effects of intensive blood-pressure control in type 2 diabetes mellitus. N Engl J Med 2010;362(17):1575−85. https://doi.org/10.1056/NEJMoa1001286.

[55] Imai K, Ratkovic M. Estimating treatment effect heterogeneity in randomized program evaluation. Ann Appl Stat 2013;7(1):443−70. https://doi.org/10.1214/12-AOAS593.

[56] Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. J Am Stat Assoc 2018;113(523): 1228−42. https://doi.org/10.1080/01621459.2017.1319839.

[57] Hahn PR, Murray JS, Carvalho CM. Bayesian regression tree models for causal inference: regularization, confounding, and heterogeneous effects (with discussion). Bayesian Anal 2020;15(3). https://doi.org/10.1214/19-BA1195.

[58] Hill J, Linero A, Murray J. Bayesian additive regression trees: a review and look forward. Annu Rev Stat Its Appl 2020;7(1):251−78. https://doi.org/10.1146/annurev-statistics-031219-041110.

[59] Athey S, Imbens G. Recursive partitioning for heterogeneous causal effects. Proc Natl Acad Sci 2016;113(27):7353−60. https://doi.org/10.1073/pnas.1510489113.

[60] Künzel SR, Sekhon JS, Bickel PJ, Yu B. Metalearners for estimating heterogeneous treatment effects using machine learning. Proc Natl Acad Sci 2019;116(10):4156−65. https://doi.org/10.1073/pnas.1804597116.

[61] Kennedy EH. Towards optimal doubly robust estimation of heterogeneous causal effects. Electron J Stat 2023;17(2):3008−49. https://doi.org/10.1214/23-EJS2157.

[62] Nie X, Wager S. Quasi-oracle estimation of heterogeneous treatment effects. Biometrika 2021;108(2):299−319. https://doi.org/10.1093/biomet/asaa076.

[63] Rose G. Sick individuals and sick populations. Int J Epidemiol 2001; 30(3):427−32. https://doi.org/10.1093/ije/30.3.427.

[64] Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses;: power and sample size for the interaction test. J Clin Epidemiol 2004;57(3):229−36. https://doi.org/10.1016/j.jclinepi.2003.08.009.

[65] Debray TPA, Moons KGM, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RHH, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. Res Synth Methods 2015;6(4):293−309. https://doi.org/10.1002/jrsm.1160.

[66] Seo M, White IR, Furukawa TA, Imai H, Valgimigli M, Egger M, et al. Comparing methods for estimating patient-specific treatment effects in individual patient data meta-analysis. Stat Med 2021;40(6): 1553−73. https://doi.org/10.1002/sim.8859.

[67] Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q 2004; 82(4):661−87. https://doi.org/10.1111/j.0887-378X.2004.00327.x.

[68] VanderWeele TJ, Knol MJ. A tutorial on interaction. Epidemiol Methods 2014;3(1):33−72. https://doi.org/10.1515/em-2013-0005.

[69] Rolling CA, Yang Y. Model selection for estimating treatment effects. J R Stat Soc Ser B Stat Methodol 2014;76(4):749−69. https://doi.org/10.1111/rssb.12043.

[70] Zhao Y, Fang X, Simchi-Levi D. Uplift modeling with multiple treatments and general response types. 2017. Available at: http://arxiv.org/abs/1705.08492. Accessed July 19, 2024.

[71] Chernozhukov V, Chetverikov D, Demirer M, Duflo E, Hansen C, Newey W. Double/debiased/neyman machine learning of treatment effects. Am Econ Rev 2017;107(5):261−5. https://doi.org/10.1257/aer.p20171038.

[72] Van Klaveren D, Steyerberg EW, Serruys PW, Kent DM. The proposed 'concordance-statistic for benefit' provided a useful metric when

modeling heterogeneous treatment effects. J Clin Epidemiol 2018;94:59−68. https://doi.org/10.1016/j.jclinepi.2017.10.021.

[73] Efthimiou O, Hoogland J, Debray TPA, Seo M, Furukawa TA, Egger M, et al. Measuring the performance of prediction models to personalize treatment choice. Stat Med 2023;42(8):1188−206. https://doi.org/10.1002/sim.9665.

[74] Maas CCHM, Kent DM, Hughes MC, Dekker R, Lingsma HF, van Klaveren D. Performance metrics for models designed to predict treatment effect. BMC Med Res Methodol 2023;23(1):165. https://doi.org/10.1186/s12874-023-01974-w.

[75] Kent DM, van Klaveren D, Paulus JK, D'Agostino R, Goodman S, Hayward R, et al. The PATH statement explanation and elaboration document. Ann Intern Med 2020;172(1):W1−25. https://doi.org/10.7326/M18-3668.

[76] Robertson SE, Leith A, Schmid CH, Dahabreh IJ. Assessing heterogeneity of treatment effects in observational studies. Am J Epidemiol 2021;190(6):1088−100. https://doi.org/10.1093/aje/kwaa235.

[77] Segal JB, Varadhan R, Groenwold RHH, Li X, Nomura K, Kaplan S, et al. Assessing heterogeneity of treatment effect in real-world data. Ann Intern Med 2023;176(4):536−44. https://doi.org/10.7326/M22-1510.

[78] Maruo K, Furukawa TA, Noma H, Imai H, Ikeda K, Yamawaki S. Qualitative treatment-subgroup interactions in the antidepressant treatment of major depression: application of QUINT to individual participant data from seven placebo-controlled randomized controlled trials. Pers Med Psychiatry 2020;21-22:100054. https://doi.org/10.1016/j.pmip.2019.100054.

Crohn's & Colitis 360, 2024, **6**, otae054
https://doi.org/10.1093/crocol/otae054
Advance access publication 4 October 2024
Observations and Research

Patient and Healthcare Professional Satisfaction, Acceptability, and Preference Experiences With Mirikizumab Administration for Ulcerative Colitis: An International Survey

David Clemow, PhD, MWC,[*] Christine Radawski, MPH,[*] Joe Milata, BSN, RN,[*] Karla Alaka, MMSc,[*] Theresa Hunter Gibble, PhD, MPH, MS,[*] Adam Schaum, MS,[*] Obi Ezennia, MPH,[*] Nicholas Martinez, MD,[†] Tibor Szaloki, MD, PhD,[‡] Yuka Ito, MD,[§] Danielle Rodriguez, PhD, MPH,[¶] Katherine Kirk, MPH[¶]

[*]Eli Lilly and Company, Indianapolis, IN, USA
[†]Gastroenterology Research of America, San Antonio, TX, USA
[‡]Javorszky Hospital, Vac, Hungary
[§]NHO Mito Medical Center, Ibaraki, Japan
[¶]Evidera, Bethesda, MD, USA

Address correspondence to: David B. Clemow, PhD, Eli Lilly and Company, Indianapolis, IN 46225, USA. Tel: 317-435-5307 (davidclemow@lilly.com).

Background: There is a need to better understand ulcerative colitis (UC) patient and healthcare provider (HCP) treatment satisfaction, acceptability, and preferences.

Methods: Two international, cross-sectional, web-based surveys were conducted among participants of a phase 3 mirikizumab study (NCT03519945). The questions captured moderate-to-severe UC patients' experience, HCPs' perception of patients' experience, and HCPs' own experience with mirikizumab administration through intravenous (IV) infusions and subcutaneous (SC) injections.

Results: Respondents included 93 patients and 42 HCPs from 11 countries. The majority of patients had UC >4 years (74.2%), were bionaive (68%), in remission at the time of the survey (63%). HCPs were primarily from the United States (57%), generally nurses (41%) or gastroenterologists (26%) with ≥6 years of experience in treating UC (57%). Most patients were "very satisfied/satisfied" (IV, 83%; SC, 91%), "completely/somewhat" accepting of mirikizumab administration (IV, 87%; SC, 97%), and agreed that improvement to their UC outweighed any administration dissatisfaction (90%). HCPs' perspectives of patients' experiences were higher: "very satisfied/satisfied" (IV, 93%; SC, 100%); "completely/somewhat" accepting (IV, 90%; SC, 98%). HCPs themselves were "very satisfied/satisfied" (IV, 81%; SC, 95%); gastroenterologists were "very satisfied" (IV, 82%; SC, 82%) more than nurses (IV, 29%; SC, 65%) who were generally at least "satisfied" (IV, 53%; SC, 35%). Two SC and monthly SC injections were "completely acceptable" by the patients (76% and 85%) and per HCPs' perceptions of patients' preferences (69% and 100%).

Conclusions: Both patients and HCPs were satisfied with and accepted mirikizumab IV induction followed by monthly maintenance SC injections. UC improvement outweighed any administration dissatisfaction.

Lay summary

Most patients with ulcerative colitis (UC) and their healthcare providers were satisfied and accepted mirikizumab intravenous infusion and subcutaneous injection, including monthly dosing with 2 injections. Over 90% of patients reported that UC improvement outweighed any administration dissatisfaction.

Key Words: administration, dosing, injection, infusion, mirikizumab

Introduction

Ulcerative colitis (UC) therapy options have significantly advanced during the past decades.[1] A large variety of treatment alternatives are available in different formulations, with different routes and frequencies of administration.[2–5] These include conventional therapy with aminosalicylates, glucocorticoids, immunomodulators, and TNF-α inhibitors as well as the most recent biologics (IL-12/23, IL-23, and integrin inhibitors) and small molecules (Janus kinase inhibitors and sphingosine-1-phosphate receptor agonists) that have improved the management of patients with UC because they addressed specific pathogenetic mechanisms.[4–8]

However, many patients do not respond to induction therapy or lose response during the maintenance treatment period.[9–11] Moreover, some of these targeted therapies may not be suitable for a specific patient due to their medical history or medication risk profile.[2,6,9,12,13] Consequently, finding more effective medications is still a therapeutic need that has not been satisfied.[14]

Received for publication: April 18, 2024. Editorial Decision: August 26, 2024

Mirikizumab is a humanized IgG4 monoclonal antibody directed against the p19 subunit of IL-23, a signaling cytokine involved in the inflammatory cascade associated with UC.[15–17] Mirikizumab was recently approved for the treatment of moderately-to-severely active UC. It is administered intravenously (IV) during induction and subcutaneously (SC) during maintenance.[18,19] Mirikizumab demonstrated significant clinical remission, compared to placebo, at week 12 of the induction trial (24.2% vs 13.3%, $P < 0.001$) and week 40 of the maintenance trial after 52 weeks of continuous treatment (49.9% vs 25.1%, $P < 0.001$), with an acceptable safety profile associated with a positive benefit-risk ratio.[20]

Despite the many years of rigorous clinical research on novel medications, there is little evidence of UC treatment administration preferences and associated satisfaction with treatment administration. Such information is important for making educated treatment decisions, particularly when there are multiple alternatives available.[1] Moreover, perceptual differences between patients and their treating physicians may result in patients having suboptimal treatment.[21,22]

To understand mirikizumab treatment administration satisfaction, acceptability, and preferences, 2 surveys were conducted: (1) a patient survey among patients with UC receiving mirikizumab and (2) a healthcare provider (HCP) survey to capture their perspectives on their patients' experiences as well as their own experiences. The study examined the differences in perceptions about treatment administration between patients and HCPs, explored differences between patients based on key clinical history and sociodemographic differences, and explored patient and HCP beliefs regarding benefit versus mirikizumab administration burden.

Methods

Study Design

This study involved 2 international, cross-sectional, non-interventional, web-based one-time surveys among patients with moderately-to-severely active UC and their treating HCPs. The optional surveys were included as a protocol addendum substudy to LUCENT-3 (AMAP), a phase 3, multicenter, open-label, 160-week, long-term extension study evaluating the efficacy and safety of mirikizumab (NCT03519945).[20] All eligible patients and HCPs were offered the opportunity to complete the surveys, however, it was not a LUCENT-3 study protocol requirement.

The patient survey evaluated patients' experiences with mirikizumab administration. The HCP survey assessed HCPs' perceptions of patients' experiences as well as the HCPs' own experiences. Both surveys were administered from October 2022 to July 2023 and took approximately 30 minutes to complete.

Participants

Participants in the LUCENT-3 study were recruited from among the completers of 2 previous multicenter, double-blind, placebo-controlled mirikizumab studies in patients with moderately-to-severely active UC who were allowed to have prior exposure to a biological agent: (1) the phase 2 induction and maintenance study (AMAC, NCT02589665),[18,23] and (2) the phase 3 withdrawal maintenance study (LUCENT-2 [AMBG], NCT03524092) in which the enrolled patients had completed a previous induction phase 3 study (LUCENT-1 [AMAN], NCT03518086; Figure 1). Mirikizumab was administered during induction as a 300 mg (20 mg/mL) intravenous (IV) infusion once every 4 weeks for 3 total infusions.[18] During maintenance, mirikizumab was administered as a 200 mg subcutaneous (SC) injection delivered in 2 consecutive 100 mg/mL injections every 4 weeks (Figures 2 and 3).

Patients could participate in the survey if they gave written informed consent, were actively enrolled and participating in LUCENT-3, were willing and able to complete the web-based survey, were able to read and enter digital responses to complete the survey, and had access to the internet and a computer, tablet, or mobile device.

HCPs could participate in the survey if they were physicians, nurses, or other site support staff responsible for administering

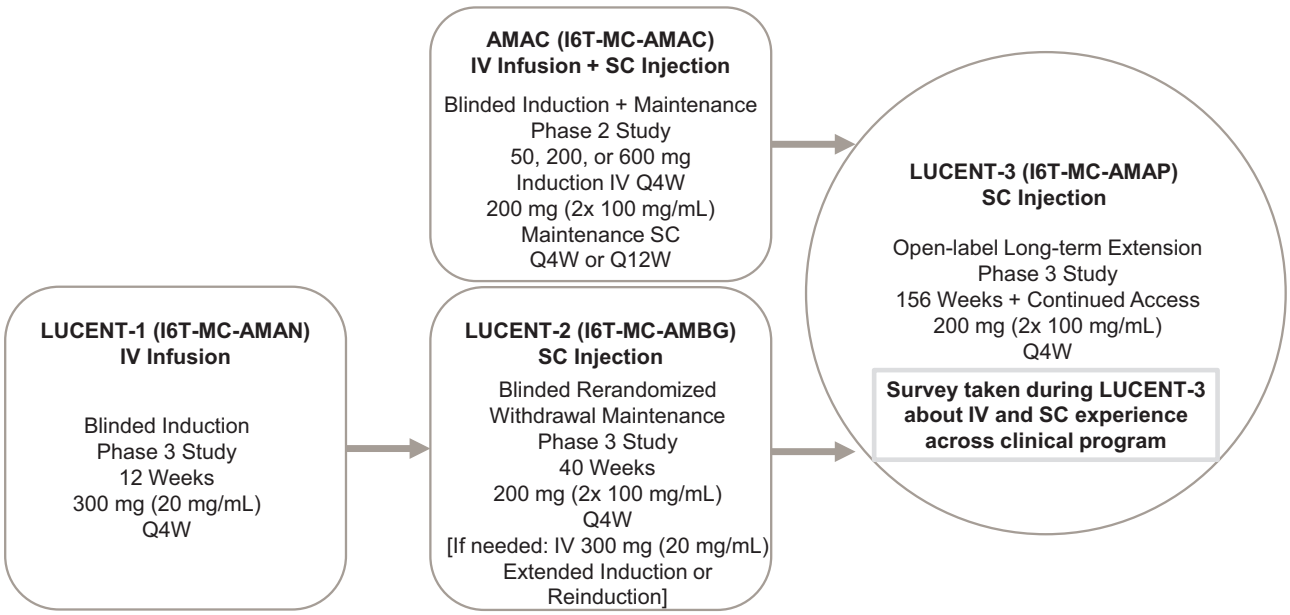


Figure 1. Patient flow and mirikizumab administration experience across the LUCENT clinical program. Abbreviations: IV, intravenous; Q4W, every 4 weeks; Q12W, every 12 weeks; SC, subcutaneous; UC, ulcerative colitis.

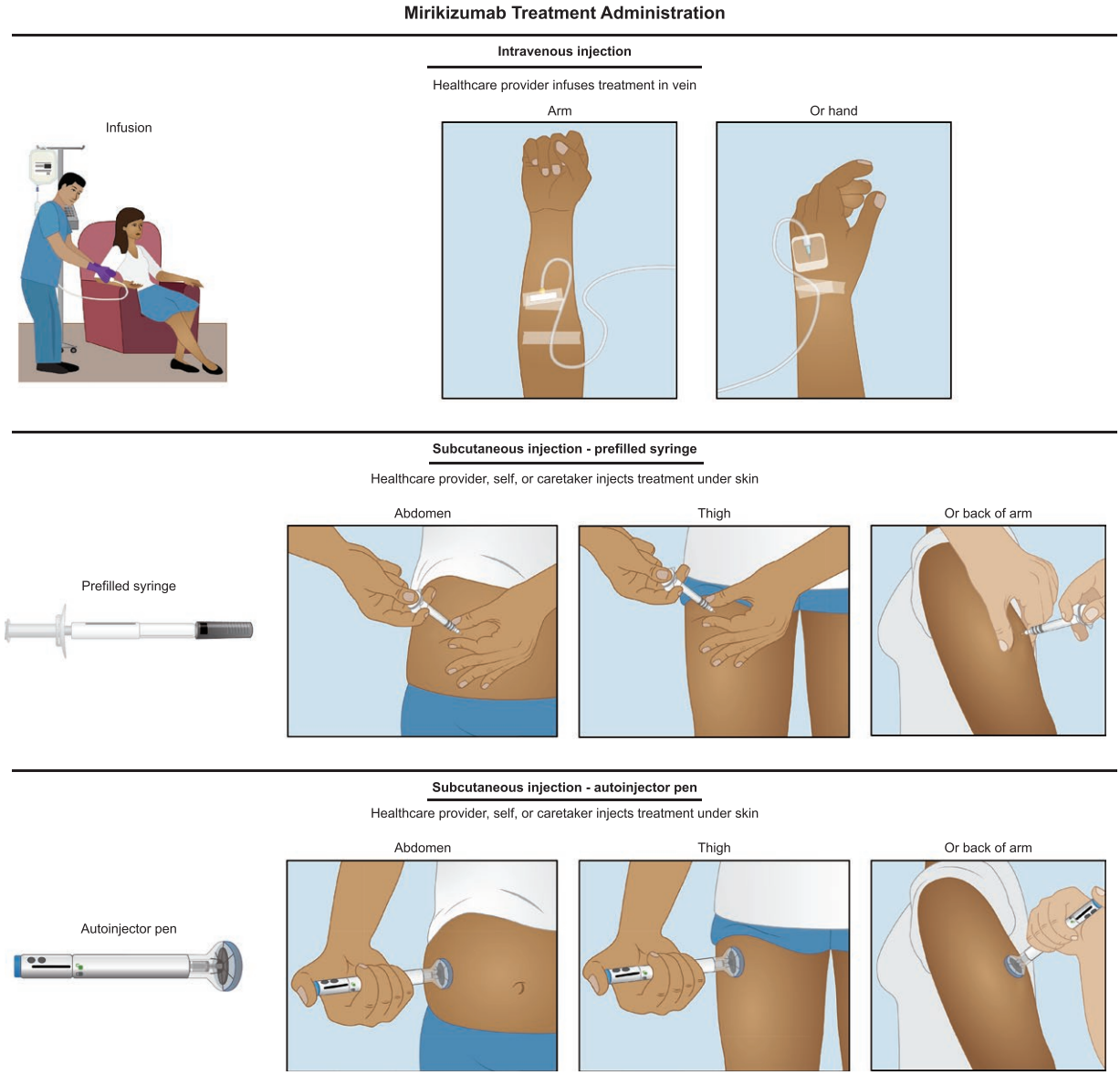


Figure 2. Illustrations for mirikizumab administration as intravenous infusion or subcutaneous injection. Infusion: 300 mg (20 mg/mL). Prefilled syringe or autoinjector pen (200 mg [2× 100 mg/mL]).

or overseeing administration of mirikizumab via IV infusion, SC injection, or both in AMAC, LUCENT-1, LUCENT-2, or LUCENT-3 studies for at least 6 months combined, and also met the criteria noted for patients.

Survey Structure and Content

The survey questions were designed to capture satisfaction with the mode of administration (IV and SC), acceptability of IV and SC, preferences for SC, and overall administration burden versus treatment satisfaction. Patients and HCPs were asked to answer the questions based on their holistic experiences across their participation in the clinical development program. Because IV infusions were not administered in LUCENT-3, patients and HCPs were asked to recall their experiences with IV infusions, administered every 4 weeks, in the induction periods of the LUCENT-3 parent studies. The SC injections' location and device were determined by study protocols across the induction, maintenance, and extension studies. Patients were able to choose to self-inject only after they had completed 7 months in LUCENT-3; until that point, all SC injections were administered by the site staff (HCPs). Moreover, in the survey, the patients were asked to rank the SC injection options—HCP, self-injection, and caregiver—that they experienced during the study; if patients experienced only one method, this was ranked first, and if they experienced all 3 options, they were asked to rank them in order of preference.

Most survey questions were multiple-choice, with responses based on 5-point Likert-type scales. Some questions included ranked response options (e.g., patients' most and least preferred location to receive an SC injection) or a list of relevant response options. The survey was originally developed in English and then translated for each country. The surveys were pretested with 3 patients and 2 HCPs and revised based on their feedback (Supplementary Material, Survey Instrument Development).

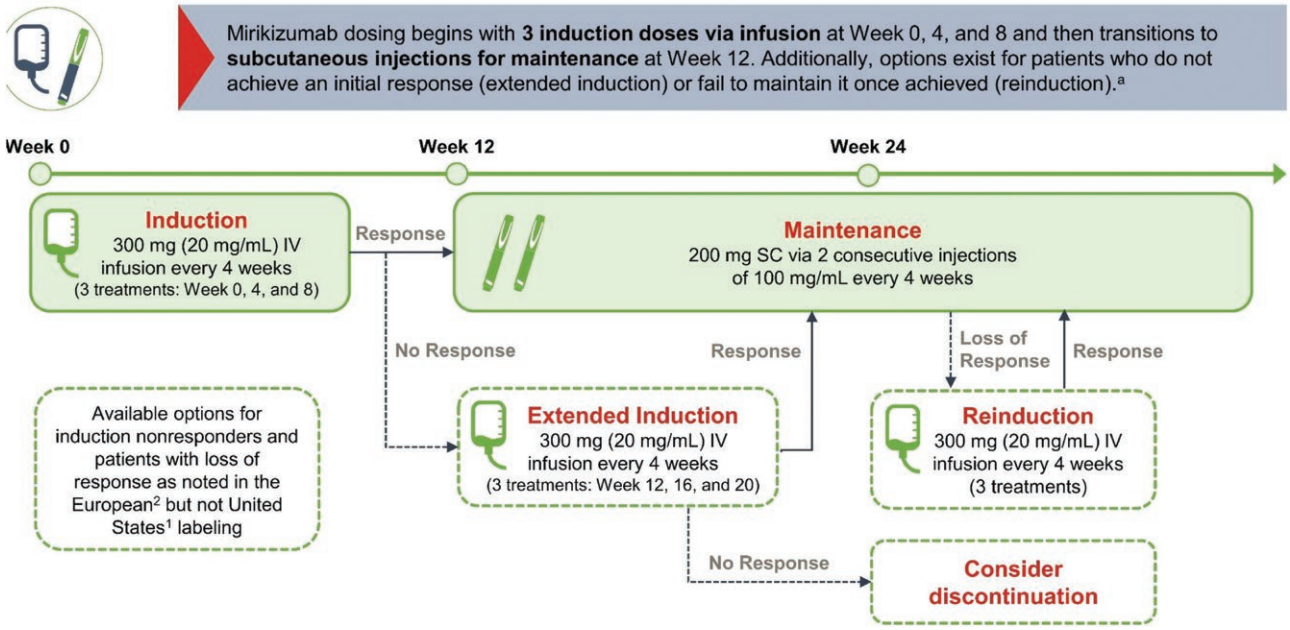


Figure 3. Mirikizumab dosing and administration. Mirikizumab IV infusion occurs over at least 30 minutes. Patients should be monitored for at least 1 hour after dosing, according to the local standard of care. Abbreviations: IV, intravenous; SC, subcutaneous. [a]Source: Mirikizumab summary of product characteristics.[19]

The patient survey (Supplementary Material, Patient survey) consisted of 28 closed-ended questions about participants satisfaction, acceptability, and preferences with different methods of study medication administration (SC or IV) as well as their satisfaction with overall treatment administration and injection device (prefilled syringe or autoinjector pen) usability. The survey did not contain questions on patients' demographic information and their clinical characteristics, as these data were captured as part of the clinical program.

The HCP survey (Supplementary Material, HCP survey) consisted of 32 closed-ended questions about the HCP's perception regarding their patients' treatment satisfaction, acceptability, preferences (Q1-Q25; Supplementary Material, HCP survey), and their personal experience administering the study medication (Q26-Q32; Supplementary Material, HCP survey). At the beginning of the HCP survey, 8 questions captured the HCPs' clinical background, including their specialty, their primary study role, the number of years they have provided care to patients with UC (excluding medical training), their primary practice setting, the percentage of working time spent actively interacting with patients (including patients with UC), their HCP experience with study medication administration methods, the number of infusions or injections they had administered or overseen during the mirikizumab clinical development program, and their country of origin.

Statistical Methods

Descriptive statistics were applied and presented for both patients' and HCPs' responses and were stratified by subgroups. The subgroup analyses were detailed and extended within the patient (Supplementary Table S1) and HCP (Supplementary Table S2) groups of survey participants. For all descriptive analyses, frequency distributions and cross-tabulations were constructed to evaluate and characterize the distribution properties of each variable assessed in the surveys. Chi-square tests were used to evaluate differences between the subgroups for categorical data when specified. Analysis of variance was used to evaluate the differences between subgroups for continuous data when specified. All statistical tests were 2-tailed and were conducted with a type I error probability fixed at 0.05. No formal hypotheses were being tested, and, therefore, no multiplicity adjustment was performed. All data entered by patients and HCPs on the web-based survey platform was provided as a clean, deidentified, fully documented dataset. For subgroup analyses, survey variables were merged with the variables from the clinical trial, including baseline demographics and patient characteristics. The analysis was performed using the SAS Enterprise Guide 7.15 HF6 (SAS Institute Inc., Cary, NC).

Ethical Considerations

Country-specific and, where applicable, site-specific institutional review board (IRB) approval was obtained prior to any data collection efforts. All participants provided written (patients) or electronic (HCPs) informed consent. Data collection complied with ISO 27001 and the European Union Data Protection Directive 95/46/EC.

Results

There were 316 patients who completed the LUCENT-2 maintenance study and entered the LUCENT-3 extension study and an estimated 183 HCPs (investigators and study coordinators) involved in treatment administration participating in the LUCENT-3 study addendum from the 11 of 35 total participating countries (Australia, Austria, Czech Republic, Germany, Hungary, Japan, Mexico, Poland, Spain, Switzerland, and United States) that were invited to voluntarily respond to the survey (Table 1). Of these, survey respondents included 93 patients and 42 HCPs.

Table 1. Patient characteristics, *n* = 93.

Age, mean years (SD)	**43 (11.8)**
Male sex, *n* (%)	51 (55)
Region, *n* (%)	
Europe	41 (44)
North America	25 (27)
Asia	20 (22)
Central or South America	5 (5)
Rest of the world	2 (2)
Race, *n* (%)	
White	65 (70)
Asian	21 (23)
American Indian or Alaska Native	5 (5)
Black or African American	2 (2)
Disease duration ≥ 4 years, *n* (%)	69 (74)
Prior bionaïve before entering mirikizumab program, *n* (%)	63 (68)
Corticosteroid or immunomodulator use, *n* (%)	57 (61)
Symptomatic remission[a] achieved prior to entering LUCENT-3, *n* (%)	59 (63)
MMS Score,[b] *n* (%)	
Moderate (MMS = 4-6)	47 (51)
Severe (MMS > 6)	43 (46)
Missing	3 (3)
Mirikizumab SC injection experience through time of survey, *n* (%)	
>2 years to ≤3 years	32 (34)
>3 years	40 (43)
Missing	21 (23)
Mirikizumab IV infusion experience through time of survey, *n* (%)	
3 IVs (12 weeks)	47 (51%)
6 IVs (24 weeks)	32 (34)
Missing	14 (15)
Mirikizumab self-injection experience through time of survey, *n* (%)	41 (44)
Fear of needles prior to taking study medication, *n* (%)	
Not at all afraid	57 (61)
A little afraid	23 (25)
Moderately afraid	8 (9)
Very afraid	2 (2)
Extremely afraid	3 (3)

[a]Symptomatic remission is based on Modified Mayo Score Stool Frequency (SF) and Rectal Bleeding (RB) components: SF = 0 or SF = 1 with a 1-point decrease in MMS from baseline; RB = 0.
[b]Ulcerative colitis severity subgroups were defined by the Modified Mayo Score: Moderate (MMS = 4-6) and Severe (MMS > 6).
Abbreviations: IV, intravenous; MMS, Modified Mayo Score; SC, subcutaneous; SD, standard deviation.

The subinvestigators may perform the endoscopic evaluations or fill in occasionally but they are more like the other partners in the practice referring patients and not the target population for this substudy.

Based upon the induction study baseline, 3 in 4 patients (69; 74%) had UC for more than 4 years. For 68% of patients, mirikizumab was their first biologic therapy, and 61% were receiving corticosteroids or immunomodulators (Table 1). Nearly all (97%) had either moderate or severe UC as indicated with modified Mayo Scores ≥4. Based upon the last visit with symptomatic remission captured before entering LUCENT-3, the majority of patients (63%) had achieved symptomatic remission when they entered LUCENT-3. The survey was taken during LUCENT-3. At the time of the survey, patients were generally experienced in receiving injections, given that 85% had received at least 3 IV infusions and 77% had received mirikizumab SC injections for at least 2 years. As expected, based on the clinical study design, experience with self-injections was limited (44%). The majority of patients who took the survey did not report any fear of needles at the clinical study baseline (61%; Table 1).

The HCPs who responded to the survey were primarily from the United States (57%) and most frequently nurses (41%) or gastroenterologists (26%; Table 2). Most HCPs were experienced (≥6 years) in treating patients with UC (57%) and provided primarily office-based patient care (71%; Table 2). When asked about time spent actively seeing patients, HCPs most frequently reported 26%-50% (*n* = 15, 36%) or 76%-100% (*n* = 13, 31%).

Table 2. HCP characteristics, *n* = 42.

Type of HCP, *n* (%)	
Nurse[a]	17 (41)
Gastroenterologist[b]	11 (26)
Internal medicine or other physician	2 (5)
Other	12 (29)
Years providing care to patients with UC (excluding medical training), *n* (%)	
0-5 years	18 (43)
6-10 years	8 (19)
11-15 years	6 (14)
16-20 years	7 (17)
≥21 years	3 (7)
Primary medical practice setting providing care to patients, *n* (%)	
Office-based	30 (71)
Hospital-based (non-university)	7 (17)
University-based	5 (12)
Percentage of working time spent actively seeing patients, *n* (%)	
0-25%	4 (10)
26%-50%	15 (36)
51%-75%	10 (24)
76%-100%	13 (31)
Country, *n* (%)	
United States	24 (57)
Australia	4 (10)
Switzerland	4 (10)
Hungary	3 (7)
Japan	2 (5)
Poland	2 (5)
Austria	1 (2)
Germany	1 (2)
Spain	1 (2)

[a]Includes nurses, nurse practitioners, or other nurse specialists.
[b]Includes gastroenterologists or physicians with gastroenterology specialty.
Abbreviations: HCP, healthcare provider; UC, ulcerative colitis.

Patients Survey Responses

IV Infusions

Satisfaction overall and by subgroups

Patients were generally either "very satisfied" or "satisfied" with the administration of IV infusion (*n* = 77, 83%; Figure 4). Only one patient reported being "dissatisfied" with IV infusions, and none of the patients reported being "very dissatisfied." Subgroup findings of interest are:

- Women were more often "very satisfied" or "satisfied" (*n* = 39, 93%) than men (*n* = 38, 76%); only a few women (*n* = 3, 7%) were neither satisfied nor dissatisfied, as opposed to one in every 4 men (*n* = 12, 24%; *p* = 0.0159).
- Most patients were "very satisfied" or "satisfied" with IV regardless of their previous experience with corticosteroids or immunomodulators (*n* = 51, 89% vs *n* = 26, 72%; *p* = 0.1104).
- Similarly, most patients were "satisfied" or "very satisfied" with IV regardless of their experience with biologics (*n* = 48, 84% vs *n* = 21, 81%; *p* = 0.4216).
- Satisfaction with IV infusion did not change over time for most patients (*n*=69, 74%). Of patients who reported satisfaction improvement over time (*n*=17, 18%), this change was more frequent among patients with <4 years' experience with UC (*n*=9, 38%) than patients with >4 years' experience (*n* = 8, 12%; *p* = 0.0374).
- Asian patients were less likely to be "very satisfied" (*n* = 4, 19%), a trend in the data that should be interpreted with caution due to being underpowered.
- With similar caution, patients were more likely to be "very satisfied" if they had Inflammatory Bowel Disease Questionnaire (IBDQ) response (≥16 improvement from baseline[24–26]; *n* = 37, 49%); had severe UC at induction baseline (*n* = 23, 54%); had self-injection experience (*n* = 21, 51%); had achieved symptomatic remission (*n* = 30, 51%); were not afraid of needles (*n* = 30, 53%). Younger (<40 years) patients (*n* = 20, 67%) chose "scheduling the IV infusions was easy" as a reason for satisfaction more than older (≥40 years) patients (*n* = 23, 45%).

