arm
Annals of Rehabilitation Medicine

Volume 50, Number 5, October 2026
pages 345-412

www.e-arm.org

eISSN: 2234-0653

arm
Annals of Rehabilitation Medicine

Official Journal of Korean Academy of Rehabilitation Medicine

Endorsed by the International Society of Physical and Rehabilitation Medicine (ISPRM)
and the Asia-Oceanian Society of Physical and Rehabilitation Medicine (AOSPRM)

Volume 50, Number 5, October 2026

Original Articles

345 Recent Trends in Pediatric Rehabilitation Utilization and Availability in South Korea: A Nationwide Population-Based Study

354 Cervical Lateral Flexion Stretching Angles Applied During Physical Therapy in Infants With Congenital Muscular Torticollis: A Video-Based Observational Study

363 Effects of Stepping Training With Reduced Visual Cues in a Virtual Environment in Chronic Stroke Patients: A Randomized Controlled Trial

376 Association of Admission Sarcopenia With 3-Month Recovery of Trunk Control and Gait in Subacute Stroke Patients With Impaired Balance

389 Reliability and Concurrent Validity of the Motor Evaluation Scale for Upper Extremity in People With Chronic Stroke

401 Longitudinal Mean HbA1c and Electrodiagnostic Severity of Distal Symmetric Sensorimotor Polyneuropathy in Type 2 Diabetes: Comparison With Index HbA1c

www.e-arm.org

eISSN: 2234-0653

Annals of Rehabilitation Medicine

Volume 50, Number 5, October 2026

Aims and Scope

Annals of Rehabilitation Medicine (*ARM*) is the official journal of the Korean Academy of Rehabilitation Medicine. It is an international, peer-reviewed open access journal, which aims to be a global leader in sharing up-to-date knowledge dedicated to the advancement of care and enhancing the function and quality of life of persons with various disabilities and chronic illnesses. As the official journal of one of the largest societies of rehabilitation medicine in Asia and Oceania, more than 3,500 physiatrists regularly receive this journal's table of contents via email, and the number is continually increasing. This journal is endorsed by the International Society of Physical and Rehabilitation Medicine (ISPRM) and the Asia-Oceanian Society of Physical and Rehabilitation Medicine (AOSPRM). International members comprise approximately half the editorial board, conducting peer-reviews of submitted manuscripts or acting as handling editors by assigning peer-review tasks to other reviewers.

The journal encompasses all aspects of physical medicine and rehabilitation, including basic, translational, clinical, and policy research, as well as rehabilitation education. Research areas covered by this journal include rehabilitation for brain disorders, spinal cord injuries, neuromuscular diseases, and dysphagia; cancer, cardiopulmonary, geriatric, pain & musculoskeletal, and pediatric rehabilitation; electrodiagnosis, orthosis & prosthesis, physical therapy, and sports medicine; and other emerging fields in rehabilitation medicine.

General Information

Annals of Rehabilitation Medicine is published in English bi-monthly on the last days of February, April, June, August, October, and December. Publication type includes original research, reviews, and other types of articles that reflect the depth and breadth of this specialty. The editorial board consists of members from 13 different countries. Articles are welcome for submission from all over the world. Full text is available from: http://www.e-arm.org/. All or part of the journal articles is indexed/tracked/covered by SCOPUS, EBSCO, PubMed Central/PubMed, KoreaMed, KoMCI, DOI/Crossref, Google Scholar, and Web of Science Core Collection as Emerging Sources Citation Index (ESCI) journal.

This journal is supported by the Korean Academy of Rehabilitation Medicine as well as the Korean Federation of Science and Technology Societies Grant funded by the Korean Government (Ministry of Education).

Open Access

Annals of Rehabilitation Medicine is available freely in electronic form at http://www.e-arm.org

Annals of Rehabilitation Medicine is an Open Access Journal. All articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2026 The Korean Academy of Rehabilitation Medicine.

Contacting the Annals of Rehabilitation Medicine

All manuscripts must be submitted online through the ARM e-submission system at https://e-arm.org. Items pertaining to manuscripts submitted for publication, as well as letters or other forms of communication regarding the editorial management of the journal, should be sent to:

Editor-in-Chief
Byung-Mo Oh, MD, PhD (Seoul National University, Korea)

Deputy Editor-in-Chief
Sung-Rae Cho (Yonsei University, Korea)

Editorial Executive Committee
Byung-Mo Oh, MD, PhD (Seoul National University, Korea)
Sung-Rae Cho (Yonsei University, Korea)
Han Gil Seo, MD, PhD (Seoul National University, Korea)
Hyun Iee Shin, MD (Chung-Ang University, Korea)
Sung Hwa Ko, MD, PhD (Pusan National University, Korea)
Na Young Kim, MD (Yonsei University, Korea)

Editorial Office
Department of Rehabilitation Medicine, Seoul National University Hospital
101 Daehak-ro, Jongno-gu, Seoul, Korea
Tel +82-10-8678-2671 Fax +82-2-6072-5244 E-mail edit@e-arm.org; edit.karm@gmail.com

Editorial Board

Editor-in-Chief

Byung-Mo Oh *Seoul National University, Korea*

Deputy Editor-in-Chief

Sung-Rae Cho *Yonsei University, Korea*

Editorial Executive Committee

Byung-Mo Oh *Seoul National University, Korea*
Sung-Rae Cho *Yonsei University, Korea*
Han Gil Seo *Seoul National University, Korea*
Hyun Iee Shin *Chung-Ang University, Korea*
Sung Hwa Ko *Pusan National University, Korea*
Na Young Kim *Yonsei University, Korea*

Section Editors

Neurological Rehabilitation

Sung-Rae Cho *Yonsei University, Korea*

Pain and Musculoskeletal Rehabilitation

Sun Jae Won *The Catholic University of Korea, Korea*

Electrodiagnosis and Neuromuscular Diseases

Dong Hwee Kim *Korea University, Korea*

Pediatric Rehabilitation

Jeong-Yi Kwon *Sungkyunkwan University, Korea*

Geriatric Rehabilitation

Jae-Young Lim *Seoul National University, Korea*

Cancer Rehabilitation

Jae Yong Jeon *University of Ulsan, Korea*

Orthosis and Prosthesis

Soo-Kyung Bok *Chungnam National University, Korea*

Associate Editors

Moon Suk Bang *Seoul National University, Korea*
Min Cheol Chang *Yeungnam University, Korea*
Won Hyuk Chang *Sungkyunkwan University, Korea*
Ian Cameron *University of Sydney, Australia*
Alberto Esquenazi *MossRehab, USA*
Michael Fredericson *Stanford University, USA*
Felipe Fregni *Harvard Medical School, USA*
Walter Frontera *University of Puerto Rico School of Medicine, Puerto Rico*
Abderrazak Hajjioui *Abdelmalek Essaâdi University, Morocco*
Der-Sheng Han *National Taiwan University, Taiwan*
Jae-Young Han *Chonnam National University, Korea*
Bo Young Hong *The Catholic University of Korea, Korea*
Ming-Yen Hsiao *National Taiwan University, Taiwan*
Shin-Ichi Izumi *Tohoku University, Japan*
Kyung-Lim Joa *Inha University, Korea*
Se Hee Jung *Seoul National University, Korea*
Ji Hee Kim *Wonkwang University, Korea*
Na Young Kim *Yonsei University, Korea*
Sang-Hun Kim *Pusan National University, Korea*
Sung Hwa Ko *Pusan National University, Korea*
Ho Jun Lee *Dongguk University, Korea*
Jang Woo Lee *National Health Insurance Service Ilsan Hospital, Korea*
Jong Hwa Lee *Dong-A University, Korea*
Sang Yoon Lee *Seoul National University, Korea*
Seung Hak Lee *University of Ulsan, Korea*
Shi-Uk Lee *Seoul National University, Korea*
Jan Lexell *Lund University, Sweden*
Jianan Li *Nanjing Medical University, China*
Leonard S W Li *University of Hong Kong, China*
Peter Lim *Baylor College of Medicine in Houston, USA*
Reynaldo Rey-Matias *St Luke's Medical Center, Philippines*
William Micheo *University of Puerto Rico, Puerto Rico*
Masahiko Mukaino *Hokkaido University Hospital, Japan*
Min-Kyun Oh *Gyeongsang National University, Korea*
Chul-Hyun Park *Sungkyunkwan University, Korea*
Dougho Park *Pohang Stroke and Spine Hospital, Korea*
Jae Hyeon Park *Hanyang University, Korea*
Jihye Park *The Catholic University of Korea, Korea*
Mooyeon Oh-Park *Albert Einstein College of Medicine, USA*
Mohamad Yatim Saari *Serdang Hospital, Malaysia*
Han Gil Seo *Seoul National University, Korea*
Hyun Iee Shin *Chung-Ang University, Korea*
Joon-Ho Shin *National Rehabilitation Center, Korea*
Gerold Stucki *University of Lucerne, Switzerland*
Simon Fuk Tan Tang *Chang Gung University, Taiwan*
Yu Hui Won *Jeonbuk National University, Korea*
Shaun Xianghu Xiong *St George's Hospital, New Zealand*
Jin A Yoon *Pusan National University, Korea*

Statistical Editor

Joongyub Lee *Seoul National University, Korea*

Ethics Editor

Ock-Joo Kim *Seoul National University, Korea*

Manuscript Editors

Yeonjin Lee *Korean Academy of Rehabilitation Medicine, Korea*
Munju Lee *INFORang Inc., Korea*

arm
Annals of Rehabilitation Medicine

This journal was supported by the Korean Federation of Science and Technology Societies Grant funded by the Korean Government (Ministry of Education).

eISSN: 2234-0653

Volume 50, Number 5, October 2026

Annals of Rehabilitation Medicine

Original Articles

345 Recent Trends in Pediatric Rehabilitation Utilization and Availability in South Korea: A Nationwide Population-Based Study
Na Yoon Yoo, Seong Woo Kim, Tae-mi Youk, Seungbeen Hong

354 Cervical Lateral Flexion Stretching Angles Applied During Physical Therapy in Infants With Congenital Muscular Torticollis: A Video-Based Observational Study
Bumcheol Kim, Jeon-Woong Kang, Dae-Hyun Jang, Jaewon Kim

363 Effects of Stepping Training With Reduced Visual Cues in a Virtual Environment in Chronic Stroke Patients: A Randomized Controlled Trial
Pei-Yun Lee, Meng-Ru Tsai, Hsien-Chang Wang, Chih-Hung Chen, Hui-Yu Tseng, Sang-I Lin

376 Association of Admission Sarcopenia With 3-Month Recovery of Trunk Control and Gait in Subacute Stroke Patients With Impaired Balance
Kyungmin Kim, Hangi Kim, Jiwon Lee, Su Jin Choi

389 Reliability and Concurrent Validity of the Motor Evaluation Scale for Upper Extremity in People With Chronic Stroke
Shamay S.M. Ng, Peiming Chen, Siyue Li, Shiu Hang Chan, Yin Tung Chan, Man Chi Cheng, Chung Him Cheung, Yau Ming Cheung, Mimi M.Y. Tse, Cynthia Y.Y. Lai

401 Longitudinal Mean HbA1c and Electrodiagnostic Severity of Distal Symmetric Sensorimotor Polyneuropathy in Type 2 Diabetes: Comparison With Index HbA1c
Songi Han, Jongkyu Kim, Sujeong Choi, Eun Sang Yoon, Ho Kyung Yoo

©2026 The Korean Academy of Rehabilitation Medicine

Original Article

Ann Rehabil Med 2026;50(5):345-353
eISSN: 2234-0653
https://doi.org/10.5535/arm.260102

arm
Annals of Rehabilitation Medicine

Recent Trends in Pediatric Rehabilitation Utilization and Availability in South Korea: A Nationwide Population-Based Study

Na Yoon Yoo, MD[1]*, Seong Woo Kim, MD[1]*, Tae-mi Youk, MS[2], Seungbeen Hong, MD[1]

[1]*Department of Physical Medicine and Rehabilitation, National Health Insurance Service Ilsan Hospital, Goyang, Korea*
[2]*Research Institute, National Health Insurance Service Ilsan Hospital, Goyang, Korea*

Objective: To analyze recent trends in the supply and demand of pediatric rehabilitation in South Korea.

Methods: We conducted a retrospective observational study using National Health Insurance Service claims data from 2019 to 2023. Pediatric patients with conditions potentially requiring rehabilitation were identified across 28 disease categories, including those in the pediatric rehabilitation medical institution pilot project. Yearly numbers of prevalent cases, therapy participation, and hospital availability were examined. Pediatric rehabilitation hospitals (PRHs) were operationally defined as hospitals providing more than 50 rehabilitation claims annually for children aged six years or younger.

Results: The number of pediatric patients with conditions potentially requiring rehabilitation increased steadily over the study period, while hospitals providing pediatric rehabilitation services and PRHs both decreased. Delayed development (R62) claims nearly doubled, consistent with trends reported globally, and may indicate more effective utilization of healthcare resources for the early detection of developmental concerns. Cerebral palsy maintained high therapy participation rates despite declining prevalence. In contrast, intellectual disability and psychological developmental disorders showed rising prevalence but consistently low therapy participation.

Conclusion: This study highlights a discrepancy between the increasing number of pediatric patients with conditions potentially requiring rehabilitation and the limited expansion of rehabilitation services in South Korea. The increase in delayed development cases and persistently low rehabilitation therapy rates among children with intellectual disabilities and psychological developmental disorders, together with regional disparities in access, suggest a need to further strengthen institutional capacity and promote equitable access to pediatric rehabilitation services.

Keywords: Child, Neurodevelopmental disorders, Rehabilitation, Prevalence, Health services accessibility

Received: May 4, 2026
Revised: September 9, 2026
Accepted: September 29, 2026

Correspondence:
Seungbeen Hong
Department of Physical Medicine and Rehabilitation, National Health Insurance Service Ilsan Hospital, 100 Ilsan-ro, Ilsandong-gu, Goyang 10444, Korea.
Tel: +82-31-900-3507
Fax: +82-31-900-0343
E-mail: sbhong@nhimc.or.kr

*These authors contributed equally to this work.