Acceptability overall and by subgroups

To the question about the acceptability of receiving medication through IV infusion, the most frequent responses were "completely acceptable" or "somewhat acceptable" (*n* = 81, 87%; Figure 4). Subgroup findings of interest are:

- Patients who found IV administration of study medication completely acceptable were older (≥40 years old; *n* = 33, 59%), men (*n* = 29, 57%), who had not experienced prior biologic failure (*n* = 35, 61%), were bionaïve (*n* = 35, 61%), had UC for <4 years (*n* = 18, 75%), had more severe UC (*n* = 26, 61%), had achieved IBDQ remission (IBDQ score ≥ 170[24–26]; *n* = 40, 56%), had achieved symptomatic remission in parent studies (*n* = 34, 58%), and had no fear of needles (*n* = 32, 56%). However, these subgroup comparisons were not statistically significant except for the comparison by disease duration (<4 years vs ≥4 years; *p* = 0.0467).
- When examining reasoning for the responses of "completely" or "somewhat acceptable" by subgroups, there were a few response option trends that were clear. Nearly all patients with higher disease severity (ie, Modified Mayo Score [MMS] score > 6; *n* = 35, 92%) and most patients with moderate disease severity (MMS score 4-6; *n* = 35, 83%) at baseline of the parent induction studies selected "I feel it helped my UC."
- Women answered "it allowed me to interact with HCPs" more often than men (*n* = 17, 45% vs *n* = 10, 23%; *p* = 0.0407).
- Patients with <4-year experience with UC more frequently selected the "length of time while receiving IV infusion was acceptable" than patients with ≥4-year experience with UC (*n* = 15, 65% vs *n* = 21, 36%; *p* = 0.0178).

SC Injections

Satisfaction overall and by subgroups

Patients were generally either "very satisfied" or "satisfied" with SC injections (*n* = 85, 91%; Figure 4). Two patients were dissatisfied with SC injections; no one was very dissatisfied. Subgroup findings of interest are:

- Although subgroup comparison differences did not meet statistical significance, patients who were most frequently "very satisfied" were bionaïve (*n* = 32, 56%), had achieved IBDQ remission (*n* = 37, 51%), had moderate disease severity at baseline (*n* = 28, 60%), had achieved symptomatic remission at the end of the maintenance study LUCENT-3 (*n* = 31, 53%), and did not have a fear of needles (*n* = 31, 54%).
- Some patients reported their satisfaction improved over time (*n* = 17, 18%). This change was statistically significant by disease duration (33% <4 years vs 13% ≥4 years; *p* = 0.0121), IBDQ remission (74% yes vs 53% no; *p* = 0.0035), symptomatic remission achieved at the end of the maintenance study LUCENT-3 (76% yes vs 57% no; *p* = 0.0195).

Acceptability overall and by subgroups

Nearly all patients (*n* = 90, 97%) found SC injections "completely" or "somewhat acceptable" (Figure 4). Two patients found SC injection "neither acceptable nor unacceptable" and one patient found it "somewhat unacceptable." When asked about the acceptability of receiving 2 SC injections, most patients found it "completely acceptable" (*n* = 71, 76%) or "somewhat acceptable" (*n* = 19, 20%). Only 3 patients selected "neither acceptable nor unacceptable" (3%; Figure 4). The patients were then asked about acceptability of receiving monthly SC injections, and most selected "completely acceptable" (*n* = 79, 85%). Fewer patients selected "somewhat acceptable" (*n* = 12, 13%) or "neither acceptable nor unacceptable" (*n* = 2, 2%; Figure 4).

Subgroup findings of interest are:

- Acceptability of monthly SC injections expressed as "completely acceptable" increased with years of injection experience from 75% (*n* = 24) among those with 2-3 years of experience to 93% (*n* = 37) among those with >3 years of experience (*p* = 0.0354).
- Asian patients favored the "somewhat acceptable" (*n* = 11, 52%) response rather than "completely acceptable"

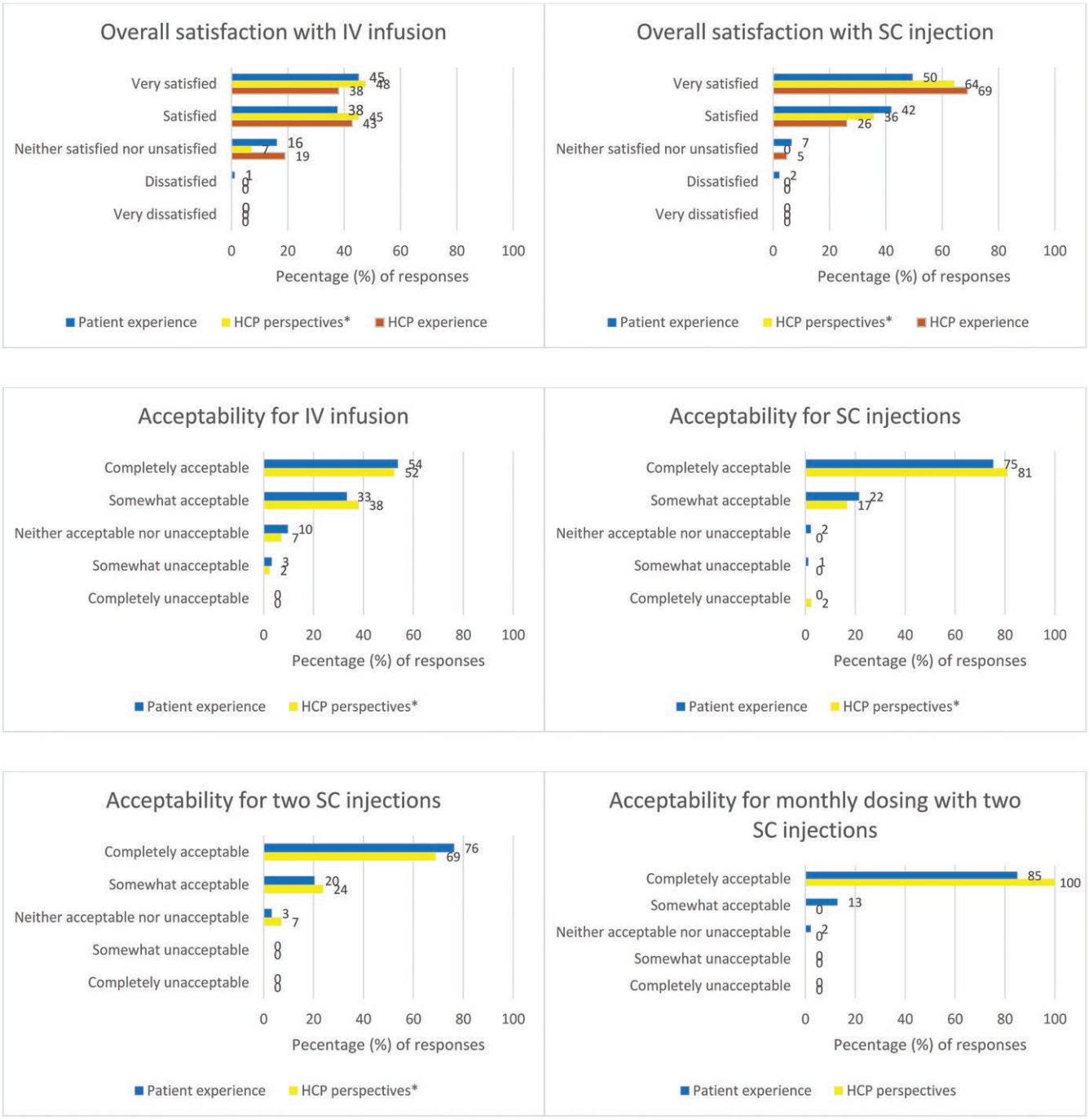


Figure 4. Satisfaction and acceptability rates associated with mirikizumab IV infusions and SC injections by patients' actual experience (n = 93), HCP perspectives of patients' experiences, and HCP own experiences (n = 42). Abbreviations: HCP, healthcare professionals; IV, intravenous; SC, subcutaneous. *HCPs' perspectives of patients' experiences.

(n = 9, 43%) that was favored by the patients from other geographic regions (p = 0.0097).

- Although comparisons were not statistically significant, those who believed that SC injections were "completely acceptable" were older (≥40 years old; n = 46, 82%), men (n = 39, 77%), did not experience prior biologic failure (n = 43, 75%), were bionaïve (n = 43, 75%), had UC for <4 years (n = 22, 92%), had moderate disease severity (n = 37, 79%), had achieved IBDQ remission (n = 56, 78%), had achieved symptomatic remission before entering the LUCENT-3 study (n = 48, 81%), and had no fear of needles (n = 46, 81%).

Administration Options and Preferences

Patients most frequently preferred to receive their SC injections from HCPs (n=78, 84%), followed by self-injection (n=41, 44%), and caregivers (n=8, 9%); however, for patients with self-injection experience, self-injection was the preference. Patients' injection-site preferences were abdomen

(*n* = 72, 77%), followed by the back of the upper arm (*n* = 32, 34%), and the thigh (*n* = 8, 9%).

The patients were then asked to rank by preference the administration options that they experienced during their clinical trial participation. The ranking data should be interpreted with caution because some patients might not have experienced all options. Those patients who only experienced receiving the injection from an HCP ranked HCPs as either first (*n* = 56, 72%) or third (*n* = 22, 28%) preferred option. However, when patients also had self-injection experience, they ranked this option as their first (*n* = 31, 76%), third (*n* = 9, 22%), or second (*n* = 1, 2%) preference. Younger patients (*n* = 15, 83%) were more likely to rank self-injection as their first preference. Almost all of those who experienced receiving the injection from a caregiver, ranked caregiver as their first (*n* = 6, 75%) preferred option. For patients who found SC injections acceptable, the reason most patients selected was "I feel it helped my UC" (*n* = 87; 97%); "I liked the self-injection option" (*n* = 18; 51%), and "I liked having the option to receive the injections at home" (*n* = 19; 54%) were chosen more by younger UC patients.

Most patients reported they "strongly agree" when asked about the convenience of using the self-injection device (*n* = 24; 59%). Thirteen patients reported "agree" (*n* = 13; 32%), while only one patient reported "neither agree nor disagree" (*n* = 1; 2%) and 3 patients reported "disagree" (*n* = 3; 7%). Moreover, when asked whether they found the device "easy to use," most patients selected "strongly agree" (*n* = 25; 61%) and "agree" (*n* = 14; 34%). The patients responded "strongly agree" (*n* = 34; 83%) and "agree" (*n* = 7; 17%) when the patients were asked whether they understood instructions to self-inject. Similarly, they responded "strongly agree" (*n* = 32; 78%) and "agree" (*n* = 7; 17%) when asked whether they understood instructions to store the medication at home.

Overall, the most reported strategy used by patients to help with SC injection administration was to "let study medication warm to room temperature before injecting" (*n* = 60; 65%), followed by "pinch skin and squeeze while injecting" (*n* = 47; 51%). Nineteen patients (20%) did not use any strategies to ease the administration burden.

Treatment Satisfaction and Administration Burden

Nearly all patients indicated that they were "very satisfied or satisfied" (98%) with the overall study medication administration (Table 3). Similarly, most patients (*n* = 88, 95%) expressed complete acceptability to the medication's administration (Table 3). Subgroup findings of interest are as follows:

- Overall satisfaction was statistically significantly higher among patients who had an IBDQ response (100% vs 86%; $p = 0.0054$).
- Complete acceptability was expressed more frequently among older patients (aged ≥ 40 years) than among younger patients (98% vs 89%, $p = 0.0369$).

With regards to the potential administration burden, almost all patients (90%) "agreed" or "strongly agreed" that the improvement in their UC outweighed any dissatisfaction they may have had with the administration of the medication (Table 3). These responses were more frequent among patients who achieved IBDQ response (92% vs 79% without response; $p = 0.0287$) and achieved symptomatic remission (93% vs 83% without remission; $p = 0.0207$) at the end of the maintenance study before entering the LUCENT-3 study when the survey was administered.

Finally, almost all (97%) patients responded that they would recommend mirikizumab to someone with UC (Table 3). The willingness to recommend mirikizumab was stronger (ie, "strongly agree" answers) among non-Asian patients (*n* = 62, 85%) than among Asian patients (*n* = 7, 35%; $p = 0.0009$ vs other regions).

HCPs Survey Responses

IV Infusions

HCPs perspectives of patients' experiences

HCPs believed that patients would be "very satisfied" or "satisfied" (*n* = 39, 93%) with IV infusions, and no one thought that there would be any dissatisfied patient (Figure 4). Subgroup findings of interest are as follows:

- HCP perspective of patient overall satisfaction with administration of IV infusion was statistically significant by HCP type ($p = 0.0048$), years providing UC care ($p = 0.0027$), and average IV infusion experience ($p = 0.0192$).
- When looking at HCP type, we found that all gastroenterologists (*n*=11, 100%) and nearly all nurses (*n*=16, 94%) believed that their patients were "very satisfied" or "satisfied"; however, the physicians' answers were grouped within "very satisfied" (*n*=10), while nurses' responses were more evenly divided between "very satisfied" (*n* = 7) and "satisfied" (*n* = 9).
- Despite this, the proportion of HCPs thinking patients were "very satisfied" or "satisfied" was similar between HCP types. However, some HCPs were potentially overestimating their patients' satisfaction from "satisfied" to "very satisfied" compared to the patients' actual satisfaction.
- HCPs with ≥16 years of providing UC care were more likely to believe that their patients were "very satisfied" (*n* = 9, 90%) with IV infusion compared with those who had <5 years of experience (*n* = 2, 11%; $p = 0.0027$).
- Similarly, HCPs who had provided >30 IV infusions (*n* = 12, 75%) were more likely to believe patients were "very satisfied" with IV infusion than HCPs who had administered <10 infusions (*n* = 4, 36%: $p = 0.0192$).
- Most HCPs said that patients' satisfaction with the administration of IV infusion "did not change over time" (*n*=32, 76%). Seven HCPs said patients' satisfaction "improved over time" mostly after the second or third infusion (*n* = 6) while one HCP believed that change occurred already after the first infusion.

Regarding acceptability, HCPs reported that they thought patients found it "completely" or "somewhat acceptable" (*n* = 38, 90%) to receive the IV infusion (Figure 4). Subgroup findings of interest are as follows:

- Gastroenterologists were more likely to select "completely acceptable" than nurses (*n* = 11, 100% vs *n* = 7, 41%; $p = 0.0008$).

Table 3. Overall experience with administration of study medication.

	Patients' experiences ($N = 93$)	HCP perspectives of patients' experiences ($N = 42$)	HCP own experiences ($N = 42$)
Improvement in UC outweighed any dissatisfaction with the administration of mirikizumab, *n* (%)			
Strongly agree	67 (72)	20 (48)	26 (62)
Agree	17 (18)	17 (41)	14 (33)
Neither agree nor disagree	4 (4)	5 (12)	1 (2)
Disagree	4 (4)	0 (0)	1 (2)
Strongly disagree	1 (1)	0 (0)	0 (0)
Overall satisfaction with mirikizumab, *n* (%)			
Very satisfied	71 (76)	27 (64)	28 (67)
Satisfied	20 (22)	13 (31)	13 (31)
Neither satisfied nor unsatisfied	2 (2)	1 (2)	1 (2.4)
Unsatisfied	0 (0)	1 (2)	0 (0)
Very unsatisfied	0 (0)	0 (0)	0 (0)
Overall acceptability of administration of mirikizumab, *n* (%)			
Completely acceptable	88 (95)	39 (93)	NA
Somewhat acceptable	3 (3)	2 (5)	NA
Neither acceptable nor unacceptable	1 (1)	1 (2)	NA
Somewhat unacceptable	1 (1)	0 (0)	NA
Completely unacceptable	0 (0)	0 (0)	NA
Recommend mirikizumab to someone with UC, *n* (%)			
Strongly agree	69 (74)	28 (67)	NA
Agree	21 (23)	13 (31)	NA
Neither agree nor disagree	2 (2)	1 (2)	NA
Disagree	1 (1)	0 (0)	NA
Strongly disagree	0 (0)	0 (0)	NA
Found SC self-injection device easy to use, *n* (%)			
Strongly agree	25 (61)	23 (55)	32 (76)
Agree	14 (34)	9 (21)	5 (12)
Neither agree nor disagree	1 (2)	8 (19)	3 (7)
Disagree	1 (2)	2 (5)	2 (5)

Abbreviations: HCP, healthcare provider; NA, not applicable; SC, subcutaneous; UC, ulcerative colitis.

- HCPs in Western Europe did not select "completely acceptable" at all and were the most likely to select "somewhat acceptable" ($n = 6$, 86%; $p = 0.0296$ vs among geographical regions).
- HCPs believed that the main reason for acceptability was that the treatment helped patients with their UC ($n = 33$, 87%).

HCPs own experiences

HCPs themselves were mainly "very satisfied" ($n = 16$, 38%) or "satisfied" ($n = 18$, 43%) with IV infusions. Few HCPs ($n = 8$, 19%) were "neither satisfied or unsatisfied" and no one was unsatisfied or very unsatisfied (Figure 4). The gastroenterologists were mostly "very satisfied" ($n = 9$, 82%) while the nurses were mostly "satisfied" ($n = 9$, 53%) and to a lesser extent "very satisfied" ($n = 5$, 29%).

SC Injections

HCPs perspectives of patients' experiences

When asked about their perspective on patients' overall satisfaction with the administration of SC injections, the HCPs believed that their patients were "very satisfied" ($n = 27$, 64%) or "satisfied" ($n = 15$, 36%; Figure 4); gastroenterologists preferentially selected "very satisfied" ($n = 9$, 82%) while nurses distributed their responses between "very satisfied" ($n = 10$, 59%) and "satisfied" ($n = 7$, 41%). Most HCPs ($n = 31$, 74%) believed that patients' satisfaction did not change over time. However, a few HCPs ($n = 9$, 21%) said their patients' satisfaction improved over time.

Overall, most HCPs perceived their patients found it "completely acceptable" ($n = 34$, 81%) or "somewhat acceptable" ($n = 7$, 17%) to receive study medication via SC injection. Most HCPs also believed that patients accepted SC injections mainly because they helped their UC ($n = 38$, 93%) and that it was easy to schedule the injections ($n = 29$, 71%).

HCPs perspective on patient acceptability of receiving 2 SC injections ("completely acceptable": $n = 29$, 69%; "somewhat acceptable": $n = 10$, 24%) were lower than patients' actual experience (Figure 4). Regarding the monthly dosing of 2 SC injections, all HCPs believed that their patients found it "completely acceptable" ($n = 42$, 100%; Figure 4).

HCPs own experiences

HCPs themselves were mainly "very satisfied" (n = 29, 69%) with SC injections. Some were "satisfied" (n = 11, 26%), only a few (n = 2, 5%) were "neither satisfied or unsatisfied" and no one was unsatisfied or very unsatisfied (Figure 4). The gastroenterologists were mostly "very satisfied" (n = 9, 82%) while the nurses were mostly "satisfied" (n = 11, 65%) and to a lesser extent "very satisfied" (n = 6, 35%).

Administration Options and Preferences

HCPs perspectives of patients' experiences

HCPs were asked to provide a ranking for all administration options. They were asked to rank what they believed were patient preferences for who (self, caregiver, or HCP) administered the SC injections. When looking at the total sample, HCPs most frequently chose "HCP" (n = 29, 69%) followed by "caregiver" (n = 22, 52%), and "self-administration" (n = 20, 48%). Based on these results, HCPs believed that patients preferred the HCP to give the injection, while patients who had performed self-injections primarily preferred the "self-injection" option, suggesting an HCP-patient disconnect on the topic of preference for who should do the injections.

HCPs were also asked to rank what they believe were the patients' preferences for the location (abdomen, thigh, and back of upper arm) of SC injections. As the first preferred option, they ranked "abdomen" (n = 26, 62%), which was aligned with the patients ranking. As the second preferred option, they reported "thigh" (n = 16, 38%) and "abdomen" (n = 14, 33%). As the third preferred option, they reported "thigh" (n = 24, 57%) and "back of upper arm" (n = 16, 38%).

HCPs own experiences

When HCPs were asked to rank their own preferences for the preferred administration method of SC injections, as the first preferred option, they most frequently ranked "SC injection, administered by the patient" (n = 18, 43%); as the second preferred option, they ranked "SC injection, administered by HCP" (n = 23, 55%); and, as the third preferred option, they ranked "SC injection, administered by caregiver" (n = 20, 48%).

Treatment Satisfaction and Administration Burden

HCPs perspectives of patients' experiences

Most HCPs believed that patients were "very satisfied" or "satisfied" with the administration of the study medication (95%) and considered it "completely" or "somewhat acceptable" (98%; Table 3). In the opinion of the HCPs, patients believed that the therapeutic advantages of the medication outweighed any dissatisfaction they had with its administration (88% "agree" and "strongly agree"; Table 3). This opinion was expressed by all gastroenterologists (n = 11, 100%) but not by all nurses (n = 9, 75%; p=0.0746). In concordance with patients' experiences, almost all HCPs (98%) agreed that their patients would recommend the study medication to someone with UC (Table 3).

Regarding HCP perspectives of patients' beliefs of injector usability, patients responded that injector device was convenient (90% "strongly agree" and "agree") and easy to use (95% "strongly agree" and "agree") more than HCPs believed they would (79% and 76% respectively). Patients also responded that the injector device instructions were understood (100% "strongly agree" and "agree") more than HCPs (91%) believed they would. Patients (54%) used injection instruction documents less than HCPs (95%) believed they did. Patients felt responses from their HCPs to their questions were more helpful (100%) than HCPs (94%) believed they were. Patients felt more confident they were using the injection device correctly (98%) and getting a full dose (100%) than HCPs (81% and 86%, respectively) believed they were.

HCPs own experiences

HCPs responses about their own opinions showed that most HCPs were "satisfied" or "very satisfied" with IV infusions (n = 34, 81%), SC injections (n = 40, 95%), and with overall mirikizumab treatment (n = 41, 98%). Most HCPs (n = 40, 95%) "agreed" or "strongly agreed" that the treatment benefits of mirikizumab outweighed any dissatisfaction with its administration (Table 3). Subgroup findings of interest are as follows:

- Subgroup analysis revealed that gastroenterologists were "very satisfied" (n = 9, 82%) with IV infusions, whereas the nurses were "satisfied" (n = 9, 53%; p = 0.0153).
- Similar differences, although not statistically significant, were seen in "strongly agree" responses that treatment benefits outweighed any dissatisfaction with the administration: 82% among gastroenterologists (n = 9) vs 53% among nurses (n = 9), p = 0.2186.
- Lastly, HCPs with longer experience providing care to patients with UC (≥16 years) were more frequently very satisfied (n = 9, 90%) with SC injections than the HCPs with less experience (n = 13, 72%) although this was not a statistically significant difference (p = 0.2090).

Regarding strategies recommended by HCPs to patients to ease the administration burden, only 79% of HCPs selected warm medication before injecting, despite the label recommending doing. Moreover, this suggestion was more frequently given by nurses than gastroenterologists (82% vs 73%; p = 0.0494). Choosing a different injection location was recommended by 69% of HCPs, but only 9% of patients noted choosing a different injection location as a method to ease the administration burden, suggesting an HCP-patient disconnect.

Discussion

This international web survey demonstrated that patients with moderately-to-severely active UC and their HCPs were predominantly satisfied and accepting of mirikizumab treatment administration, whether administered via IV infusion or SC injection including monthly SC dosing with 2 injections. Gastroenterologists were more likely than nurses to think patients were "very satisfied" with and "completely acceptable" of IV infusion and SC injections. HCPs overall were more likely to overestimate the proportion of patients with the highest degree of satisfaction or acceptance of IV infusions and SC injections. However, the overall satisfaction range responses (very satisfied + satisfied) and overall acceptability range responses (completely acceptable + acceptable) for IV infusion and SC injection were similar across patients' experiences, HCPs perspectives of patients' experiences, and HCPs' personal experiences. HCPs and patients were closely

aligned that the abdomen was the preferred site for injections. HCPs were less likely than patients to respond that patients fully understood administration instructions and indicate that the self-injection option was easy and convenient. There was also a disconnect between patients and HCPs regarding methods for helpfulness with SC injection in which HCPs recommended choosing a different injection site, but almost no patients selected this as an option.

Importantly, patients and HCPs believed that the benefits of mirikizumab therapy outweighed any potential burdens associated with the administration method. Patients were satisfied with and accepted both the IV and SC routes of administration because they felt mirikizumab treatment helped their UC. These results offer some insight into patients' preferences regarding UC treatment attributes and the treatment administration burden they are willing to accept for therapeutic success. They suggest acceptance of the administration burden for therapies that offer clinically relevant therapeutic benefits. This is in alignment with previous studies showing that the administration route is of lesser concern to patients and their HCPs than treatment effectiveness and safety profile.[27–31] For example, a conjoint analysis assessing biologics treatment preferences among patients with IBD showed that patients, either naïve or experienced in biologics, rated route of administration as the third important treatment characteristic after efficacy and safety. Another study on biologic naïve patients with moderate-to-severe UC found that patients primarily cared for long-lasting effectiveness, and they considered of "no real importance" or "completely irrelevant" the route of administration (25.3%) and the dosing frequency (32.3%).[28] There are data available from 4 discrete choice experiments among patients with IBD that included route of administration as one of the treatment attributes assessed.[29–32] Two of those studies found that the administration route was not among patients' primary considerations.[29,31] In the third study, some patients were willing to accept 10.3% (95% CI, 6.6%-14.0%) added risk to replace IV administration at a hospital with an injection at home.[30] The most recently published study reported that patients with UC preferred oral or SC administration over IV ($P < 0.001$).[32] The current data does suggest a similar patient preference for SC injection over IV infusion: satisfied or very satisfied (IV, 83%; SC, 91%) and completely acceptable or somewhat acceptable (IV, 87%; SC, 97%). Finally, in a structured interview with patients of a UK hospital with IBD, patients with prior biological therapy experience were more receptive to SC or IV therapies than bionaïve patients.[33]

For several survey questions, HCPs' and patients' responses were not aligned. For example, the HCP perspective of patient overall satisfaction with IV infusions was elevated compared with what was reported by patients. Similarly, HCPs believed that patients preferred the HCP to give the SC injections, whereas patients reported that they preferred to self-inject. In addition, the HCP survey demonstrated a disconnect between gastroenterologists and nurses. This was observed with patient satisfaction for IV infusion and SC injection, patient acceptability for IV infusion, and overall HCP satisfaction with the administration of study medication. While gastroenterologists' responses tended to aggregate around "very satisfied" or "completely" acceptable," nurses' responses were divided between these 2 categories and were closer to patients' responses. For example, gastroenterologists were "very satisfied" (IV, 82%; SC, 82%) more than nurses (IV, 29%; SC, 65%). This discrepancy appears to be influenced by the finding that nurses were less likely to respond that patients felt their treatment "helped my ulcerative colitis" (IV: gastroenterologist 100% vs nurse 81.3%; SC: gastroenterologist 100% vs nurse 87.5%). Nurses were also less likely to believe patients found the self-injection device easy to use (strongly agree: gastroenterologist 81.8% vs nurse 41.2%) or that patients felt confident using the SC self-injection device (strongly agree: gastroenterologist 81.8% vs nurse 58.8%). The similarity of nurses' responses to patients' responses was previously shown in a Spanish study assessing the satisfaction of patients with IBD with healthcare services received.[34] Although not specific to treatment administration preferences like the current study, the Spanish study highlighted the essential role of nurses in the management of patients with IBD that brings them closer to patients and ultimately the understanding of their needs. A role has been suggested to involve them in IBD care management as patients' educators.[35]

Considerations, Strengths, and Limitations

When interpreting these findings, the study's limitations and strengths should be acknowledged. Strengths include that this web-based survey provides the first evidence regarding the level of patients' and HCPs' satisfaction and acceptability of mirikizumab treatment administration in UC. Moreover, beyond some reports on satisfaction with conventional therapies,[28,36] there is no apparent research evidence on satisfaction with other biological treatments. This is also the first research to look at patient and HCP drug administration preferences for a specific medicine used to treat UC.

Among the study's strengths is that its findings are not geographically limited, as participants were recruited from varying countries. The fact that the treatment under evaluation was administered within a clinical trial helps to overcome disparities in prescription patterns, pharmaceutical costs, reimbursement, and general therapeutic approaches between countries.

The main study limitation was its small sample size. Because of this, the study was underpowered, and all comparative results and subgroup analyses should be interpreted with caution. These findings may only apply to patients with UC who meet the eligibility criteria to enroll in a phase 3 study in UC; such criteria usually exclude many patients attending routine clinical practice. Patients' participation in LUCENT-3 was voluntary so it was unlikely and expected that no patients in the survey would have found the study medication completely unacceptable. Since the questionnaire was administered a long time after the induction IV administration, patients who discontinued during or shortly after induction treatment may have had different opinions of mirikizumab administration than those who did not discontinue, and those insights thus may not have been captured. Thus, patients who found the study medication administration unacceptable might have discontinued study participation before LUCENT-3; however, study discontinuation data do not suggest this was an issue.[18,20,37,38] Of note, only 3.8% of mirikizumab-treated patients discontinued the LUCENT-1 induction study due to any reason—1.7% due to adverse event; 0.6% due to lack of efficacy; 0.3% due to withdrawal by subject.[20] During the LUCENT-2 maintenance study: amongst both mirikizumab induction responders and non-responders 25% discontinued

due to any reason—1.7% due to adverse event, 19.0% due to lack of efficacy, and 1.6% due to withdrawal by subject.[20]

The currently described survey was added to an ongoing extension study (LUCENT-3). The patient population in LUCENT-3 represented a subset of patients who were originally in the induction and maintenance studies, LUCENT-1 and LUCENT-2. This survey was conducted as a protocol amendment offered to a subset of participating countries. Participation in the survey was voluntary. Because only a subset of eligible patients and HCPs completed the questionnaire, this could have introduced bias into the results. Similarly, since the questionnaire was administered to a patient population that successfully entered the extension study (LUCENT-3), the majority of patients were experiencing clinical benefit from treatment, introducing bias regarding acceptance of the administration route versus acceptance due to the clinical improvement presented. Many experienced patients had transitioned to a hybrid of on-site and remote study visits and dosing, which decreased their access to this site-based survey, and this may also have added bias. Additionally, the voluntary nature of survey participation could have introduced selection bias, which could have leaned toward more satisfied patients.

The survey study did not allow open-ended responses, and no adverse event questions were asked; therefore, there is no way to align adverse events and survey respondents nor ascertain how these might have influenced patient satisfaction and acceptability. Of note, 0.4% of mirikizumab-treated patients reported infusion-site reactions during IV induction, 8.7% reported injection-site reactions during maintenance treatment, and 5.5% during the first year of extension treatment.[20,39]

Future discrete choice experiment research designs may give more detailed information about patients' preference drivers and the trade-offs they are willing to make between different treatment attributes.

Conclusion

For the administration of mirikizumab, this study revealed some aspects of discordance between patients' experiences and HCPs' perspectives of patients' experiences, such as patient preference for self-administration, as well as some differences amongst HCP subgroups such as gastroenterologists and nurses. However, overall, both HCPs and patients reported satisfaction with and acceptance of mirikizumab IV and SC administration. Nevertheless, it may be important for HCPs to fully understand patients' treatment administration preferences and perspectives. Importantly, most patients felt that UC improvement outweighed any administration dissatisfaction. These findings may aid patients and HCPs in their treatment choices if they are seeking information on prior patients' experiences.

Supplementary Data

Supplementary data are available at *Crohn's & Colitis 360* online.

Acknowledgments

Stephanie Woerner (study execution), Richard Moses (medical peer review), Deborah Fisher (medical peer review), and Jordan Johns (stat peer review) from Eli Lilly and Company; Patrick Daniele (stat analysis, Evidera). Athanasia Benekou (Evidera) and Phil Leventhal (Evidera) provided medical writing services, which were funded by Eli Lilly and Company, in accordance with Good Publication Practice (GPP) guidelines (Good Publication Practice (GPP) Guidelines for Company-Sponsored Biomedical Research: 2022 Update. Ann Intern Med).