© 2026 by Korean Academy of Rehabilitation Medicine

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Pediatric rehabilitation is a cornerstone of child health, enabling children with developmental challenges to achieve their fullest potential. Across diverse healthcare systems worldwide, its role has increasingly been recognized as a priority. In South Korea, heightened attention to childhood rehabilitation needs has created new opportunities to expand services and promote equitable access. With demand continuing to rise, the academic community and government have focused on assessing existing structures and advancing evidence-based strategies for improvement.

Building on these academic discussions, researchers have worked to strengthen the infrastructure for pediatric rehabilitation. Park et al. [1] conducted a comprehensive assessment of national supply and demand for pediatric rehabilitation and proposed a structured delivery system model, highlighting the importance of establishing public children's rehabilitation institutions and reinforcing the existing healthcare system. In a subsequent study, they emphasized the necessity of tailoring public institutions to regional characteristics—such as population distribution and healthcare accessibility—and provided detailed operational and designation criteria for each area [2]. Building on these foundations, Kim et al. [3] delineated the specific roles of public children's rehabilitation hospitals and medical centers, offering strategic recommendations to ensure the stable and effective functioning of the delivery system.

Reflecting both academic recommendations and societal needs, the government prioritized the development of regional pediatric rehabilitation infrastructure through major policy initiatives, including the 2017 *Plan for Strengthening National Health Insurance Coverage* and the 2018 *5th Comprehensive Policy Plan for Persons with Disabilities* [4,5]. During this period, a pronounced regional imbalance, due to the limited number of private-sector-led pediatric rehabilitation hospitals (PRHs), became a major social concern. In response, the government underscored the necessity of establishing regional hub hospitals from a public healthcare perspective. Accordingly, in 2018 the Ministry of Health and Welfare launched a public call for the construction of regional public children's rehabilitation hospitals and medical centers, based on the *Public Health and Medical Services Act* and the *Act on Welfare of Persons with Disabilities* [5].

To operationalize these policy directions, the Ministry introduced the pediatric rehabilitation medical institution pilot project. This initiative aimed to expand service capacity by integrating public efforts with existing private-sector resources, while testing the feasibility of a structured delivery system. The pilot project was reported to be effective in studies by Park et al. [6], with outcomes including increased treatment volume at designated centers, reduced financial burden on families, and revitalized community-linked services. Continuity and accessibility of care were also enhanced, and in certain regions, functional cooperation between public children's rehabilitation hospitals/medical centers and designated institutions strengthened local rehabilitation infrastructure. Nevertheless, disparities in service quality among designated institutions, staffing difficulties, and limitations in the reimbursement system were identified as persistent challenges that require ongoing policy refinement.

This study aims to analyze recent trends in the supply and demand of pediatric rehabilitation over a five-year period (2019–2023) using National Health Insurance Service (NHIS) claims data. By examining both dimensions, the study seeks to provide a comprehensive understanding of current challenges and opportunities within the pediatric rehabilitation system. The findings are expected to generate evidence that informs policy development, supports the rational allocation of resources, and guides the refinement of the delivery system, thereby contributing to improved accessibility and quality of care for children with disabilities.

METHODS

This study was designed as a retrospective observational study using NHIS claims data. Pediatric patients under 18 years of age in South Korea with conditions potentially requiring specialized rehabilitation were identified based on the data from a five-year period between 2019 and 2023. A total of 28 disease groups were considered conditions potentially requiring specialized rehabilitation, comprising the 26 disease groups included in the pediatric rehabilitation medical institution pilot project, as well as intellectual disability (ID) and disorders of psychological development. Detailed disease categories and their corresponding International Classification of Diseases, Tenth Revision codes are listed in Supplementary Table S1.

We identified the total number of pediatric patients with each of the 28 conditions from 2019 to 2023. Each condition was defined as diagnosed when at least two claims were recorded within a given year. Patients with incomplete or inaccurate data regarding age, sex, insurance premium quartile, or place of res-

idence were excluded from the analysis. Furthermore, for each year, we examined the number and proportion of patients with prevalent cases who had at least one claim for specialized rehabilitation during that year. Specialized rehabilitation comprised physical therapy, occupational therapy, pool therapy, rehabilitation counseling/social work services, and botulinum toxin injections, with corresponding procedure fee codes and drug active ingredient codes provided in Supplementary Table S2.

We analyzed annual trends in the total number of hospitals providing rehabilitation therapy to pediatric prevalent cases. To identify hospitals consistently delivering pediatric rehabilitation therapy, we operationally defined PRHs as hospitals that claimed specialized rehabilitation therapy more than 50 times annually for children aged six years or younger [1]. The annual total number of PRHs was also examined to evaluate the availability of hospitals with a greater focus on pediatric rehabilitation.

Additionally, to assess regional healthcare utilization patterns in pediatric rehabilitation, we analyzed the use of medical institutions within individuals' areas of residence by medical service area. Medical service areas are intermediate-scale areas established based on local residents' healthcare utilization patterns and accessibility. A total of 70 medical service areas are considered the basic units for the provision of essential local medical services and the allocation of medical resources [7]. We defined an individual's primary medical institution as the medical institution visited for the greatest number of days during the corresponding year. We subsequently assessed, by medical service area, the proportion of individuals whose primary medical institution was located within their area of residence.

The Institutional Review Board of National Health Insurance Service Ilsan Hospital reviewed and approved this study (NHIMC 2025-11-009). The ethics committee waived the need for written consent.

RESULTS

From 2019 to 2023, the number of patients with conditions potentially requiring pediatric rehabilitation increased from 188,663 to 246,699 (Fig. 1). Although the absolute number of patients receiving rehabilitation therapy rose slightly, the proportion declined as the overall patient population expanded.

Fig. 2 illustrates temporal trends in both patient numbers and hospitals providing pediatric rehabilitation. While the patient population grew, hospitals offering rehabilitation therapy decreased by 17.8% over five years (Fig. 2A), and the number of PRHs also declined (Fig. 2B).

The number of patients diagnosed with delayed development (DD, R62) increased markedly from 47,587 in 2019 to 96,149 in 2023, nearly doubling over five years. Since 2022, patients with DD have accounted for more than one-third of all pediatric patients with conditions potentially requiring rehabilitation (Supplementary Table S3). Although the absolute number of

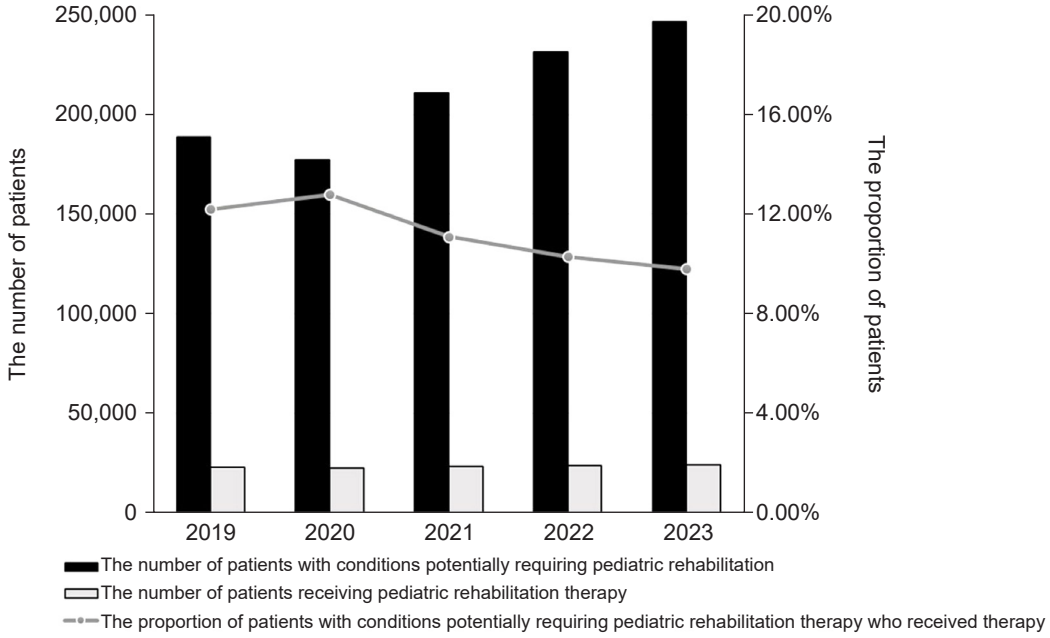

Fig. 1. Annual trends in the number of patients with conditions potentially requiring pediatric rehabilitation and receiving therapy.

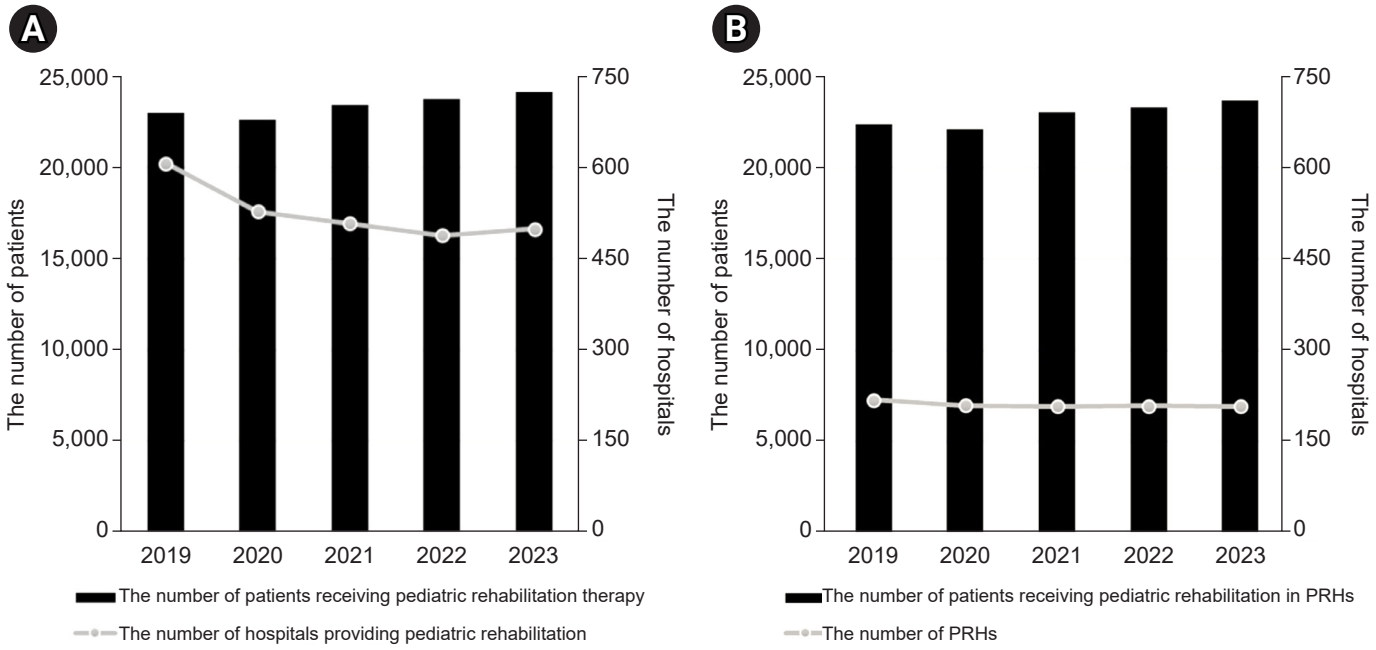

Fig. 2. (A) Annual trends in the number of hospitals providing pediatric rehabilitation services and treated pediatric patients. (B) Annual trends in the number of pediatric rehabilitation hospitals (PRHs) and treated pediatric patients.

patients with DD receiving rehabilitation therapy rose slightly, the proportion declined due to the sharp increase in overall prevalence (Fig. 3), reflecting the broader trend observed in Fig. 1.

Excluding DD (R62), the ten most prevalent disease categories in 2023 are shown in Fig. 4A. Patient numbers for other disorders of the nervous system (G90–99), disorders of psychological development (F80–89), ID (F7), and prematurity (P05–07) increased steadily, with growth rates of 45.45%, 36.97%, 29.72%, and 22.50%, respectively (Fig. 4B). In contrast, cerebral palsy (CP, G80–83) and perinatal respiratory and cardiovascular disorders (P20–24) declined by 15.81% and 15.43%. No consistent trends were observed for epilepsy, musculoskeletal malformations, nerve disorders, or nervous system malformations (G40–41, Q65–79, G50–59, Q00–07).

Supplementary Table S3 summarizes rehabilitation therapy rates across disease categories. CP consistently demonstrated the highest rates throughout the study period, followed by systemic atrophies primarily affecting the central nervous system (G10–12) and other perinatal disorders (P90–94). In contrast, disorders of psychological development (F80–89) and ID, despite their large patient populations, showed rehabilitation therapy rates below 10%.

Supplementary Table S4 presents the proportion of individuals whose primary medical institution was located within their area of residence from 2019 to 2023. In 2023, the highest proportions were observed in Jeonju (92.1%), Suncheon (91.1%), Wonju (90.9%), Changwon (89.4%), and Jeju (88.8%). In contrast, the proportion was 0.0% in Pocheon, Yeongwol, Geochang, Jeongeup, and Yeonggwang, followed by Sangju (2.3%), Namwon (2.4%), and Donghae (3.8%). No discernible temporal trend was observed in these proportions across the areas during the study period.

DISCUSSION

This study analyzed NHIS claims data from 2019 to 2023 to examine trends in pediatric patients with conditions potentially requiring pediatric rehabilitation and their rehabilitation service utilization. During this period, the number of pediatric patients with conditions potentially requiring pediatric rehabilitation increased consistently, despite an approximately 11% decline in the national pediatric population, from 7,928,907 in 2019 to 7,077,206 in 2023 [8]. Several factors may explain this trend, including advanced maternal age and rising rates of high-risk pregnancies; increased public participation in the national developmental screening program; evolving parental awareness; and improved accessibility of therapeutic services [9-12]. These factors collectively suggest that both demographic shifts and healthcare system changes contributed to the observed increase.

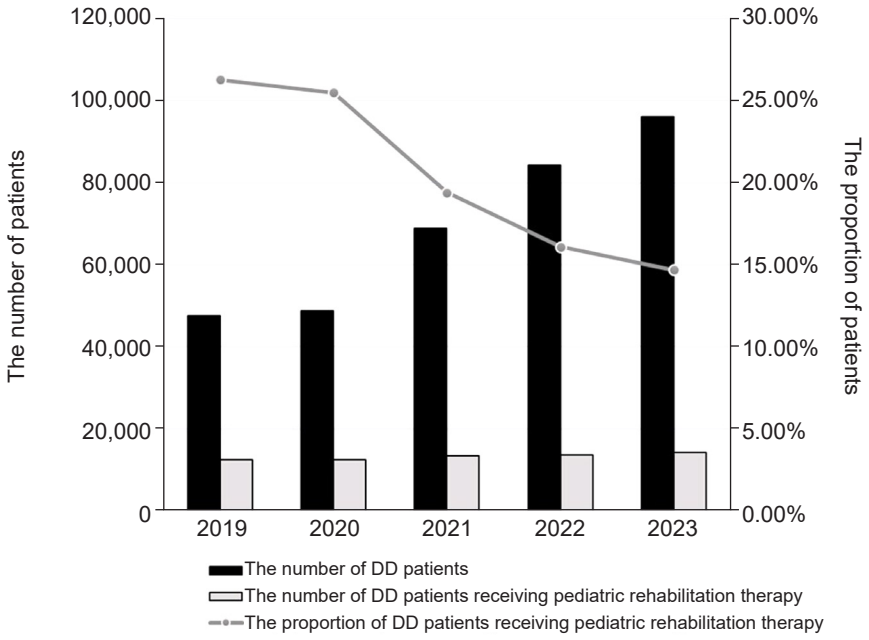

Fig. 3. Annual trends in the number of patients with delayed development (DD) and receiving rehabilitation therapy.

In Fig. 1, the absolute number of patients receiving pediatric rehabilitation therapy remained similar over the 5-year study period, whereas the number of children with DD receiving rehabilitation therapy in Fig. 3 slightly increased. Nevertheless, because the total number of individuals increased more substantially, the proportion of patients receiving rehabilitation therapy showed a relative decrease in both figures. This may be because the denominator increased as the total number of patients increased, resulting in a lower proportion of patients receiving rehabilitation therapy. Meanwhile, as demonstrated in Fig. 2, the number of hospitals providing pediatric rehabilitation services remained limited, with a declining trend over the study period. Thus, this may suggest that the proportion of patients receiving rehabilitation therapy did not increase because the current supply in pediatric rehabilitation has not kept pace with the demand. Additionally, individuals who could not be served by hospitals may have shifted to community-based developmental rehabilitation services. However, as this study was based on NHIS claims data, the analysis was limited to hospital- or clinic-based rehabilitation services. Further studies incorporating data on community-based service provision may better represent the actual utilization of pediatric rehabilitation services.

Consistent with the limited supply of pediatric rehabilitation services described above, Fig. 2 shows that the number of hospitals providing pediatric rehabilitation services and PRHs decreased during the study period. Hospital counts fell from 608 in 2019 to 528 in 2020, remaining around 500 thereafter (509, 489, and 500). Similarly, PRHs decreased from 219 in 2019 to 210 in 2020, with subsequent counts fluctuating minimally between 208 and 209. These findings indicate that the availability of pediatric rehabilitation facilities did not increase despite the growing number of children with conditions potentially requiring rehabilitation services. The number of pediatric patients receiving rehabilitation also declined in 2020 compared with 2019 (Supplementary Table S3), a reduction likely attributable to the coronavirus disease 2019 (COVID-19) pandemic. Previous studies reported that outpatient utilization decreased by 15.7% and hospital admissions by 11.6% in 2020, while specialized pediatric care fell by 42.7%, returning to pre-pandemic levels in 2022 [13]. Another study found outpatient visits decreased by 13.4% following the spread of COVID-19, with pediatric outpatient visits showing a substantial decline of 37.5% [14]. These findings suggest that pediatric rehabilitation services may be vulnerable to external shocks such as the COVID-19 pandemic, warranting further consideration of resilience in service delivery.

A prior study using NHIS claims data from 2014 reported that conditions originating in the perinatal period accounted for the largest proportion of pediatric patients with conditions potentially requiring rehabilitation (38.5%), followed by

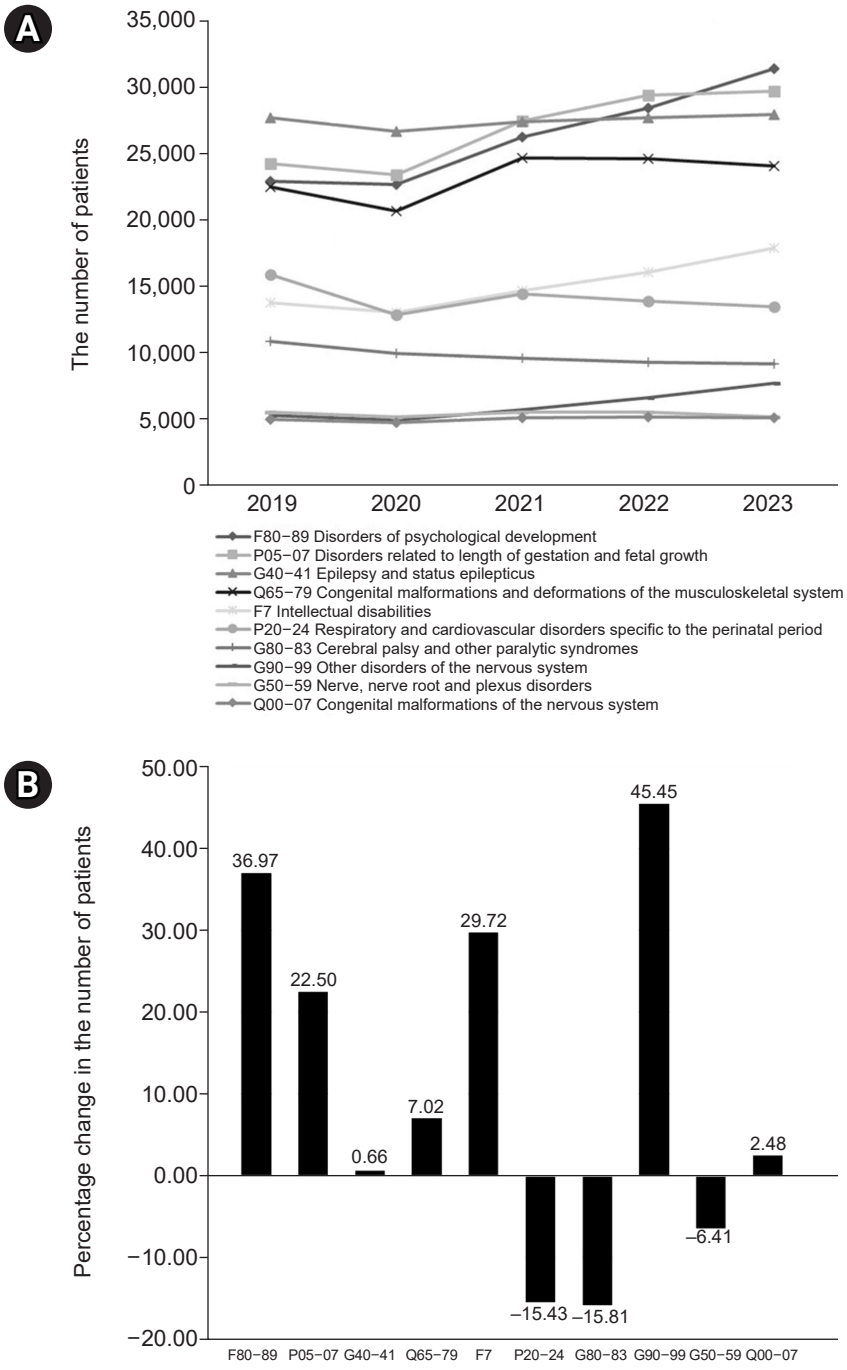

Fig. 4. Annual trends in the number of pediatric patients diagnosed with conditions potentially requiring rehabilitation therapy (by International Classification of Diseases, Tenth Revision diagnosis). (A) Temporal trends in the top ten disease categories ranked by prevalence. (B) Percentage change in the number of patients from 2019 to 2023.

congenital malformations, deformations and chromosomal abnormalities (30.2%), epilepsy (9.3%), and DD (8.8%) [15]. Although disease group classifications differed from those in our analysis, the nearly twofold increase in DD diagnoses observed in 2023 relative to 2019 suggests a substantial shift in pediatric rehabilitation practice. Recent NHIS Health Screening Statistical Yearbooks show a growing proportion of children who require detailed assessment or require follow-up examination,

reflecting strengthened detection of developmental concerns at the population level [16]. Epidemiological studies have reported similar increases in developmental delay diagnoses, underscoring heightened recognition of neurodevelopmental issues in pediatric populations [17,18]. Additionally, in clinical practice, DD is frequently coded as a presumed diagnosis prior to a definitive diagnosis, most often under R62. The observed increase in DD claims may be interpreted as an indication that clinical pathways for the early detection of developmental delay are being utilized more effectively.

Conversely, CP, a representative disability associated with substantial social costs during childhood and adolescence, showed a 16% reduction, from 10,854 patients in 2019 to 9,138 in 2023. Data from the Australian Cerebral Palsy Register similarly demonstrate a decline in prevalence, from 2.1 per 1,000 live births in 1995–1996 to 1.4 per 1,000 in 2015–2016, and our findings suggest a comparable downward trend in South Korea [19]. Despite this decline, CP remained the second most common condition after DD among children receiving rehabilitation therapy and was the only disorder with consistently high rehabilitation participation rates, exceeding 70%. Moreover, the cost of rehabilitation for children with CP has continued to rise, imposing a growing economic burden on families and healthcare systems [20]. These findings underscore that, although the number of patients with CP is decreasing, it remains a condition requiring sustained rehabilitation, and rehabilitation therapy for CP continues to account for the largest share of pediatric rehabilitation services.

Our results show that the prevalence of ID (F7) and disorders of psychological development (F80–89) has risen steadily over the five-year study period. This observation is consistent with prior NHIS claims analyses, which documented increases in ID and pervasive developmental disorders between 2007 and 2019 [10], as well as other studies reporting notable growth in autism spectrum and language disorders [11]. Fombonne [12] has attributed these upward trends to greater social awareness and improved diagnostic accessibility. Despite this rising prevalence, our findings reveal that, as of 2023, only 6.40% of patients with ID and 6.50% of those with psychological developmental disorders were receiving rehabilitation. This discrepancy highlights a mismatch between growing demand and limited service provision within the national health insurance system. Addressing this gap will require policy-level adjustments to expand coverage and ensure equitable access to rehabilitation therapy for these patient groups.

As shown in Fig. 4A, prematurity (P05–07) increased over the study period, whereas respiratory and cardiovascular disorders specific to the perinatal period (P20–24) and birth trauma (P10–15) declined (Supplementary Table S3). These divergent trends likely reflect the rise in high-risk pregnancies and maternal factors contributing to prematurity, alongside advances in obstetric management and neonatal care that have reduced other perinatal complications. The Korean Neonatal Network reported favorable outcomes following two years of quality improvement initiatives launched in 2017, including reductions in endotracheal intubation, chronic lung disease, and sepsis [21]. Similarly, longitudinal studies of very low birth weight infants have documented improved survival rates and declining incidence of major complications such as sepsis and intraventricular hemorrhage [22,23].

Regional disparities in access to pediatric rehabilitation services have previously been reported. The report showed that approximately 60% of inpatient pediatric rehabilitation services were provided in the Seoul metropolitan area, including Seoul, Gyeonggi, and Incheon, and that many patients received rehabilitation therapy in Seoul despite not residing there [15]. In this study, we assessed, by medical service area, the proportion of individuals whose primary medical institution for pediatric rehabilitation was located within their area of residence. The results suggest that regional disparities in access to pediatric rehabilitation services persisted during 2019–2023. Public children's rehabilitation hospitals and medical centers, as well as institutions participating in the first-phase pediatric rehabilitation medical institution pilot project, were designated from among institutions that had already been providing pediatric rehabilitation services. Therefore, the observed pattern may be attributable, at least in part, to the limited expansion of pediatric rehabilitation services in regions where such services had not previously been provided. This may also explain why the proportion of individuals whose primary medical institution was located within their area of residence did not show an increasing trend between 2019 and 2023.

Limitations

First, this study was designed as a retrospective observational analysis, which enabled identification of trends but did not allow assessment of specific causes or establishment of causal relationships. Second, the diagnostic categories analyzed were clinically heterogeneous, and rehabilitation needs may vary substantially among patients with the same diagnostic code.

Consequently, the rehabilitation participation rates observed in this study should be interpreted with caution, as they may partly reflect differences in the underlying indications for rehabilitation in addition to unmet needs or barriers to access. Third, further studies with longer-term follow-up and more recent data are needed to determine whether the implementation and expansion of public children's rehabilitation hospitals/medical centers and the pediatric rehabilitation medical institution pilot project increased the utilization of medical institutions within individuals' areas of residence. Fourth, in the present study, we operationally defined PRHs as hospitals that claimed specialized rehabilitation therapy more than 50 times annually for children aged six years or younger, based on the previous report. However, the number of specialized hospitals may differ according to the threshold selected for the annual number of rehabilitation therapy claims. Finally, the observation period overlapped with the COVID-19 pandemic, a major external factor that disrupted the healthcare system and reduced hospital utilization, potentially contributing to delayed diagnoses and lower rehabilitation uptake. These limitations should be acknowledged when interpreting the results and designing future policy interventions.

Conclusion

In conclusion, this study reveals a growing mismatch between the increasing number of pediatric patients with conditions potentially requiring rehabilitation and the limited expansion of rehabilitation services in South Korea. The sharp increase in DD claims and persistently low therapy participation among children with intellectual disabilities and psychological developmental disorders highlight the need for policy-level improvements. Moving forward, strengthening institutional capacity, and ensuring equitable access will be essential to building a resilient and future-oriented pediatric rehabilitation system.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

This study was supported by the 2025 National Health Insurance Service Ilsan Hospital grant (NHIMC-2025-PR-006).

AUTHOR CONTRIBUTION

Conceptualization: Kim SW. Methodology: Hong S. Formal analysis: Youk T. Funding acquisition: Kim SW. Project administration: Hong S. Visualization: Yoo NY, Kim SW. Writing – original draft: Yoo NY. Writing – review and editing: Hong S. Approval of final manuscript: all authors.