Authors' Contributions

David Clemow: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Writing—review & editing; Christine Radawski: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Writing—review & editing; Joe Milata: Conceptualization; Investigation; Methodology; Validation; Writing—review & editing; Karla Alaka: Conceptualization; Investigation; Methodology; Writing—review & editing; Theresa Hunter Gibble: Conceptualization; Methodology; Writing—review & editing; Adam Schaum: Conceptualization; Investigation; Methodology; Supervision; Writing—review & editing; Obi Ezennia: Data curation; Investigation; Writing—review & editing; Nicholas Martinez: Writing—review & editing; Tibor Szaloki: Writing—review & editing; Yuka Ito: Writing—review & editing; Danielle Rodriguez: Formal analysis; Investigation; Methodology; Validation; Writing—original draft; Writing—review & editing; Katherine Kirk: Formal analysis; Investigation; Methodology; Validation; Writing—original draft; Writing—review & editing. All authors approved the final version of the manuscript.

Funding

Eli Lilly and Company was the current study sponsor, funded the original research as well as the medical writing and editing services.

Conflicts of Interest

David Clemow, Christine Radawski, Joe Milata, Karla Alaka, Theresa Hunter Gibble, Adam Schaum, and Obi Ezennia are employees of Eli Lilly at the time the study was conducted and the manuscript developed. Danielle Rodriguez and Katherine Kirk are employees of Evidera, which was contracted by Eli Lilly for work relating to this study. Tibor Szaloki is an employee of Javorszky Hospital and was contracted by Eli Lilly to assist as an external collaborator for this study. Nicholas Martinez is an employee of Gastroenterology Research of America and was contracted by Eli Lilly to assist as an external collaborator for this study. Yuka Ito is an employee of NHO Mito Medical Center and was contracted by Eli Lilly to assist as an external collaborator for this study. Medical writing support was provided by Athanasia Benekou, Principal Medical Writer, from Evidera's Medical Writing and Healthcare Communications, and funded by Eli Lilly.

Data Availability

Eli Lilly and Company provides access to all individual participant data collected during the study, after anonymization.

Data are available to request 6 months after the indication studied has been approved in the USA and EU and after primary publication acceptability, whichever is later. No expiration date of data requests is currently set once data are made available. Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement. Data and documents, including the study protocol, statistical analysis plan, and study report will be provided in a secure data sharing environment. For details on submitting a request, see the instructions provided at [www.vivli.org].

References

1. Koliani-Pace JL, Haron AM, Zisman-Ilani Y, Thompson KD, Siegel CA. Patients' perceive biologics to be riskier and more dreadful than other IBD medications. *Inflamm Bowel Dis*. 2020;26(1):141-146. doi:10.1093/ibd/izz121
2. Bretto E, Ribaldone DG, Caviglia GP, Saracco GM, Bugianesi E, Frara S. Inflammatory bowel disease: emerging therapies and future treatment strategies. *Biomedicines*. 2023;11(8):2249. doi:10.3390/biomedicines11082249
3. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG clinical guideline: ulcerative colitis in adults. *Am J Gastroenterol*. 2019;114(3):384-413. doi:10.14309/ajg.0000000000000152
4. Guo M, Wang X. Pathological mechanism and targeted drugs of ulcerative colitis: a review. *Medicine (Baltim)*. 2023;102(37):e35020. doi:10.1097/MD.0000000000035020
5. Gros B, Kaplan GG. Ulcerative colitis in adults: a review. *JAMA*. 2023;330(10):951-965. doi:10.1001/jama.2023.15389
6. Aggarwal A, Sabol T, Vaziri H. Update on the use of biologic therapy in ulcerative colitis. *Curr Treat Options Gastroenterol*. 2017;15(1):155-167. doi:10.1007/s11938-017-0120-8
7. NIDDK. National Institute of Diabetes and Digestive and Kidney Diseases: Treatment for Ulcerative Colitis 2020. Accessed October 5, 2023. https://www.niddk.nih.gov/health-information/digestive-diseases/ulcerative-colitis/treatment#:~:text=How%20do%20doctors%20treat%20symptoms%20and%20complications%20of,fluids%20and%20electrolytes%20to%20prevent%20and%20treat%20dehydration
8. Chao YS, Loshak H. *CADTH rapid response reports: biologics versus immunomodulators for the treatment of ulcerative colitis: a review of comparative clinical effectiveness and cost-effectiveness*. Published April 17, 2029. Accessed February 13, 2024. https://www.ncbi.nlm.nih.gov/books/NBK549363/pdf/Bookshelf_NBK549363.pdf
9. Roda G, Jharap B, Neeraj N, Colombel JF. Loss of response to anti-TNFs: definition, epidemiology, and management. *Clin Transl Gastroenterol*. 2016;7(1):e135. doi:10.1038/ctg.2015.63
10. Ho GT, Chiam P, Drummond H, Loane J, Arnott IDR, Satsangi J. The efficacy of corticosteroid therapy in inflammatory bowel disease: analysis of a 5-year UK inception cohort. *Aliment Pharmacol Ther*. 2006;24(2):319-330. doi:10.1111/j.1365-2036.2006.02974.x
11. D'Haens G, Lindsay JO, Panaccione R, Schreiber S. Ulcerative colitis: shifting sands. *Drugs R D*. 2019;19(2):227-234. doi:10.1007/s40268-019-0263-2
12. Singh JA, Wells GA, Christensen R, et al. Adverse effects of biologics: a network meta-analysis and Cochrane overview. *Cochrane Database Syst Rev*. 2011;2011(2):CD008794. doi:10.1002/14651858.CD008794.pub2
13. Hanzel J, Hulshoff MS, Grootjans J, D'Haens G. Emerging therapies for ulcerative colitis. *Expert Rev Clin Immunol*. 2022;18(5):513-524. doi:10.1080/1744666X.2022.2069562
14. Allocca M, Furfaro F, Fiorino G, Gilardi D, D'Alessio S, Danese S. Can IL-23 be a good target for ulcerative colitis? *Best Pract Res Clin Gastroenterol*. 2018;32-33 (Epub 2018 May 23):95-102. doi:10.1016/j.bpg.2018.05.016
15. Sewell GW, Kaser A. Interleukin-23 in the pathogenesis of inflammatory bowel disease and implications for therapeutic intervention. *J Crohns Colitis*. 2022;16(Supplement_2):ii3-ii19. doi:10.1093/ecco-jcc/jjac034
16. Verstockt B, Salas A, Sands BE, et al.; Alimentiv Translational Research Consortium (ATRC). IL-12 and IL-23 pathway inhibition in inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol*. 2023;20(7):433-446. doi:10.1038/s41575-023-00768-1
17. Steere B, Beidler C, Martin A, Bright S, Kikly K, Benschop RJ. Generation and characterization of Mirikizumab, a humanized monoclonal antibody targeting the p19 subunit of IL-23. *J Pharmacol Exp Ther*. 2023;387(2):180-187. doi:10.1124/jpet.122.001512
18. Sandborn WJ, Ferrante M, Bhandari BR, et al. Efficacy and safety of mirikizumab in a randomized phase 2 study of patients with ulcerative colitis. *Gastroenterology*. 2020;158(3):537-549. doi:10.1053/j.gastro.2019.08.043
19. Omvoh (mirikizumab): EPAR - product information. Accessed December 18, 2023. https://www.ema.europa.eu/en/documents/product-information/omvoh-epar-product-information_en.pdf
20. D'Haens G, Dubinsky M, Kobayashi T, et al. Mirikizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med*. 2023;388(26):2444-2455. doi:10.1056/nejmoa2207940
21. Peyrin-Biroulet L, Sandborn W, Sands BE, et al. Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): determining therapeutic goals for treat-to-target. *Am J Gastroenterol*. 2015;110(9):1324-1338. doi:10.1038/ajg.2015.233
22. Schreiber S, Panés J, Louis E, Holley D, Buch M, Paridaens K. Perception gaps between patients with ulcerative colitis and healthcare professionals: an online survey. *BMC Gastroenterol*. 2012;12:108. doi:10.1186/1471-230X-12-108
23. Sandborn WJ, Ferrante M, Bhandari BR, et al. Efficacy and safety of continued treatment with mirikizumab in a phase 2 trial of patients with ulcerative colitis. *Clin Gastroenterol Hepatol*. 2022;20(1):105-115.e14. doi:10.1016/j.cgh.2020.09.028
24. Irvine EJ. Development and subsequent refinement of the inflammatory bowel disease questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 1999;28(4):S23-S27. doi:10.1097/00005176-199904001-00003
25. Irvine EJ, Zhou Q, Thompson AK. The short inflammatory bowel disease questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators. Canadian Crohn's Relapse Prevention Trial. *Am J Gastroenterol*. 1996;91(8):1571-1578.
26. Dubinsky M, Rice A, Yarlas A, et al. Systematic literature review: ability of the IBDQ-32 to detect meaningful change in ulcerative colitis health indicators. *Inflamm Bowel Dis*. 2023:izad282. doi:10.1093/ibd/izad282
27. Almario CV, Keller MS, Chen M, et al. Optimizing selection of biologics in inflammatory bowel disease: development of an online patient decision aid using conjoint analysis. *Am J Gastroenterol*. 2018;113(1):58-71. doi:10.1038/ajg.2017.470
28. Peyrin-Biroulet L, Van Assche G, Sturm A, et al. Treatment satisfaction, preferences and perception gaps between patients and physicians in the ulcerative colitis CARES study: a real world-based study. *Dig Liver Dis*. 2016;48(6):601-607. doi:10.1016/j.dld.2016.01.013
29. Wickramasekera N, Coates E, Barr A, et al. Patient preferences for treatment in steroid resistant ulcerative colitis - a discrete-choice experiment. *Scand J Gastroenterol*. 2022;57(7):797-806. doi:10.1080/00365521.2022.2036808
30. Louis E, Siegel CA, James B, Heidenreich S, Krucien N, Ghosh S. Patients with inflammatory bowel disease have heterogeneous treatment preferences that are largely determined by the avoidance of abdominal pain and side effects [P-POWER IBD Study]. *J Crohns Colitis*. 2023;17(2):231-239. doi:10.1093/ecco-jcc/jjac130

31. Schubert S, Picker N, Cavlar T, Knop J, Kahraman A, Mohl W. Inflammatory bowel disease patients' treatment preferences using a discrete choice experiment technique: the InPuT study. *Adv Ther*. 2022;39(6):2889-2905. doi:10.1007/s12325-022-02143-z
32. Fiorino G, Bent-Ennakhil N, Varriale P, Braegger F, Hoefkens E. Patient preferences for treatment attributes in inflammatory bowel disease: results from a large survey across seven european countries using a discrete choice experiment. *Inflamm Bowel Dis*. 2024. doi:10.1093/ibd/izae015
33. Denesh D, Carbonell J, Kane JS, Gracie D, Selinger CP. Patients with inflammatory bowel disease (IBD) prefer oral tablets over other modes of medicine administration. *Expert Rev Gastroenterol Hepatol*. 2021;15(9):1091-1096. doi:10.1080/17474124.2021.1898944
34. Casellas F, Vera I, Ginard D, Torrejón A. Inflammatory bowel disease patient's satisfaction with healthcare services received. Physicians' and nurses' perceptions. *Rev Esp Enferm Dig*. 2013;105(7):385-391. doi:10.4321/s1130-01082013000700003
35. Prasad SS, Potter M, Keely S, Talley NJ, Walker MM, Kairuz T. Roles of healthcare professionals in the management of chronic gastrointestinal diseases with a focus on primary care: a systematic review. *JGH Open* 2020;4(2):221-229. doi:10.1002/jgh3.12235
36. Coates E, Wickramasekera N, Barr A, et al. Patient preferences and current practice for adults with steroid-resistant ulcerative colitis: POPSTER mixed-methods study. *Health Technol Assess*. 2022;26(41):1-118. doi:10.3310/rhxr5192
37. Sands BE, Peyrin-Biroulet L, Kierkus J, et al. Efficacy and safety of mirikizumab in a randomized phase 2 study of patients with Crohn's disease. *Gastroenterology*. 2022;162(2):495-508. doi:10.1053/j.gastro.2021.10.050
38. Blauvelt A, Kimball AB, Augustin M, et al. Efficacy and safety of mirikizumab in psoriasis: results from a 52-week, double-blind, placebo-controlled, randomized withdrawal, phase III trial (OASIS-1). *Br J Dermatol*. 2022;187(6):866-877. doi:10.1111/bjd.21743
39. Sands BE, D'Haens G, Clemow DB, et al. Two-year efficacy and safety of mirikizumab following 104 weeks of continuous treatment for ulcerative colitis: results from the LUCENT-3 open-label extension study. *Inflamm Bowel Dis*. 2024;30(6):1044-1045. doi:10.1093/ibd/izae024

Online October 19, 2024
Journal of Neuroendovascular Therapy 2024; 18: 313–320 DOI: 10.5797/jnet.oa.2024-0054

Original Article

Clinical Outcomes of Endovascular Coil Embolization for Ruptured Middle Cerebral Artery Aneurysms

Takao Koiso,[1,2] Yoji Komatsu,[1] Daisuke Watanabe,[1] Hisayuki Hosoo,[2] Masayuki Sato,[2] Yoshiro Ito,[2] Tomoji Takigawa,[3] Mikito Hayakawa,[4] Aiki Marushima,[2] Wataro Tsuruta,[5] Noriyuki Kato,[6] Kazuya Uemura,[7] Kensuke Suzuki,[3] Akio Hyodo,[3] Eichi Ishikawa,[2] and Yuji Matsumaru[2,4]

Objective: Middle cerebral artery (MCA) aneurysms are difficult to treat with coil embolization (CE) due to their location and shape, but the number of CE-treated MCA has gradually increased as treatment techniques have improved. However, the outcomes of CE for ruptured MCA aneurysms are poorly understood. This study aimed to evaluate the outcomes of CE for ruptured MCA aneurysms.

Methods: We retrospectively analyzed the medical records of patients with aneurysmal subarachnoid hemorrhages (aSAH) that were treated with CE between 2013 and 2020, and compared the differences in outcomes depending on aneurysm location.

Results: A total of 468 patients with aSAH were included: 39 patients had ruptured MCA aneurysms (group M), and 429 had ruptured aneurysms at other sites (group O). There were no significant differences between the background characteristics of the 2 groups. Also, there were no significant intergroup differences in occlusion status, the frequency of complications such as ischemia, hemorrhaging, rebleeding, retreatment, or the modified Rankin Scale score at discharge. However, intracerebral hemorrhage (ICH) removal was required significantly more frequently in group M than in group O (10.3% vs. 0.5%, p = 0.0006). By case-matching analysis, there were no significant differences in these outcomes. All MCA cases that needed removal had more than 36 ml of hematoma volume. Logistic regression analysis showed that the existence of ICH at onset was a poor prognostic factor for ruptured MCA aneurysms.

Conclusion: CE for ruptured MCA aneurysms produced acceptable outcomes in selected cases. However, the indications for CE in patients with ICH should be carefully considered.

Keywords middle cerebral artery aneurysm, subarachnoid hemorrhage, coil embolization, intracerebral hematoma

[1]*Department of Neurosurgery, Hitachi General Hospital, Hitachi, Ibaraki, Japan*
[2]*Department of Neurosurgery & Stroke, Institute of Clinical Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan*
[3]*Department of Neurosurgery, Dokkyo Medical University Saitama Medical Center, Koshigaya, Saitama, Japan*
[4]*Division of Stroke Prevention and Treatment, Faculty of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan*
[5]*Department of Endovascular Neurosurgery, Toranomon Hospital, Tokyo, Japan*
[6]*Department of Neurosurgery, Mito Medical Center, Mito, Ibaraki, Japan*
[7]*Department of Neurosurgery, Tsukuba Medical Center Foundation, Tsukuba, Ibaraki, Japan*

Received: June 3, 2024; Accepted: September 3, 2024
Corresponding author: Takao Koiso. Depertment of Neurosurgery, Hitachi General Hospital, 2-1-1, Jonancho, Hitachi, Ibaraki 317-0077, Japan
Email: s0201534@hotmail.co.jp

Introduction

Endovascular treatment for cerebral aneurysms has become more common. However, in many institutions, clipping is still the first-line treatment for middle cerebral artery (MCA) aneurysms.[1,2] MCA bifurcation aneurysms are often wide-necked, which makes endovascular treatment difficult.[3–5] On the other hand, the difficulty of clipping increases for short M1 segments, the superior wall type of M1 segment, or larger size. It was reported that symptomatic complications occurred after clipping for M1 segment aneurysms in 8.7%–33.3% of patients, and such complications were even more common in cases involving ruptured aneurysms.[6–9] Recently, the number of reports in which coil embolization (CE) was performed for unruptured MCA aneurysms has gradually been increasing as treatment techniques and devices have improved.[10–12] For unruptured MCA aneurysms, good outcomes may be achieved using adjunctive techniques such as stent-assist

coiling.[12–14] However, stents have not been approved for ruptured aneurysms in Japan. Furthermore, stents are difficult to use for ruptured aneurysms because of the use of antithrombotics in the perioperative period. Compared to unruptured MCA aneurysms, the outcomes of endovascular CE for ruptured MCA aneurysms are unclear. This study aimed to evaluate the clinical outcomes of CE for ruptured MCA aneurysms.

Materials and Methods

Patient population

This was a retrospective multicenter cohort study conducted by 6 institutions in Japan. This was approved by the institutional ethics committee (H30-137) and complied with the conditions laid out by the Declaration of Helsinki. Opt-out consent was employed, and the requirement to obtain informed consent was waived by Institutional Review Board. The data set consisted of the medical records of consecutive subarachnoid hemorrhage (SAH) patients who were treated with endovascular management (EM) during the period from January 2013 through April 2020.

Surgical clipping and endovascular treatment were performed at all institutions. At each institution, the treatment modality was determined after considering both surgery and endovascular treatment on a case-by-case basis. All endovascular procedures were performed under general anesthesia. During the procedure, the activated clotting time was controlled above 200 seconds by heparin. The following data were collected from medical records: age, sex, World Federation of Neurosurgical Societies (WFNS) grades,[15] modified Rankin Scale (mRS) scores before onset,[16] aneurysm maximum size, dome/neck ratio,[17] aspect ratio,[17] the presence of intracerebral hemorrhage (ICH), the presence of intra-aneurysmal thrombosis, and the percentage of adjunctive technique.

The degree of aneurysm occlusion after the initial CE was classified as follows: total exclusion of the aneurysm from the circulation was defined as complete occlusion (CO), limited residual filling at the junction with the parent vessel was defined as a neck remnant (NR), and residual filling within the coil interstices or at the aneurysm's perimeter was defined as body filling (BF).

The primary endpoint was a favorable outcome, defined as mRS 0–2 at discharge. The secondary endpoints were the degree of aneurysm occlusion, periprocedural hemorrhagic events, periprocedural ischemic events, rebleeding after the procedure, retreatment for aneurysm, and removal of ICH after CE. Periprocedural hemorrhagic events included intraprocedural aneurysmal perforation, blood vessel perforation, and enlargement of the ICH after the procedure. All ischemic strokes, whether symptomatic or not, related to the procedure or delayed cerebral ischemia (DCI) were counted as ischemic events. DCI was defined as cerebral infarction due to vasospasm detected by magnetic resonance imaging within 14 days after the procedure. "Retreatment" included endovascular treatment or clipping for treated aneurysms. All radiological and clinical outcomes were determined by 2 or more neurosurgeons at each institution who did not know the object of this study.

Statistical analysis

To clarify the characteristics of CE for ruptured MCA aneurysms, we compared the clinical factors and outcomes of patients who underwent this procedure with those of patients who underwent CE for ruptured aneurysms located at other sites. To identify the clinical factors that influenced the clinical outcomes of patients who underwent CE for ruptured MCA aneurysms, univariate and multivariate logistic regression analyses were performed. The variables that exhibited significance in the univariate analyses were included in the multivariate analysis.

For baseline variables, summary statistics are presented (frequencies and percentages for categorical data and medians and interquartile ranges [IQR] for continuous data). Fisher's exact test was used to analyze categorical data, and the Wilcoxon rank-sum test was used to analyze continuous data.

A case-matched study was conducted to reduce bias due to differences in patient background between the 2 groups. Patient selection was performed employing the propensity score matching method for clinical factors (age, sex, WFNS grade, mRS score before onset, the existence of ICH, and the existence of intra-aneurysmal thrombosis).

All comparisons were planned, and all tests were 2-sided. p-Values of less than 0.05 were considered to indicate a significant difference. All statistical analyses were performed using JMP (Japanese version 12 for Windows; SAS Institute Inc., Cary, NC, USA).

Results

After excluding 86 patients whose SAH was caused by dissection or fusiform aneurysms, 13 patients who were treated >14 days after onset, a patient with an arteriovenous malformation-related aneurysm, and 37 patients for

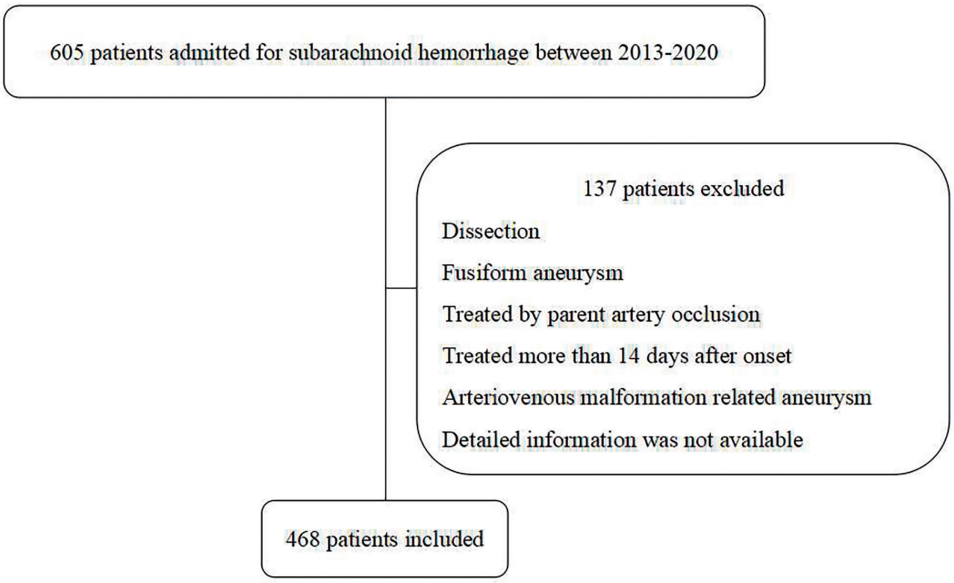


Fig. 1 Schematic drawing of the patient selection method.

whom detailed information was not obtained, 468 patients were identified (**Fig. 1**). There were 170 cases of MCA aneurysms that underwent clipping during the same period. Neuroimaging follow-up was performed in 258 patients (55.1%) after discharge.

Table 1 summarizes the clinical characteristics of all patients and the 2 groups; patients with ruptured MCA aneurysms were placed in group M (39 patients), while those whose ruptured aneurysms were located at other sites were included in group O (429 patients). Thirty-five (89.7%) of the 39 cases in group M were bifurcation aneurysms. There were no significant intergroup differences in clinical characteristics, including age, sex, the proportion of patients with WFNS grades 1–3, the proportion of patients with mRS scores of 0–2 before onset, the maximum aneurysm diameter, dome/neck ratio, aspect ratio, the frequency of ICH, and the presence/absence of intraaneurysmal thromboses. Adjunctive techniques were used less frequently in group M than in group O (43.6% vs. 60.6%, p = 0.04).

The patients' outcomes are summarized in **Table 2**. The proportion of patients with mRS scores of 0–2 at discharge was 61.5% in MCA cases, and this result was not significantly different from those of aneurysms at other sites (p = 0.51). The proportion of BF cases did not differ significantly between the 2 groups (p = 0.11). There were no significant differences in the frequencies of periprocedural hemorrhagic events (p = 0.07) or ischemic events (p = 0.45) between the 2 groups. One of the 2 M1 aneurysms resulted in procedure-related cerebral ischemia. One of the 2 distal MCA aneurysms required the removal of ICH after CE. In addition, the frequencies of rebleeding and retreatment after the first procedure did not differ significantly (both p = 1.00). On the other hand, the percentage of patients requiring post-treatment ICH removal was significantly higher in group M than in group O (10.3% vs. 0.5%, p = 0.0006). Three of 8 (37.5%) cases of ruptured MCA aneurysms with ICH had hematoma enlargement. These 3 cases required ICH removal after CE. A remaining patient who required hematoma removal had a hematoma with midline shift before CE, but CE was preceded by hematoma removal and the patient died. On the other hand, 9 (18.8%) of 48 cases with ruptured aneurysms with ICH at other sites had enlarged hematomas, and 2 (4.2%) of these cases required hematoma removal. The percentage of patients with ICH that required hematoma removal due to enlargement after CE was significantly higher for MCA aneurysm than for other sites (p = 0.01). The mortality rates did not differ significantly between the 2 groups (p = 0.56).

After case matching, each of the 39 patients was selected. There were no significant differences in patient characteristics (**Table 3**) and outcomes of CE (**Table 4**) between the 2 groups.

In the multivariate analysis of patients with ruptured MCA aneurysms (**Table 5**), only the presence of ICH

Table 1 Summary of the clinical characteristics of 468 patients in which aneurysmal SAH were treated with interventional radiology

	Total	Group M	Group O	p-Value
No. of patients	468	39 (8.3%)	429 (91.7%)	
Age, years				
Median	67.0	63.0	67.0	0.17
IQR	54.0–77.0	59.0–70.0	53.0–78.0	
Sex, female	336 (71.8%)	26 (66.7%)	310 (72.3%)	0.46
WFNS grade, 1–3	333 (71.1%)	28 (71.8%)	305 (71.1%)	1.00
mRS score before onset, 0–2	453 (97.0%)	69 (94.5%)	385 (97.5%)	0.25
Location				
MCA	39 (8.3%)			
M1 superior wall	2 (0.4%)			
Bifurcation	2 (0.4%)			
Distal MCA	35 (7.5%)			
ICA	176 (37.6%)			
Acom	63 (26.1%)			
A1	2 (0.8%)			
dACA	7 (2.9%)			
BA	65 (13.9%)			
VA	28 (6.0%)			
AN maximum size, mm				
Median	5.7	5.9	5.7	0.86
IQR	4.0–7.8	3.6–7.8	4.0–7.9	
Dome/neck ratio				
Median	1.54	1.63	1.54	0.92
IQR	1.23–2.06	1.20–2.12	1.24–2.05	
Aspect ratio				
Median	1.50	1.41	1.50	0.59
IQR	1.13–1.95	1.11–2.00	1.13–1.94	
ICH, yes	56 (12.0%)	8 (18.0%)	48 (11.2%)	0.12
Intra-aneurysmal thrombosis	10 (2.1%)	1 (2.6%)	9 (2.1%)	0.58
Adjunctive technique, yes	277 (59.2%)	17 (43.6%)	260 (60.6%)	0.04
Balloon	245 (52.4%)	14 (35.9%)	231 (53.8%)	
Double catheter	44 (9.4%)	6 (15.4%)	38 (8.9%)	
Stent	23 (4.9%)	1 (2.6%)	23 (5.4%)	
Follow-up period, median days (IQR)	61.0 (26.0–548.0)	491 (88.0–1416.0)	55 (25.0–439.8)	0.0002

Acom, anterior communicating artery; AN, aneurysm; BA, basilar artery; dACA, distal anterior cerebral artery; ICA, internal carotid artery; ICH, intracerebral hemorrhage; IQR, interquartile range; MCA, middle cerebral artery; mRS, modified Rankin Scale; SAH, subarachnoid hemorrhage; VA, vertebral artery; WFNS, World Federation of Neurosurgical Societies

was found to be associated with poor outcomes (odds ratio [OR]: 9.43, p = 0.03). There was a total of 8 patients with ICH in the MCA aneurysm group. Of these, 4 patients did not require hematoma removal after interventional radiology (IVR), with hematoma volumes of 5.3, 14.9, 24.1, and 32.7 ml, respectively. The remaining 4 patients required hematoma removal, with hematoma volumes of 36.7, 37.0, 67.4, and 124.7 ml, respectively. All patients with ruptured MCA aneurysm who needed removal had more than 36 ml of hematoma volume. The median hematoma volume was 19.5 and 52.5 ml, showing a significant difference (p = 0.03).

Discussion

In this study, we examined the treatment outcome of CE for ruptured MCA aneurysms, mainly composed of bifurcation aneurysms. We also compared the clinical characteristics and outcomes of such procedures with those of CE for ruptured aneurysms at other locations. The favorable outcome at discharge after CE was 61.5% for MCA aneurysms, which was not significantly different from the 55.9% of those for aneurysms at other sites. The rate of periprocedural complications, rebleeding rate, and retreatment rate did not differ significantly between the groups.

Table 2 Outcomes of 468 patients in which aneurysmal SAH was treated with interventional radiology

	Total	Group M	Group O	p-Value
Degree of occlusion				
BF	150 (32.1%)	17 (43.6%)	133 (31.0%)	0.11
NR	174 (37.2%)	12 (30.8%)	162 (37.8%)	
CO	144 (30.7%)	10 (25.6%)	134 (31.2%)	
Periprocedural hemorrhagic events	28 (6.0%)	5 (12.8%)	23 (5.4%)	0.07
Enlargement of ICH	12 (2.6%)	3 (7.7%)	9 (2.1%)	
Perforation	16 (3.4%)	2 (5.1%)	14 (3.3%)	
Periprocedural ischemic events	57 (12.2%)	6 (15.4%)	51 (11.9%)	0.45
Rebleeding after procedure	13 (2.8%)	1 (2.6%)	12 (2.8%)	1.00
Retreatment for aneurysm	29 (6.2%)	2 (5.1%)	27 (6.3%)	1.00
Removal of ICH	6 (1.3%)	4 (10.3%)	2 (0.5%)	0.0006
mRS score @ discharge				
0–2	264 (56.4%)	24 (61.5%)	240 (55.9%)	0.51
3–5	163 (34.6%)	13 (33.3%)	150 (35.0%)	
6	41 (8.7%)	2 (5.1%)	39 (9.1%)	0.56

BF, body filling; CO, complete occlusion; ICH, intracerebral hemorrhage; mRS, modified Rankin Scale; NR, neck remnant; SAH, subarachnoid hemorrhage

Table 3 Results of the univariate and multivariate analyses of the risk factors associated with poor outcomes in patients with ruptured MCA AN

	Outcome at discharge		Univariate		Multivariate	
	mRS score: 0–2 (good)	mRS score: 3–6 (poor)	OR (95% CI)	p-Value	OR (95% CI)	p-Value
No. of patients	24 (61.5%)	15 (38.5%)				
Age, median (IQR)	61.0 (50.5–67.0)	67.0 (60.0–75.0)	1.06 (0.99–1.15)	0.11		
Sex, female	13 (54.2%)	13 (86.7%)	5.50 (1.01–29.85)	0.04	4.06 (0.62–43.50)	0.13
WFNS grade, 4–5	3 (12.5%)	8 (53.3%)	8.00 (1.64–38.79)	0.01	4.55 (0.77–30.82)	0.05
AN maximum size (IQR)	5.8 (3.7–7.2)	5.9 (3.3–13.0)	1.13 (0.95–1.36)	0.46		
Aspect ratio (IQR)	1.4 (1.1–1.6)	1.6 (1.1–2.5)	2.84 (0.82–18.61)	0.33		
Dome/neck ratio (IQR)	1.8 (1.2–2.2)	1.4 (1.2–1.9)	0.59 (0.19–1.57)	0.44		
ICH	1 (4.2%)	6 (40.0%)	15.33 (1.61–145.90)	0.008	9.43 (1.18–201.30)	0.03
Result of occlusion, BF	10 (41.7%)	7 (46.7%)	1.23 (0.33–4.49)	1.00		

AN, aneurysm; BF, body filling; CI, confidence interval; ICH, intracerebral hemorrhage; IQR, interquartile range; MCA, middle cerebral artery; mRS, modified Rankin Scale; OR, odds ratio; WFNS, World Federation of Neurosurgical Societies

Based on these results, CE for ruptured MCA aneurysms was acceptable. On the other hand, the presence of ICH was found to be a prognostic factor of CE for MCA aneurysm. Post-treatment hematoma removal was required more often in cases of MCA aneurysm cases than in cases of other sites.

There were a few studies that reported the outcomes of CE for ruptured MCA aneurysms. In a case series of ruptured MCA aneurysms, 58.5% of CE patients had a good prognosis, which was not so different from the findings of the present study.[18] On the other hand, to our knowledge, there were no reports of CE for ruptured MCA aneurysm with ICH. A retrospective cohort study reported that the ICH rate was higher in cases with MCA aneurysm than in cases of aneurysm at other sites ($p < 0.001$, adjusted OR: 7.04).[19] They also revealed that the ICH volume was greater in cases with MCA aneurysm (median; 32 ml vs. 5 ml, $p < 0.0001$).[19] In addition, a previous study reported that IVR was associated with a higher incidence of enlargement of ICH than clipping.[20] Antithrombotic drugs are usually used during IVR, and this increases the risk of enlargement of ICH. As expected, a prospective cohort study reported that the hematoma enlargement was associated with poor prognosis.[21] Furthermore, a correlation between delayed hematoma removal and poor prognosis has been reported.[22] In this study, ICH due to MCA aneurysm often enlarged after CE and needed to be removed. In the case of MCA aneurysm with ICH, treatment options should be considered, such as hematoma removal before CE.

Procedure-related complications were prognostic factors, and because the MCA aneurysm was located distally, the risk of complications was expected to be high. In a

Table 4 Summary of the clinical characteristics of 78 case-matched patients in which aneurysmal SAH were treated with interventional radiology

	Total	Group M	Group O	p-Value
No. of patients	78	39	39	
Age, years				
Median	64.0	63.0	65.0	0.70
IQR	58.0–70.0	59.0–70.0	55.0–71.0	
Sex, female	52 (66.7%)	26 (66.7%)	26 (66.7%)	1.00
WFNS grade, 1–3	61 (78.2%)	28 (71.8%)	33 (84.6%)	0.27
mRS score before onset, 0–2	78 (100%)	39 (100%)	39 (100%)	–
AN maximum size, mm				
Median	5.8	5.9	5.6	0.70
IQR	3.9–8.2	3.6–7.8	4.2–8.6	
Dome/neck ratio				
Median	1.59	1.63	1.55	0.56
IQR	1.19–2.01	1.20–2.12	1.13–2.00	
Aspect ratio				
Median	1.43	1.41	1.45	0.82
IQR	1.09–1.79	1.11–2.00	1.04–1.77	
ICH, yes	15 (19.2%)	8 (18.0%)	7 (18.0%)	1.00
Intra-aneurysmal thrombosis	2 (2.6%)	1 (2.6%)	1 (2.6%)	1.00
Adjunctive technique, yes	46 (59.0%)	17 (43.6%)	29 (74.4%)	0.01
Balloon	40 (51.3%)	14 (35.9%)	26 (66.7%)	
Double catheter	10 (12.8%)	6 (15.4%)	4 (10.3%)	
Stent	4 (5.1%)	1 (2.6%)	3 (7.7%)	

AN, aneurysm; ICH, intracerebral hemorrhage; IQR, interquartile range; mRS, modified Rankin Scale; SAH, subarachnoid hemorrhage; WFNS, World Federation of Neurosurgical Societies

Table 5 Outcomes of 78 case-matched patients in which aneurysmal SAH were treated with interventional radiology

	Total	Group M	Group O	p-Value
Degree of occlusion				
BF	150 (32.1%)	17 (43.6%)	8 (20.5%)	0.05
NR	174 (37.2%)	12 (30.8%)	16 (41.0%)	
CO	144 (30.7%)	10 (25.6%)	15 (38.5%)	
Periprocedural hemorrhagic events	28 (6.0%)	5 (12.8%)	2 (5.1%)	0.43
Enlargement of ICH	12 (2.6%)	3 (7.7%)	0	
Perforation	16 (3.4%)	2 (5.1%)	2 (5.1%)	
Periprocedural ischemic events	57 (12.2%)	6 (15.4%)	9 (23.1%)	0.57
Rebleeding after procedure	13 (2.8%)	1 (2.6%)	0	1.00
Retreatment for aneurysm	29 (6.2%)	2 (5.1%)	3 (7.7%)	1.00
Removal of ICH	6 (1.3%)	4 (10.3%)	0	0.12
mRS score @ discharge				
0–2	264 (56.4%)	24 (61.5%)	27 (69.2%)	0.63
3–5	163 (34.6%)	13 (33.3%)	11 (35.0%)	
6	41 (8.7%)	2 (5.1%)	1 (2.6%)	1.00

BF, body filling; CO, complete occlusion; ICH, intracerebral hemorrhage; mRS, modified Rankin Scale; NR, neck remnant; SAH, subarachnoid hemorrhage

multicenter retrospective study of ruptured MCA aneurysm, CE-related symptomatic ischemia was 5.3%.[23] This result was less than the post-clipping ischemia (19.8%) in the same study ($p = 0.01$).[23] Another previous study of CE for ruptured MCA aneurysms reported a procedure-related complications rate of 5.1% and a disease-related complications rate of 8.5%.[18] In the current study, perioperative hemorrhagic and ischemic complications occurred after CE for a ruptured MCA aneurysm in 10.3% and 15.4% of cases, respectively. These results were more common than in previous reports but were thought to be due to differences in patient backgrounds and assessment methods. In addition, postoperative ischemia was often difficult to determine whether they were associated with the procedure or with SAH itself. Furthermore, it was often difficult to assess the presence or absence of symptoms due to

complications in patients with SAH. In this study, all cases with infarction on postoperative MRI were counted as ischemic complications. To reduce bias as much as possible, this study used case-matching analysis, and there was no difference in complication rates. This result suggested that CE for MCA aneurysms can be treated as safely as CE for other sites.

Rerupture would also affect the prognosis. Incomplete occlusion was reported to be one of the risk factors for rerupture.[24] The CO rate at 1 year after CE for ruptured MCA aneurysms was reported to be 41.7%.[25] In the current study, the CO rate was 25.6%, and this was lower than those found in previous studies. This was attributed to the fact that the differences in evaluators and the timing of evaluations had an impact. In our study, the CO rate for ruptured aneurysms at other sites was not significantly different from that for the ruptured MCA aneurysms. In a prospective study of CE for 72 MCA aneurysms, including both ruptured and unruptured aneurysms, the retreatment rate after 1 year was 9.7%.[26] In addition, a few retrospective observational studies have reported that the retreatment rate after CE was 4.9%–10.4%.[15,27–29] The rebleeding rate after CE for ruptured MCA aneurysms was similarly low, at 6.1%.[25] In this study, the rebleeding and retreatment rates were 2.6% and 5.1%, respectively, for ruptured MCA aneurysms, and these were not different from past reports. Furthermore, the MCA aneurysms had a longer follow-up period than the aneurysms at other sites; nevertheless, the rebleeding and retreatment rates for the MCA aneurysms were comparable to those of the aneurysms at other sites. This suggested that the recurrence-preventing effects of CE for ruptured MCA aneurysms were at least as good as those of CE for ruptured aneurysms at other sites.

It was expected that the outcome of IVR for MCA ruptured aneurysms would be inferior to that of other sites, but contrary to expectations, the outcome in this study was not different. Clipping may have been chosen in cases where the surgeon deemed it difficult to treat by CE, but the results suggested that good outcomes could be achieved for those cases deemed CE-eligible. On the other hand, many cases of MCA aneurysm with ICH had additional hematoma removal after CE, indicating the need for caution in treatment selection.

Limitations

This was a retrospective multicenter study, and there were differences in the treatment strategy, the timing of imaging studies, and the assessment of complications among the facilities. A case-matching study was conducted to reduce any bias due to differences in patient background, and there was no statistical difference in treatment outcomes between MCA aneurysms and other sites. However, no comparison with clipping was performed because the data on clipping cases were not available. Hence, our results may not apply to all ruptured MCA aneurysms. In addition, the post-treatment complications, retreatment, and rebleeding rates may have been underestimated because some patients were in poor condition and could not be evaluated with post-treatment imaging. In fact, 55% of cases did not have image follow-up.

Conclusion

The outcomes of CE for ruptured MCA aneurysms did not differ significantly from those of CE for ruptured aneurysms at other sites, suggesting outcomes of CE for ruptured MCA aneurysms were acceptable. On the other hand, for MCA aneurysm with ICH, the risk of hematoma enlargement after CE was high, and treatment selection should be done carefully.

Disclosure Statement

The authors declare that they have no conflicts of interest.

References

1) Berro DH, L'Allinec V, Pasco-Papon A, et al. Clip-first policy versus coil-first policy for the exclusion of middle cerebral artery aneurysms. *J Neurosurg* 2019; 133: 1124–1131.
2) Nussbaum ES, Madison MT, Goddard JK, et al. Microsurgical treatment of unruptured middle cerebral artery aneurysms: a large, contemporary experience. *J Neurosurg* 2018; 130: 1498–1504.
3) Dashti R, Hernesniemi J, Niemelä M, et al. Microneurosurgical management of middle cerebral artery bifurcation aneurysms. *Surg Neurol* 2007; 67: 441–456.
4) Jayaraman MV, Do HM, Versnick EJ, et al. Morphologic assessment of middle cerebral artery aneurysms for endovascular treatment. *J Stroke Cerebrovasc Dis* 2007; 16: 52–56.
5) Kalb S, Spetzler RF. Middle Cerebral artery bifurcation aneurysms: when and how to treat asymptomatic unruptured aneurysms. *World Neurosurg* 2015; 84: 620–622.
6) Ha SK, Lim DJ, Kang SH, et al. Analysis of multiple factors affecting surgical outcomes of proximal middle cerebral artery aneurysms. *Clin Neurol Neurosurg* 2011; 113: 362–367.

7) Iwama T, Yoshimura S, Kaku Y, et al. Considerations in the surgical treatment of superior-wall type aneurysm at the proximal (M1) segment of the middle cerebral artery. *Acta Neurochir (Wien)* 2004; 146: 967–972; discussion, 72.
8) Park DH, Kang SH, Lee JB, et al. Angiographic features, surgical management and outcomes of proximal middle cerebral artery aneurysms. *Clin Neurol Neurosurg* 2008; 110: 544–551.
9) Paulo MS, Edgardo S, Fernando M, et al. Aneurysms of the middle cerebral artery proximal segment (M1) anatomical and therapeutic considerations revision of a series. Analysis of a series of the pre bifurcation segment aneurysms. *Asian J Neurosurg* 2010; 5: 57–63.
10) Bhogal P, AlMatter M, Bäzner H, et al. Flow diversion for the treatment of MCA bifurcation aneurysms-a single centre experience. *Front Neurol* 2017; 8: 20.
11) Goertz L, Liebig T, Siebert E, et al. Woven endobridge embolization versus microsurgical clipping for unruptured anterior circulation aneurysms: a propensity score analysis. *Neurosurgery* 2021; 88: 779–784.
12) Hagen F, Maurer CJ, Berlis A. Endovascular treatment of unruptured MCA bifurcation aneurysms regardless of aneurysm morphology: short- and long-term follow-up. *AJNR Am J Neuroradiol* 2019; 40: 503–509.
13) Eboli P, Ryan RW, Alexander JE, et al. Evolving role of endovascular treatment for MCA bifurcation aneurysms: case series of 184 aneurysms and review of the literature. *Neurol Res* 2014; 36: 332–338.
14) Pierot L, Klisch J, Cognard C, et al. Endovascular WEB flow disruption in middle cerebral artery aneurysms: preliminary feasibility, clinical, and anatomical results in a multicenter study. *Neurosurgery* 2013;73:27–34; discussion34-5.
15) Report of World Federation of Neurological Surgeons Committee on a Universal Subarachnoid Hemorrhage Grading Scale. *J Neurosurg* 1988; 68: 985–986.
16) van Swieten JC, Koudstaal PJ, Visser MC, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988; 19: 604–607.
17) Kanat A, Aydin Y. Selection of cerebral aneurysms for treatment using guglielmi detachable coils: the Preliminary University of Illinois at Chicago Experience. *Neurosurgery* 1999; 45: 670–674.
18) Hagen F, Berlis A, Skalej M, et al. Endovascular treatment of ruptured middle cerebral artery bifurcation aneurysms. A retrospective observational study of short- and long-term follow-up. *Cardiovasc Intervent Radiol* 2021; 44: 587–595.
19) Darkwah Oppong M, Skowronek V, Pierscianek D, et al. Aneurysmal intracerebral hematoma: risk factors and surgical treatment decisions. *Clin Neurol Neurosurg* 2018; 173: 1–7.
20) Jabbarli R, Reinhard M, Roelz R, et al. Intracerebral hematoma due to aneurysm rupture: are there risk factors beyond aneurysm location? *Neurosurgery* 2016; 78: 813–820.
21) Brouwers HB, Chang Y, Falcone GJ, et al. Predicting hematoma expansion after primary intracerebral hemorrhage. *JAMA Neurol* 2014; 71: 158–164.
22) Wang WM, Jiang C, Bai HM. New insights in minimally invasive surgery for intracerebral hemorrhage. *Front Neurol Neurosci* 2015; 37: 155–165.
23) Sturiale CL, Scerrati A, Ricciardi L, et al. Clipping versus coiling for treatment of middle cerebral artery aneurysms: a retrospective Italian multicenter experience. *Neurosurg Rev* 2022; 45: 3179–3191.
24) Li K, Guo Y, Zhao Y, et al. Acute rerupture after coil embolization of ruptured intracranial saccular aneurysms: a literature review. *Interv Neuroradiol* 2018; 24: 117–124.
25) Guinto-Nishimura GY, Ramírez-Andrade JJ, Nathal E, et al. Treatment of middle cerebral artery aneurysms: a comparative study and proposed treatment algorithm. *Cir Cir* 2022; 90(S1): 84–91.
26) De Leacy R, Bageac DV, Siddiqui N, et al. Safety and long-term efficacy outcomes for endovascular treatment of wide-neck bifurcation aneurysms of the middle cerebral artery: insights from the SMART Registry. *Front Neurol* 2022; 13: 830296.
27) Brinjikji W, Lanzino G, Cloft HJ, et al. Endovascular treatment of middle cerebral artery aneurysms: a systematic review and single-center series. *Neurosurgery* 2011; 68: 397–402; discussion, 402.
28) Vendrell JF, Costalat V, Brunel H, et al. Stent-assisted coiling of complex middle cerebral artery aneurysms: initial and midterm results. *AJNR Am J Neuroradiol* 2011; 32: 259–263.
29) Zhou Y, Yang PF, Li Q, et al. Stent placement for complex middle cerebral artery aneurysms. *J Stroke Cerebrovasc Dis* 2014; 23: 1447–1456.

Current Oncology

MDPI

Article

Prognostic Factors for Patients with Small-Cell Lung Cancer Treated with Chemoimmunotherapy: A Retrospective Multicenter Study

Takashi Hatori [1,2], Takeshi Numata [2], Toshihiro Shiozawa [1,*], Manato Taguchi [3], Hirofumi Sakurai [4], Tomohiro Tamura [5], Jun Kanazawa [6], Hiroaki Tachi [7], Kyoko Kondo [8], Kunihiko Miyazaki [8], Norihiro Kikuchi [9], Koichi Kurishima [10], Hiroaki Satoh [11] and Nobuyuki Hizawa [1]

1 Department of Respiratory Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba 305-8575, Ibaraki, Japan; s2330424@u.tsukuba.ac.jp (T.H.)
2 Divisions of Respiratory Medicine, Mito Medical Center, Ibarakimachi 311-3193, Ibaraki, Japan
3 Division of Respiratory Medicine, Moriya Daiichi General Hospital, Moriya 302-0102, Ibaraki, Japan
4 Division of Respiratory Medicine, Ibaraki Seinan Medical Center Hospital, Sakai 306-0433, Ibaraki, Japan
5 Respiratory Center, Ibaraki Prefectural Central Hospital, Kasama 310-8555, Ibaraki, Japan
6 Department of Respiratory Medicine, National Hospital Organization, Ibaraki Higashi National Hospital, Tokai-Village 319-1113, Ibaraki, Japan
7 Divisions of Respiratory Medicine, Hitachi General Hospital, Hitachi 317-0077, Ibaraki, Japan
8 Division of Respiratory Medicine, Ryugasaki Saiseikai Hospital, Ryugasaki 301-0854, Ibaraki, Japan
9 Division of Respiratory Medicine, Kasumigaura Medical Center, Tsuchiura 300-8585, Ibaraki, Japan
10 Division of Respiratory Medicine, Tsukuba Medical Center Hospital, Tsukuba 305-8558, Ibaraki, Japan
11 Division of Respiratory Medicine, Mito Kyodo General Hospital, Mito 310-0015, Ibaraki, Japan
* Correspondence: t-shiozawa@md.tsukuba.ac.jp; Tel.: +81-29-853-3144; Fax: +81-29-853-7886

Abstract: Background: This study aimed to investigate prognostic factors for predicting the survival of patients with extensive-disease-stage small-cell lung cancer treated with chemoimmunotherapy. Methods: Patients were classified according to overall survival (OS): favorable corresponded to an OS $\geq$ 24 months, moderate corresponded to an OS of 6–24 months, and poor corresponded to an OS < 6 months. Multivariate Cox regression analyses were used to evaluate prognostic factors. Results: Of 130 patients, the proportions of performance status decline and liver metastasis were significantly higher in the poor-prognosis group. With regard to the laboratory findings, neutrophil/lymphocyte ratios and albumin levels differed significantly among the groups. Multivariate analysis showed that the independent prognostic factors for OS were liver metastasis and decreased albumin levels (<3.5 mg/dL). After classifying the patients into three groups according to the quantities of these prognostic factors, the OS differed significantly among the groups (18.3 vs. 13.5 vs. 3.8 months; $p < 0.001$). The incidence of immune-related adverse events (irAEs) was higher in patients without these prognostic factors than in those with both (36% vs. 5%; $p = 0.01$). Conclusion: Liver metastasis and decreased albumin levels are independent unfavorable prognostic factors. Patients with both prognostic factors showed unfavorable OS; however, patients without these factors may have a favorable prognosis but be at greater risk of irAEs.

Keywords: small cell lung cancer; immunotherapy; prognosis; albumin; liver metastasis; immune-related adverse event

check for updates

Citation: Hatori, T.; Numata, T.; Shiozawa, T.; Taguchi, M.; Sakurai, H.; Tamura, T.; Kanazawa, J.; Tachi, H.; Kondo, K.; Miyazaki, K.; et al. Prognostic Factors for Patients with Small-Cell Lung Cancer Treated with Chemoimmunotherapy: A Retrospective Multicenter Study. *Curr. Oncol.* **2024**, *31*, 6502–6511. https://doi.org/10.3390/curroncol31110482

Received: 27 September 2024
Revised: 19 October 2024
Accepted: 21 October 2024
Published: 23 October 2024

1. Introduction

Small-cell lung cancer (SCLC) is characterized by a more rapid progression than other histological types and is highly associated with smoking [1]. SCLC can be classified into two stages in terms of the choice of chemoradiotherapy or chemotherapy: limited disease (LD) and extensive disease (ED). LD-SCLC is defined as disease that is confined to the thorax, including the ipsilateral hilar, bilateral mediastinal, and bilateral supraclavicular lymph

nodes, for which treatment with thoracic irradiation is feasible. In contrast, ED-SCLC is defined as having distant metastasis or extending beyond the area for which thoracic radiation treatment is feasible. Because of the rapid tumor growth, more than two-thirds of patients with SCLC are diagnosed with ED. Combination chemotherapy with platinum and etoposide (ETP) has been positioned as the standard regimen for treating ED-SCLC. Although first-line chemotherapy shows high chemosensitivity, most patients relapse soon after treatment completion; thus, the prognosis of ED-SCLC is poor, with a median overall survival (OS) of less than 12 months [2,3]. Moreover, little progress has been made in treating ED-SCLC for over 20 years.

Recent pivotal trials using immune checkpoint inhibitors (ICIs) have led to a paradigm shift in ED-SCLC treatment. Impower133, a randomized, phase III trial, investigated the efficacy and safety of adding atezolizumab, an anti-programmed death ligand-1 (PD-L1) monoclonal antibody, to a combined carboplatin (CBDCA) and ETP regimen. Both progression-free survival (PFS) and OS were longer in the atezolizumab plus CBDCA and ETP therapy group compared with the control group treated with a placebo plus CBDCA and ETP (OS: hazard ratio [HR] = 0.75; 95% confidence interval [CI], 0.54–0.91; $p = 0.007$; PFS: HR = 0.77; 95% CI, 0.62–0.96, $p = 0.02$) [4]. In the phase III CASPIAN trial, a combination of platinum (cisplatin or CBDCA) and ETP plus durvalumab, an anti-PD-L1 monoclonal antibody, significantly prolonged OS compared to the control group, which was treated with platinum and ETP alone (HR = 0.73; 95% CI, 0.59–0.91; $p = 0.0047$) [5]. Thus, chemoimmunotherapy with combined platinum and ETP plus ICIs is the new standard treatment in the first-line setting of ED-SCLC.

A novel insight from these trials is that some patients treated with chemoimmunotherapy are long-term responders. In the IMpower133 trial, a survival difference of 13% was observed at 18 months between the atezolizumab group and the placebo group (34% in the atezolizumab group vs. 21% in the placebo group) [6]. In the CASPIAN trial, the 3-year OS rate in the durvalumab arm was approximately 17.6%, while it was only 5.8% in the control arm [7]. Although these results suggest that some patients will achieve a durable survival benefit with the addition of ICIs, others will experience relapses, and their prognosis will remain dismal. Therefore, identifying the prognostic factors helpful in predicting the survival of ED-SCLC patients treated with chemoimmunotherapy regimens is required. In non-small-cell lung cancer (NSCLC), tumor mutation burden (TMB) and the expression status of PD-L1 have been reported to be predictive factors of ICI treatment [8–10]; however, in contrast to the findings for NSCLC, previous studies showed that neither TMB nor PD-L1 expression status could predict the efficacy of ICI treatment for ED-SCLC [4,11,12].

Based on this background, we conducted a multi-institutional retrospective cohort study including 11 institutions in Ibaraki prefecture, Japan. This study was conducted to investigate prognostic factors helpful in predicting survival for patients with ED-SCLC treated with chemoimmunotherapy regimens.

2. Materials and Methods

2.1. Ethical Approval

The protocol of this study was approved by the institutional review board of Tsukuba University Hospital (approval number: R04-048). Owing to the retrospective nature of the analysis conducted, the requirement of informed consent from patients was waived. Instead, opt-out statements were published on the websites of each participating institution. This study did not receive funding from any for-profit or not-for-profit organizations or funding agencies.

2.2. Patients and Data Collection

Between September 2019 and May 2022, patients who were undergoing a chemoimmunotherapy regimen (Impower133 or CASPIAN regimen) were enrolled in this study. The data cut-off date was 30 September 2022. The inclusion criteria of this study were (1) pathologically diagnosed SCLC and (2) having undergone at least one cycle of a chemoimmunotherapy regimen. To evaluate clinical characteristics affecting survival, we classified enrolled patients into three groups based on previous reports [13,14]. Briefly, patients with an OSs of more than 24 months, 6–24 months, and less than 6 months were classified into favorable-, moderate-, and poor-prognosis groups, respectively.

We collected clinical data at the initiation of the chemoimmunotherapy regimen as baseline. The collected data included the following: age; sex; performance status (PS); smoking status (current, former, or never smoker); clinical stage; the presence of metastases in the brain, liver, or bone; and the history of radiotherapy before chemoimmunotherapy. We also collected the following laboratory data: counts of white blood cells and their fractions and levels of hemoglobin, platelets, albumin, lactate dehydrogenase (LDH), and progastrin-related protein (ProGRP). The cut-off values were defined as the upper limit of the normal value. The absolute number ratio of neutrophils to lymphocytes was calculated as the neutrophil-to-lymphocyte ratio (NLR), and its cut-off was defined as 5.0, as reported previously [15,16].

Each attending physician evaluated antitumor response using CT scans of the chest and abdomen and magnetic resonance imaging of the head. These evaluations were made according to the Response Evaluation Criteria in Solid Tumors version 1.1. Adverse events (AEs) were evaluated according to the Common Terminology Criteria for Adverse Events version 5.1.

2.3. Statistical Analysis

Chi-squared or Fisher's exact tests were applied in group comparisons of categorical variables. Continuous variables were compared using Mann–Whitney U tests. The Kaplan–Meier method was used to estimate median PFS and OS, and the log-rank test was applied to compare survivals among groups. The Cox regression model was used to investigate prognostic factors. Variables for multivariate analysis were selected based on their clinical significance and the results of univariate analysis. Hazard ratios in the multivariate analysis are reported with their 95% CIs. Statistical analyses were performed using IBM SPSS statistics (version 24.0) for Windows (IBM Corp., Armonk, NY, USA). All tests were two-sided, with p-values < 0.05 considered to indicate statistical significance.

3. Results

3.1. Patient Characteristics and Efficacy

In total, 130 patients from 11 institutions were enrolled. Table 1 shows the patients' characteristics at baseline. The median age was 71 years old (range, 42–85). Most patients were current or former smokers. Approximately 25% of the patients had a PS ranging from 2 to 4. The liver was the most frequent metastasis site (40/130, 31%). Regarding platinum doublet regimens, most patients received a CBDCA-based regimen. With respect to baseline characteristics, there were statistically significant differences in liver metastases and a PS decline at 2–4 across the groups, both of which were more frequent in the poor-prognosis group (p = 0.03 for liver metastasis, p = 0.02 for PS).

The median follow-up period was 9.3 months (95% CI, 6.7–12.9). In the entire population, the objective response rate and disease control rate were 61% (95% CI, 52–69) and 82% (95% CI, 75–89), respectively. The estimated median PFS was 6.4 months (95% CI, 5.8–7.1), and the 1-year PFS rate was 24%. The estimated median OS was 14.4 months (95% CI, 11.2–17.6). The 1-year OS rate was 55%.

Table 1. Patient characteristics.

	Total (n = 130)	Poor (n = 35)	Moderate (n = 77)	Favorable (n = 18)	p Value
Age	71 (42–85)	73 (46–82)	69 (42–84)	70 (62–85)	0.67
Gender					
Male	103	29	61	13	
Female	27	6	16	5	0.67
Smoking status					
Never	4	1	3	0	
Former or current	126	34	74	18	0.69
Performance status					
0	27	1	22	4	
1	71	19	41	11	
2	23	9	12	2	
3–4	9	6	2	1	0.02
Clinical stage					
IIIB-IIIC	10	3	6	1	
IVA	36	6	22	8	
IVB	80	26	47	7	
Recurrent	4	0	2	2	0.10
Metastatic site					
Brain	29	12	13	4	0.12
Liver	40	17	20	3	0.03
Bone	25	5	15	5	0.73
Prior radiotherapy					
Yes	27	6	16	5	
No	103	29	61	13	0.67
Platinum doublet regimen					
Cisplatin and etoposide	6	1	5	0	
Carboplatin and etoposide	124	34	72	18	0.56
Immune checkpoint inhibitors					
Atezolizumab	98	25	57	16	
Durvalumab	32	10	20	2	0.34

3.2. Laboratory Data at Baseline

We compared the laboratory data at baseline among the groups to identify the laboratory parameters affecting OS (Figure 1). The NLR was significantly lower in the favorable-prognosis group compared with that for the other two groups (Figure 1A). A statistically significant difference was also observed in serum albumin levels between the poor- and favorable-prognosis groups (p = 0.03, Figure 1D). There was no significant difference in other laboratory data among the groups.

3.3. Univariate and Multivariate Analyses

Table 2 presents the results of the univariate and multivariate analyses of OS. The univariate analysis showed that the presence of liver metastasis and albumin levels that have decreased to less than 3.5 mg/dL were associated with poor OS, while an NLR < 5 was significantly associated with favorable OS. Multivariate analysis demonstrated that liver metastases (p = 0.002, hazard ratio [HR]: 2.03, 95% CI: 1.25–3.30) and decreased albumin levels (p = 0.02, HR: 1.84, 95% CI: 1.14–2.96) were independent unfavorable prognostic factors associated with OS.

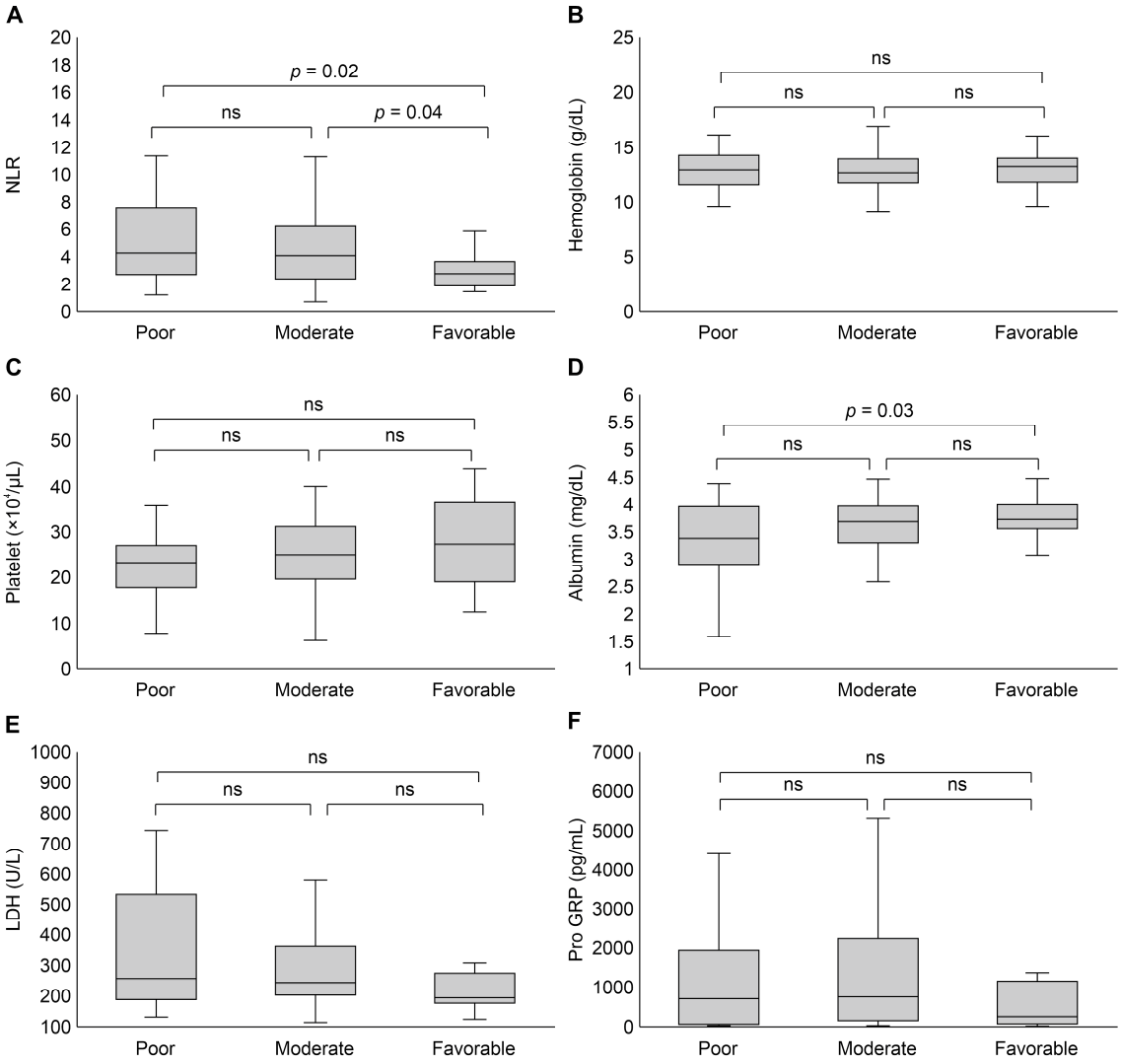


Figure 1. Group comparison of pretreatment laboratory data. (**A**) NLR, (**B**) hemoglobin, (**C**) platelet, (**D**) albumin, (**E**) LDH, and (**F**) ProGRP. Abbreviations: LDH, lactate dehydrogenase; NLR, neutrophil/lymphocyte ratio; ns, not significant; ProGRP, progastrin-releasing peptide.

Table 2. Results of univariate and multivariate analyses of overall survival.

	Univariate		Multivariate	
	HR (95% CI)	***p* Value**	**HR (95% CI)**	***p* Value**
Age (≥75 vs. <75)	1.43 (0.87–2.35)	0.16	1.66 (1.01–2.76)	0.06
PS (0–1 vs. 2–3)	0.89 (0.53–1.43)	0.59	0.79 (0.71–2.00)	0.50
Liver metastasis	1.81 (1.13–2.91)	0.01	2.03 (1.25–3.30)	0.002
NLR (<5.0 vs. ≥5.0)	0.60 (0.37–0.97)	0.04	0.97 (0.68–2.55)	0.07
Alb (<3.5 vs. ≥3.5)	2.02 (1.30–3.14)	<0.001	1.84 (1.14–2.96)	0.02

Abbreviations: CI, confidence interval; PS, performance status; NLR, neutrophil-to-lymphocyte ratio; Alb, albumin.

3.4. Number of Prognostic Factors and the Impact on Survival

We then tested the impact of the number of applicable prognostic factors on survival (Figure 2). The OS of patients with both liver metastasis and decreased albumin levels was 3.8 months, which was significantly shorter than that of patients with none or only one of these factors ($p < 0.001$). The OS of patients without these factors tended to be longer than that of patients with one of these factors.