ACKNOWLEDGMENTS

This study used the National Health Information Database (NHIS-2026-03-1-021), made by the National Health Insurance Service.

SUPPLEMENTARY MATERIALS

Supplementary materials can be found via https://doi.org/10.5535/arm.260102.

ORCID

Na Yoon Yoo, https://orcid.org/0009-0001-0286-1811
Seong Woo Kim, https://orcid.org/0000-0002-1548-8147
Tae-mi Youk, https://orcid.org/0000-0002-4273-3777
Seungbeen Hong, https://orcid.org/0000-0002-8222-2920

REFERENCES

1. Park JH, Kim SW, Ko SE, Kwon JY, Lee JS, Lee SY, et al. A study on strategies for establishing a rehabilitation medical delivery system for children with cerebral palsy and other disabilities. Ministry of Health and Welfare; 2016 Dec.
2. Park JH, Kim SW, Ko SE, Kwon JY, Lee JS, Yang SS, et al. A study on strategies for expanding pediatric rehabilitation medical services. Ministry of Health and Welfare; 2017 Dec.
3. Kim MY, Ko SE, Han SH, Bang MS, Lee JS, Yang SS, et al. Strategies for enhancing pediatric rehabilitation services through public children's rehabilitation institutions. Ministry of Health and Welfare; 2018 Dec.
4. The 5th comprehensive policy plan for persons with disabilities [Internet]. Ministry of Health and Welfare; 2018 [cited 2026 Jan 27]. Available from: https://www.mohw.go.kr/board.es?mid=a10401000000&bid=0008&tag=&act=view&list_no=344647
5. Guidelines for the pilot project on designation and operation of pediatric rehabilitation medical institutions [Internet]. Health In-

surance Review & Assessment Service; 2020 [cited 2026 Jan 27]. Available from: https://www.hira.or.kr/bbsDummy.do?pgmid=HIRAA020002000100&brdScnBltNo=4&brdBltNo=8263#none
6. Park JG, Jeon IH, Kim SR, Seol YS. A study on evaluation of the performance of the pilot project for designated operation of children's rehabilitation medical institutions and estimate to expand the designated size. Health Insurance Review & Assessment Service; 2022 Dec.
7. Comprehensive plan for the development of public health and medical services to realize an inclusive nation with no regional disparities in essential healthcare [Internet]. Ministry of Health and Welfare; 2018 [cited 2026 Aug 22]. Available from: https://www.mohw.go.kr/board.es?mid=a10503010100&bid=0027&tag=&act=view&list_no=346233&cg_code=
8. KOSIS: Korean Statistical Information Service [Internet]. Ministry of Data and Statistics [cited 2025 Dec 13]. Available from: https://kosis.kr/index/index.do
9. Boyle CA, Boulet S, Schieve LA, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, et al. Trends in the prevalence of developmental disabilities in US children, 1997-2008. Pediatrics 2011;127:1034-42.
10. Kim SW, Jeon HR, Jung HJ, Kim JA, Song JE, Kim J. Clinical characteristics of developmentally delayed children based on interdisciplinary evaluation. Sci Rep 2020;10:8148.
11. Rah SS, Hong SB, Yoon JY. Prevalence and incidence of developmental disorders in Korea: a nationwide population-based study. J Autism Dev Disord 2020;50:4504-11.
12. Fombonne E. Editorial: the rising prevalence of autism. J Child Psychol Psychiatry 2018;59:717-20.
13. Yoo KJ, Lee Y, Lee S, Friebel R, Shin SA, Lee T, et al. The road to recovery: impact of COVID-19 on healthcare utilization in South Korea in 2016-2022 using an interrupted time-series analysis. Lancet Reg Health West Pac 2023;41:100904.
14. Byun H, Kang D, Go SI, Kim HI, Hahm JR, Kim RB. The impact of the COVID-19 pandemic on outpatients of internal medicine and pediatrics: a descriptive study. Medicine (Baltimore) 2022;101:e28884.
15. Kim SW, Kim JY, Jeon HR, Youk TM, Park JH, Kwon JY, et al. Analysis of utilization patterns of rehabilitation medical institutions for children with cerebral palsy and developmental disorders. National Health Insurance Service Ilsan Hospital; 2017 Dec.
16. 2024 National health screening statistical yearbook [Internet]. National Health Insurance Service; 2025 [cited 2026 Mar 5]. Available from: https://www.nhis.or.kr/nhis/together/wbhaec07000m01.do?mode=view&articleNo=11008126&article.offset=0&articleLimit=10
17. Li Q, Li Y, Zheng J, Yan X, Huang J, Xu Y, et al. Prevalence and trends of developmental disabilities among US children and adolescents aged 3 to 17 years, 2018-2021. Sci Rep 2023;13:17254.
18. Global report on children with developmental disabilities: from the margins to the mainstream [Internet]. UNICEF; 2023 [cited 2026 Mar 22]. Available from: https://www.unicef.org/documents/global-report-children-developmental-disabilities
19. Smithers-Sheedy H, Waight E, Goldsmith S, Reid S, Gibson C, Watson L, et al. Declining trends in birth prevalence and severity of singletons with cerebral palsy of prenatal or perinatal origin in Australia: a population-based observational study. Dev Med Child Neurol 2022;64:1114-22.
20. Kim SW, Jeon HR, Youk T, Kim J. Cost of rehabilitation treatment of patients with cerebral palsy in Korea. Ann Rehabil Med 2018;42:722-9.
21. Lee SM. Collaborative quality improvement in neonatal intensive care units. Perinatology 2025;36:9-14.
22. Song IG, Lim JW, Chang YS; Korean Neonatal Network. Improved outcomes of very low birth weight infants in Korea: 2015-2016 vs. 2021-2022 from the Korean neonatal network. J Korean Med Sci 2025;40:e258.
23. Jeon GW, Lee JH, Oh M, Chang YS. Serial short-term outcomes of very-low-birth-weight infants in the Korean Neonatal Network from 2013 to 2020. J Korean Med Sci 2022;37:e229.

Original Article

Ann Rehabil Med 2026;50(5):354-362
eISSN: 2234-0653
https://doi.org/10.5535/arm.260098

arm
Annals of Rehabilitation Medicine

Cervical Lateral Flexion Stretching Angles Applied During Physical Therapy in Infants With Congenital Muscular Torticollis: A Video-Based Observational Study

Bumcheol Kim, PhD, Jeon-Woong Kang, MD, Dae-Hyun Jang, MD, PhD, Jaewon Kim, MD, PhD

Department of Physical Medicine and Rehabilitation, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Incheon, Korea

Objective: To quantify the cervical lateral flexion stretching angles attained and tolerated during routine physical therapy in infants aged 12 months or younger with congenital muscular torticollis (CMT), and to compare them between two therapeutic positions and across age groups.

Methods: Infants with CMT were recorded once, at the first physical therapy session after diagnosis. Therapists stretched in both directions in supported forward-facing and supine positions. Angles were computed from video frames between the inter-pupillary and inter-glenohumeral vectors; the direction of flexion was determined visually from the recordings. The primary analysis used the therapeutic direction, lateral flexion away from the head tilt.

Results: Eighty-two infants were analyzed. In the therapeutic direction the angle was greater in the supported forward-facing position (41.9±8.0°, n=79) than supine (31.7±9.1°, n=54); among 51 infants measured in both, the mean paired difference was 10.8° (95% confidence interval, 8.1–13.5; p<0.001; Cohen's d=1.12). In exploratory comparisons, no difference was detected between the therapeutic and opposite directions (p=0.39, p=0.73) or across age groups (p>0.05). Inter-rater reliability of annotation was high (intraclass correlation coefficient, 0.98).

Conclusion: A greater lateral flexion angle was attained in the supported forward-facing position than supine; exploratory analyses detected no difference across age groups or between stretching directions. These findings provide descriptive values for therapist-applied stretching in this cohort; their relationship with clinical outcomes and safety requires further study.

Keywords: Torticollis, Infant, Range of motion, articular, Physical therapy modalities, Video recording

Received: May 12, 2026
Revised: September 13, 2026
Accepted: October 12, 2026

Correspondence:
Jaewon Kim
Department of Physical Medicine and Rehabilitation, Incheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 56 Dongsu-ro, Bupyeong-gu, Incheon 21431, Korea.
Tel: +82-32-280-6601
Fax: +82-32-280-5040
E-mail: jw2356@naver.com

© 2026 by Korean Academy of Rehabilitation Medicine

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Congenital muscular torticollis (CMT) is a prevalent congenital musculoskeletal disorder of infancy characterized by abnormal head and neck posture. It typically presents as lateral head tilt to one side with rotation of the chin to the opposite side and arises from various etiological factors [1,2]. The reported global incidence in children ranges from 0.3% to 1.9%, with considerable variation across populations [3].

Most cases are of muscular origin, affecting the sternocleidomastoid (SCM) muscle; less common causes include cervical spine abnormalities, ocular disorders such as strabismus, and, rarely, central nervous system pathology [4]. On the basis of ultrasonographic and physical findings, CMT is commonly classified into postural, muscular, and SCM-mass types, and limitation of passive range of motion (PROM) is a hallmark finding except in the postural type [1,5-8].

First-line treatment is physical therapy, with reported success rates of 61% to 99%, and early diagnosis and intervention are crucial to optimal outcomes [7,9-11]; if untreated, CMT may lead to craniofacial deformity, cervical spine dysplasia, restricted cervical motion, and asymmetric or delayed motor development [4,12-14].

Within physical therapy, PROM stretching exercises—achieved through specific positioning and handling techniques—form the cornerstone of management and target cervical rotation and lateral flexion; they are particularly important in infants with an SCM mass or muscle tightness [1,15,16]. Of the two directions, lateral flexion is the more difficult to standardize: its endpoint is judged from the therapist's perception of tissue resistance and the infant's tolerance rather than from a measured value, and neither the angle applied nor an intended target angle is routinely documented, so the magnitude of the stretch actually delivered in practice remains essentially uncharacterized.

Such therapeutic stretching angles are distinct from normative range of motion in healthy infants and from PROM measured during a brief assessment, and values obtained under one of these conditions should not be assumed to apply to the others.

The 2024 clinical practice guideline of the American Physical Therapy Association gives cervical passive stretching its strongest recommendation and specifies how passive motion should be measured at examination, but it does not specify how far the neck should be laterally flexed during a treatment session or in which position the stretch should be performed; the magnitude of the applied stretch, as distinct from its frequency, therefore remains undefined [1].

However, few studies have quantified the cervical lateral flexion angles that are actually applied and tolerated during therapeutic stretching sessions in infants aged 12 months or younger with CMT, and evidence on whether these angles differ across age in infancy is lacking.

We therefore aimed to quantify the cervical lateral flexion stretching angles applied during routine physical therapy sessions across age groups in infants aged 12 months or younger with CMT, and to compare these angles between two therapeutic positions—supported forward-facing and supine.

METHODS

Infants were enrolled prospectively into an observational cohort; the analysis of stretching angles reported here is cross-sectional, as each infant contributed a single recorded session. The study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement [17]. The study was approved by the Institutional Review Board of Incheon St. Mary's Hospital (No. OC23OISI0070). Written informed consent was obtained from a parent of every participant, including consent for the release of any potentially identifiable photographs and data. All procedures were performed in accordance with relevant guidelines and regulations.

Study participants

Infants who were diagnosed with CMT and required physical therapy in the Department of Physical Medicine and Rehabilitation of a tertiary hospital between July 2023 and May 2024 were enrolled. Inclusion criteria were (1) age 0 to 12 months; (2) a clinical diagnosis of CMT, established by a pediatric rehabilitation physician on the basis of at least one of the following three findings: a sustained lateral head tilt, a limitation of passive cervical rotation or lateral flexion relative to the contralateral side, or an SCM abnormality on ultrasonography (US); (3) a clinical indication for physical therapy; (4) normal neurological examination; (5) normal cervical X-ray findings; (6) no clinical suspicion of strabismus; (7) availability of neck US; and (8) full-term birth. The direction of the head tilt was documented for every infant; the SCM on the side toward which the head was tilted was regarded as the affected, shortened muscle and

determined the direction of stretching. Exclusion criteria were (1) history of trauma; (2) recent rapid onset or progression of torticollis; (3) suspected or confirmed concurrent disease; and (4) visual or auditory impairment.

Participants were classified into two groups on the basis of US findings: the SCM-mass group, which had a discrete abnormality of the SCM such as a mass, and the non-SCM-mass group, which had no such abnormality. Postural and muscular CMT were not distinguished within the non-SCM-mass group. That distinction rests on the presence of a PROM limitation, which could not be assessed reliably in every infant because cooperation was limited, particularly in older infants with stranger anxiety; US was therefore used as the sole, observer-independent basis for grouping.

Study procedures

Five physical therapists affiliated with the tertiary hospital conducted the physical therapy sessions. All had at least 5 years of experience in the physical therapy management of CMT and followed the same institutional approach, under which stretching occupied at least 10 minutes of each 30-minute session and prolonged low-load holds were the intended target; the number of repetitions and the duration of individual therapeutic holds were not prescribed or recorded. Therapist identity was not recorded in the study dataset, so therapist-level variability could not be quantified. Sessions were recorded with a smartphone (Samsung Galaxy A24; Samsung Electronics) capturing a frontal view of the infant. The smartphone was mounted on a tripod approximately 1 m from the infant, centered on the infant's midline and aligned perpendicular to the frontal plane. The same setup was used for all recordings, and the default camera magnification was retained throughout; no zoom was applied. Recorded maneuvers were cervical lateral flexion stretching performed in both directions in (1) a supported forward-facing (football-hold) position and (2) the supine position. Stretching in both directions forms part of routine practice at our institution. For analysis, the direction that elongates the affected SCM—lateral flexion away from the side toward which the head was tilted—was designated the therapeutic direction and used for the primary analysis; the opposite direction was analyzed separately. End range was defined as the point at which the therapist perceived firm tissue resistance and further lateral flexion could not be achieved without compensatory trunk or shoulder movement. During the recorded maneuver, the end-range position was held for 2 to 3 seconds to allow capture of the maximum angle; this brief hold served the recording and does not describe the duration of therapeutic holds during the remainder of the session. Crying alone did not exclude a frame: frames were retained if the stretch was nevertheless carried through its full range, and were excluded only when poor cooperation prevented this. Each infant was recorded once, at the first physical therapy session following diagnosis, so that recordings were obtained at a uniform point in the treatment course and before any effect of therapy could have accumulated. The technique recorded was identical to that used in routine sessions. Treatment was intended to continue until clinical resolution, defined as the absence of a visible head tilt together with a difference of 5° or less between sides in passive cervical rotation and lateral flexion, at which point direct therapy was discontinued. Treatment consisted of 30-minute physical therapy sessions one to two times per week, with daily 30-minute home exercises taught to caregivers. Adverse events were not monitored under a formal protocol; infants were observed by the treating therapist during and immediately after each session, and caregiver-reported events were recorded at subsequent outpatient visits.