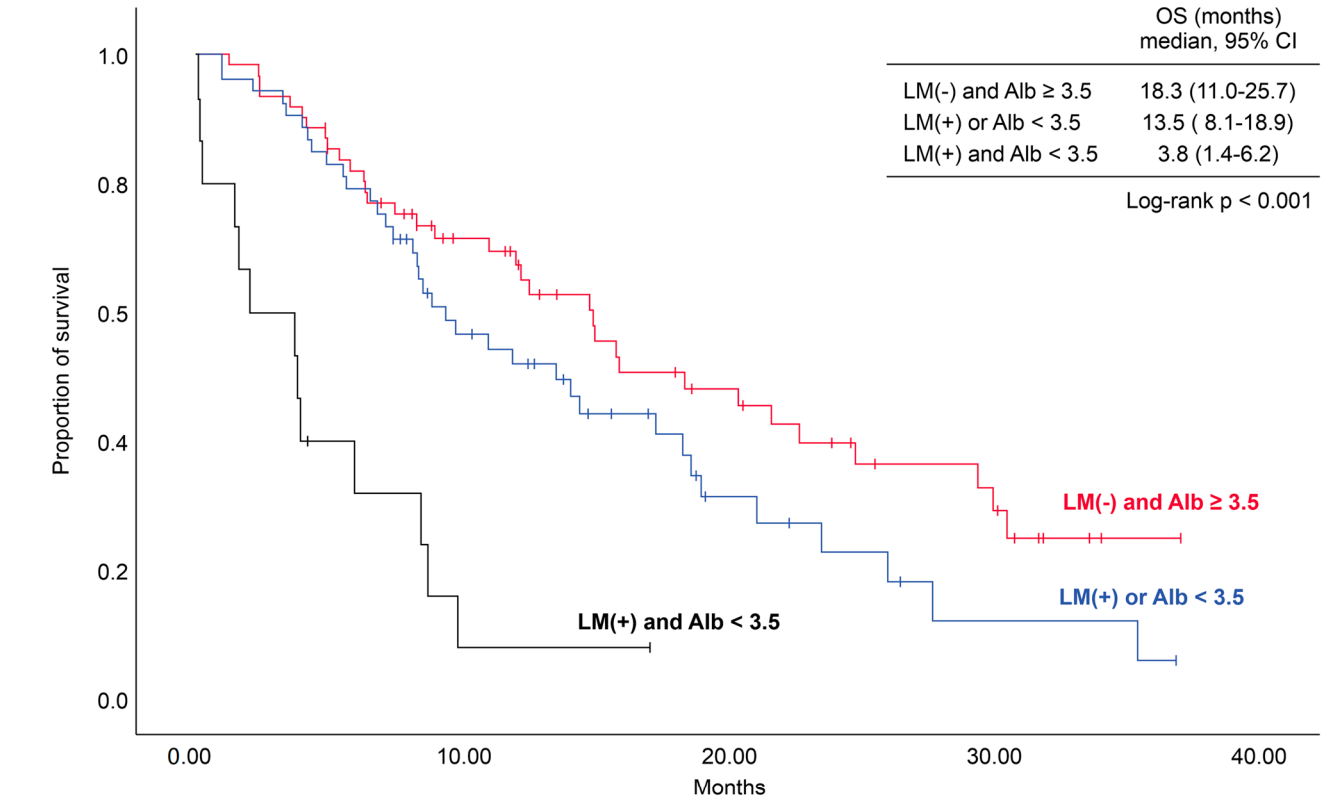


Figure 2. Kaplan–Meier survival curve according to the presence or absence of the prognostic factors. The following is a description of what each line indicates: red, patients without LM and Alb ≥ 3.5 mg/dL; blue, patients with LM or Alb < 3.5 mg/dL; black, patients with both LM and Alb < 3.5 mg/dL. Abbreviations: Alb, albumin; LM, liver metastasis; OS, overall survival.

3.5. Number of Prognostic Factors and the Impact on Immune-Related Adverse Event (irAE) Incidence

We further evaluated the relationship between the number of applied prognostic factors and the incidence of irAEs (Figure 3). In the overall population, pneumonitis was the most frequent irAE, at 6%, followed by thyroid dysfunction (4%) and hepatitis (3%). The frequencies of all grades of irAEs in patients with neither liver metastasis nor decreased albumin levels, either one, or both were 36%, 18%, and 5%, respectively, with a statistically significant difference among the groups ($p = 0.01$).

3.6. Subsequent Chemotherapy

At the cut-off date, 113 patients discontinued their chemoimmunotherapy regimens (97 because of disease progression and 16 because of AEs). Among them, 62 patients received subsequent chemotherapy: 39 patients were administered amrubicin; 19 patients received platinum doublet regimens, including the re-administration of their first-line regimens; and 4 patients received single-agent chemotherapy, such as irinotecan and nogitecan treatment.

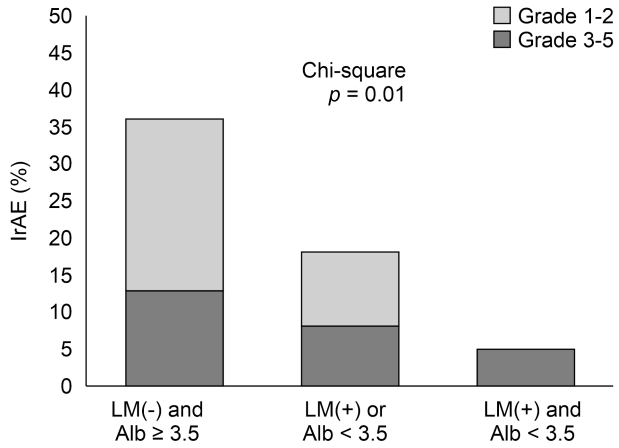


Figure 3. Comparison of the proportion of irAE incidence across the groups. Abbreviations: Alb, albumin; LM, liver metastasis; irAE, immune-related adverse event.

4. Discussion

In this study, we investigated prognostic factors for predicting the survival of patients with ED-SCLC undergoing chemoimmunotherapy regimens. Our results showed that liver metastasis and decreased albumin levels were independent unfavorable prognostic factors for OS. When the patients were classified based on the number of these prognostic factors, the OS differed significantly among groups. Additionally, the incidence of irAEs was more frequent in patients to which these factors did not apply, suggesting that the evaluation of these prognostic factors could also help predict the risk of irAEs. Because both liver metastasis and serum albumin levels are clinical parameters routinely assessed in diagnostic workups, the results of this study can be easily applied in clinical practice.

This study reveals that liver metastasis is an independent unfavorable prognostic factor of ED-SCLC. In regard to NSCLC, several previous studies reported that patients with liver metastasis had inferior responses to ICI regimens compared with patients with other-organ metastasis [17–20]. It remains unclear, however, whether liver metastasis affects the survival of patients with ED-SCLC treated with chemoimmunotherapy regimens. A previous study on a large ED-SCLC cohort reported that OS for patients with liver metastases was inferior to that for patients without liver metastasis (9.0 vs. 12.0 months, $p < 0.001$) [21]; however, this study was conducted before ICI-containing regimens were introduced for treating ED-SCLC. The results of this study indicate the necessity of paying attention to the presence of liver metastasis as an unfavorable prognostic factor for chemoimmunotherapy for ED-SCLC as well as NSCLC.

Serum albumin is a nutritional parameter commonly associated with cancer cachexia. Several previous studies reported prognostic significances of pretreatment albumin levels as prognostic factors of both NSCLC and SCLC [22–24]; however, most of these studies investigated the significance of albumin levels combined with other laboratory parameters such as albumin-to-alkaline phosphatase ratios, albumin-to-fibrinogen ratios, and albumin/globulin ratios. In this cohort study, we evaluated the significance of albumin alone, considering the simplicity of recording this parameter in daily clinical practice, and identified it as an unfavorable prognostic factor. We thus propose that albumin is a useful laboratory parameter for predicting survival in ED-SCLC.

In our study, we did not identify an independent favorable prognostic factor for patients with ED-SCLC treated with chemoimmunotherapy regimens. NLRs did not reveal statistically significant differences in multivariate analysis, although the NLR was lower in the favorable-prognosis group compared with that of moderate- or poor-prognosis groups. One way to interpret the results is based on the cut-off value of the NLR. Several studies considered the prognostic implication of the NLR in terms of both NSCLC and SCLC [15,16,25,26]; however, the cut-off values vary across studies. Moreover, both neutrophils and lymphocytes are affected by several cancer-associated events, such as

corticosteroid use as part of palliative treatment and coexisting infections. Further research is thus needed to identify the optimal NLR cut-off value for predicting the OS of patients with ED-SCLC.

We also examined the association between the number of prognostic factors and the incidence of irAEs. The incidence of irAEs was greater for patients without unfavorable prognostic factors, suggesting an association between the risk of irAEs and favorable prognosis. In NSCLC, the incidence of irAEs has been reported to be a favorable prognostic factor of ICI treatment [27–29]. In contrast, conflicting results have been reported regarding the association between irAEs and the outcome of ICI treatment for patients with ED-SCLC. Yokoo et al. assessed the development of irAEs and treatment efficacy for 40 patients with ED-SCLC treated with ICIs and platinum-plus-ETP regimens [30]. The median OSs were comparable between the irAE and non-irAE groups (27.6 vs. 24.9 months; $p = 0.268$). Nishimura et al. compared survival between patients with ED-SCLC who developed irAEs ($n = 23$) and those who did not ($n = 67$), and the median OS was longer for the patients with irAEs than those without irAEs (22 vs. 9.3 months, $p = 0.013$) [31]. Although it remains unclear whether the incidence of irAE is a prognostic factor in ED-SCLC, our results underscore the need for physicians to be aware of the risk of irAEs, especially in patients without liver metastasis and decreased albumin levels.

The limitations of the current study include the following. First, the current study was a retrospective cohort analysis with a limited sample size; hence, the validation of our findings using a prospective cohort is required. Second, the follow-up period in this study was relatively short. Third, the frequency of radiological evaluations varied among cases due to the retrospective nature of this study. Fourth, the proportion of patients with poor PS was relatively low, which may have influenced the results of this study. Finally, there is also a limitation regarding tumor markers. Neuron-specific enolase (NSE) is a useful tumor marker of SCLC, similar to ProGRP; however, NSE data were insufficient in this study because the NSE level was measured in a limited number of patients.

5. Conclusions

Liver metastasis and albumin levels at baseline were identified as independent unfavorable prognostic factors for patients with ED-SCLC undergoing chemoimmunotherapy regimens. Moreover, the number of prognostic factors was associated with both survival and the incidence of irAEs. Patients with both prognostic factors showed unfavorable OS, while patients without these factors were suggested to have a favorable prognosis but to be at a greater risk of irAEs.

Author Contributions: Conceptualization, T.H. and T.S.; methodology, T.S.; Investigation, T.S., T.N., M.T., H.S. (Hirofumi Sakurai), T.T., J.K., H.T., K.K. (Kyoko Kondo), K.M., N.K., K.K. (Koichi Kurishima) and H.S. (Hiroaki Satoh); data curation, T.H.; writing—original draft preparation, T.H.; writing—review and editing, T.S.; visualization, T.S.; supervision, H.S. (Hiroaki Satoh). and N.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the institutional review board of Tsukuba University Hospital (approval number: R04-048, date of approval 15 September 2022) for studies involving humans.

Informed Consent Statement: Owing to the retrospective nature of the analysis, the requirement to obtain informed consent from patients was waived. Instead, opt-out statements were published on the websites of each participating institution.

Data Availability Statement: The data used in the present study are available from the corresponding author upon request.

Acknowledgments: The authors thank all patients and the institutions for their cooperation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rudin, C.M.; Brambilla, E.; Faivre-Finn, C.; Sage, J. Small-cell lung cancer. *Nat. Rev. Dis. Primers* **2021**, *7*, 3. [CrossRef] [PubMed]
2. Farago, A.F.; Keane, F.K. Current standards for clinical management of small cell lung cancer. *Transl. Lung Cancer Res.* **2018**, *7*, 69–79. [CrossRef] [PubMed]
3. O'sullivan, D.E.; Cheung, W.Y.; Syed, I.A.; Moldaver, D.; Shanahan, M.K.; Bebb, D.G.; Sit, C.; Brenner, D.R.; Boyne, D.J. Real-World Treatment Patterns, Clinical Outcomes, and Health Care Resource Utilization in Extensive-Stage Small Cell Lung Cancer in Canada. *Curr. Oncol.* **2021**, *28*, 3091–3103. [CrossRef] [PubMed]
4. Horn, L.; Mansfield, A.S.; Szczesna, A.; Havel, L.; Krzakowski, M.; Hochmair, M.J.; Huemer, F.; Losonczy, G.; Johnson, M.L.; Nishio, M.; et al. First-Line Atezolizumab plus Chem-otherapy in Extensive-Stage Small-Cell Lung Cancer. *N. Engl. J. Med.* **2018**, *379*, 2220–2229. [CrossRef] [PubMed]
5. Paz-Ares, L.; Dvorkin, M.; Chen, Y.; Reinmuth, N.; Hotta, K.; Trukhin, D.; Statsenko, G.; Hochmair, M.J.; Özgüroğlu, M.; Ji, J.H.; et al. Durvalumab plus platinum-etoposide versus platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): A randomised, controlled, open-label, phase 3 trial. *Lancet* **2019**, *39*, 1929–1939. [CrossRef]
6. Liu, S.V.; Reck, M.; Mansfield, A.S.; Mok, T.; Scherpereel, A.; Reinmuth, N.; Garassino, M.C.; De Castro Carpeno, J.; Califano, R.; Nishio, M.; et al. Updated Overall Survival and PD-L1 Subgroup Analysis of Patients with Extensive-Stage Small-Cell Lung Cancer Treated with Atezolizumab, Carboplatin, and Etoposide (IMpower133). *J. Clin. Oncol.* **2021**, *39*, 619–630. [CrossRef]
7. Paz-Ares, L.; Chen, Y.; Reinmuth, N.; Hotta, K.; Trukhin, D.; Statsenko, G.; Hochmair, M.; Özgüroğlu, M.; Ji, J.; Garassino, M.; et al. Durvalumab, with or without tremelimumab, plus platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer: 3-year overall survival update from CASPIAN. *ESMO Open* **2022**, *7*, 100408. [CrossRef]
8. Yi, M.; Jiao, D.; Xu, H.; Liu, Q.; Zhao, W.; Han, X.; Wu, K. Biomarkers for predicting efficacy of PD-1/PD-L1 inhibitors. *Mol. Cancer* **2018**, *17*, 129. [CrossRef]
9. Mino-Kenudson, M.; Schalper, K.; Cooper, W.; Dacic, S.; Hirsch, F.R.; Jain, D.; Lopez-Rios, F.; Tsao, M.S.; Yatabe, Y.; Beasley, M.B.; et al. Predictive Biomarkers for Immunotherapy in Lung Cancer: Perspective from the International Association for the Study of Lung Cancer Pathology Committee. *J. Thorac. Oncol.* **2022**, *17*, 1335–1354. [CrossRef]
10. Russano, M.; La Cava, G.; Cortellini, A.; Citarella, F.; Galletti, A.; Di Fazio, G.R.; Santo, V.; Brunetti, L.; Vendittelli, A.; Fioroni, I.; et al. Immunotherapy for Metastatic Non-Small Cell Lung Cancer: Therapeutic Advances and Biomarkers. *Curr. Oncol.* **2023**, *30*, 2366–2387. [CrossRef]
11. Paz-Ares, L.; Garassino, M.C.; Chen, Y.; Reinmuth, N.; Hotta, K.; Poltoratskiy, A.; Trukhin, D.; Hochmair, M.J.; Özgüroğlu, M.; Ji, J.H.; et al. Durvalumab +/- tremelimumab + platinum-etoposide in extensive-stage small-cell lung cancer (CASPIAN): Outcomes by PD-L1 expression and tissue tumor mutational burden. *Clin. Cancer Res.* **2023**, *30*, 824–835. [CrossRef] [PubMed]
12. El Sayed, R.; Blais, N. Immunotherapy in Extensive-Stage Small Cell Lung Cancer. *Curr. Oncol.* **2021**, *28*, 4093–4108. [CrossRef] [PubMed]
13. Plaja, A.; Moran, T.; Carcereny, E.; Saigi, M.; Hernández, A.; Cucurull, M.; Domènech, M. Small-Cell Lung Cancer Long-Term Survivor Patients: How to Find a Needle in a Haystack? *Int. J. Mol. Sci.* **2021**, *16*, 13508. [CrossRef] [PubMed]
14. Wu, Y.; Zhang, J.; Zhou, W.; Yuan, Z.; Wang, H. Prognostic factors in extensive-stage small cell lung cancer patients with or-gan-specific metastasis: Unveiling commonalities and disparities. *J. Cancer Res. Clin. Oncol.* **2024**, *150*, 74. [CrossRef]
15. Imai, H.; Wasamoto, S.; Tsuda, T.; Nagai, Y.; Kishikawa, T.; Masubuchi, K.; Osaki, T.; Miura, Y.; Umeda, Y.; Ono, A.; et al. Using the neutrophil-to-lymphocyte ratio to predict the outcome of individuals with nonsquamous non-small cell lung cancer receiving pembrolizumab plus platinum and pemetrexed. *Thorac. Cancer* **2023**, *14*, 2567–2578. [CrossRef]
16. Mirili, C.; Guney, I.B.; Paydas, S.; Seydaoglu, G.; Kapukaya, T.K.; Ogul, A.; Gokcay, S.; Buyuksimsek, M.; Yetisir, A.E.; Karaalioglu, B.; et al. Prognostic significance of neutrophil/lymphocyte ratio (NLR) and correlation with PET–CT metabolic parameters in small cell lung cancer (SCLC). *Int. J. Clin. Oncol.* **2018**, *24*, 168–178. [CrossRef]
17. Tumeh, P.C.; Hellmann, M.D.; Hamid, O.; Tsai, K.K.; Loo, K.L.; Gubens, M.A.; Rosenblum, M.; Harview, C.L.; Taube, J.M.; Handley, N.; et al. Liver Metastasis and Treatment Outcome with Anti-PD-1 Monoclonal Antibody in Patients with Melanoma and NSCLC. *Cancer Immunol. Res.* **2017**, *5*, 417–424. [CrossRef]
18. Takeuchi, E.; Kondo, K.; Okano, Y.; Kunishige, M.; Kondo, Y.; Kadota, N.; Machida, H.; Hatakeyama, N.; Naruse, K.; Ogino, H.; et al. Early mortality factors in immune checkpoint inhibitor monotherapy for advanced or metastatic non-small cell lung cancer. *J. Cancer Res. Clin. Oncol.* **2022**, *149*, 3139–3147. [CrossRef]
19. Campos-Balea, B.; Carpeño, J.d.C.; Massutí, B.; Vicente-Baz, D.; Parente, D.P.; Ruiz-Gracia, P.; Crama, L.; Dols, M.C. Prognostic factors for survival in patients with metastatic lung adenocarcinoma: An analysis of the SEER database. *Thorac. Cancer* **2020**, *11*, 3357–3364. [CrossRef]
20. Wang, Q.; Fang, Y.; Li, C.; Leong, T.L.; Provencio, M.; Oh, I.-J.; Zhang, Z.; Su, C. Differential organ-specific tumor response to first-line immune checkpoint inhibitor therapy in non-small cell lung cancer—A retrospective cohort study. *Transl. Lung Cancer Res.* **2023**, *12*, 312–321. [CrossRef]
21. Ma, X.; Zhang, Z.; Chen, X.; Zhang, J.; Nie, J.; Da, L.; Hu, W.; Tian, G.; Wu, D.; Han, J.; et al. Prognostic factor analysis of patients with small cell lung cancer: Real-world data from 988 patients. *Thorac. Cancer* **2021**, *12*, 1841–1850. [CrossRef] [PubMed]
22. Zhou, S.; Wang, H.; Jiang, W.; Yu, Q.; Zeng, A. Prognostic Value of Pretreatment Albumin-to-Alkaline Phosphatase Ratio in Extensive-Disease Small-Cell Lung Cancer: A Retrospective Cohort Study. *Cancer Manag. Res.* **2020**, *12*, 2015–2024. [CrossRef] [PubMed]

23. Li, Q.; Li, L.; Wang, Y.; Xu, C.; Zou, J. The prognostic value of pretreatment albumin-to-fibrinogen ratio in small cell lung cancer patients receiving first-line platinum-based chemotherapy. *Heliyon* **2023**, *9*, e19225. [CrossRef] [PubMed]
24. Nakanishi, Y.; Masuda, T.; Yamaguchi, K.; Sakamoto, S.; Horimasu, Y.; Mimae, T.; Nakashima, T.; Miyamoto, S.; Tsutani, Y.; Iwamoto, H.; et al. Albumin–globulin ratio is a predictive biomarker of antitumor effect of anti-PD-1 antibody in patients with non-small cell lung cancer. *Int. J. Clin. Oncol.* **2019**, *25*, 74–81. [CrossRef] [PubMed]
25. Nasu, I.; Kondo, M.; Uozumi, R.; Takada, S.; Nawata, S.; Iihara, H.; Okumura, Y.; Takemoto, M.; Mino, K.; Sasaki, T.; et al. Prognostic Model of Baseline Medications plus Neutrophil-to-lymphocyte Ratio in Patients with Advanced Non-small-cell Lung Cancer Receiving Immune Checkpoint In-hibitor plus Platinum Doublet: A Multicenter Retrospective Study. *J. Cancer* **2023**, *14*, 676–688. [CrossRef]
26. Bi, H.; Ren, D.; Xiao, Y.; Zhou, Y.; Yi, B.; Han, W.; Shao, Y.; Wang, J.; Zhang, C.; Wang, H. Prognostic implications of neutrophil-to-lymphocyte ratio in patients with extensive-stage small cell lung cancer receiving chemoimmunotherapy: A multicenter, real-world study. *Thorac. Cancer* **2024**, *15*, 559–569. [CrossRef]
27. Haratani, K.; Hayashi, H.; Chiba, Y.; Kudo, K.; Yonesaka, K.; Kato, R.; Kaneda, H.; Hasegawa, Y.; Tanaka, K.; Takeda, M.; et al. Association of Immune-Related Adverse Events with Nivolumab Efficacy in Non–Small-Cell Lung Cancer. *JAMA Oncol.* **2018**, *4*, 374–378. [CrossRef]
28. Cook, S.; Samuel, V.; Meyers, D.E.; Stukalin, I.; Litt, I.; Sangha, R.; Morris, D.G.; Heng, D.Y.C.; Pabani, A.; Dean, M.; et al. Immune-Related Adverse Events and Survival Among Patients with Metastatic NSCLC Treated with Immune Checkpoint Inhibitors. *JAMA Netw. Open* **2024**, *7*, e2352302. [CrossRef]
29. Capella, M.P.; Pang, S.A.; Magalhaes, M.A.; Esfahani, K. A Review of Immunotherapy in Non-Small-Cell Lung Cancer. *Curr. Oncol.* **2024**, *31*, 3495–3512. [CrossRef]
30. Yokoo, K.; Kitamura, Y.; Suzuki, K.; Morikawa, K.; Sawai, T.; Honda, H.; Kudo, S.; Yamada, G. Relationship between immune-related adverse events and treatment effectiveness in extensive disease small cell lung cancer. *Thorac. Cancer* **2023**, *14*, 2251–2258. [CrossRef]
31. Nishimura, T.; Fujimoto, H.; Fujiwara, T.; Ito, K.; Fujiwara, A.; Yuda, H.; Itani, H.; Naito, M.; Kodama, S.; Furuhashi, K.; et al. Impact of immune-related adverse events on survival outcomes in extensive-stage small cell lung cancer patients treated with immune checkpoint inhibitors. *Cancer Med.* **2024**, *13*, e7188. [CrossRef] [PubMed]

Received: 25 June 2024 | Revised: 19 September 2024 | Accepted: 3 October 2024

DOI: 10.1002/cncr.35611

ORIGINAL ARTICLE

Outcomes of pregnancy in patients with chronic myeloid leukemia in the era of tyrosine kinase inhibitors

Takeshi Kondo MD, PhD[1] | Eri Matsuki MD, PhD[2] | Tomoiku Takaku MD, PhD[3,4] | Naoki Watanabe MD, PhD[4] | Chikashi Yoshida MD, PhD[5] | Masaya Okada MD, PhD[6] | Kazunori Murai MD, PhD[7] | Takashi Kodama MD, PhD[8] | Naoto Takahashi MD, PhD[9] | Shinya Kimura MD, PhD[10] | Itaru Matsumura MD, PhD[11] | for the Preg-CML/Japan Study Investigators

[1]Blood Disorders Center, Aiiku Hospital, Sapporo, Hokkaido, Japan

[2]Division of Hematology, Department of Medicine, Keio University School of Medicine, Tokyo, Japan

[3]Department of Hematology, Saitama Medical University, Iruma district, Saitama, Japan

[4]Department of Hematology, Juntendo University School of Medicine, Tokyo, Japan

[5]Department of Hematology, NHO Mito Medical Center, Ibaraki, Japan

[6]First Department of Internal Medicine, Kansai Medical University Medical Center, Moriguchi, Osaka, Japan

[7]Department of Hematology, Iwate Prefectural Central Hospital, Morioka, Japan

[8]Department of Reproductive Medicine, Hiroshima Prefectural Hospital, Minami-ku, Hiroshima, Japan

[9]Department of Hematology, Nephrology and Rheumatology, Akita University Graduate School of Medicine, Akita, Japan

[10]Division of Hematology, Respiratory Medicine and Oncology, Department of Internal Medicine, Saga University, Saga, Japan

[11]Department of Hematology and Rheumatology, Kindai University Faculty of Medicine, Osaka, Japan

Correspondence
Takeshi Kondo, Blood Disorders Center, Aiiku Hospital, 2-1, South 4, West 25, Chuo-ku, Sapporo, Hokkaido, Japan.
Email: kondo@aiiku-hp.or.jp

Funding information
Pfizer Health Research Foundation

Abstract

Background: Young female patients with chronic myeloid leukemia (CML) often face challenges becoming pregnant due to the teratogenicity of tyrosine kinase inhibitors (TKIs).

Methods: The authors conducted a nationwide survey of female patients with CML who experienced pregnancy between 2002 and 2020.

Results: Information for 70 pregnancies in 49 patients was obtained. There were three types of pregnancies: CML onset during pregnancy (n = 9), unplanned pregnancy mostly during treatment with a TKI (n = 25), and planned pregnancy during treatment-free remission (TFR) or treatment with interferon-alpha (IFN-α) (n = 36). The median duration from CML diagnosis to pregnancy in patients with planned pregnancy was significantly longer than that in patients with unplanned pregnancy (10.6 years vs. 4.1 years, $p < .001$). In 48 pregnancies that resulted in childbirth, TFR

A complete list of participating investigators in the Study of Pregnancy in Patients with CML in JAPAN (Preg-CML/JAPAN study) is provided in the Supporting Information.

This study is registered in the University Medical Information Network (UMIN000042762).

and treatment with IFN-α were chosen in 26 and 17 pregnancies, respectively. Sustained major or deeper molecular response was observed in 18 of 26 pregnancies with TFR. The patients who fulfilled the requirements for TKI therapy discontinuation by European LeukemiaNet recommendations achieved a TFR rate of 77% in pregnancy. Treatment with IFN-α might be effective for patients who are in complete cytogenetic response or deeper response (response rate, 76%).

Conclusion: Pregnancy by TFR or treatment with IFN-α could be a safe and feasible way for patients with CML. However, a substantial duration of treatment with a TKI before conception may be needed for planned pregnancy. Planning and evaluation for pregnancy should be considered at the time of CML onset for female patients with childbearing potential.

KEYWORDS
chronic myeloid leukemia, interferon-alpha, pregnancy, treatment discontinuation, treatment-free remission, tyrosine kinase inhibitor

INTRODUCTION

The introduction of tyrosine kinase inhibitors (TKIs) has greatly improved the survival of patients with chronic myeloid leukemia (CML), aligning their life expectancy with that of the general population.[1,2] Many CML patients achieve and maintain a deep molecular response (DMR), making them potential candidates for treatment-free remission (TFR).[3,4] TFR is defined as maintaining a major molecular response (MMR) or DMR after stopping TKI therapy without needing to resume treatment.

Although the median age at diagnosis of CML is approximately 55 years, CML also affects women of childbearing age.[5] TKIs have potential risks of fetal abnormalities and spontaneous abortion, particularly when administered during the first trimester as a period of organogenesis.[6,7] In addition, a TKI can be transferred to the fetus before the establishment of the blood–placental barrier within 16 gestational weeks. Among TKIs, dasatinib can cross the blood–placental barrier in the second trimester and could thus affect fetal development.[8–10]

Therefore, TKIs are fundamentally contraindicated for pregnant women. Planned pregnancy during TFR is an ideal way to avoid TKI exposure. However, not all conceptions are planned, not all patients fulfill the eligibility criteria for TFR, and CML develops in some patients during pregnancy. To address these issues, we conducted a nationwide survey in Japan regarding pregnancy outcomes in patients with CML.

MATERIALS AND METHODS

Study design

We designed a retrospective observational study focusing on female patients with CML who experienced conception and pregnancy. We performed a nationwide survey of all hematology centers certified by the Japanese Society of Hematology. The survey consisted of two steps. The initial survey was performed to obtain information on 1) the number of female patients aged 45 years or younger who were diagnosed as having CML in the chronic phase (CML-CP) between 2002 and 2020, and 2) whether these patients became pregnant or not. The second survey was performed for the hematology centers in which there were pregnant patients with CML to obtain information on treatment, conception and pregnancy outcomes. Detailed methods are described in the Supporting Information. The clinical study was approved by the institutional review boards of Aiiku Hospital and each hospital participating in the second survey. This study is registered in the University Medical Information Network (UMIN000042762).

Evaluation of disease status

The disease status of CML was basically assessed according to the treatment response criteria defined by the European LeukemiaNet (ELN).[11] In Japan, *BCR::ABL1* mRNA levels were measured by the transcription-mediated amplification and hybridization protection (TMA) method until March 2015 and by standardized *BCR::ABL1* quantification on the international scale thereafter. An undetectable *BCR::ABL1* level by the TMA method (TMA-UD) was equivalent to MMR or a deeper response.[12,13] Because the specific recommendations for TFR were not established at the time when the patients became pregnant, TKI discontinuation is referred to as TFR if MMR or deeper response was achieved before TKI discontinuation. Failure of TFR was defined as loss of MMR according to the previous TFR study.[14]

RESULTS

Patient characteristics

We sent the initial survey questionnaires to 463 hematological centers across Japan and received replies from 206 centers (recovery

rate: 44.5%). Between 2002 to 2020, 853 female patients aged 45 years or younger were diagnosed as having CML-CP. Among them, 78 patients in 55 centers became pregnant and 775 young female patients did not become pregnant (pregnancy rate of 9.1%). The reasons for 363 patients not becoming pregnant were as follows: 289 patients (79.6%) had no desire to bear children, 49 patients (13.5%) were advised against pregnancy by their physicians, and 25 patients (6.9%) did not become pregnant despite attempting to conceive. Thirty-six centers participated in the second survey, providing information on 70 pregnancies in 49 patients (Figure S1).

There were three types of pregnancies: CML onset during pregnancy (nine pregnancies in nine patients), unplanned pregnancy during treatment or stopping TKI treatment (25 pregnancies in 24 patients), and planned pregnancy where TKI was discontinued or replaced by interferon-alpha (IFN-α) before conception (36 pregnancies in 24 patients) (Table 1).

The use of TKIs before conception was categorized into three groups: imatinib alone (23 pregnancies in 13 patients), imatinib followed by second-generation TKIs (2G-TKIs) (18 pregnancies of 12 patients), and primarily 2G-TKIs (20 pregnancies of 18 patients) (details are presented in Figure S2).

Twenty-two of 25 unplanned pregnancies and 18 of 36 planned pregnancies were the first pregnancies after diagnosis of CML. Times of pregnancy varied from one to five, and times of childbirth ranged from zero to three (Figure 1). The median durations from CML onset to the first pregnancy were 9.1 years in planned pregnancies and 4.1 years in unplanned pregnancies. The duration from CML onset to the first pregnancy in planned pregnancies was significantly longer than that in unplanned pregnancies (unpaired t-test, $p = .002$) (Table 1; Figure S3A). Similarly, the median duration from CML onset to all pregnancies in planned pregnancies was significantly longer than that in unplanned pregnancies (10.6 years vs. 4.1 years, Mann-Whitney test, $p < .001$) (Table 1; Figure S3B). As a result, the median age at pregnancy in patients with planned pregnancy was 36 years, which was the oldest among the three types of pregnancies (Kruskal–Wallis test, $p = .0218$). The patient's age at pregnancy was significantly older for patients with planned pregnancy than for patients with unplanned pregnancy (median age: 36 years vs. 28 years, Mann-Whitney test, $p < .001$) and patients with CML onset during pregnancy (median age: 36 years vs. 28 years, Mann-Whitney test, $p = .003$) (Figure S3C).

Types of conception and outcomes of pregnancy

Outcomes of pregnancy according to types of conception are summarized in Table 2 and Figure 2.

Nine patients developed CML during pregnancy (Table S1). Elective abortion was performed in two pregnancies in the first trimester, and seven pregnancies resulted in live births (Table 2). Two patients in the first trimester and one patient in the third trimester had no drug treatment until delivery. Three patients commenced treatment with IFN-α. Two patients were treated with imatinib after either no treatment or treatment with IFN-α in the third trimester and delivered while in a complete hematologic response (Figure 2A; Table S1).