Data collection and frame selection

Video recordings were segmented into still frames at 1-second intervals. Frames in which the cervical region was not clearly visible, or in which the extent of lateral flexion could not be determined, were excluded. Only frames in which the infant's face and trunk were oriented directly toward the camera (true frontal view) were retained. Frames in which lateral flexion had reached the maximum available range during stretching were used for analysis.

Image analysis

An image-based analysis was conducted in three stages: (1) region of interest (ROI) annotation, (2) ROI coordinate extraction, and (3) cervical lateral flexion stretching angle measurement.

Stage 1 — ROI annotation

Snapshots were extracted from videos using Python (version 3.10.6) and the opencv-contrib-python library (version 4.9.0.80) and processed in ImageJ (version 1.53a; National Institutes of Health). An annotation tool was used to mark ROIs at the key anatomical landmarks—the anterior glenohumeral joints of both shoulders and the pupils of both eyes (Fig. 1). Infants remained clothed during recording, so the glenohumeral landmarks were identified from the surface contour of the shoulder through

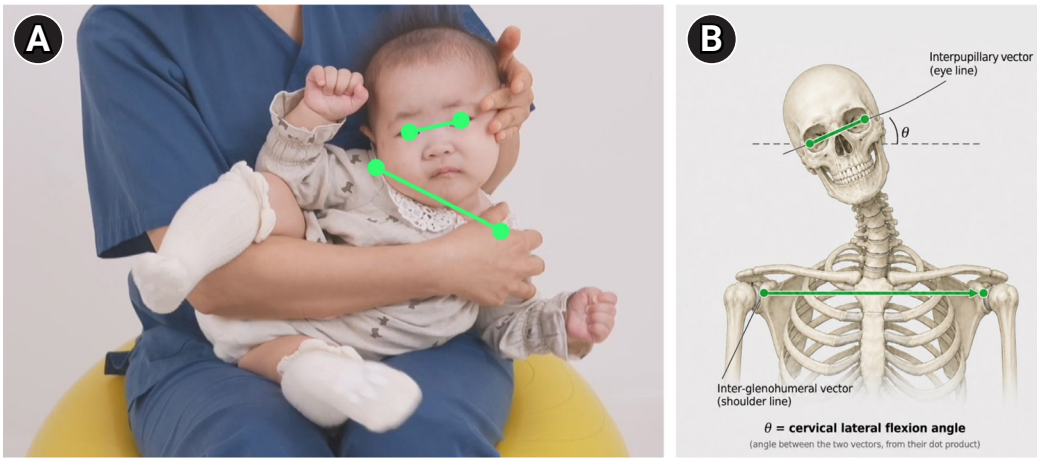

Fig. 1. Measurement of the cervical lateral flexion angle. (A) Region of interest annotation in ImageJ (National Institutes of Health) at the pupils and the anterior glenohumeral joints. (B) Schematic of the calculation: the angle is computed from the dot product of the vector connecting the two pupils and the vector connecting the two anterior glenohumeral joints. Written informed consent for image publication was obtained from a parent of the patient.

clothing rather than by direct visualization. Where a shoulder landmark was obscured in the target frame—which occurred frequently, as the infant's arm or the therapist's supporting hand commonly covered a shoulder—adjacent frames from the same sequence were referenced to approximate the position of the glenohumeral joint. Frames handled in this way were not recorded separately, so their number cannot be reported and a sensitivity analysis excluding them could not be performed.

Stage 2 — ROI coordinate extraction

ROI coordinates were exported using Python (version 3.10.6) and the read-roi library (version 1.6.0). Extracted coordinates were overlaid on the lateral-flexion frames to confirm correct positioning, cross-verified against the original images, and any discrepancies were corrected manually to ensure data integrity.

Stage 3 — Cervical lateral flexion angle measurement

Vectors were defined between the left and right pupil coordinates and between the left and right glenohumeral coordinates using Python (version 3.10.6) and NumPy (version 1.24.3). The cervical lateral flexion angle was computed as $\theta=\arccos[(u \cdot v)/(\|u\| \|v\|)]$, where $u=(u1, u2)$ is the vector connecting the two pupil landmarks and $v=(v1, v2)$ the vector connecting the two glenohumeral landmarks in image coordinates; θ is bounded to the interval [0°, 180°], and because all observed values were below 90°, no additional handling of reflex angles was required. The direction of lateral flexion in each frame (left or right) was determined visually from the recordings at frame extraction. Landmark annotation was performed blinded to the infants' clinical information, including the side of the head tilt. Each frame was then classified as showing lateral flexion in the therapeutic or the opposite direction according to the side of the head tilt, which had been ascertained in advance from the medical records. Frame eligibility was assessed independently of the treating therapists by a rehabilitation medicine physician and a physical therapist, who reviewed the recordings and selected frames by consensus; both were blinded to the infants' age group and clinical characteristics. When multiple frames met the inclusion criteria for a single infant in a given direction and position, the frame yielding the maximum angle in that direction was selected. Inter-rater reliability of landmark annotation was assessed on 30 frames, annotated independently by two trained physical therapists working from the same frames; intraclass correlation coefficients (ICCs) were computed on the resulting angle values. Reliability of frame selection, intra-rater reliability, and test–retest reliability were not assessed.

Statistical analysis

Descriptive statistics are presented as mean±standard deviation for continuous variables and as frequency (percentage) for categorical variables. The primary outcome was the maximum angle attained in the therapeutic direction. Differences in this angle between the supported forward-facing and supine positions were examined with a paired t-test restricted to infants in whom both positions were successfully recorded, with the Wilcoxon signed-rank test as a sensitivity analysis. The angle attained in the op-

posite direction was compared with the therapeutic direction within infants by paired t-test; this comparison was regarded as exploratory. A further sensitivity analysis used the maximum angle attained in either direction, irrespective of the therapeutic direction. Differences across age groups (0–2, 3–5, 6–8, and 9–12 months) within each position were examined with one-way analysis of variance. Because the number of infants in the older age groups was small, these comparisons were regarded as exploratory and were not powered to exclude a clinically relevant age effect. Normality was assessed with the Shapiro–Wilk test and homogeneity of variance with Levene's test. An exploratory, descriptive analysis of the association between stretching angle and treatment duration used Pearson correlation, separately for each position; because only four of the 56 infants with a documented resolution date had an SCM mass, no multivariable modeling was performed. Inter-rater reliability was assessed by ICC using a two-way mixed-effects model with absolute agreement, average measures, pooled across both treatment positions rather than separately for each; ICC>0.75 was considered indicative of high reliability. Statistical analyses were performed in Python version 3.12.3 with the SciPy (version 1.17.1) and statsmodels (version 0.14.6) libraries. A two-sided p-value<0.05 was considered statistically significant.

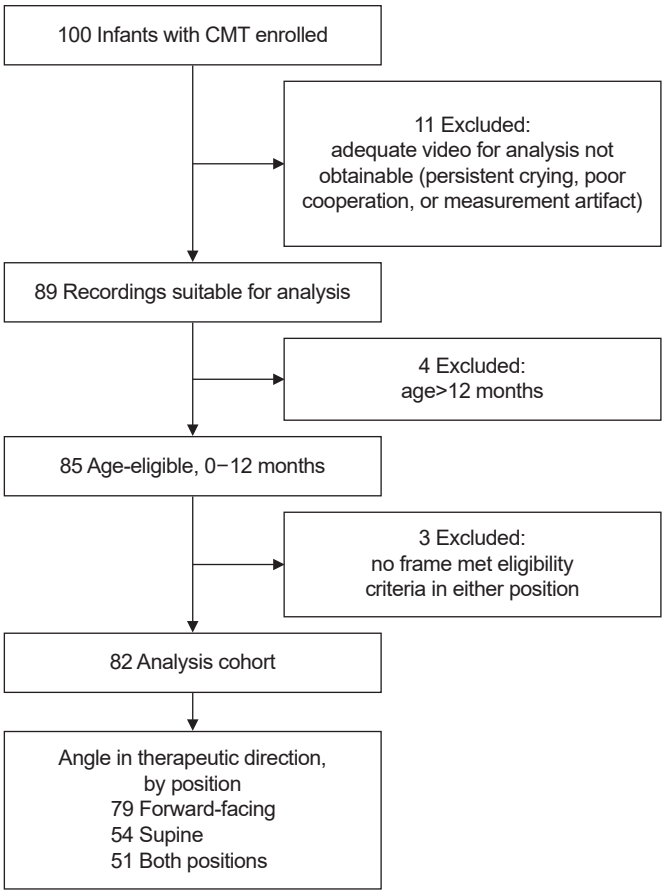

Fig. 2. Participant flow diagram. CMT, congenital muscular torticollis.

RESULTS

Participant characteristics

Of 100 infants enrolled, recordings suitable for analysis were obtained from 89; in the remaining 11, adequate video could not be obtained because of persistent crying, poor cooperation, or measurement artifact. Four were excluded because age at recording exceeded 12 months and three because no frame met the eligibility criteria in either position, leaving 82 infants for analysis (Fig. 2). A forward-facing measurement in the therapeutic direction was obtained in 79 infants and a supine measurement in 54; 51 infants had measurements in both positions, 28 in the forward-facing position only, and 3 in the supine position only. The cohort comprised 47 males (57.3%) and 35 females (42.7%), with a mean age at recording of 4.8±2.7 months. Left-sided torticollis was more prevalent (n=46, 56.1%) than right-sided (n=36, 43.9%). An SCM mass was identified in 9 infants (11.0%); the remaining 73 (89.0%) had no detectable SCM lesion. Clinical resolution was documented in 56 of the 82 infants (68.3%) during the study period, in 4 of 9 infants with an SCM mass and in 52 of 73 without; one infant with an SCM mass underwent surgical release, and 25 infants were lost to follow-up before documented resolution. Treatment duration among the 56 infants with a documented resolution date averaged 153±106 days (median, 133; range, 21–497). Duration is not reported separately by mass status, because it can be calculated only in infants who reached resolution and would therefore exclude the majority of infants with a mass; the end of follow-up was not recorded for infants who had not reached resolution, so a time-to-event analysis was not possible. No adverse events were reported by the treating therapists during or immediately after any session, or by caregivers at subsequent outpatient visits; adverse events were not ascertained under predefined criteria. Demographic and clinical characteristics are summarized in Table 1.

Inter-rater reliability

Inter-rater reliability for the cervical lateral flexion angle measurement was high, with an ICC of 0.98 (95% confidence inter-

val [CI], 0.97–0.99) calculated on 30 images pooled across both treatment positions. Reliability was not assessed separately by position.

Cervical lateral flexion stretching angle by treatment position

In the therapeutic direction, the mean cervical lateral flexion stretching angle was 41.9±8.0° (95% CI, 40.1–43.7; n=79) in the supported forward-facing position and 31.7±9.1° (95% CI, 29.2–34.2; n=54) in the supine position (Table 2; Supplementary Table S1). Among the 51 infants measured in both positions, the angle was greater in the supported forward-facing position, with a mean paired difference of 10.8° (95% CI, 8.1–13.5; t=8.00, p<0.001; Cohen's d=1.12); the Wilcoxon signed-rank test gave the same result (p<0.001). In an exploratory comparison, no difference was detected between the angle attained in the therapeutic direction and that attained in the opposite direction, in either the supported forward-facing position (opposite direction 40.9±9.7°, n=77; mean within-infant difference 0.9°; 95% CI, -1.1 to 2.9; p=0.39; n=77) or the supine position (opposite direction 30.2±9.7°, n=42; mean difference 0.8°; 95% CI, -3.7 to 5.3; p=0.73; n=36). The angles attained in each direction are summarized in Table 2. In a sensitivity analysis using the maximum angle attained in either direction, irrespective of the therapeutic direction, the angle was likewise greater in the supported forward-facing position (44.9±7.9°, n=79) than in the supine position (35.8±7.1°, n=60), with a mean paired difference of 9.1° (95% CI, 6.5–11.7; p<0.001; n=57). A supine measurement in the therapeutic direction could not be obtained in 28 of the 82 infants (34.1%); missingness was distributed across age groups (0–2 months, 4 of 14; 3–5 months, 12 of 39; 6–8 months, 8 of 22; 9–12 months, 4 of 7) rather than concentrated in older infants, and forward-facing angles did not differ between infants in whom a supine measurement was obtained (42.6±7.5°, n=51) and those in whom it was not (40.6±8.8°, n=28; p=0.33).

Table 1. Demographic and clinical characteristics of study participants (n=82)

Characteristic	Value
Sex	
Male	47 (57.3)
Female	35 (42.7)
Age at recording (mo)	4.8±2.7
0–2	14 (17.1)
3–5	39 (47.6)
6–8	22 (26.8)
9–12	7 (8.5)
Side of cervical tilt	
Left	46 (56.1)
Right	36 (43.9)
SCM mass	
Absent	73 (89.0)
Present	9 (11.0)
Treatment course	
Resolution documented	56 (68.3)
SCM mass absent	52 (71.2)
SCM mass present	4 (44.4)
Duration to resolution (day)[a]	153±106

Values are presented as number (%) or mean±standard deviation.
Percentages for resolution by mass status are within the respective group.
One infant with an SCM mass underwent surgical release, and 25 infants were lost to follow-up before documented resolution.
SCM, sternocleidomastoid.
[a]Calculated in the 56 infants with a documented resolution date.

Cervical lateral flexion stretching angle by age

Within each position, no statistically significant difference in the cervical lateral flexion stretching angle was detected across age groups (supported forward-facing, F[3,75]=0.30, p=0.83; supine, F[3,50]=0.51, p=0.68) (Table 3), although the small numbers in the older age groups mean these comparisons were underpowered. Distributional and variance assumptions were satisfied (Shapiro–Wilk and Levene tests, all p≥0.05).

In an exploratory, descriptive analysis among the 56 infants with a documented date of resolution, neither the forward-facing angle (r=-0.07, p=0.62, n=53) nor the supine angle (r=-0.04, p=0.79, n=42) was associated with treatment duration. Because

Table 2. Cervical lateral flexion stretching angle by direction of stretching

Position	Direction	n	Angle (°)	Within-infant difference (95% CI)	p-value
Supported forward-facing	Therapeutic	79	41.9±8.0	0.9 (-1.1 to 2.9); n=77	0.39
	Opposite	77	40.9±9.7		
Supine	Therapeutic	54	31.7±9.1	0.8 (-3.7 to 5.3); n=36	0.73
	Opposite	42	30.2±9.7		

Values are presented as mean±standard deviation.
The therapeutic direction is lateral flexion away from the side of the head tilt, which elongates the affected sternocleidomastoid.
Within-infant differences and p-values are from paired t-tests restricted to infants measured in both directions.
CI, confidence interval.

Table 3. Differences in cervical lateral flexion stretching angle in the therapeutic direction across age groups, by treatment position

Position	Age (mo)	n	Angle (°)	df	F	p-value
Supported forward-facing position	0–2	12	43.5±8.9	3, 75	0.30	0.83
	3–5	38	41.5±8.4			
	6–8	22	41.2±7.4			
	9–12	7	43.2±7.3			
Supine position	0–2	10	33.9±11.6	3, 50	0.51	0.68
	3–5	27	30.2±8.2			
	6–8	14	32.9±9.7			
	9–12	3	32.0±4.2			

Values are presented as mean±standard deviation.
p-values are from one-way analysis of variance.
df, degrees of freedom; F, F statistic.

only four of these infants had an SCM mass, no multivariable analysis was performed, and the correlations were not corrected for multiple comparisons.

DISCUSSION

This study quantified the cervical lateral flexion angles attained and tolerated during routine physical therapy in infants aged 12 months or younger with CMT, and compared them across ages and between two therapeutic positions. Two principal findings emerged. First, no statistically significant difference in stretching angle was detected across age groups. Second, a greater lateral flexion angle was attained in the supported forward-facing position than in the supine position.

In the therapeutic direction the mean stretching angle was 41.9±8.0° in the supported forward-facing position and 31.7±9.1° in the supine position, with a mean paired difference of 10.8°, corresponding to a large effect size (Cohen's d=1.12). These results indicate that a greater lateral flexion angle was attained in the supported forward-facing position. Whether this translates into a greater therapeutic effect cannot be determined from the present data. No threshold of lateral flexion has been shown to change outcome in CMT, so the difference is best interpreted as a consistent positional tendency rather than a clinically decisive margin. In the exploratory comparison of directions, no difference was detected between the therapeutic and the opposite direction, although the supine interval was wide and a modest difference in that position cannot be excluded. This may suggest that the angle reached during a session is bounded by the infant's tolerance rather than by the restriction itself, but both the direction classification and this interpretation are exploratory. The supine position was particularly difficult once infants were able to roll, whereas cooperation was better in the supported forward-facing position, where the righting reaction can also be used to engage opposing muscles in infants older than 3 months [1].

Values reported in previous studies are not directly comparable with ours, because they were obtained under different conditions. The 2024 clinical practice guideline gives a normative passive cervical lateral flexion of approximately 70° in healthy infants, measured at examination with an arthrodial protractor and limited chiefly by contact between the head and the shoulder [1]. Öhman and Beckung [18] measured passive cervical range of motion in 38 healthy infants examined at 2, 4, 6, and 10 months of age and reported a mean lateral flexion of 70°, regarding 65° or more as good. Their measurement, however, required three examiners—one to hold the child, one to position the goniometer, and one to perform the maneuver—using a large analog protractor, which makes high-precision measurement difficult, and was performed during assessment rather than during a therapeutic session. Their values may therefore be higher than those obtained in a real-world treatment context. By contrast, Taylor and Norton [19], who measured during treatment, considered an angle >40° to be excellent, which aligns closely with our findings. A strength of the present study is that therapeutic angles were measured in actual patients during treatment rather than during assessment. The differences across studies suggest that angles obtained during brief assessment may differ meaningfully from those attained and tolerated during a sustained therapeutic stretch, and the angles reported here should therefore be compared only with values obtained during treatment. Further research with similarly objective measurement is warranted.

Clinical resolution was documented in 56 of the 82 infants (68.3%), after a mean treatment duration of 153±106 days (median, 133; range, 21–497). This is broadly consistent with the

duration reported by Emery (3 to 6 months) [10], although the comparison is limited because duration could be determined only in infants who reached resolution. No adverse events were reported during treatment, immediately afterwards, or at subsequent outpatient visits. Because adverse events were not ascertained under predefined criteria and the study was not designed as a safety trial, this supports the feasibility of the angles observed but does not establish them as safe thresholds.

Measuring from video frames with landmark annotation, rather than by goniometry during handling, allowed the angle to be determined after the session without interrupting treatment, and inter-rater agreement for annotation could be documented. Snapshots were extracted at one-second intervals with Python and the opencv-contrib-python library, landmarks were annotated in ImageJ, coordinates were exported with the read-roi library, and angles were computed from the resulting vectors. Whether this approach is more accurate than goniometry was not tested, and the sources of error inherent to two-dimensional video are set out below.

Several limitations must be acknowledged, and they can be grouped into four themes. First, generalizability is limited. The study was conducted at a single tertiary center without a healthy comparison group; the non-SCM-mass group combined postural and muscular CMT, because PROM could not be assessed reliably in every infant; plagiocephaly, SCM thickness or mass size, and a graded measure of head-tilt severity were not recorded systematically; and the sample was small in the older age groups, so the age comparisons were underpowered and should not be interpreted as evidence of equivalence. Second, the video-based measurement carries sources of error that could not be quantified. The angle was derived from two-dimensional frontal video and is sensitive to cervical rotation, trunk rotation, asymmetrical shoulder elevation, and camera position, with frame retention judged visually and no quantitative threshold for residual rotation; the glenohumeral landmarks were inferred from surface contour in clothed infants and, when a shoulder was obscured by the therapist's hand, approximated from adjacent frames that were not flagged; selecting the frame with the maximum angle may have favored the upper end of the range achieved in each session; the direction of lateral flexion was determined visually and was not independently re-adjudicated; and the reported reliability reflects landmark annotation only, pooled across positions, with frame selection performed once by consensus, intra-rater and test–retest reliability not assessed, and therapist identity not recorded, so neither the reproducibility of the full measurement process nor therapist-level variability is established. Third, a supine measurement in the therapeutic direction was unavailable in 28 of 82 infants; although missingness was distributed across age groups and forward-facing angles did not differ by supine availability, unmeasured differences cannot be excluded, and the supine comparison between directions rested on 36 infants. Finally, the observed angles were not linked to treatment effectiveness or safety: the exploratory duration analysis was limited to the infants with a documented resolution date, the end of follow-up was not recorded for those lost to follow-up, adverse events were not ascertained under predefined criteria, and cervical rotation, another key component of CMT intervention, was not analyzed. Future work should link therapeutic angles to clinical outcomes and address rotation as well as lateral flexion.

Conclusion

This video-based observational study quantified cervical lateral flexion angles achieved during routine physical therapy in infants aged 12 months or younger with CMT. In the direction that elongates the affected SCM, a greater angle was attained in the supported forward-facing position (41.9±8.0°) than in the supine position (31.7±9.1°); in exploratory analyses, no statistically significant differences were detected between age groups or between the therapeutic and the opposite stretching direction. These findings provide descriptive values for the angles attained and tolerated during therapist-applied stretching in this cohort; their relationship with clinical outcomes and with safety requires further study.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (No. RS-2023-00243053).

AUTHOR CONTRIBUTION

Conceptualization: Kim J. Methodology: Kim J, Kim B, Kang JW. Software: Kim B. Investigation: Kang JW, Kim J. Data cura-

tion: Kim B. Formal analysis: Kim B, Kim J. Validation: Kim J. Visualization: Kim B, Kang JW. Writing – original draft: Kim J, Kim B. Writing – review and editing: Kim B, Kang JW, Kim J. Funding acquisition: Kim J. Project administration: Kim J, Jang DH. Supervision: Kim J. Approval of final manuscript: all authors.

DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during the current study are not publicly available due to participant confidentiality and institutional data-sharing policies, but are available from the corresponding author on reasonable request and with approval of the Institutional Review Board of Incheon St. Mary's Hospital.

SUPPLEMENTARY MATERIALS

Supplementary materials can be found via https://doi.org/10.5535/arm.260098.

ORCID

Bumcheol Kim, https://orcid.org/0009-0006-1784-5706
Jeon-Woong Kang, https://orcid.org/0009-0003-0441-2141
Dae-Hyun Jang, https://orcid.org/0000-0001-8293-084X
Jaewon Kim, https://orcid.org/0000-0002-3921-0190

REFERENCES

1. Sargent B, Coulter C, Cannoy J, Kaplan SL. Physical therapy management of congenital muscular torticollis: a 2024 evidence-based clinical practice guideline from the American Physical Therapy Association Academy of Pediatric Physical Therapy. Pediatr Phys Ther 2024;36:370-421.
2. Sargent B, Kaplan SL, Coulter C, Baker C. Congenital muscular torticollis: bridging the gap between research and clinical practice. Pediatrics 2019;144:e20190582.
3. Porter SB, Blount BW. Pseudotumor of infancy and congenital muscular torticollis. Am Fam Physician 1995;52:1731-6.
4. Kaplan SL, Coulter C, Fetters L. Physical therapy management of congenital muscular torticollis: an evidence-based clinical practice guideline: from the Section on Pediatrics of the American Physical Therapy Association. Pediatr Phys Ther 2013;25:348-94.
5. Hulbert KF. Congenital torticollis. J Bone Joint Surg Br 1950;32-B:50-9.
6. Tatli B, Aydinli N, Caliskan M, Ozmen M, Bilir F, Acar G. Congenital muscular torticollis: evaluation and classification. Pediatr Neurol 2006;34:41-4.
7. Cheng JC, Tang SP, Chen TM, Wong MW, Wong EM. The clinical presentation and outcome of treatment of congenital muscular torticollis in infants--a study of 1,086 cases. J Pediatr Surg 2000;35:1091-6.
8. Choi KE, Lee HC, Youn SY, Chun JM, Shin SM, Han BH, et al. Clinical characteristics and courses of congenital muscular torticollis. Korean J Pediatr 2009;52:1273-8.
9. Cheng JC, Wong MW, Tang SP, Chen TM, Shum SL, Wong EM. Clinical determinants of the outcome of manual stretching in the treatment of congenital muscular torticollis in infants. A prospective study of eight hundred and twenty-one cases. J Bone Joint Surg Am 2001;83:679-87.
10. Emery C. The determinants of treatment duration for congenital muscular torticollis. Phys Ther 1994;74:921-9.
11. Seo SJ, Kim JH, Joh YH, Park DH, Lee IJ, Lim H, et al. Change of facial asymmetry in patients with congenital muscular torticollis after surgical release. J Craniofac Surg 2016;27:64-9.
12. Hung NN, Anh LT. A comparison of outcome of age at time surgery between younger and older than 8 years old in children with congenital muscular torticollis. Open Access Libr J 2017;4:e4105.
13. Hussein MA, Yun IS, Park H, Kim YO. Cervical spine deformity in long-standing, untreated congenital muscular torticollis. J Craniofac Surg 2017;28:46-50.
14. Hussein MA, Yun IS, Lee DW, Park H, Kim YO. Cervical spine dysmorphism in congenital muscular torticollis. J Craniofac Surg 2018;29:925-9.
15. van Vlimmeren LA, Helders PJ, van Adrichem LN, Engelbert RH. Torticollis and plagiocephaly in infancy: therapeutic strategies. Pediatr Rehabil 2006;9:40-6.
16. Kaplan SL, Coulter C, Sargent B. Physical therapy management of congenital muscular torticollis: a 2018 evidence-based clinical practice guideline from the APTA Academy of Pediatric Physical Therapy. Pediatr Phys Ther 2018;30:240-90.
17. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453-7.
18. Öhman AM, Beckung ER. Reference values for range of motion and muscle function of the neck in infants. Pediatr Phys Ther 2008;20:53-8.
19. Taylor JL, Norton ES. Developmental muscular torticollis: outcomes in young children treated by physical therapy. Pediatr Phys Ther 1997;9:173-8.

Original Article

Ann Rehabil Med 2026;50(5):363-375
eISSN: 2234-0653
https://doi.org/10.5535/arm.260003

arm
Annals of Rehabilitation Medicine

Effects of Stepping Training With Reduced Visual Cues in a Virtual Environment in Chronic Stroke Patients: A Randomized Controlled Trial

Pei-Yun Lee, PT, PhD[1,2], Meng-Ru Tsai, MS[3], Hsien-Chang Wang, PhD[4], Chih-Hung Chen, MD[5], Hui-Yu Tseng, MD, MS[6], Sang-I Lin, PT, PhD[7]

[1]*Department of Physical Therapy, College of Medicine, National Cheng Kung University, Tainan, Taiwan*
[2]*Physical Therapy Center, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan*
[3]*Division of Physical Therapy, Department of Physical Medicine and Rehabilitation, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan*
[4]*Department of Digital Contents Design, College of Informatics and Design, Chang Jung Christian University, Tainan, Taiwan*
[5]*Department of Neurology, College of Medicine, National Cheng Kung University, Tainan, Taiwan*
[6]*Department of Rehabilitation, Tainan Hospital, Ministry of Health and Welfare, Tainan, Taiwan*
[7]*Graduate Institute of Long-term Care, College of Health and Welfare, MacKay Medical University, New Taipei, Taiwan*

Objective: To take a step and appropriately place the foot in the environment is an ability that may be impaired after stroke. This study aimed to determine if a virtual reality (VR) stepping training with gradual removal of on-line visual cues about the feet or target would be superior to the same training with more visual cues for persons with chronic stroke.

Methods: Participants were randomly assigned to either a control group (CON, n=19) or an experimental group (EXP, n=19) to receive stepping training within a fully immersive VR environment. During practice, images of the feet and the target were presented to CON, while these images were gradually removed for EXP. The outcome measures included walking speed and gait symmetry ratio, error distance and variation during targeted stepping, and toe clearance and landing distance during obstacle crossing stepping.