For 25 unplanned pregnancies, six were conceived at less than MMR and 19 with MMR or DMR. One pregnancy occurred during TKI discontinuation due to nonadherence and another during TFR clinical trial. The other 23 pregnancies were established during TKI treatment. Information for gestational age at positive pregnancy test was available in 24 pregnancies; the median was 5 (range, 2–20) weeks. Elective abortion was chosen in eight pregnancies. There were three spontaneous abortions and 14 live births. Elective abortion was chosen by four of six patients with no MMR (67%) and by four of 19 patients with MMR or DMR (21%) (Table 2). Although this difference was not statistically significant, the response to treatment could be a determinant for continuing pregnancy. Of the 14 live births, no treatment was chosen in one patient with nonadherence by the patient's self-decision, TFR was initiated for 11 pregnancies, and IFN-α treatment was commenced for two patients. Loss of MMR occurred in four cases during pregnancy. No treatment was performed in two patients until delivery, and treatment with either a TKI or IFN-α was initiated in two patients (Figure 2B).

Planned pregnancies were established in 36 pregnancies including 20 pregnancies in TFR and 16 pregnancies during IFN-α treatment. Thirty-three pregnancies were conceived in MMR or DMR status, whereas MMR was lost at the time of a positive pregnancy test in three patients. There were seven spontaneous abortions and one stillbirth, and one pregnancy was ongoing at the time of investigation (Table 2). Therefore, 27 pregnancies resulted in live births. Loss of MMR was observed in five patients during pregnancy. No treatment was introduced in one patient until delivery, and treatment with either imatinib or IFN-α was initiated in four patients after loss of MMR (Figure 2C).

Treatment discontinuation for pregnancy and delivery

Treatment discontinuation for pregnancy was conducted in 26 pregnancies for which there were 26 childbirths. Because all of the patients had achieved MMR or deeper response at the time of TKI discontinuation, the discontinuation of treatment is referred as TFR here. TFR was attempted in 11 unplanned pregnancies at the time of positive pregnancy test and in 15 planned pregnancies in which TKI was discontinued before pregnancy (Table 3). The disease status at the time of a positive pregnancy test varied from complete cytogenetic response (CCyR) to DMR: one patient in CCyR, four patients in MMR, 19 patients in DMR, and two patients in TMA-UD.

One patient became pregnant during TFR for a planned pregnancy but lost MMR at the time of a positive pregnancy test, and that patient was treated with imatinib after the second trimester until delivery (Figure 2C). Three of the four patients who were in MMR at the time of a positive pregnancy test lost MMR during pregnancy. In the 19 patients who were in DMR at the time of a positive pregnancy test, loss of MMR during pregnancy occurred in three patients. MMR was lost in one of the two patients who were in TMA-UD at the time of a positive

TABLE 1 Characteristics of patients.

Demographic characteristics	Total	CML onset during pregnancy	Unplanned pregnancy	Planned pregnancy
Pregnant patients, n1				
1st pregnancy	49	9	22	18
2nd pregnancy	14	0	3	11
3rd pregnancy	5	0	0	5
4th pregnancy	1	0	0	1
5th pregnancy	1	0	0	1
Total pregnancies, n2	70	9	25	36
Age at CML diagnosis, median (range), year	26 (9–40)	28 (21–39)	24 (9–40)	26 (15–34)
Duration from diagnosis to the first pregnancy, median (range), year	5.8 (0.8–16.1)	Not applicable	4.1 (0.8–14.2)[a]	9.1 (4.2–16.1)
TKI used before pregnancy, n2				
Imatinib	23	0	8	15
Imatinib + 2G-TKI	18	0	6	12
2G-TKI	20	0	11	9
Age at first pregnancy, median (range), year	34 (21–42)	28 (21–39)	27 (21–42)	36 (27–42)
Duration from diagnosis to pregnancy, median (range), year	7.2 (0.8–16.6)	Not applicable	4.1 (0.8–16.2)[a]	10.6 (1.7–16.6)
Age at pregnancy, median (range), year	33 (21–42)	28 (21–39)	28 (21–42)	36 (27–42)
Procedure of fertilization, n2				
Natural conception	57	9	25	23
ART	13	0	0	13
Treatment at the time of positive pregnancy test, n2				
None	10	9	1	0
TKI	23	0	23	0
IFN	16	0	0	16
TFR	21	0	1	20
Disease status at pregnancy, n2				
Without MMR	18	9	6	3
MMR/DMR	52	0	19	33
Outcome of pregnancy, n2				
Spontaneous abortion	10	0	3	7
Elective abortion	10	2	8	0
Stillbirth	1	0	0	1
Live birth	48	7	14	27
During pregnancy	1	0	0	1
Initial treatment during pregnancy, n3				
None	6	5	1	0
IFN	16	2	2	12
TFR	26	0	11	15

TABLE 1 (Continued)

Demographic characteristics	Total	CML onset during pregnancy	Unplanned pregnancy	Planned pregnancy
Treatment at delivery, n3				
None	4	3	1	0
Imatinib	4	2	1	1
IFN	19	2	2	15
TFR	21	0	10	11
Disease status at delivery, n3				
Without MMR	14	7	4	3
MMR/DMR	34	0	10	24

Note: n1, number of patients; n2, number of pregnancy; n3, number of live birth.

Abbreviations: ART, assisted reproductive technology; CML, chronic myeloid leukemia; DMR, deep molecular response; IFN, interferon; MMR, major molecular response; TFR, treatment-free remission; 2G-TKI, second-generation tyrosine kinase inhibitor.

[a]Data for one patient with unplanned pregnancy was missing.

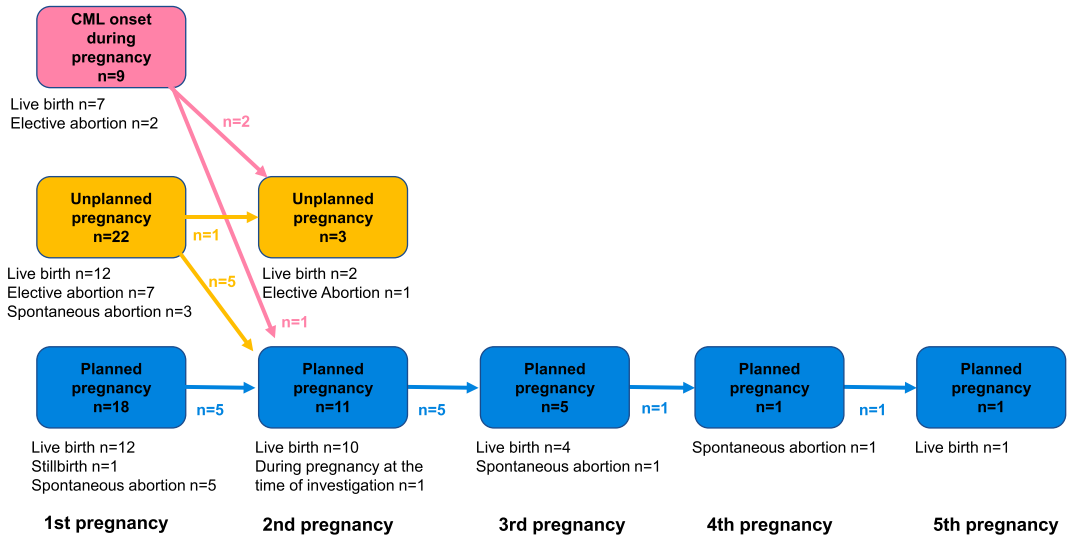


FIGURE 1 Pregnancy pattern and outcome of pregnancy in patients with chronic myeloid leukemia (CML). There were three types of pregnancy: CML onset during pregnancy, unplanned pregnancy during the clinical course of CML and planned pregnancy with treatment-free remission (TFR) or treatment with interferon-alpha (IFN-α) before conception. Times of pregnancy varied from 1 to 5. Times of childbirth varied from 0 to 3.

pregnancy test. As a result, loss of MMR was observed in eight of the 26 pregnancies: no treatment was introduced in three pregnancies and treatment with either IFN-α or imatinib was commenced in five pregnancies. Thus, TFR was sustained in 18 of 26 pregnancies, giving a 69% TFR rate. The median durations from diagnosis to pregnancy and sustained TFR rates by prior TKIs use were 12.0 years and 67% for imatinib (n = 8, data missing for one patient), 11.2 years and 71% for imatinib and subsequent 2G-TKIs (n = 7), and 4.6 years and 70% for primary 2G-TKI use (n = 10). Although sustained TFR rates were similar among the three groups, the duration from diagnosis to pregnancy in patients treated with 2G-TKIs was significantly shorter than the durations in the other two groups (Figure S4).

The TFR rate in patients with DMR at the time of a positive pregnancy test was significantly higher than that in patients with MMR (84% vs. 25%, Fisher's exact test, p = .040). We also examined the association of duration of treatment and DMR with successful TFR for pregnancy. Information was available for 24 pregnancies in 22 patients. Requirements for discontinuation of TKI therapy were proposed by the ELN recommendations: duration of TKI therapy >5 years (>4 years for 2G TKI) and duration of DMR (MR4 or better) >2 years.[4] Thirteen patients with pregnancies fulfilled these requirements and the TFR rate during pregnancy was 77%. On the other hand, the TFR rate was 55% in 11 pregnancies that did not meet the requirements.

Clinical use of IFN-α for pregnancy and delivery

Treatment with IFN-α was primarily used in 17 pregnancies including three patients with onset of CML during pregnancy, two unplanned pregnancies, and 12 planned pregnancies (Figure 2; Table 4).

TABLE 2 Outcomes of pregnancies.

Situation of pregnancy	Pregnancy (*n*)	Response at pregnancy (*n*)	Treatment at pregnancy (*n*)	Outcome of pregnancy (*n*)	Initial treatment during pregnancy with childbirth (*n*)	Treatment at delivery (*n*)	Response at delivery (*n*)
CML onset during pregnancy	9	No MMR (9)	No treatment (9)	Live birth (7)	None (4)	None (3)	No MMR (7)
						IMA (1)	
				Elective abortion (2)	IFN (3)	IFN (2)	
						IMA (1)	
Unplanned pregnancy	25	No MMR (6)	Nonadherence (1)	Live birth (1)	None (1)	None (1)	No MMR (1)
			TKI (5)	Live birth (1)	IFN (1)	IFN (1)	MMR/ DMR (1)
				Elective abortion (4)			
		MMR/ DMR (19)	TKI (18)	Live birth (12)	TFR (11)	TFR (2)	No MMR (3)
				Spontaneous abortion (3)		IFN (1)	
				Elective abortion (3)		TFR (8)	MMR/ DMR (8)
			TFR (1)	Elective abortion (1)	IFN (1)	IMA (1)	MMR/ DMR (1)
Planned pregnancy	36	No MMR (3)	IFN (1)	Spontaneous abortion (1)	TFR (1)	IMA (1)	No MMR (1)
			TFR (2)	Live birth (1)			
				Spontaneous abortion (1)			
		MMR/ DMR (33)	IFN (15)	Live birth (12)	IFN (12)	IFN (12)	MMR/ DMR (12)
				Stillbirth (1)			
				Spontaneous abortion (2)			
			TFR (18)	Live birth (14)	TFR (14)	TFR (11)	No MMR (1)
							MMR/ DMR (10)
				During pregnancy (1)		IFN (3)	No MMR (1)
				Spontaneous abortion (3)			MMR/ DMR (2)

Abbreviations: ART, assisted reproductive technology; CML, chronic myeloid leukemia; DMR, deep molecular response; IFN, interferon-α; IMA, imatinib; MMR, major molecular response; TFR, treatment-free remission.

Treatment with IFN-α was commenced as a primary treatment for the three patients with CML onset during pregnancy, and there was no hematologic response at the time of delivery. In the two unplanned pregnancies, IFN-α was chosen as a primary treatment. One patient in CCyR at the time of a positive pregnancy test achieved MMR at the time of delivery, and another patient in DMR at the time of a positive pregnancy test lost MMR during treatment with IFN-α. In the 12 planned pregnancies, treatment with IFN-α had been initiated before conception, and MMR or DMR was sustained until delivery. As a result, treatment with IFN-α resulted in maintenance or achievement of MMR or DMR in 13 of 17 pregnancies (76%).

Treatment with IFN-α was used as a salvage treatment for loss of MMR during pregnancy in four patients in whom TFR was attempted, including one patient with an unplanned pregnancy and three patients with planned pregnancies (Figure 2B,C; Table 4). MMR or DMR was achieved again by treatment with IFN-α in two patients. In addition, TFR was attempted in two patients before conception; however, they lost MMR before pregnancy (Figure 2C). After achieving DMR or TMA-UD again, treatment with IFN-α was initiated for patients with planned pregnancy, and MMR or TMA-UD was sustained in those patients during pregnancy. As a result, MMR or DMR was achieved again or was sustained in four of the six patients

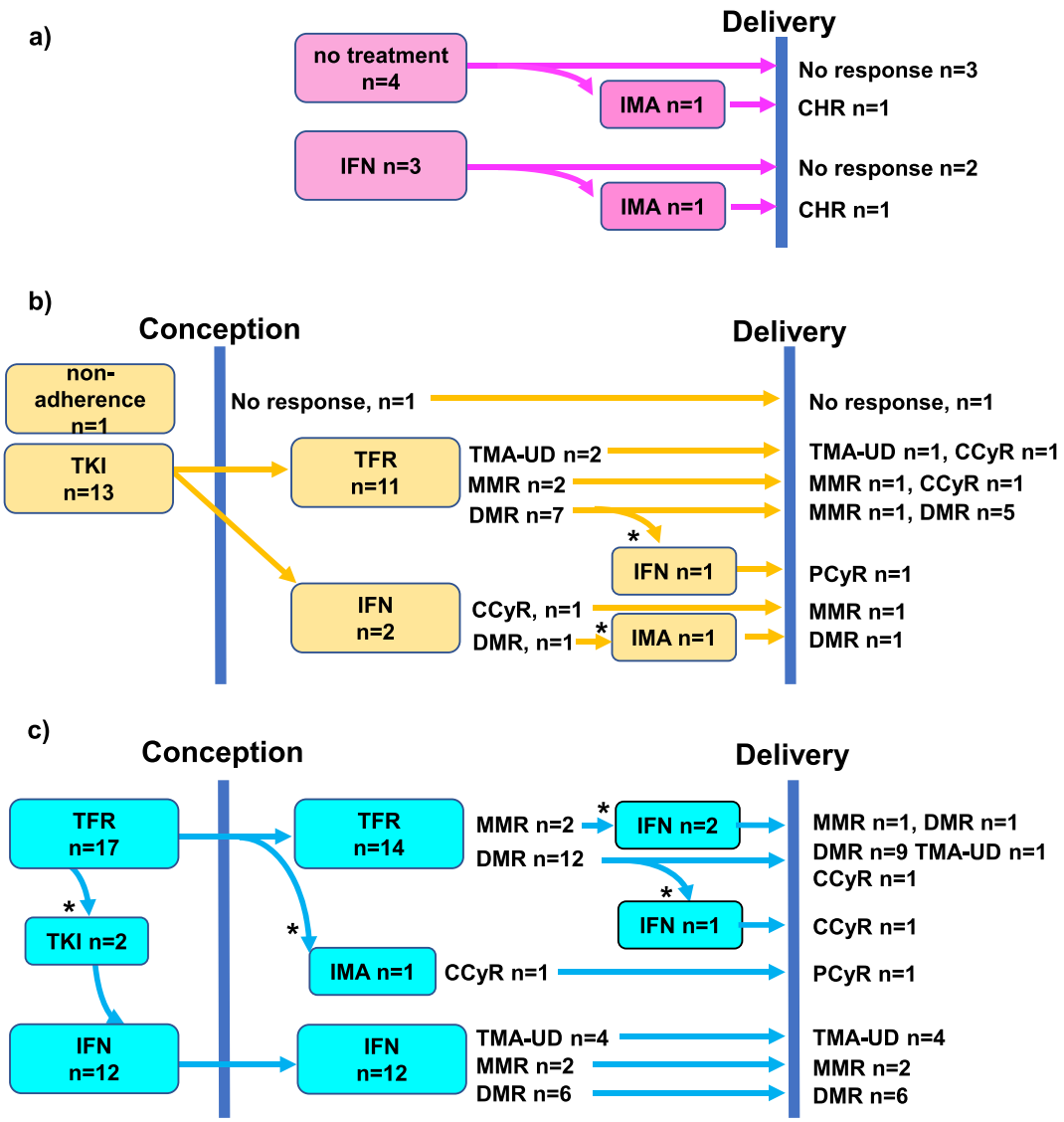


FIGURE 2 Treatment and response of CML during pregnancy and at delivery. (A) Cases of CML onset during pregnancy. (B) Cases with unplanned pregnancy. (C) Cases with planned pregnancy. *Treatment was changed because loss of MMR was observed. CCyR, complete cytogenetic response; CML, chronic myeloid leukemia; CHR, complete hematologic response; DMR, deep molecular response; IFN, interferon-α; IMA, imatinib; MMR, major molecular response; PCyR, partial cytogenetic response; TFR, treatment-free remission; TKI, tyrosine kinase inhibitor; TMA-UD, undetectable level of *BCR::ABL1* mRNA by the TMA method (equivalent to MMR or deeper response).

(Table S2). The durations of TFR were 62 days and 95 days in two patients for whom treatment with IFN-α was ineffective and 171 days or longer in four patients for whom treatment with IFN-α was effective.

Pregnancy by assisted reproductive technology

Thirteen planned pregnancies were achieved via assisted reproductive technology (ART) in 10 patients (Table 1). Seven patients had cryopreserved eggs or embryos at CML diagnosis, resulting in seven pregnancies of five patients during TFR or while on IFN-α after achieving DMR with TKIs. In contrast, egg retrieval and ART were performed during TKI discontinuation after achieving DMR with TKIs in six pregnancies of five patients. The median durations from TKI discontinuation to the first positive pregnancy test via ART were 131 (range, 18–308) days for patients with cryopreserved eggs at the time of CML diagnosis, 390 (range, 104–2136) days for patients who underwent egg retrieval during TKI discontinuation, and 226 (range, 20–1655) days for patients who conceived naturally (Table 5).

Complications of pregnancy

There were 10 spontaneous abortions including seven in planned pregnancy and three in unplanned pregnancy (Table 6). In planned pregnancies, spontaneous abortions occurred in patients older than 25 years of age who were treated with IFN-α or TFR at conception.

TABLE 3 TFR attempt for pregnancy in patients with CML.

Situation of pregnancy	Patient (*n*)	Pregnancy (*n*)	Treatment at pregnancy	Response at pregnancy (*n*)	Initial treatment during pregnancy	Treatment at delivery (*n*)	Response at delivery (*n*)
Unplanned pregnancy	10	11	TKI	MMR (2)	TFR	TFR (2)	CCyR (1)
							MMR (1)
				DMR (7)	TFR	TFR (6)	MMR/DMR (6)
						IFN (1)[a]	PCyR (1)
				TMA-UD (2)	TFR	TFR (2)	CCyR (1)
							TMA-UD (1)
Planned pregnancy	15	15	TFR	CCyR (1)	TFR	IMA (1)[a]	PCyR (1)
				MMR (2)	TFR	IFN (2)[a]	MMR/DMR (2)
				DMR (12)	TFR	TFR (11)	CCyR (1)
							MMR/DMR (10)
						IFN (1)[a]	CCyR (1)

Abbreviations: CCyR, complete cytogenetic response; CML, chronic myeloid leukemia; DMR, deep molecular response; IMA, imatinib; IFN, interferon-α; TFR, treatment-free remission; MMR, major molecular response; PCyR, partial cytogenetic response; TMA-UD, undetectable level of BCR::ABL1 mRNA by the TMA method (equivalent to MMR or deeper response); TKI, tyrosine kinase inhibitor.

[a]Treatment was changed because loss of MMR was observed.

TABLE 4 Efficacy of IFN-α for treatment of CML during pregnancy.

Situation of pregnancy	Patient number (*n*)	Pregnancy number (*n*)	Treatment at pregnancy (*n*)	Response at pregnancy (*n*)	Initial treatment during pregnancy	Treatment at delivery (*n*)	Response at delivery (*n*)
CML onset during pregnancy	3	3	No treatment (3)	No response (3)	IFN	IFN (2)	No response (2)
						IMA (1)	CHR (1)
Unplanned pregnancy	3	3	DAS (3)	CCyR (1)	IFN	IFN (1)	MMR (1)
				DMR (1)	IFN	IMA (1)[a]	DMR (1)
				DMR (1)	TFR	IFN (1)[a]	PCyR (1)
Planned pregnancy	9	15	IFN (12)	MMR/DMR (12)	IFN	IFN (12)	MMR/DMR (12)
			TFR (3)	MMR/DMR (3)	TFR	IFN (3)[a]	CCyR (1)
							MMR/DMR (2)

Abbreviations: CCyR, complete cytogenetic response; CHR, complete hematologic response; CML, chronic myeloid leukemia; DAS, dasatinib; DMR, deep molecular response; IFN, interferon-α; IMA, imatinib; MMR, major molecular response; PCyR, partial cytogenetic response; TFR, treatment-free remission.

[a]Treatment was changed because loss of MMR was observed.

TABLE 5 Characteristics according to the conception method in planned pregnancy.

Planned pregnancy	Patients, *n*	Pregnancy, *n*	Duration from CML diagnosis to pregnancy, median (range), days	Duration from TKI discontinuation to first pregnancy, median (range), days	Age at onset, median (range), years	Age at pregnancy, median (range), years	Live birth, *n*
Natural conception	17	23	2776 (631–5878)	226 (20–1655)	25 (15–34)	34 (27–40)	16
ART (egg preservation)	5	7	3936 (1585–5017)	131 (18–308)	24 (23–30)	35 (34–42)	5
ART (no egg preservation)	5	6	4547 (3165–6073)	390 (104–2136)	27 (20–28)	40 (30–42)	5

Abbreviations: ART, assisted reproductive technology; CML, chronic myeloid leukemia; TKI, tyrosine kinase inhibitor.

The rate of spontaneous abortion in patients aged 35 years and older tended to be higher than in patients aged 25 to 34 years (20% vs. 7.1%, Fisher's exact test, $p = .234$). In unplanned pregnancies, spontaneous abortions occurred in patients aged 24 years or younger who were treated with a TKI. There were 16 unplanned pregnancies except for pregnancies with elective abortions during treatment with a TKI, and spontaneous abortions occurred only in patients for whom dasatinib was administered at the time of a positive pregnancy test (Table 7).

Other than spontaneous abortion, hypertension and diabetes occurred in four pregnancies each (Table 6). Complications of newborns included one stillbirth and one hypothyroidism in patients with planned pregnancy that were not exposed to TKIs throughout pregnancy. Although treatment with imatinib was introduced in four patients in the second trimester or later, there were no complications in both patients and newborns.

Disease status after pregnancy

Information on the disease status after pregnancy was available for 47 patients. Forty-four patients remained in the first CP: Two patients were in CCyR during either treatment discontinuation or treatment with IFN-α, and 42 patients were in MMR or DMR at the time of the last follow-up. One patient in CCyR continued treatment discontinuation after delivery for breast-feeding, and another patient lost MMR during treatment with IFN-α for wishing subsequent pregnancy. There were three patients who progressed to the advanced phase of CML after pregnancy (Table S3): one patient progressed to the accelerated phase (AP), and two patients progressed to the blastic phase (BP) after pregnancy. One patient who progressed to BP did not achieve CCyR after 3 years of treatment with imatinib. In the other two patients who progressed to AP or BP, treatment with a TKI was interrupted due to nonadherence.

TABLE 6 Pregnancy age and complications during pregnancy.

Age at pregnancy, years	20–24	25–29	30–34	35–39	40+
Delivery, *n*	3	10	16	15	5
Spontaneous abortion, *n*	3	1	1	3	2
Maternal complications, *n*					
Diabetes	0	2	2	0	0
Hypertension	0	0	1	1	2
Others[a]	0	1	2	2	0
Newborn complications, *n*	0	0	1	0	1

Note: Three spontaneous abortions were observed in patients of 20–24 years of age who were treated with dasatinib. Other abortions were observed in patients of 25 years of age or older who were treated with IFN-α or were in TFR.

Abbreviations: IFN-α, interferon-alpha; TFR, treatment-free remission.

[a]Maternal complications other than diabetes and hypertension were as follows: threatened premature labor ($n = 1$), preterm premature rupture of membrane ($n = 1$), depression ($n = 1$), hypothyroidism ($n = 1$), and exacerbation of Behçet's disease ($n = 1$).

TABLE 7 TKI use at the time of positive pregnancy test and spontaneous abortion.

Outcome of pregnancy	Imatinib	Dasatinib	Nilotinib	Bosutinib
Delivery, *n*	6	5	1	1
Spontaneous abortion, *n*	0	3	0	0

Abbreviation: TKI, tyrosine kinase inhibitor.

DISCUSSION

This study presents outcomes of 70 pregnancies in CML patients during the TKI era. Several reports about CML and pregnancy have been published but most were based on treatment with imatinib.[7,15–21] Our study included for 70 pregnancies in 49 patients including 38 patients who were treated with at least one 2G-TKI, reflecting the current treatment landscape for CML.[22,23]

There were 48 live births from 70 pregnancies including seven live births of nine pregnancies in patients with CML onset during pregnancy, 14 live births from 25 unplanned pregnancies, and 27 live births from 36 planned pregnancies. Elective abortion was chosen for eight of 25 unplanned pregnancies including four of six pregnancies in patients without MMR and four of 19 pregnancies in patients with MMR or DMR. In addition to exposure of the embryo to a TKI, disease status at the time of a positive pregnancy test could be a key reason for continuing unplanned pregnancies.

In cases with CML onset during pregnancy, seven of nine pregnancies resulted in live births. Among them, three patients were observed until delivery and four patients were treated with either IFN-α or imatinib. IFN-α treatment takes time to show an effect, making its efficacy limited for patients with CML diagnosed during pregnancy.[24–26] Recent reports suggest that the necessity of treatment for patients with CML diagnosed during pregnancy should be carefully evaluated based on clinical factors such as gestational age, the degree of increase in white blood cell and platelet counts, and the risk of disease progression.[27,28]

TFR was attempted in 26 pregnancies with live births either before conception for planned pregnancies or after a positive pregnancy test in unplanned pregnancies. Notably, the duration from diagnosis to pregnancy was significantly shorter in patients primarily treated with a 2G-TKI than in patients primarily treated with imatinib. Patients who fulfilled the requirements for TKI discontinuation by ELN 2020 recommendations achieved a TFR rate of 77% in pregnancy. Thus, TFR attempts according to the ELN 2020 recommendations could be applicable to pregnancy. Although the duration from CML diagnosis to first conception in planned pregnancy was twice as long as that in unplanned pregnancy, using 2G-TKI as first-line treatment may shorten the duration for planned pregnancy.

Furthermore, ART is a possible option not only to resolve the problem of less fertility associated with aging but also to minimize the time needed to achieve pregnancy. Consequently, egg or embryo preservation should be proactively considered for patients of childbearing potential at the time of CML diagnosis if they have a childbearing desire. It may also be possible to attempt TFR after conception throughout the pregnancy (stop TKI at the first pregnancy test) in planned pregnancy. The feasibility of this approach warrants consideration.

Additionally, three patients who lost MMR during pregnancy remained treatment-free until delivery, at which point their disease status was CCyR. As previously suggested,[29] a treatment-free pregnancy could be a viable option for patients in CCyR after losing MMR during pregnancy.

Administration of IFN-α is an alternative treatment for TKI-free pregnancy.[24,25] Our results suggest that IFN-α may be effective for patients who have achieved CCyR or deeper response. Additionally, IFN-α could be serve as a salvage therapy for patients in whom TFR has failed.

Disease progression is a major concern of TKI therapy discontinuation in patients with pregnancy. Three patients progressed to AP or BP after pregnancy. Insufficient response to treatment with a TKI or nonadherence seemed to be the main reason for disease progression. Other patients generally showed optimal responses after delivery. Thus, patients with uncontrolled disease status should refrain from becoming pregnant.

Spontaneous abortions occurred in 10 pregnancies, including seven planned pregnancies and three unplanned pregnancies. The spontaneous abortions in planned pregnancies seemed to be associated with maternal age.[30,31] Notably, the median duration from CML onset to pregnancy in patients with planned pregnancy was 10.6 years, which would have led to an advanced maternal age. On the other hand, spontaneous abortions in unplanned pregnancies occurred in patients who were treated with dasatinib at the time of a positive pregnancy test. Dasatinib might have an effect on an early embryo possibly due to its off-target effect.[32] Complications of pregnancy other than spontaneous abortion, including diabetes and hypertension, were observed, but their frequencies did not seem to be higher than those in the general population.[33,34]

Our study has several limitations. First, the number of patients was small, especially for statistical analysis. Second, less than 10% of the female patients aged 45 years or younger became pregnant, suggesting that our analysis reflects the situation of pregnancy in selected patients in female patients with childbearing potential. Third, there were no data for patients who were treated with ponatinib or asciminib before becoming pregnant.

In conclusion, TFR could be a reasonable option for patients who achieve sustained DMR and desire childbearing. Treatment with IFN-α might be another option for pregnant patients who have achieved CCyR or a deeper response. Planned pregnancy by TFR or treatment with IFN-α could be a safe and feasible way for patients with CML. However, a substantial duration of treatment with TKI before conception may be needed. Planning and evaluation for pregnancy and childbirth, including cryopreservation of eggs or embryos, should be considered at the time of CML onset for female patients with childbearing potential.

AUTHOR CONTRIBUTIONS

Takeshi Kondo: Conceptualization, methodology, data curation, investigation, validation, funding acquisition, project administration, writing–original draft, writing–review and editing, and resources. **Eri Matsuki**: Conceptualization, writing–review and editing, resources, and validation. **Tomoiku Takaku**: Conceptualization, writing–review and editing, and validation. **Naoki Watanabe**: Conceptualization, resources, and writing–review and editing. **Chikashi Yoshida**: Conceptualization, methodology, resources, writing–review and editing, and validation. **Masaya Okada**: Conceptualization, resources, writing–review and editing, and validation. **Kazunori Murai**: Conceptualization, writing–review and editing, resources, and validation. **Takashi Kodama**: Writing–original draft, writing–review and editing, and validation. **Naoto Takahashi**: Supervision, resources, and writing–review and editing. **Shinya Kimura**: Supervision, resources, and writing–review and editing. **Itaru Matsumura**: Supervision, resources, and writing–review and editing.

ACKNOWLEDGMENTS

We express our warmest gratitude and sincere respect to the patients and their children. We also express special thanks to the colleagues who treated and consulted patients for treating CML and supporting their child-bearing desire. We thank Ms. Keiko Tanaka and Mr. Hidenori Komatsu of the Aiiku Hospital for secretarial assistance. The clinical study was approved by the institutional review boards of Aiiku Hospital and each hospital participating in the second survey and performed in accordance with the Declaration of Helsinki. The opt-out approach was used to obtain consent from the patients in the research. This work was supported by a grant from Pfizer Health Research Foundation to Takeshi Kondo.

CONFLICT OF INTEREST STATEMENT

Takeshi Kondo has received honoraria from Astellas Pharma, AbbVie, Amgen, Kyowa Kirin, Nippon Shinyaku, Ono Pharmaceutical, Otsuka Pharmaceutical, Novartis, Pfizer, and Bristol-Myers Squibb; and consultancy fees from Otsuka Pharmaceutical. Eri Matsuki received honoraria from Novartis. Tomoiku Takaku received honoraria from Novartis, Otsuka Pharmaceutical, and Pfizer; and research funding from Bristol-Myers Squibb, Otsuka Pharmaceutical, and ThinkCyte. Chikashi Yoshida received honoraria from AbbVie, Bristol-Myers Squibb, Chugai Pharmaceutical, Janssen, Novartis, Nippon Shinyaku, Ono Pharmaceutical, Otsuka Pharmaceutical, Pfizer, and Takeda Pharmaceutical; and research funding from Bristol-Myers Squibb. Naoto Takahashi received honoraria from Novartis and Otsuka Pharmaceutical; and research funding from Novartis, Asahi Kasei Pharma, and Otsuka Pharmaceutical. Shinya Kimura has received honoraria from Otsuka Pharmaceutical, Novartis, Pfizer, and Bristol-Myers

Squibb; and research funding from Otsuka Pharmaceutical, Novartis, Bristol-Myers Squibb, Pfizer, and Ohara Pharmaceutical. Itaru Matsumura received speakers bureau fees from Takeda Pharmaceutical, Daiichi Sankyo, Pfizer, SymBio Pharmaceuticals, Astellas Pharma, Ono Pharmaceutical, Novartis, Chugai Pharmaceutical, AbbVie, Janssen, Bristol-Myers Squibb (Celgene), Amgen, AstraZeneca, and Otsuka Pharmaceutical; research funding from Chugai Pharmaceutical, Kyowa Kirin, Sumitomo Pharma, Takeda Pharmaceutical, Astellas Pharma, Ono Pharmaceutical, Sanofi, Mitsubishi Tanabe Pharma, Nippon Shinyaku, Eisai, MSD, Asahi Kasei Pharma, AbbVie, Janssen, Taiho Pharmaceutical, Shionogi, Teijin Pharma, Nippon Boehringer Ingelheim, Nippon Pharmaceutical, Daiichi Sankyo, Nippon Kayaku, CSL Behring, Mundipharma, Ayumi Pharmaceutical, Eli Lily, Actelion Pharmaceuticals, and Amgen; and consultancy fees from Otsuka Pharmaceutical. The other authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are not openly available to preserve the privacy of the patients. They can be obtained from the corresponding author on reasonable request and permission.

ORCID

Takeshi Kondo https://orcid.org/0000-0001-7455-5824

REFERENCES

1. Bower H, Björkholm M, Dickman PW, Höglund M, Lambert PC, Andersson TM. Life expectancy of patients with chronic myeloid leukemia approaches the life expectancy of the general population. *J Clin Oncol*. 2016;34(24):2851-2857. doi:10.1200/jco.2015.66.2866
2. Maas CCHM, van Klaveren D, Ector GICG, et al. The evolution of the loss of life expectancy in patients with chronic myeloid leukaemia: a population-based study in the Netherlands, 1989-2018. *Br J Haematol*. 2022;196(5):1219-1224. doi:10.1111/bjh.17989
3. Mahon FX, Réa D, Guilhot J, et al. Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. *Lancet Oncol*. 2010;11:1029-1035. doi:10.1016/s1470-2045(10)70233-3
4. Hochhaus A, Baccarani M, Silver RT, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia*. 2020;34(4):966-984. doi:10.1038/s41375-020-0776-2
5. Hoffmann VS, Baccarani M, Hasford J, et al. The EUTOS population-based registry: incidence and clinical characteristics of 2904 CML patients in 20 European Countries. *Leukemia*. 2015;29(6):1336-1343. doi:10.1038/leu.2015.73
6. Apperley J. CML in pregnancy and childhood. *Best Pract Res Clin Haematol*. 2009;22(3):455-474. doi:10.1016/j.beha.2009.09.008
7. Pye SM, Cortes J, Ault P, et al. The effects of imatinib on pregnancy outcome. *Blood*. 2008;111(12):5505-5508. doi:10.1182/blood-2007-10-114900
8. Russell MA, Carpenter MW, Akhtar MS, Lagattuta TF, Egorin MJ. Imatinib mesylate and metabolite concentrations in maternal blood, umbilical cord blood, placenta and breast milk. *J Perinatol*. 2007;27(4):241-243. doi:10.1038/sj.jp.7211665
9. Cortes JE, Abruzzese E, Chelysheva E, Guha M, Wallis N, Apperley JF. The impact of dasatinib on pregnancy outcomes. *Am J Hematol*. 2015;90(12):1111-1115. doi:10.1002/ajh.24186
10. Berveiller P, Andreoli A, Mir O, et al. A dramatic fetal outcome following transplacental transfer of dasatinib. *Anti Cancer Drugs*. 2012;23:754-757. doi:10.1097/cad.0b013e328352a8fe
11. Baccarani M, Cortes J, Pane F, et al. Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *J Clin Oncol*. 2009;27(35):6041-6051. doi:10.1200/jco.2009.25.0779
12. Langabeer SE, Gale RE, Harvey RC, Cook RW, Mackinnon S, Linch DC. Transcription-mediated amplification and hybridisation protection assay to determine BCR-ABL transcript levels in patients with chronic myeloid leukaemia. *Leukemia*. 2002;16(3):393-399. doi:10.1038/sj.leu.2402392
13. Nakamae H, Yoshida C, Miyata Y, et al. A new diagnostic kit, ODK-1201, for the quantitation of low major BCR-ABL mRNA level in chronic myeloid leukemia: correlation of quantitation with major BCR-ABL mRNA kits. *Int J Hematol*. 2015;102(3):304-311. doi:10.1007/s12185-015-1826-9
14. Rousselot P, Charbonnier A, Cony-Makhoul P, et al. Loss of major molecular response as a trigger for restarting tyrosine kinase inhibitor therapy in patients with chronic-phase chronic myelogenous leukemia who have stopped imatinib after durable undetectable disease. *J Clin Oncol*. 2014;32(5):424-430. doi:10.1200/jco.2012.48.5797
15. Ault P, Kantarjian H, O'Brien S, et al. Pregnancy among patients with chronic myeloid leukemia treated with imatinib. *J Clin Oncol*. 2006;24:1204-1208.
16. Zhou L, You JH, Wu W, Li JM, Shen ZX, Wang AH. Pregnancies in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitor. *Leuk Res*. 2013;37(10):1216-1221. doi:10.1016/j.leukres.2013.07.020
17. Alizadeh H, Jaafar H, Rajnics P, Khan MI, Kajtár B. Outcome of pregnancy in chronic myeloid leukaemia patients treated with tyrosine kinase inhibitors: short report from a single centre. *Leuk Res*. 2015;39(1):47-51. doi:10.1016/j.leukres.2014.10.002
18. Mukhopadhyay A, Dasgupta S, Kanti Ray U, Gharami F, Bose CK, Mukhopadhyay S. Pregnancy outcome in chronic myeloid leukemia patients on imatinib therapy. *Ir J Med Sci*. 2015;184(1):183-188. doi:10.1007/s11845-014-1084-5
19. Lasica M, Willcox A, Burbury K, et al. The effect of tyrosine kinase inhibitor interruption and interferon use on pregnancy outcomes and long-term disease control in chronic myeloid leukemia. *Leuk Lymphoma*. 2019;60(7):1796-1802. doi:10.1080/10428194.2018.1551533
20. Madabhavi I, Sarkar M, Modi M, Kadakol N. Pregnancy outcomes in chronic myeloid leukemia: a single center experience. *J Glob Oncol*. 2019;5:1-11. doi:10.1200/jgo.18.00211
21. Dou X, Qin Y, Huang X, Jiang Q. Planned pregnancy in female patients with chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy. *Oncologist*. 2019;24(11):e1141-e1147.-7. doi:10.1634/theoncologist.2019-0109
22. Assi R, Kantarjian H, Keating M, et al. Management of chronic myeloid leukemia during pregnancy among patients treated with a tyrosine kinase inhibitor: a single-Center experience. *Leuk Lymphoma*. 2021;62(4):909-917. doi:10.1080/10428194.2020.1849672
23. Abruzzese E, Aureli S, Bondanini F, et al. Chronic myeloid leukemia and pregnancy: when dreams meet reality. state of the art, management and outcome of 41 cases, nilotinib placental transfer. *J Clin Med*. 2022;11(7):1801. doi:10.3390/jcm11071801
24. Yazdani Brojeni P, Matok I, Garcia Bournissen F, Koren G. A systematic review of the fetal safety of interferon alpha. *Reprod Toxicol*. 2012;33(3):265-268. doi:10.1016/j.reprotox.2011.11.003
25. Balsat M, Etienne M, Elhamri M, Hayette S, Salles G, Thomas X. Successful pregnancies in patients with BCR-ABL-positive leukemias

treated with interferon-alpha therapy during the tyrosine kinase inhibitors era. *Eur J Haematol*. 2018;101(6):774-780. doi:10.1111/ejh.13167
26. Talpaz M, Kantarjian HM, McCredie KB, Keating MJ, Trujillo J, Gutterman J. Clinical investigation of human alpha interferon in chronic myelogenous leukemia. *Blood*. 1987;69(5):1280-1288. doi:10.1182/blood.v69.5.1280.1280
27. Chelysheva E, Apperley J, Turkina A, et al. Chronic myeloid leukemia diagnosed in pregnancy: management and outcome of 87 patients reported to the European LeukemiaNet international registry. *Leukemia*. 2024;38(4):788-795. doi:10.1038/s41375-024-02183-0
28. Robertson HF, Milojkovic D, Butt N, et al. Expectations and outcomes of varying treatment strategies for CML presenting during pregnancy. *Br J Haematol*. Published online May 2, 2024. doi:10.1111/bjh.19491
29. Abruzzese E, Mauro M, Apperley J, Chelysheva E. Tyrosine kinase inhibitors and pregnancy in chronic myeloid leukemia: opinion, evidence, and recommendations. *Ther Adv Hematol*. 2020;11:2040620720966120. doi:10.1177/2040620720966120
30. Frederiksen LE, Ernst A, Brix N, et al. Risk of adverse pregnancy outcomes at advanced maternal age. *Obstet Gynecol*. 2018;131(3):457-463. doi:10.1097/aog.0000000000002504
31. Nybo Andersen AM, Wohlfahrt J, Christens P, Olsen J, Melbye M. Maternal age and fetal loss: population based register linkage study. *BMJ*. 2000;320(7251):1708-1712. doi:10.1136/bmj.320.7251.1708
32. Al-Asmakh M, Bawadi H, Hamdan M, et al. Dasatinib and PD-L1 inhibitors provoke toxicity and inhibit angiogenesis in the embryo. *Biomed Pharmacother*. 2021;134:111134. doi:10.1016/j.biopha.2020.111134
33. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. *Obstet Gynecol*. 2013;122:1122-1131.
34. ACOG Practice Bulletin. No. 190: gestational diabetes mellitus. *Obstet Gynecol*. 2018;131:e49-e64.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kondo T, Matsuki E, Takaku T, et al. Outcomes of pregnancy in patients with chronic myeloid leukemia in the era of tyrosine kinase inhibitors. *Cancer*. 2025; e35611. doi:10.1002/cncr.35611

Original Article

AO SPINE

Global Spine Journal
2025, Vol. 15(4) 2274–2285

DOI: 10.1177/21925682241297587
journals.sagepub.com/home/gsj
Sage

Influence of Diabetes Mellitus on Neurological Recovery in Older Patients With Cervical Spinal Cord Injury Without Bone Injury: A Retrospective Multicenter Study

Kazuki Takeda, MD[1,2], Kota Watanabe, MD[1], Satoshi Nori, MD[1], Junichi Yamane, MD[1,3], Hitoshi Kono, MD[1,4], Noriaki Yokogawa, MD[5], Takeshi Sasagawa, MD[5,6], Hiroaki Nakashima, MD[7], Naoki Segi, MD[7], Toru Funayama, MD[8], Fumihiko Eto, MD[9], Takeo Furuya, MD[10], Atsushi Yunde, MD[10], Hideaki Nakajima, MD[11], Tomohiro Yamada, MD[12,13], Tomohiko Hasegawa, MD[12], Yoshinori Terashima, MD[14,15], Ryosuke Hirota, MD[14], Hidenori Suzuki, MD[16], Yasuaki Imajo, MD[17], Shota Ikegami, MD[18], Hitoshi Tonomura, MD[19], Munehiro Sakata, MD[19,20], Ko Hashimoto, MD[21], Kenichi Kawaguchi, MD[22], Nobuyuki Suzuki, MD[23], Hiroshi Uei, MD[24,25], Kazuo Nakanishi, MD[26], Hidetomi Terai, MD[27], Gen Inoue, MD[28], Katsuhito Kiyasu, MD[29], Yoichi Iizuka, MD[30], Koji Akeda, MD[31], Haruki Funao, MD[32,33,34], Yasushi Oshima, MD[35], Takashi Kaito, MD[36], Toshitaka Yoshii, MD[37], Masayuki Ishihara, MD[38], Seiji Okada, MD[36], Shiro Imagama, MD[7], and Satoshi Kato, MD[5]

[1] Department of Orthopaedic Surgery, Keio University School of Medicine, Tokyo, Japan
[2] Department of Orthopaedic Surgery, Japanese Red Cross Shizuoka Hospital, Shizuoka, Japan
[3] Department of Orthopaedic Surgery, National Hospital Organization Murayama Medical Center, Tokyo, Japan
[4] Department of Orthopaedic Surgery, Keiyu Orthopedic Hospital, Tatebayashi-shi, Japan
[5] Department of Orthopaedic Surgery, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Japan
[6] Department of Orthopaedic Surgery, Toyama Prefectural Central Hospital, Toyama, Japan
[7] Department of Orthopedic Surgery, Nagoya University, Graduate School of Medicine, Nagoya, Japan
[8] Department of Orthopaedic Surgery, NHO Mito Medical Center, Ibaraki, Japan
[9] Department of Orthopaedic Surgery, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan
[10] Department of Orthopaedic Surgery, Graduate school of Medicine, Chiba University, Chiba, Japan
[11] Department of Orthopaedics and Rehabilitation Medicine, Faculty of Medical Sciences University of Fukui, Fukui, Japan
[12] Department of Orthopaedic Surgery, Hamamatsu University School of Medicine, Hamamatsu City, Japan
[13] Department of Orthopaedic Surgery, Nagoya Kyoritsu Hospital, Nagoya-shi, Japan
[14] Department of Orthopaedic Surgery, Sapporo Medical University, Sapporo, Japan
[15] Department of Orthopaedic Surgery, Matsuda Orthopedic Memorial Hospital, Sapporo, Japan
[16] Department of Orthopaedic Surgery, Yamaguchi University Graduate School of Medicine, Ube City, Japan
[17] Department of Orthopaedic Surgery, Tokuyama Central Hospital, Yamaguchi, Japan
[18] Department of Orthopaedic Surgery, Shinshu University School of Medicine, Matsumoto, Japan
[19] Department of Orthopaedics, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan
[20] Department of Orthopaedics, Saiseikai Shiga Hospital, Shiga, Japan
[21] Department of Orthopaedic Surgery, Tohoku University Graduate School of Medicine, Sendai, Japan
[22] Department of Orthopaedic Surgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
[23] Department of Orthopaedic Surgery, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan

Corresponding Author:
Kota Watanabe, Department of Orthopaedic Surgery, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan.
Email: kw197251@keio.jp

Abstract

Study Design: Retrospective study.

Objectives: To investigate the impact of diabetes mellitus on neurological recovery and determine the relationship between moderate-severe diabetes and neurological recovery in patients with cervical spinal cord injury (CSCI) without bone injury.

Methods: A retrospective study was conducted on 389 consecutive patients aged ≥65 years with CSCI without bone injury across 33 medical institutes. The patients were divided into a nondiabetic group (n = 270) and a diabetic group (n = 119). Neurological outcomes were compared between the two groups through propensity score matching. The impact of moderate-severe diabetes (defined as hemoglobin A1c ≥ 7.0% or requiring insulin treatment) on neurological recovery was evaluated through multiple linear regression analysis.

Results: Propensity score matching revealed no significant differences between the diabetic and nondiabetic groups in terms of American Spinal Injury Association (ASIA) impairment scale grade and mean total ASIA motor scores (AMS) at 6 months post-injury. Multiple linear regression analysis indicated that age on admission (B = −0.34; 95% confidence interval [CI], −0.59 to −0.08; $P = 0.01$), dementia (B = −16.50; 95% CI, −24.99 to −8.01; $P < 0.01$), and baseline total AMS (B = −0.62; 95% CI, −0.72 to −0.51; $P < 0.01$) were negative predictors of neurological recovery at 6 months post-injury. The presence of moderate-severe diabetes did not influence neurological recovery at 6 months post-injury.

Conclusions: Diabetic patients with CSCI without bone injury achieved improvements in neurological function comparable to those of nondiabetic patients. Moderate-severe diabetes did not affect neurological recovery in patients with CSCI without bone injury.

Keywords

cervical spinal cord injury without bone injury, diabetes mellitus, moderate-severe diabetes mellitus, blood glucose level, American spinal injury association impairment scale grade, American spinal injury association motor scores, older patients, propensity score matching, multiple linear regression analysis, prognostic factors, dementia

Introduction

Cervical spinal cord injury (CSCI) is one of the most devastating spinal injuries, and can cause severe permanent neurological deficits. A nationwide survey in Japan demonstrated that the overall rate of CSCI among cases of traumatic spinal cord injury (SCI) was 88.1%.[1] Among the CSCI cases, 70.7% did not suffer a bone injury resulting from minimal trauma, the older population. The incidence of CSCI without bone injury is anticipated to increase with the growing aging population.

Although several prognostic factors for neurological recovery in patients with CSCI without bone injury have been identified, most studies have focused on imaging features or surgical interventions. Preexisting conditions such as cervical spondylosis, ossification of the posterior longitudinal ligament (OPLL), and signal intensity (SI) changes in the spinal cord on magnetic resonance imaging (MRI) are relevant prognostic factors for neurological recovery in patients with CSCI without bone injury.[2-4] Surgical intervention is also reported as a prognostic factor for neurological recovery in patients with CSCI without bone injury.[5,6] Recently, Nakajima et al revealed that factors such as body mass index (BMI), OPLL, SI changes on MRI, American Spinal Injury Association (ASIA) impairment scale (AIS) on admission, comorbidity of dementia/delirium, and post-injury pneumonia were independent prognostic factors for the recovery of walking ability in patients with CSCI without major bone injury.[7]

[24] Department of Orthopaedic Surgery, Nihon University Hospital, Tokyo, Japan
[25] Department of Orthopaedic Surgery, Nihon University School of Medicine, Tokyo, Japan
[26] Department of Orthopedics, Traumatology and Spine Surgery, Kawasaki Medical School, Okayama, Japan
[27] Department of Orthopaedic Surgery, Osaka Metropolitan University Graduate School of Medicine, Osaka, Japan
[28] Department of Orthopaedic Surgery, Kitasato University School of Medicine, Sagamihara, Japan
[29] Department of Orthopaedic Surgery, Kochi Medical School, Kochi University, Nankoku, Japan
[30] Department of Orthopaedic Surgery, Gunma University, Graduate School of Medicine, Maebashi, Japan
[31] Department of Orthopaedic Surgery, Mie University Graduate School of Medicine, Tsu City, Japan
[32] Department of Orthopaedic Surgery, School of Medicine, International University of Health and Welfare, Chiba, Japan
[33] Department of Orthopaedic Surgery, International University of Health and Welfare Narita Hospital, Chiba, Japan
[34] Department of Orthopaedic Surgery and Spine and Spinal Cord Center, International University of Health and Welfare Mita Hospital, Tokyo, Japan
[35] Department of Orthopaedic Surgery, The University of TokyoHospital, Tokyo, Japan
[36] Department of Orthopaedic Surgery, Osaka UniversityGraduate School of Medicine, Osaka, Japan
[37] Department of Orthopaedic Surgery, Tokyo Medical and Dental University, Tokyo, Japan
[38] Department of Orthopaedic Surgery, Kansai Medical University Hospital, Osaka, Japan

Diabetes mellitus is one of the most common comorbidities which is well known to negatively affect perioperative complications in spine surgery, including decreased spinal fusion rates and increased surgical site infections.[8-10] Additionally, numerous studies have suggested that diabetes is a risk factor for poor surgical outcomes in patients with lumbar disc degenerative disease.[11,12] Regarding cervical spondylotic myelopathy (CSM), some studies have reported preferable surgical outcomes in diabetic,[13-15] while others diabetes contributes to poor surgical outcomes.[16,17] However, the influence of diabetes on CSCI without bone injury also remains unclear because of the few studies.[7] Several animal studies supported that diabetes negatively affected neurological outcomes in patients with CSCI without bone injury.[18,19] Furthermore, no previous studies have evaluated the relationship between diabetes severity and neurological recovery in patients with CSCI without bone injury.

This multicenter large-cohort study aimed to investigate the impact of diabetes on post-injury complications and neurological recovery, and the relationship between moderate-severe diabetes and neurological recovery in patients with CSCI without bone injury.

Materials and Methods

Participants

The multicenter study was conducted by the Japan Association of Spine Surgeons with Ambition and included 1512 consecutive patients aged ≥65 years with cervical spine/SCIs from 33 medical institutes between 2010 and 2020. The original dataset used in the study has also been utilized in other studies.[7,20,21] Of the 1512 patients, 614 (40.6%) were diagnosed with CSCI without bone injury. This condition was defined as a CSCI with no evidence of spinal fractures or dislocations on radiography or computed tomography (CT).[22] Any patients with missing values in diabetes, baseline AIS grade, baseline ASIA motor score (AMS), AIS grade at 6 months after injury, or AMS at 6 months after injury were excluded from the study. Patients with baseline AIS grade A were also excluded because the probability of the AIS grade converting to grades C–E is very low.[23] A total of 389 patients followed up for at least 6 months were included in the present study (Figure 1). The 6-month time period was chosen based on previous clinical trials that demonstrated that neurological recovery after traumatic SCI mainly occurs within the first 6-9 months.[24,25] Patients were divided into two groups: nondiabetic (n = 270) and diabetic group (n = 119). The diabetic group consisted of patients with hemoglobin A1c (HbA1c) ≥6.5%, those undergoing diabetes treatment with oral agents or insulin or both, or those previously diagnosed with diabetes by diabetologists by the clinical practice guidelines for diabetes.[5] National Glycohemoglobin Standardization Program method was used for measuring HbA1c. Patients with an HbA1c level ≥7.0% or those requiring insulin treatment were categorized into the moderate-severe diabetic group (n = 55).[26,27]

Data Collection

Demographic variables, including age at injury, sex, BMI, and medical comorbidities, were collected. Medical comorbidities included diabetes, hypertension, cardiovascular disease, cerebrovascular disease, rheumatoid arthritis, osteoporosis, respiratory disease, renal disease, Parkinson's disease, dementia, and a history of surgery for musculoskeletal disorders. For diabetic patients, the HbA1c levels at admission and medication details, including insulin, were also documented. Cervical OPLL and diffuse idiopathic skeletal hyperostosis (DISH) were assessed by spinal radiography and CT. SI changes in the cervical spinal cord were evaluated using T2-weighted sagittal and axial MRI at the time of injury. Senior spinal surgeons and physical therapists at each center evaluated the neurological status on admission, at discharge, and 6 months post-injury using the AIS and AMS. The indications for surgery and steroid therapy were determined by the attending spinal surgeons at each institute. Complications during hospitalization, such as motor or sensory neurological deterioration, cerebral infarction, delirium, dysphagia, respiratory failure, pulmonary embolism, pneumonia, and renal infection, were also documented.

Statistical Analysis

The level of statistical significance was set at $P < 0.05$. The R Statistical Package version 2.6.0 (R Foundation for Statistical Computing, Vienna, Austria) was used for all statistical analyses. Continuous variables were analyzed using the unpaired *t*-test, Welch's *t*-test, or Wilcoxon rank-sum test, as appropriate. Categorical variables were analyzed using the χ^2 test or Fisher's exact test. Propensity score matching was conducted to compare neurological outcomes between the nondiabetic and diabetic groups. A multivariate logistic regression model was used to calculate the propensity scores. The moderator variables were age, sex, BMI, cervical OPLL, cervical DISH, SI changes on MRI, medical comorbidities, baseline AIS grade, baseline AMS score, steroid therapy, and surgical intervention for CSCI. To adjust for baseline characteristics and comorbidities, 1-to-1 matching with fixed caliper widths (0.15) was performed without replacement. Each nondiabetic case was matched with a corresponding case in the diabetic group with the same propensity score. Standardized differences were used to measure covariate balance, with a standardized difference <10% indicating a negligible difference between both groups. Additionally, patients with an HbA1c level ≥7.0% or those requiring insulin treatment were classified as the moderate-severe diabetes group. Clinically relevant variables (age, sex, BMI, cervical DISH, SI changes on MRI, baseline AIS grade, baseline AMS, and surgical intervention), along with variables with significance level <0.05, as determined by univariate analysis (blood glucose level on admission, diabetes, rheumatoid arthritis, dementia, post-injury complication of dysphagia, cervical OPLL, and

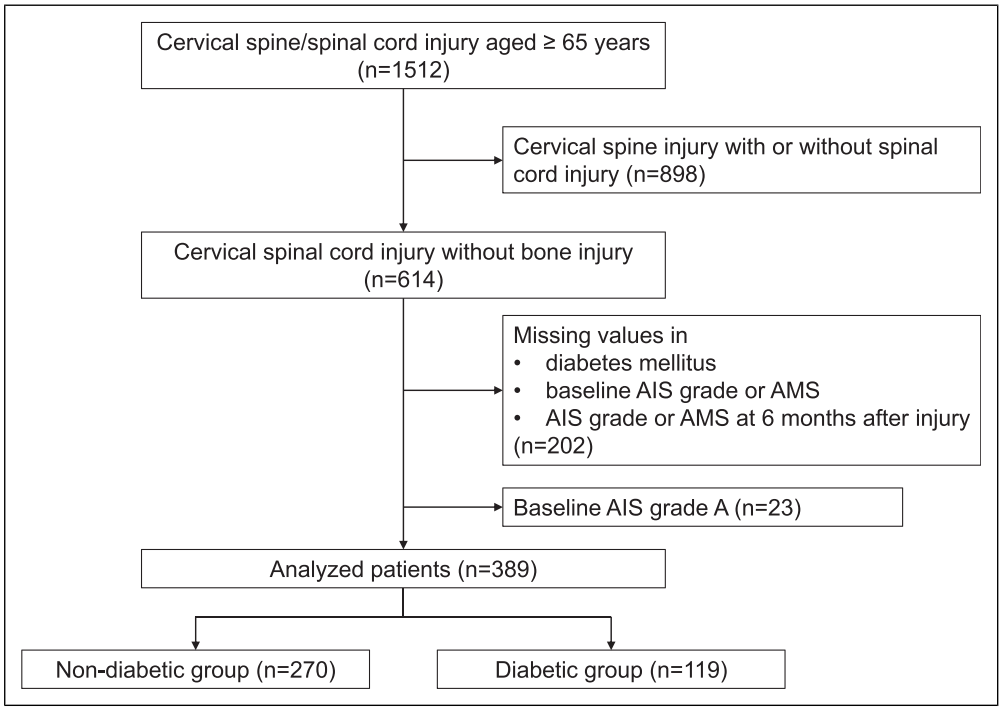


Figure 1. Flowchart of the multicenter cohort. ASIA, American Spinal Injury Association; AIS, American Spinal Injury Association impairment scale.

steroid therapy for CSCI), were included in the multiple linear regression analysis with listwise deletion of missing data. This was performed to assess the influence of moderate-severe diabetes on the degree of improvement in total AMS at 6 months post-injury.

Results

Comparison of Baseline Characteristics and Medical Comorbidities Between the Nondiabetic and Diabetic Groups

A total of 389 patients were included in the present study (male, n = 279, 71.7%; female, n = 110, 28.3%; mean age at the time of injury, 74.7 ± 6.4 years). The blood glucose level on admission in the diabetic group was significantly higher than that in the nondiabetic group (Table 1). The prevalence of dementia was significantly higher in the diabetic group than in the nondiabetic group. The mean baseline total AMS and lower extremity AMS scores were comparable between the two groups, but both were slightly lower in the diabetic group.

Comparison of Post-Injury Complications and Neurological Outcomes Between the Nondiabetic and Diabetic Groups

The mean total AMS at discharge in the diabetic group was significantly lower than that of the nondiabetic group (Table 2). Post-injury complications during hospitalization did not differ significantly between the groups. The AIS grade at 6 months post-injury was significantly poorer in the diabetic group than that in the nondiabetic group. The mean total, upper extremity, and lower extremity AMS in the diabetic group were significantly lower than those in the nondiabetic group.

Propensity Score–Matched Comparison of Baseline Characteristics and Comorbidities Between the Nondiabetic and Diabetic Groups

After propensity score matching, 96 patients were included in each group. The standardized difference in the moderator variables in the matched cohort was <10%. No significant differences in baseline characteristics or comorbidities were identified between the groups (Table 3).

Propensity Score–Matched Comparison of Post-Injury Complications and Neurological Outcomes Between the Nondiabetic and Diabetic Groups

No significant differences between the groups in terms of post-injury complications during hospitalization were observed. Similarly, the mean total AMS score at the time of discharge was not significantly different between the groups (Table 4). The AIS grade, mean total AMS, and upper and lower extremity AMS at 6 months post-injury were also not significantly different between the groups. The degree of

Table 1. Comparison of Baseline Characteristics and Medical Comorbidities.

Variables	Non-diabetic (n = 270)	Diabetic (n = 119)	OR [95% CI]	*P* Value
Demographic variables				
Age (years)	74.4 ± 6.4	75.5 ± 6.3	-	0.08**
Sex (male, %)	188 (69.6)	91 (76.5)	0.8 [0.8-2.4]	0.17**
BMI	22.3 ± 3.3	22.3 ± 3.7	-	0.98**
Blood glucose level (mg/dl)	125.1 ± 36.3	172.6 ± 71.9	-	9.8×10^{-14}**
HbA1c (%)	-	7.1 ± 1.1	-	-
Medical comorbidities (%)				
Hypertension	128 (47.4)	66 (55.5)	1.4 [0.9-2.2]	0.15′
Cardiovascular disease	37 (13.7)	16 (13.4)	1.0 [0.5-1.9]	0.93′
Cerebrovascular disease	20 (7.4)	7 (5.9)	0.8 [0.3-2.0]	0.58′
Rheumatoid arthritis	1 (0.4)	3 (2.5)	6.9 [0.5-364.4]	0.09″
Osteoporosis	15 (5.6)	5 (4.2)	0.7 [0.2-2.2]	0.57″
Respiratory disease	12 (4.4)	2 (1.7)	0.4 [0.04-1.7]	0.24″
Renal disease	12 (4.4)	6 (5.0)	1.1 [0.3-3.4]	0.80′
Parkinson Disease	5 (1.9)	1 (0.8)	0.4 [0.01-4.1]	0.67″
Dementia	8 (3.0)	10 (8.4)	3.0 [1.0-9.0]	0.02′
Surgical histroy for musculoskeletal disorders	34 (12.6)	17 (14.3)	1.2 [0.6-2.2]	0.66′
Cervical OPLL (%)	84 (31.1)	43 (36.1)	1.3 [0.8-2.0]	0.33′
Cervical DISH (%)	23 (8.5)	14 (11.8)	1.4 [0.7-3.0]	0.31′
SI Changes on MRI (%)	219 (81.1)	101 (84.9)	1.5 [0.8-3.0]	0.22′
Steroid therapy for SCI (%)	60 (22.2)	19 (16.0)	0.7 [0.4-1.2]	0.16′
Surgical intervention (%)	136 (50.4)	71 (60.0)	1.5 [0.9-2.3]	0.09′
Baseline AIS grade (%)				
A	Omit	Omit		
B	7 (2.6)	7 (5.9)		
C	91 (33.7)	44 (37.0)		
D	172 (63.7)	68 (57.1)	-	0.19‴
Total AMS	70.9 ± 26.2	66.2 ± 27.5	-	0.10**
Upper extremity AMS	32.2 ± 13.6	30.6 ± 13.5	-	0.28**
Lower extremity AMS	38.9 ± 15.1	35.5 ± 16.5	-	0.06**

Means and standard deviations.
*unpaired t-test, **Wilcoxon rank-sum test, ***Welch's t-test.
'chi-squared test, "Fisher's exact test.
BMI body mass index, HbA1c glycated hemoglobin levels, OPLL ossification of posterior longitudinal ligament, DISH diffuse idiopathic skeletal hyperostosis, MRI magnetic resonance imaging, SI signal intensity, SCI spinal cord injury, AIS American Spinal Injury Association impairment scale, AMS American spinal injury association motor score, OR odds ratio, CI confidence interval.

improvement in the mean total, upper extremity, and lower extremity AMS at 6 months post-injury demonstrated no significant differences between the two groups.

Comparison of Baseline Characteristics, Medical Comorbidities, and Post-Injury Complications Between Nondiabetic And Moderate-Severe Diabetic Groups

Of the 119 patients with diabetes, 55 (46.2%) were categorized into the moderate-severe diabetic group. The prevalence rates of rheumatoid arthritis, dementia, cervical OPLL, and post-injury dysphagia in the moderate-severe diabetic group were significantly higher than those in the nondiabetic group (Table 5). The prevalence of steroid therapy for CSCI upon admission in the moderate-severe diabetic group was significantly lower than that in the nondiabetic group.

Influence of Moderate-Severe Diabetes on Neurological Recovery in CSCI Without Bone Injury

To identify the influence of moderate-severe diabetes on the degree of improvement in the mean total AMS at 6 months post-injury, we conducted a multiple linear regression analysis. Adjusted R^2 was 0.52. Age, dementia, and baseline total AMS negatively influence neurological recovery at 6 months post-injury (Table 6). Blood glucose levels at admission or moderate-severe diabetes did not affect neurological recovery 6 months post-injury.

Discussion

To the best of our knowledge, this is the first multicenter large-cohort study using propensity score matching to demonstrate

Table 2. Comparison of Post-injury Complications and Neurological Outcomes.