Results: For walking, there were no significant between- or within-group differences. For targeted stepping, EXP significantly reduced the unaffected limb error distance (p=0.013), but not CON (p=0.238). For obstacle crossing stepping, when the unaffected limb led, EXP significantly reduced the landing distance in both the affected and unaffected limb (p=0.002), but not CON (p=0.595).

Conclusion: Gradual removal of visual cues during stepping training in a VR environment resulted in better accuracy in the unaffected limb targeted stepping and in both limbs during the unaffected limb leading obstacle crossing. This type of training may be considered for inclusion in stroke rehabilitation, with careful consideration of the amount of visual cues provided.

Keywords: Stroke rehabilitation, Virtual reality, Gait, Feedback, Locomotion

Received: December 16, 2025
Revised: February 5, 2026
Accepted: August 10, 2026

Correspondence:
Sang-I Lin
Graduate Institute of Long-term Care, College of Health and Welfare, MacKay Medical University, No. 46, Section 3, Zhongzheng Rd, Putou Village, Sanzhi District, New Taipei 252, Taiwan.
Tel: +886-2-26360303
Fax: +886-2-26360303
E-mail: lin31@mail.ncku.edu.tw

© 2026 by Korean Academy of Rehabilitation Medicine

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Stroke is a highly prevalent health problem and the leading cause of disability in adults globally [1]. Although disability after stroke is often a multifactorial problem, limited walking ability is one of its primary underlying causes. Approximately 20% of stroke survivors never regain their walking ability [2]. Among ambulatory stroke survivors, gait impairments, including slow walking speeds, asymmetrical gait and poor gait adaptability, are common [3,4]. Gait adaptability refers to the ability to modify the stepping patterns to meet the task goals and environmental demands [5]. This ability is needed in many different daily activities, such as crossing obstacles/potholes, stepping onto curbs/stairs/targets, walking in cluttered environment or on uneven terrain. The success rate of obstacle clearing has been found to be significantly lower after stroke [4,6]. Furthermore, atypical patterns of foot placement, trajectory and lower limb muscle activations have also been noted in persons with stroke [3,7]. Atypical changes in the lower limb muscle activations and stepping patterns, which could lead to postural instability, have also been found in people with stroke while walking on uneven terrains or turning [8].

The sub-tasks involved in gait adaptability primarily include taking a step to properly place the foot in the environment and, at the same time, maintaining forward propulsion force and balance [9]. Foot placement (or stepping) is important for walking because it defines the individual's position in the environment and the base of support, and governs the biomechanical relationships between the individual and the environment [10]. The accuracy of foot placement during treadmill walking has been reported to be affected by stroke [8]. When taking a step in standing while well supported by a balance frame, both the ipsilateral and contralateral foot of people with stroke showed significantly greater stepping error, compared to the controls [11]. These stroke-related impairments in foot placement may affect gait adaptability and community walking. To improve the ability to place the foot properly, it is intuitive to plan a training regimen that includes repeated stepping onto a target. However, information about the effects of targeted stepping training on gait adaptability remains scarce for people with stroke.

In the training of walking ability after stroke, feedback is often used. Feedback obtained during movement execution (i.e., on-line or concurrent feedback, knowledge of performance) and after the completion of the motor task (i.e., post-task or terminal feedback, knowledge of results) have both been found to improve the learning of the upper limb movement control [12]. However, it has been noted that on-line feedback may hinder acquisition, retention and transfer of motor learning because of a greater reliance on the feedback in healthy adults [13,14]. Removing on-line visual feedback during motor learning has been shown to lead to greater reliance on feedforward control of arm movement in healthy adults [15]. It has also been proposed that more feedback does not necessarily lead to better learning in healthy adults [16,17]. These notions make it important to understand how the availability of on-line visual cues during training would affect motor learning in persons with stroke.

Manipulating visual cues as a balance or gait training strategy has long been used in stroke rehabilitation. For people with stroke, augmented visual cues on the ground or treadmill could help to improve balance, mobility, and gait speed [18,19], while deprivation of visual cues during training also was found to improve balance and cadence [20]. These findings suggested that more visual cues during training may not always lead to better outcomes. The development of virtual reality (VR) technology allows for more extensive use of on-line visual cues in stroke rehabilitation. A VR environment uses computer software to simulate a real-world environment that the user can interact with using a human-machine interface. When used in rehabilitation, VR provides augmented on-line feedback, allows for repetitive, task-oriented practice in an enriched, controlled, and safe environment, and facilitates interactions with the environment. For stroke rehabilitation, there is moderate to strong evidence showing that training with VR improves the recovery of gait and balance in stroke patients compared to training without VR [21-23]. It is unclear, however, how on-line visual cues should be provided to lead to better recovery in a VR environment.

Gait adaptability not only is needed for independent walking at home and in the community, it has also been found to be related to risk of falls and limitation in social participation in older adults or persons with stroke [24,25]. Previous randomized controlled studies showed significantly better obstacle clearance ability, composite functional gait ability, or self-perceived community walking ability with task-oriented gait adaptability training [26,27]. None of the above studies focused on targeted stepping or manipulated the availability of the on-line visual cues in their interventions. The purpose of this study was to determine if a VR targeted stepping training with gradual removal of on-line visual cues about the foot or target would be superior to the same training with more on-line visual cues for people with chronic stroke. It was hypothesized that the former would lead to better stepping performance.

METHODS

This study was a single-blind randomized controlled trial. Ethical approval was obtained from the Institutional Review Board of the National Cheng Kung University Hospital (protocol number: A-ER-108-183). The study was registered at ClinicalTrials.gov Protocol Registration and Results System (trial number NCT04436237). Written consents were obtained from all participants before participation.

Participants and setting

People with stroke were recruited from the departments of rehabilitation and neurology in local hospitals. Based on a previous study using a VR environment with treadmill training and real-world video recordings to show improvement in walking speed in chronic stroke patients [28], the F test with an effect size of 0.27, an alpha of 0.05, and a power of 0.8 was used to estimate the sample size in G*Power 3.1.9.4. The total sample size was 30 subjects. In consideration of a possible high dropout rate (30%) due to coronavirus disease 2019, 39 subjects would be needed. The inclusion criteria were (1) first-time unilateral stroke with onset time over six months; (2) ability to walk independently over 10 meters with or without an assistive device; (3) ability to follow the experimental procedures; and (4) ability to see a 2-centimeter circle on the floor in standing. The exclusion criteria were having (1) nearly normal gait pattern by visual inspection; (2) other neurological or musculoskeletal disease affecting walking ability; (3) hemineglect; and (4) intolerance of using VR device, such as headache, blurred vision, or vomiting.

All the assessments and training were conducted in a motion lab in a university. After a baseline assessment, the participants were randomly allocated into control group (CON, n=20) or experimental group (EXP, n=19) to receive foot targeted movement training in sitting and standing with full or gradual removal of on-line visual cues, respectively. Randomization order was created by a computer randomization program and kept by one of the experimenters. The allocation sequence was given to another experimenter in a sealed envelope to be opened after the pre-test. The participants were not aware of the training program received by other participants. The assessors were not blinded to group allocation.

Clinical tests for sample characterization

All the participants went through a series of clinical cognitive and sensorimotor function tests for sample characterization. The Mini Mental State Examination (MMSE) was used to assess the cognitive function. The Fugl-Meyer lower extremity motor scale (FMLE-motor) was used to assess the motor function of the lower limb. Plantar cutaneous sensitivity of bilateral first metatarsal heads (big toe) at the plantar side was assessed using the Semmes-Weinstein monofilaments (Semmes-Weinstein Aesthesiometer). Grip strength of the unaffected side was measured by a JAMAR® Hand Dynamometer (Sammons Preston).

Walking and stepping assessments

The assessments included preferred walking, targeted stepping, and obstacle crossing stepping tests. The SIMI MOTION® (SIMI Reality Motion Systems GmbH) with a six-camera system was used for the assessment of the task performance. Four reflective markers were attached to bilateral 2nd metatarsal heads and heels. The participants practiced all the tasks twice prior to data collection.

Because stepping is a critical subtask of walking, it is reasonable to assume that better control of foot placement (stepping) might lead to a more symmetrical foot placement. Furthermore, it has been reported that greater spatiotemporal asymmetry was related to slower gait speed in people with stroke [29]. Thus, walking performance, including gait speed and symmetry, was measured to determine whether stepping training could result in an improvement in gait symmetry and speed. For the walking test, the participant was instructed to walk barefoot on an eight-meter walkway with preferred speed once. After the walking test, the reflective markers that were originally placed on the 2nd metatarsal heads were moved to the big toes for the stepping tests.

The targeted stepping test was to take two consecutive steps from a quiet standing position. There were two targets, each serving as the target for one step (Fig. 1A). Target-1 (100 cm wide and 0.9 cm long coloured rectangle adhesive tape) was placed one-step length in front of the big toe of the leading limb. Target-2 was placed parallel to and 20 cm away from target-1. The participant was asked to look at the target for 3 seconds to become visually aware of the target location. The participant would then be asked to look straight ahead and take one step to place the big toe as close to the posterior edge of target-1 as possible, followed by stepping with the other leg, aiming at target-2. Thus, the visual cues of the target and the feet were not available during the targeted stepping test.

The obstacle crossing stepping test was to take a step to cross over an 8 cm (H)×5 cm (L)×90 cm (W) obstacle, and then re-

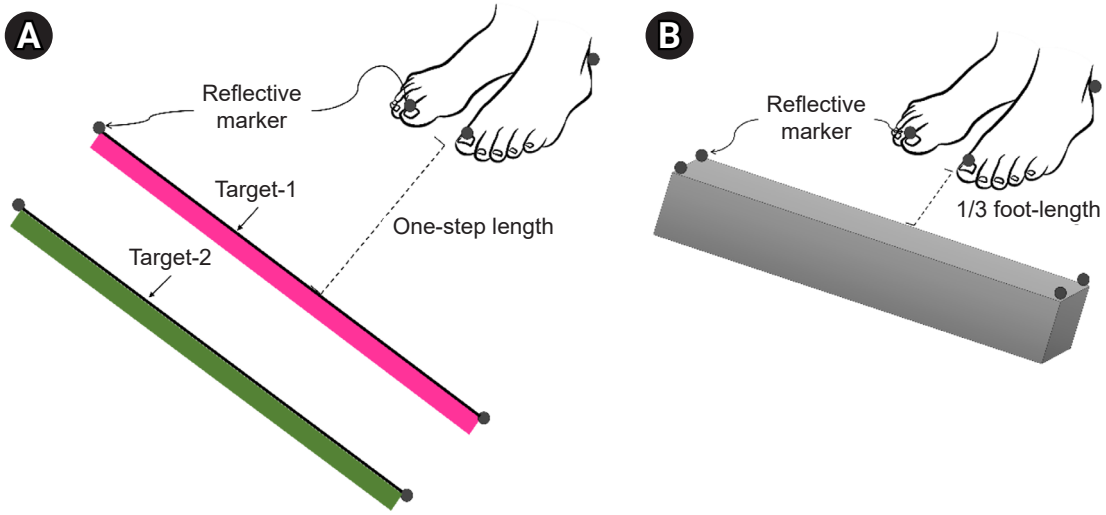

Fig. 1. Experimental setup for targeted stepping (A) and obstacle crossing stepping (B) tests. For the targeted stepping test, the instruction was to take two consecutive steps, each to place the big toe at the posterior edge (black line) of the target. For the obstacle crossing stepping test, the instruction was to cross the obstacle to stand neutrally with the heels as close to the posterior edge of the obstacle as possible.

sume the neutral standing posture. The test was conducted with each leg serving as the leading limb once at random order. The participant was allowed to observe the obstacle before assuming a neutral standing posture (standing upright and looking straight ahead) with the big toes one-third of the foot-length behind the obstacle (Fig. 1B). The instruction was to continue looking straight ahead, and then step over the obstacle to resume the neutral standing posture with both the leading and trailing heels landing as close to the posterior edge of the obstacle as possible. Any leg-obstacle contact would be considered a fail and the test would be repeated.

Interventions

The training was targeted stepping practice in sitting and standing in a total-immersive virtual environment, 30 minutes/session, 2 sessions/week for two weeks for both groups. Each session consisted of eight 20-trial blocks and thus a total of 640 trials were practiced. The HTC VIVE™ VR system was used to construct the virtual environment. The participant wore a VIVE head-mounted display to watch the virtual scenes and a tracker was placed on each foot to monitor the position of the foot which would be inserted into the virtual environment. In the virtual environment, the participant would be situated on the meadow in a park and could see the feet and a target (a coin) (Fig. 2). The training session consisted of two parts: (1) sagittal plane training where the participant was required to touch the target at a specific height with the big toe; and (2) horizontal plane training where the participant was required to step onto a coin on the meadow ground with the big toe. For a successful trial, i.e. the big toe 'touched' the target, the color of the target would change. For a failed trial, the image of the big toe, the target and a straight line connecting these two would appear.

The training was conducted in three postures sequentially, sitting, quiet standing, and taking a step in standing (Fig. 3). In the taking-a-step training, the participant would take a step forward, and then use the trailing foot to reach the target. Both limbs were trained. For the EXP, the training was divided into two levels with the on-line visual cues about the feet and target gradually removed (Fig. 2). The CON received the same targeted movement training, except that the images of the target and the feet remained visible for the entire training session. The same visual feedback about the task performance was provided as the EXP.