Variables	Non-diabetic (n = 270)	Diabetic (n = 119)	OR [95% CI]	*P* Value
Hospitalization (Days)	50.0 ± 68.3	65.3 ± 91.6	-	0.65**
Complications (%)				
Neurological deterioration (motor)	5 (1.9)	1 (0.8)	0.4 [0.01-4.1]	0.67″
Neurological deterioration (sensory)	3 (1.1)	1 (0.8)	0.7 [0.01-9.4]	0.75″
Cerebral infarction	1 (0.4)	1 (0.8)	2.3 [0.03-178.0]	0.52″
Delirium	9 (3.3)	7 (5.9)	1.8 [0.6-5.6]	0.25′
Dysphagia	2 (0.7)	4 (3.4)	4.6 [0.6-51.6]	0.07″
Respiratory failure	0 (0)	1 (0.8)	Inf [0.1-Inf]	0.31″
Pulmonary embolism	0 (0)	1 (0.8)	Inf [0.1-Inf]	0.31″
Pneumonia	8 (3.0)	2 (1.7)	0.6 [0.1-2.8]	0.73″
Renal infection	23 (8.5)	8 (6.7)	0.8 [0.3-1.8]	0.53′
Time point of discharge				
Total AMS	84.8 ± 17.3	78.1 ± 23.1	-	0.01**
6 months after the injury				
AIS grade (%)				
A	0 (0)	1 (0.8)		
B	0 (0)	2 (1.7)		
C	21 (7.8)	17 (14.3)		
D	206 (76.3)	86 (73.1)		
E	43 (15.9)	12 (10.1)	-	0.01″
Total AMS	88.9 ± 15.6	81.9 ± 21.8	-	1.1×10^{-3}**
Improvement of total AMS	17.9 ± 20.9	15.7 ± 16.2	-	0.96**
Upper extremity AMS	43.0 ± 8.5	40.3 ± 10.3	-	0.01**
Improvement of upper extremity AMS	10.8 ± 11.6	9.6 ± 8.7	-	0.96**
Lower extremity AMS	45.9 ± 8.2	41.6 ± 12.6	-	7.8×10^{-5}**
Improvement of lower extremity AMS	7.0 ± 12.3	6.1 ± 10.0	-	0.9**

Means and standard deviations.
*unpaired t-test, **Wilcoxon rank-sum test, ***Welch's t-test.
'chi-squared test, "Fisher's exact test.
AIS American spinal injury association impairment scale, AMS American spinal injury association motor score, OR odds ratio, CI confidence interval.

the influence of diabetes on neurological recovery in patients with CSCI without bone injury. Propensity score matching revealed no significant differences in the degree of improvement in AMS at 6 months post-injury between the nondiabetic and diabetic groups. Furthermore, multiple linear regression analysis indicated that moderate-severe diabetes did not affect the degree of improvement in AMS. The aforementioned findings provide valuable insights for decision-making regarding the medical treatment of patients with CSCI without bone injury.

Influence of Diabetes on Neurological Outcomes in CSCI Without Bone Injury

Previous reports have emphasized that diabetes leads to poor neurological outcomes following surgical intervention in patients with lumbar degenerative disc disease.[11,12] However, the influence of diabetes on neurological recovery in patients with cervical spine disorders remains controversial. Kim et al reported that although diabetic patients with CSM could benefit from cervical laminoplasty, their rate of recovery was expected to be lower than those without diabetes.[28] Machino et al conducted a prospective cohort study of more than 500 patients with CSM and concluded that both diabetic and nondiabetic patients with CSM experienced similar benefits from cervical laminoplasty.[14] Dokai et al and Nori et al noted that CSM patients with diabetes experienced improvements in neurological function as a result of posterior decompression surgery to the same extent as those without diabetes.[13,15] Although numerous studies have evaluated the influence of diabetes on neurological recovery in patients with CSM, few comparative studies have assessed this in patients with CSCI. Kobayakawa et al conducted a human cohort study of 206 patients with SCI, focusing on the relationship between blood glucose concentration on admission and functional outcomes, and supplemented by mouse model experiments.[18] They reported that hyperglycemia on admission exacerbated secondary injury, resulting in poor functional outcomes after SCI, regardless of whether the patient had diabetes. In the present study, a propensity score matching

Table 3. Propensity Score-matched Comparison of Baseline Characteristics and Medical Comorbidities.

Variables	Non-diabetic (n = 96)	Diabetic (n = 96)	*P* Value	Standardized Difference
Demographic variables				
Age (years)	74.7 ± 6.6	74.9 ± 6.2	0.70**	0.04
Sex (male, %)	73 (76.0)	74 (77.1)	1′	0.03
BMI	22.5 ± 3.1	22.3 ± 3.8	0.58***	0.05
Blood gluicose level (mg/dl)	125.6 ± 33.7	171.4 ± 64.9	8.8×10^{-8}**	-
HbA1c (%)	-	7.1 ± 1.1	-	
Medical comorbidities (%)				
Hypertension	48 (50.0)	49 (51.0)	1′	0.02
Cardiovascular disease	12 (12.5)	13 (13.5)	1′	0.03
Cerebrovascular disease	6 (6.2)	6 (6.2)	1′	<0.001
Rheumatoid arthritis	1 (1.0)	1 (1.0)	1″	<0.001
Osteoporosis	5 (5.2)	5 (5.2)	1″	<0.001
Respiratory disease	2 (2.1)	2 (2.1)	1″	<0.001
Renal disease	6 (6.2)	4 (4.2)	0.75″	0.09
Parkinson Disease	2 (2.1)	1 (1.0)	1″	0.08
Dementia	6 (6.2)	4 (4.2)	0.75″	0.09
Surgical histroy for	12 (12.5)	12 (12.5)	1′	<0.001
Musculoskeletal disorders				
Cervical OPLL (%)	33 (34.4)	35 (36.5)	0.88′	0.04
Cervical DISH (%)	11 (11.5)	11 (11.5)	1′	<0.001
SI Changes on MRI (%)	85 (88.5)	83 (86.5)	0.83′	0.06
Steroid therapy for SCI (%)	16 (16.7)	18 (18.8)	0.85′	0.06
Surgical intervention (%)	56 (58.3)	58 (60.4)	0.88′	0.04
Baseline AIS grade (%)				
A	-	-		
B	5 (5.2)	7 (7.3)		
C	34 (35.4)	32 (33.3)		
D	57 (59.4)	57 (59.4)	0.87″	0.09
Total AMS	65.8 ± 27.9	67.2 ± 28.0	0.7**	0.05
Upper extremity AMS	30.4 ± 13.3	31.3 ± 13.7	0.58**	0.07
Lower extremity AMS	35.4 ± 16.9	35.9 ± 16.6	0.8**	0.03

Means and standard deviations.
*unpaired t-test, **Wilcoxon rank-sum test, ***Welch's t-test.
'chi-squared test, "Fisher's exact test.
BMI body mass index, HbA1c glycated hemoglobin levels, OPLL ossification of posterior longitudinal ligament, DISH diffuse idiopathic skeletal hyperostosis, MRI magnetic resonance imaging, SI signal intensity, SCI spinal cord injury, AIS American Spinal Injury Association impairment scale, AMS American spinal injury association motor score.

demonstrated that AIS grade, AMS, and degree of improvement in AMS were not significantly different between two groups. Several possible factors influence the effects of diabetes on neurological recovery in the studies. First, the previous study included patients with CSCI as well as those with thoracic and lumbar SCI. Second, the sample size of the previous study was relatively small. Third, the previous study mainly focused on nondiabetic patients with SCI and hyperglycemia on admission. The study was strictly limited to patients with CSCI without bone injury and utilized a multicenter large-cohort analysis. We evaluated the differences in neurological recovery between the nondiabetic and diabetic groups using propensity score matching, which can accurately match baseline characteristics, including neurological status, on admission.

Influence of Moderate-Severe Diabetes on Neurological Recovery

There has been only one study has investigated relationship between severity of diabetes and neurological recovery in patients with SCI.[18] The study reported that the HbA1c was negatively associated with AMS and recovery rate of AMS.[18] However, the severity of diabetes is defined by not only HbA1c but also by the requirement for insulin treatment. In line with previous studies, including the Clinical Practice Guideline for Diabetes, we classified patients with an HbA1c $\geq$ 7.0%, which is the treatment target for preventing diabetes-related complications, and those requiring insulin treatment as having moderate-severe diabetes.[26,27] Multiple linear regression analysis revealed that age on admission,

Table 4. Propensity Score-matched Comparison of Post-injury Complications and Neurological Outcomes.

Variables	Non-diabetic (n = 96)	Diabetic (n = 96)	OR [95% CI]	*P* Value
Hospitalization (Days)	46.5 ± 51.1	65.4 ± 94.1	-	0.86**
Complications (%)				
Neurological deterioration (motor)	2 (2.1)	1 (1.0)	0.5 [0.01-9.6]	0.62″
Neurological deterioration (sensory)	1 (1.0)	1 (1.0)	1.0 [0.01-78.4]	1″
Cerebral infarction	0 (0)	1 (1.0)	Inf [0.03-Inf]	1″
Delirium	3 (3.2)	5 (5.2)	1.7 [0.3-11.1]	0.72″
Dysphagia	2 (2.1)	3 (3.1)	1.5 [0.2-18.3]	1″
Respiratory failure	0 (0)	1 (1.0)	Inf [0.03-Inf]	1″
Pulmonary embolism	0 (0)	1 (1.0)	Inf [0.03-Inf]	1″
Pneumonia	5 (5.2)	2 (2.1)	0.4 [0.04-2.4]	0.28″
Renal infection	9 (9.5)	6 (6.2)	0.6 [0.2-2.1]	0.43′
Time point of discharge				
Total AMS	81.9 ± 18.9	78.5 ± 23.7	-	0.51**
6 months after the onset				
AIS grade (%)				
A	0 (0)	1 (1.0)		
B	0 (0)	2 (2.1)		
C	11 (11.5)	14 (14.6)		
D	75 (78.1)	70 (72.9)		
E	10 (10.4)	9 (9.4)	-	0.56**
Total AMS	86.4 ± 17.2	82.0 ± 22.5	-	0.28**
Improvement of total AMS	20.6 ± 21.7	14.9 ± 16.2	-	0.14**
Upper extremity AMS	41.9 ± 9.0	40.3 ± 10.8	-	0.39**
Improvement of upper extremity AMS	11.6 ± 10.9	9.0 ± 8.7	-	0.19**
Lower extremity AMS	44.5 ± 9.1	41.8 ± 12.7	-	0.17**
Improvement of lower extremity AMS	9.1 ± 14.1	5.9 ± 10.1	-	0.29**

Means and standard deviations.
*unpaired t-test, **Wilcoxon rank-sum test, ***Welch's t-test.
'chi-squared test, "Fisher's exact test.
AIS American spinal injury association impairment scale, AMS American spinal injury association motor score, OR odds ratio, CI confidence interval.

preexisting dementia, and baseline total AMS negatively influenced neurological recovery, whereas blood glucose level and moderate-severe diabetes did not impact neurological recovery. These results indicated that tight glycemic control by diabetologists during the acute phase of CSCI might lead to better neurological outcomes even in moderate-severe diabetic patients who not only had conservative treatment but also underwent the surgery. Several reports have suggested that early surgery for CSCI is beneficial, and we believe that clinicians should not consider the sufficient justification to deny the spine surgery for CSCI with diabetes alone.[5,29] The present study did not evaluate the transition in blood glucose concentration during post-injury hospitalization. Further study is mandatory to reveal the impact of tight glycemic control during the acute phase of CSCI on neurological recovery in patients with CSCI. Age is widely recognized as a negative prognostic factor for neurological recovery in patients with SCI.[30] Jakob et al reported that elderly SCI patients have difficulties in translating an improvement of neurological deficit into function even after discharge from the rehabilitation center.[31] Clinicians should consider the personalized rehabilitation approaches which focus on training of daily living activities and ensure that patients are motivated to apply the skills they have acquired. A systematic review of cognitive function after SCI demonstrated a strong correlation between cognitive impairment and SCI and identified cognitive impairment as a predictor of poor social participation, including rehabilitation post-discharge.[32,33] In patients with dementia, the quality of rehabilitation might decline, which could contribute to reduced neurological improvement. Systematic literature review based on prospective studies also revealed that diabetes mellitus is likely to increase the risk of cognitive impairment such as Alzheimer's disease, which has frequently been attributed to cerebrovascular disease.[34] While diabetes can be managed with the tight glycemic control, interventions for dementia are often more challenging. Therefore, a multidisciplinary CSCI care approach should focus on preventing prolonged immobility and complications such as infections (pneumonia, renal infection and surgical site infection) that can exacerbate preexisting cognitive impairment. Nakajima et al reported similar findings, asserting that the AIS grade on admission and the presence of dementia/delirium were

Table 5. Comparison of Baseline Characteristics, Medical Comorbidities and Post-injury Complications Between Non-diabetic and Moderate-Severe Diabetic Groups.

Variables	Non-diabetic (n = 270)	Moderate-Severe Diabetic (n = 55)	OR [95% CI]	*P* Value
Demographic variables				
Age (years)	74.4 ± 6.4	74.9 ± 6.2	-	0.46**
Sex (male, %)	188 (69.6)	45 (81.8)	2.0 [0.9-4.6]	0.10′
BMI	22.3 ± 3.3	22.1 ± 3.6	-	0.75*
Blood gluicose level (mg/dl)	125.1 ± 36.3	206.6 ± 84.9	-	5.9×10^{-15}**
HbA1c (%)	-	7.8 ± 1.1	-	-
Medical comorbidities (%)				
Hypertension	128 (47.4)	31 (56.4)	1.4 [0.8-2.7]	0.3′
Cardiovascular disease	37 (13.7)	8 (14.5)	1.1 [0.4-2.5]	1′
Cerebrovascular disease	20 (7.4)	4 (7.3)	1.0 [0.2-3.1]	1″
Rheumatoid arthritis	1 (0.4)	3 (5.5)	15.3 [1.2-810.3]	0.02″
Osteoporosis	15 (5.6)	2 (3.6)	0.6 [0.1-2.9]	0.75″
Respiratory disease	12 (4.4)	1 (1.8)	0.4 [0.01-2.8]	0.7″
Renal disease	12 (4.4)	2 (3.6)	0.8 [0.1-3.8]	1″
Parkinson Disease	5 (1.9)	0 (0)	0 [0-5.4]	0.59″
Dementia	8 (3.0)	7 (12.7)	4.7 [1.4-15.7]	5.4×10^{-3}′
Surgical history for musculoskeletal disorders	34 (12.6)	7 (12.7)	1.0 [0.4-2.5]	1′
Cervical OPLL (%)	84 (31.1)	26 (47.3)	2.0 [1.1-3.7]	0.03′
Cervical DISH (%)	23 (8.5)	5 (9.1)	1.1 [0.3-3.1]	0.8″
SI Changes on MRI (%)	219 (81.1)	48 (87.3)	1.5 [0.6-4.2]	0.46′
Steroid therapy for CSCI (%)	60 (22.2)	4 (7.3)	0.3 [0.1-0.8]	9.0×10^{-3}″
Surgical intervention (%)	136 (50.4)	31 (56.4)	1.3 [0.7-2.4]	0.46′
Baseline AIS grade (%)				
A	Omit	Omit		
B	7 (2.6)	2 (3.6)		
C	91 (33.7)	20 (36.4)		
D	172 (63.7)	33 (60.0)	-	0.73″
Total AMS	70.9 ± 26.2	67.9 ± 26.2	-	0.35**
Upper extremity AMS	32.2 ± 13.6	31.7 ± 13.0	-	0.71**
Lower extremity AMS	38.9 ± 15.1	36.3 ± 16.1	-	0.24**
Complications (%)				
Neurological deterioration (motor)	5 (1.9)	1 (1.8)	1.0 [0.02-9.0]	1″
Neurological deterioration (sensory)	3 (1.1)	1 (1.8)	1.6 [0.03-20.8]	0.53″
Cerebral infarction	1 (0.4)	1 (1.8)	4.9 [0.1-388.2]	0.31″
Delirium	9 (3.3)	5 (9.1)	2.9 [0.7-10.0]	0.07″
Dysphagia	2 (0.7)	3 (5.5)	7.6 [0.9-93.2]	0.04″
Respiratory failure	0 (0)	0 (0)	-	-
Pulmonary embolism	0 (0)	0 (0)	-	-
Pneumonia	8 (3.0)	1 (1.8)	0.6 [0.01-4.7]	1″
Renal infection	23 (8.5)	2 (3.6)	0.4 [0.04-1.7]	0.28″

Means and standard deviations.
*unpaired t-test, **Wilcoxon rank-sum test, ***Welch's t-test.
'chi-squared test, "Fisher's exact test.
BMI body mass index, HbA1c glycated hemoglobin levels, OPLL ossification of posterior longitudinal ligament, DISH diffuse idiopathic skeletal hyperostosis, MRI magnetic resonance imaging, SI signal intensity, SCI spinal cord injury, AIS American Spinal Injury Association impairment scale, AMS American spinal injury association motor score, OR odds ratio, CI confidence interval.

independent prognostic factors for walking recovery (for patients whose baseline AIS grades A-C converted to AIS grades D-E) among patients with CSCI without bone injury, using a comparable patient population from the same original dataset,[7] and they concluded that diabetes did not influence neurological recovery in patients with CSCI without bone injury. Conversely, Nori et al, utilizing a similar patient population from the same original dataset, applied multiple linear regression analysis, which revealed that diabetes adversely impacted postoperative changes in total AMS among patients with CSCI without bone injury who had undergone surgery.[20] The discrepancy in results between the studies may be attributable to different patient selection criteria. Specifically, the present study and that of Nakajima et al included patients with CSCI without bone injury who underwent both surgery and conservative treatment, whereas the study of Nori

Table 6. Multiple Linear Regression Analysis of the Degree of Improvement in Total AMS at 6 months After Injury.

		95% CI			
	B	Lower	Upper	VIF	*P* Value
Demographic variables					
Age	−0.34	−0.59	−0.08	1.12	0.01*
Sex	−1.25	−4.70	2.20	1.04	0.48
BMI	0.34	−0.13	0.82	1.04	0.16
Blood glucose level	0.01	−0.03	0.05	1.55	0.56
Medical comorbidities					
Diabetes mellitus	−2.48	−7.56	2.60	1.64	0.34
Dementia	−16.50	−24.99	−8.01	1.17	1.7×10^{-4}*
Rheumatoid arthritis	2.73	−15.28	20.75	1.09	0.77
Cervical OPLL	−1.10	−4.69	2.49	1.21	0.55
Cervical DISH	0.30	−5.43	6.03	1.11	0.92
SI Changes on MRI	−0.74	−5.17	3.69	1.14	0.74
Steroid therapy for SCI	2.94	−1.24	7.11	1.11	0.17
Surgical intervention	−1.02	−4.33	2.28	1.14	0.54
Baseline AIS grade [C]	7.39	−3.14	17.92	3.01	0.17
Baseline AIS grade [D]	8.04	−4.26	20.34	3.01	0.20
Baseline total AMS	−0.62	−0.72	−0.51	3.05	4.3×10^{-25}*
Complications					
Dysphagia	−9.08	−20.94	2.78	1.17	0.13

*Statistically significant.
BMI body mass index, OPLL ossification of posterior longitudinal ligament, DISH diffuse idiopathic skeletal hyperostosis, MRI magnetic resonance imaging, SI signal intensity, SCI spinal cord injury, AIS American Spinal Injury Association impairment scale, AMS American spinal injury association motor score, VIF variance inflation factor.

et al focused solely on those who underwent surgery. The findings suggest that patients with CSCI without bone injury who have undergone surgery are more likely to develop moderate-severe diabetes.

Limitations and Strength

The primary limitation of this study is its retrospective nature; thus, the study has some missing values. The present study had the retrospective nature and the analyses were conducted using a heterogeneous cohort without a control group, which was not a comprehensive survey. Since the clinical data were collected from 33 high-volume trauma centers, it was difficult to obtain the completely accurate medical records. The present study excluded patients with missing values in diabetes, baseline AIS grade, baseline AMS, AIS grade at 6 months after injury, or AMS at 6 months after injury, which can cause the selection bias. The present study excluded the patients with baseline AIS grade A as well even though that could cause selection bias. The previous study demonstrated the probability of the AIS grade converting to grades C–E is very low in SCI patients with AIS grade A.[23] We excluded 23 CSCI patients without bone injury who were classified into AIS grade A on admission, and 83% of them were AIS grade A or B at final follow-up. This result was comparable with the previous study and indicated that it was difficult to evaluate the neurological recovery especially motor function. Furthermore, it was difficult to completely compare treatment outcomes due to the biases that the surgical indications and procedures, and rehabilitation approaches differed across the facilities and there was no standardized treatment approach. Additionally, the survey was conducted in Japan, which has the unique background of being the most advanced aging society in the world. However, as the pace of population aging has been accelerating drastically worldwide, the results of the present study might be generalized beyond the specific population. Regarding the missing values, smoking and drinking histories, which could potentially influence the neurological outcomes, were excluded from the analysis due to the high proportion of missing data. Although the proportion of missing values in other demographic variables that evaluated in the present study were all less than 5% and pairwise deletion was used in each statistical analysis, that could potentially cause the bias. Preexisting conditions such as CSM and SI changes in the spinal cord on MRI have not evaluated because of the retrospective nature of the present study. Evaluation of cervical and global spinal alignments, which might have potentially impact on the neurological outcomes, was not performed. Follow-up period was relatively short at 6 months post-injury to evaluate cases that show delayed recovery following CSCI. It is common for CSCI patients to be transferred to the rehabilitation hospitals immediately after acute-phase treatment in Japan, resulting in a high proportion of missing AMS at 12 months post-injury (approximately 30%),

which also had to be excluded from the analysis. Neurological outcomes at 6-month post-injury were chosen, since the previous clinical trials demonstrated that neurological recovery after traumatic SCI mainly occurs within the first 6-9 months.[24,25] Further studies are mandatory to determine the influence of diabetes on long-term neurological recovery in patients with CSCI without bone injury. The transition in blood glucose concentration during post-injury hospitalization was not observed in the present study. Kobayakawa et al reported that controlling the blood glucose concentration during the acute phase of SCI could prevent exacerbation of the pathophysiology and improve motor function in hyperglycemic mice post-SCI.[18] Tight glycemic control by diabetologists during the acute phase of CSCI might lead to better neurological outcomes in diabetic patients with CSCI without bone injury in the present study. Despite these limitations, this is the first multicenter large-cohort study to demonstrate the influence of diabetes on neurological recovery in patients with CSCI without bone injury.

Conclusion

We investigated the influence of diabetes, particularly moderate-severe diabetes, on neurological recovery in patients with CSCI without bone injury. Diabetic patients with CSCI without bone injury experienced improvements in neurological function comparable to those of nondiabetic patients. Furthermore, moderate-severe diabetes did not affect neurological recovery.

Acknowledgments

The authors acknowledge the contributions of the members of the 33 participating institutions in the assistance with data collection.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical Statement

Ethical Approval

The study received ethical approval from the institutional review boards of the participating hospitals and was approved by the institutional review board of Keio University School of Medicine representative facility (Keio University School of Medicine, No. 20200233). We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during the course of this research.

Informed Consent

Informed consent was obtained from all the participants.

ORCID iDs

Kazuki Takeda https://orcid.org/0000-0003-3857-4985
Kota Watanabe https://orcid.org/0000-0002-4830-4690
Takeshi Sasagawa https://orcid.org/0000-0002-3849-0178
Hiroaki Nakashima https://orcid.org/0000-0002-0039-9678
Naoki Segi https://orcid.org/0000-0001-9681-2422
Atsushi Yunde https://orcid.org/0000-0003-3202-056X
Hideaki Nakajima https://orcid.org/0000-0001-8260-7401
Tomohiro Yamada https://orcid.org/0000-0002-7220-7321
Hidenori Suzuki https://orcid.org/0000-0002-3156-0591
Yasuaki Imajo https://orcid.org/0000-0003-1291-745X
Shota Ikegami https://orcid.org/0000-0001-6404-5249
Ko Hashimoto https://orcid.org/0000-0002-9644-054X
Hidetomi Terai https://orcid.org/0000-0001-9183-3363
Gen Inoue https://orcid.org/0000-0001-6500-9004
Yasushi Oshima https://orcid.org/0000-0003-4696-1846
Toshitaka Yoshii https://orcid.org/0000-0003-3511-9020
Masayuki Ishihara https://orcid.org/0000-0001-6062-6767
Shiro Imagama https://orcid.org/0000-0002-6951-8575

References

1. Miyakoshi N, Suda K, Kudo D, et al. A nationwide survey on the incidence and characteristics of traumatic spinal cord injury in Japan in 2018. *Spinal Cord*. 2021;59(6):626-634. doi:10.1038/s41393-020-00533-0
2. Ishida Y, Tominaga T. Predictors of neurologic recovery in acute central cervical cord injury with only upper extremity impairment. *Spine*. 2002;27(15):1652-1658. doi:10.1097/00007632-200208010-00011.
3. Kato H, Kimura A, Sasaki R, et al. Cervical spinal cord injury without bony injury: a multicenter retrospective study of emergency and critical care centers in Japan. *J Trauma*. 2008;65(2):373-379. doi:10.1097/TA.0b013e31817db11d
4. Oichi T, Oshima Y, Okazaki R, Azuma S. Preexisting severe cervical spinal cord compression is a significant risk factor for severe paralysis development in patients with traumatic cervical spinal cord injury without bone injury: a retrospective cohort study. *Eur Spine J*. 2016;25(1):96-102. doi:10.1007/s00586-015-4142-4
5. OSCIS investigatorsChikuda H, Koyama Y, et al. Effect of early vs delayed surgical treatment on motor recovery in incomplete cervical spinal cord injury with preexisting cervical stenosis: a randomized clinical trial. *JAMA Netw Open*. 2021;4(11):e2133604. doi:10.1001/jamanetworkopen.2021.33604
6. Uribe J, Green BA, Vanni S, Moza K, Guest JD, Levi AD. Acute traumatic central cord syndrome–experience using surgical decompression with open-door expansile cervical laminoplasty. *Surg Neurol*. 2005;63(6):505-510. doi:10.1016/j.surneu.2004.09.037
7. Nakajima H, Yokogawa N, Sasagawa T, et al. Prognostic factors for cervical spinal cord injury without major bone injury in elderly patients. *J Neurotrauma*. 2022;39(9-10):658-666. doi:10.1089/neu.2021.0351
8. Browne JA, Cook C, Pietrobon R, Bethel MA, Richardson WJ. Diabetes and early postoperative outcomes following lumbar

fusion. *Spine*. 2007;32(20):2214-2219. doi:10.1097/BRS.0b013e31814b1bc0

9. Golinvaux NS, Varthi AG, Bohl DD, Basques BA, Grauer JN. Complication rates following elective lumbar fusion in patients with diabetes: insulin dependence makes the difference. *Spine*. 2014;39(21):1809-1816. doi:10.1097/BRS.0000000000000506
10. Tang H, Zhu J, Ji F, Wang S, Xie Y, Fei H. Risk factors for postoperative complication after spinal fusion and instrumentation in degenerative lumbar scoliosis patients. *J Orthop Surg Res*. 2014;9(1):15. doi:10.1186/1749-799X-9-15
11. Airaksinen O, Herno A, Turunen V, Saari T, Suomlainen O. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. *Spine*. 1997;22(19):2278-2282. doi:10.1097/00007632-199710010-00016
12. Arinzon Z, Adunsky A, Fidelman Z, Gepstein R. Outcomes of decompression surgery for lumbar spinal stenosis in elderly diabetic patients. *Eur Spine J*. 2004;13(1):32-37. doi:10.1007/s00586-003-0643-7
13. Dokai T, Nagashima H, Nanjo Y, Tanida A, Teshima R. Surgical outcomes and prognostic factors of cervical spondylotic myelopathy in diabetic patients. *Arch Orthop Trauma Surg*. 2012;132(5):577-582. doi:10.1007/s00402-011-1449-4
14. Machino M, Yukawa Y, Ito K, et al. Impact of diabetes on the outcomes of cervical laminoplasty: a prospective cohort study of more than 500 patients with cervical spondylotic myelopathy. *Spine*. 2014;39(3):220-227. doi:10.1097/BRS.0000000000000102
15. Nori S, Nagoshi N, Yoshioka K, et al. Diabetes does not adversely affect neurological recovery and reduction of neck pain after posterior decompression surgery for cervical spondylotic myelopathy: results from a retrospective multicenter study of 675 patients. *Spine*. 2021;46(7):433-439. doi:10.1097/BRS.0000000000003817
16. Kawaguchi Y, Matsui H, Ishihara H, Gejo R, Yasuda T. Surgical outcome of cervical expansive laminoplasty in patients with diabetes mellitus. *Spine*. 2000;25(5):551-555. doi:10.1097/00007632-200003010-00004
17. Machino M, Yukawa Y, Ito K, et al. Risk factors for poor outcome of cervical laminoplasty for cervical spondylotic myelopathy in patients with diabetes. *J Bone Joint Surg Am*. 2014;96(24):2049-2055. doi:10.2106/JBJS.N.00064
18. Kobayakawa K, Kumamaru H, Saiwai H, et al. Acute hyperglycemia impairs functional improvement after spinal cord injury in mice and humans. *Sci Transl Med*. 2014;6(256):256ra137. doi:10.1126/scitranslmed.3009430
19. Chen Z, Guo H, Lu Z, Sun K, Jin Q. Hyperglycemia aggravates spinal cord injury through endoplasmic reticulum stress mediated neuronal apoptosis, gliosis and activation. *Biomed Pharmacother*. 2019;112:108672. doi:10.1016/j.biopha.2019.108672
20. Nori S, Watanabe K, Takeda K, et al. Influence of the timing of surgery for cervical spinal cord injury without bone injury in the elderly: a retrospective multicenter study. *J Orthop Sci*. 2023;29:480-485. doi:10.1016/j.jos.2023.01.004
21. Nori S, Watanabe K, Takeda K, et al. Does surgery improve neurological outcomes in older individuals with cervical spinal cord injury without bone injury? A multicenter study. *Spinal Cord*. 2022;60(10):895-902. doi:10.1038/s41393-022-00818-6
22. Maeda T, Ueta T, Mori E, et al. Soft-tissue damage and segmental instability in adult patients with cervical spinal cord injury without major bone injury. *Spine*. 2012;37(25):E1560-E1566. doi:10.1097/BRS.0b013e318272f345
23. Kirshblum S, Snider B, Eren F, Guest J. Characterizing natural recovery after traumatic spinal cord injury. *J Neurotrauma*. 2021;38(9):1267-1284. doi:10.1089/neu.2020.7473
24. Bracken MB, Shepard MJ, Holford TR, et al. Administration of methylprednisolone for 24 or 48 hours or tirilazad mesylate for 48 hours in the treatment of acute spinal cord injury. Results of the third national acute spinal cord injury randomized controlled trial. National acute spinal cord injury study. *JAMA*. 1997;277(20):1597-1604.
25. Geisler FH, Coleman WP, Grieco G, Poonian D, Sygen Study Group. The Sygen multicenter acute spinal cord injury study. *Spine*. 2001;26(24 Suppl):S87-S98. doi:10.1097/00007632-200112151-00015
26. Araki E, Goto A, Kondo T, et al. Japanese clinical practice guideline for diabetes 2019. *J Diabetes Investig*. 2020;11(4):1020-1076. doi:10.1111/jdi.13306
27. Mori T, Nagata T, Nagata M, Fujimoto K, Fujino Y, Mori K. Diabetes severity measured by treatment control status and number of anti-diabetic drugs affects presenteeism among workers with type 2 diabetes. *BMC Publ Health*. 2021;21(1):1865. doi:10.1186/s12889-021-11913-3
28. Kim HJ, Moon SH, Kim HS, et al. Diabetes and smoking as prognostic factors after cervical laminoplasty. *J Bone Joint Surg Br*. 2008;90(11):1468-1472. doi:10.1302/0301-620X.90B11.20632
29. Inoue T, Suzuki S, Endo T, Uenohara H, Tominaga T. Efficacy of early surgery for neurological improvement in spinal cord injury without radiographic evidence of trauma in the elderly. *World Neurosurg*. 2017;105:790-795. doi:10.1016/j.wneu.2017.06.070
30. Burns SP, Golding DG, Rolle WA Jr., Graziani V, Ditunno JF. Recovery of ambulation in motor-incomplete tetraplegia. *Arch Phys Med Rehabil*. 1997;78(11):1169-1172. doi:10.1016/s0003-9993(97)90326-9
31. Jakob W, Wirz M, van Hedel HJ, Dietz V, EM-SCI Study Group. Difficulty of elderly SCI subjects to translate motor recovery–"body function"–into daily living activities. *J Neurotrauma*. 2009;26(11):2037-2044. doi:10.1089/neu.2008.0824
32. Craig A, Nicholson Perry K, Guest R, Tran Y, Middleton J. Adjustment following chronic spinal cord injury: determining factors that contribute to social participation. *Br J Health Psychol*. 2015;20(4):807-823. doi:10.1111/bjhp.12143
33. Sachdeva R, Gao F, Chan CCH, Krassioukov AV. Cognitive function after spinal cord injury: a systematic review. *Neurology*. 2018;91(13):611-621. doi:10.1212/WNL.0000000000006244
34. Kopf D, Frolich L. Risk of incident Alzheimer's disease in diabetic patients: a systematic review of prospective trials. *J Alzheimers Dis*. 2009;16(4):677-685. doi:10.3233/JAD-2009-1011