Outcome measures

For preferred walking, the outcome variables were gait speed and symmetry ratios (affected side divided by the unaffected side) [30] for step length, single limb support time, and swing time. For targeted stepping tests, the outcome variables were error distance and variation. Error distance was defined as the distance on the horizontal plane between the end position of the big toe and the target to represent the accuracy of foot placement. Variation was

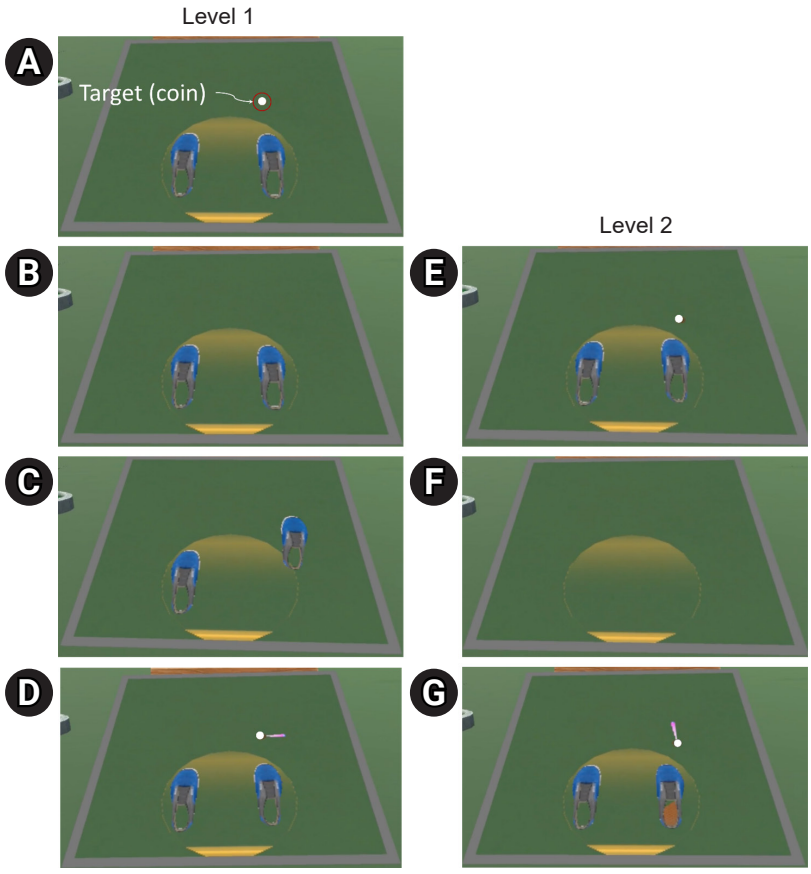

Fig. 2. The set-up of the total immersive virtual reality environment during training for the experimental group. There were two levels in the training. In level 1, the target and feet would be visible for 3 seconds (A) before the image of the target disappeared (B). The participant would then place the big toe onto the target (C). The foot then returned to the original position (D) and the feedback about the task performance would be provided. In level 2, the target and feet would be visible for 3 seconds (E) and the images of both the feet and target would disappear (F). The participant would then place the big toe onto the target. The foot then returned to the original position and the feedback about the task performance would be provided (G). In this example, the feedback was a line between the target and the big toe to indicate a failed trial (G).

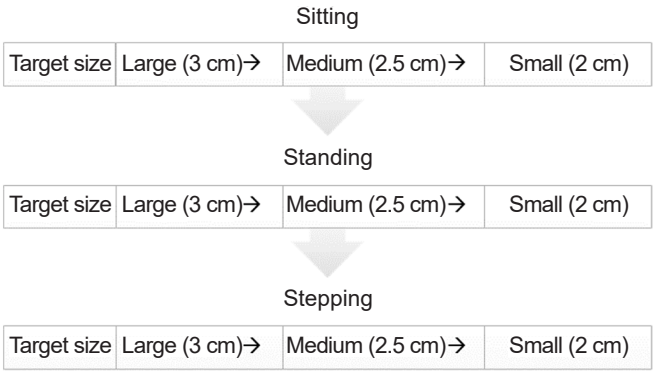

Fig. 3. Progression of the targeted stepping training. In each training posture, if the target was reached 10 times successfully out of 20 trials, the size of the target would decrease. When the smallest-sized target was completed, the next training posture would be adopted.

the mean distance between each data point of the big toe and a straight line formed by the starting and end position of the big toe. This variable was used to represent the extent to which the foot motion deviated from a straight line. For obstacle crossing stepping test, the outcome variables were toe clearance and landing distance. Toe clearance was defined as the highest point of the big toe during crossing. This variable was used as an indication of the obstacle crossing ability. Landing distance was defined as the distance between the heel at landing and the posterior edge of the obstacle. Because the instruction was to land the foot as close to the posterior edge as possible, thus a smaller landing distance would be indicative of better stepping accuracy. MATLAB software (version R2013a; MathWorks, Inc) was used to compute the data collected from SIMI MOTION® analysis system.

Statistical analysis

For baseline sample characteristics, independent t, Mann–Whitney U, and chi square tests were used for between group comparisons for continuous, ordinal and categorical variables, respectively. To compare the training effects between the two groups, group×time repeated multivariate measures analysis of variance (RMANOVA) was conducted for preferred walking symmetry ratio (step length, single limb support time, swing time). For preferred walking speed, univariate RMANOVA was conducted. For targeted stepping, group×time RMANOVAs (error distance and variation) were performed for the affected and unaffected sides separately. For obstacle crossing stepping, the data from the trailing and leading limb of each trial were analyzed separately. The data for the affected and unaffected sides were also analyzed separately. Thus, a total of four group×time RMANOVAs (toe clearance and landing distance) were conducted for obstacle crossing stepping. The Fisher's least significant difference test was used for post-hoc pair-wise analysis, and the p-values were adjusted for multiple comparisons. Associations in changes (post-minus-pre) between gait speed and symmetry index (for the walking test), as well as between affected and unaffected limbs for targeted stepping and obstacle crossing were analyzed using Pearson correlation. The significance level was set at 0.05.

RESULTS

During the period between July 2019 and July 2021, 170 persons with chronic stroke were contacted. Among them, 39 fulfilled the eligibility screening and consented to participate. One participant in the CON withdrew from the experiment after one training session due to time limitation. Thus a total of 19 participants in each group completed the pre- and post-tests and interventions. None of the participants reported any adverse effects.

Basic characteristics

The two groups did not differ significantly in age, sex, stroke onset duration and side, MMSE, FMLE-motor, affected or unaffected plantar sensation, or unaffected grip strength at the baseline assessment (Table 1).

Walking and stepping performance

Preferred walking

The results for walking speed and the symmetry ratio are presented in Table 2. For walking speed, univariate RMANOVA showed nonsignificant time (F=0.134, p=0.717) or group (F=0.083, p=0.775) main effect or interaction (F=0.866, p=0.359). For symmetry ratio, multivariate RMANOVA also showed nonsignificant time (F=0.554, p=0.649) or group (F=0.826, p=0.49) main effect or interaction (F=0.544, p=0.656). Changes in walking speed after the intervention were not significantly correlated with changes in symmetry ratio in either group (r=-0.031 to -0.250, all p>0.05).

Targeted stepping

Fig. 4 shows representative trajectories of the targeted stepping test for the affected (Fig. 4A) and unaffected (Fig. 4B) foot, both before and after the intervention of a participant in the EXP. Fig. 5 shows the trajectories of the affected (Fig. 5A) and unaffected (Fig. 5B) foot motion during targeted stepping of individual subject, as well as averages, in EXP before (light blue) and after (light red) the intervention. The error distance and variation of

Table 1. Basic characteristics

Characteristic	Control (n=19)	Experimental (n=19)	p-value
Age (yr)	57.2±10.0	53.0±10.3	0.205
Sex, female	5	4	0.5
Body mass index (kg/m^2)	25.7±4.4	27.5±9.2	0.448
Stroke			
Side (right/left)	12/7	10/9	0.372
Duration (mo)	33.4±33.7	55.6±22.9	0.325
Mini Mental State Examination	28.1±2.7	27.3±2.7	0.394
Fugl-Meyer lower extremity motor scale	19.5±4.9	18.2±6.3	0.506
Grip strength (kg)[a]	20.1±9.2	22.0±9.4	0.534
Plantar sensitivity big toe (dB)			
Affected (contralateral) side	4.12±0.69	4.31±0.88	0.563
Unaffected (ipsilateral) side	3.68±0.47	3.85±0.54	0.499

Values are presented as mean±standard deviation or number only.
[a]Unaffected (ipsilateral) side. Affected side not tested because most participants were unable to grip the dynamometer.

Table 2. Walking, stepping and targeted movement performance before and after training

	Control		Experimental	
	Pre-test	Post-test	Pre-test	Post-test
Self-paced walking				
Speed (cm/s)	43.5±21.3	42.6±18.3	39.9±24.0	41.9±22.5
Symmetry ratio				
Step length	0.268±0.224	0.606±1.107	0.497±0.514	0.454±0.945
Single limb support time	0.209±0.170	0.186±0.127	0.278±0.191	0.252±0.175
Swing time	0.365±0.560	0.259±0.238	0.494±0.476	0.414±0.508
Targeted stepping (cm)				
Affected limb				
Error distance	2.93±3.32	4.28±3.53	2.84±2.25	2.75±2.03
Variation	0.116±0.035	0.109±0.045	0.119±0.038	0.111±0.029
Unaffected limb				
Error distance[a)]	3.73±2.52	5.05±3.64	4.61±3.19	1.85±1.96
Variation	0.077±0.043	0.098±0.029	0.101±0.046	0.109±0.041
Obstacle crossing stepping (cm)				
Affected limb leading				
Leading limb				
Toe clearance	9.46±3.64	11.15±4.91	9.90±4.54	10.05±4.23
Landing distance	9.27±4.24	9.21±4.28	10.04±4.63	6.45±2.91
Trailing limb				
Toe clearance	12.78±4.57	12.54±2.98	11.32±3.96	12.57±4.64
Landing distance	12.12±5.27	13.65±7.10	12.97±6.98	10.03±7.02
Unaffected limb leading				
Leading limb				
Toe clearance	15.01±4.19	15.63±3.54	15.28±4.95	15.44±5.00
Landing distance[a)]	7.61±3.64	8.31±4.24	9.23±3.21	6.11±3.13
Trailing limb				
Toe clearance	9.96±5.94	10.13±3.85	8.11±4.05	8.60±4.91
Landing distance[a)]	12.32±3.74	14.02±7.79	13.97±3.86	8.21±4.90

Values are presented as mean±standard deviation.
[a)]Significantly different between pre- and post-test for the experimental group.

targeted stepping performance are presented in Table 2. For the affected side, multivariate RMANOVA showed nonsignificant time (F=1.529, p=0.232) or group (F=0.664, p=0.522) main effect or interaction (F=0.862, p=0.432). For the unaffected side, multivariate RMANOVA showed significant group×time interactions (F=3.876, p=0.031). Univariate analysis showed significant differences in error distance (F=7.691, p=0.009) but not variation (F=3.219, p=0.082). Post hoc analysis showed that for EXP, error distance significantly reduced in post-test (t=2.761, p=0.013), but not for CON (t=1.223, p=0.238). Changes in error distance or variation in the unaffected limb after the intervention were not significantly correlated with changes in the affected limb in either group (r=-0.340 to 0.373, all p>0.05).

Obstacle crossing stepping

Fig. 6 shows representative trajectories of the unaffected limb leading obstacle crossing stepping test for the affected (Fig. 6A) and unaffected (Fig. 6B) foot, both before and after the intervention of a participant in EXP. The toe clearance and landing distance of obstacle crossing stepping are presented in Table 2. When the unaffected leg was the leading limb, multivariate RMANOVA for the leading limb showed nonsignificant time (F=2.456, p=0.1) or group (F=0.06, p=0.941) main effect. Although the multivariate analysis showed a nonsignificant interaction (F=3.146, p=0.056), univariate analysis showed a significantly smaller landing distance after intervention (F=6.075, p=0.019). Post hoc analysis showed that for EXP, landing distance significantly reduced in post-test (t=3.575, p=0.002), but not for CON (t=0.542, p=0.595). In this condition, multivariate RMANOVA for the trailing limb (i.e., affected limb) showed significant interaction (F=4.653, p=0.016). Univariate analysis showed significant differences in landing distance (F=9.459, p=0.004) but not toe clearance (F=0.02, p=0.88). Post hoc analysis showed that for EXP, landing distance significantly reduced

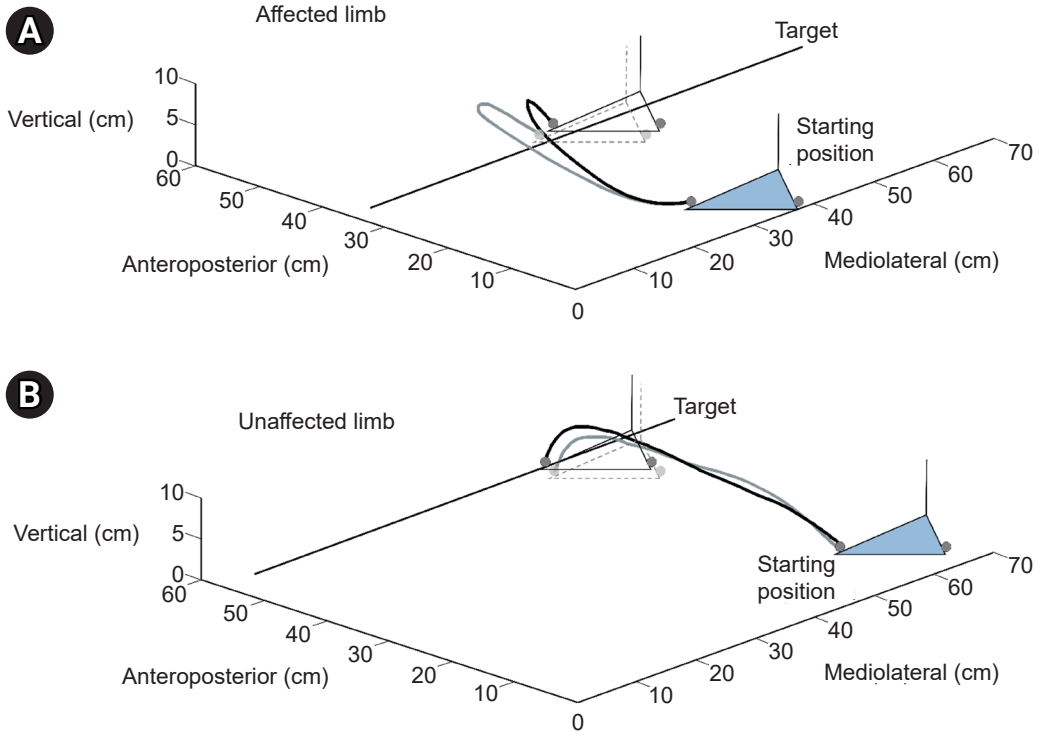

Fig. 4. Representative trajectories for the targeted stepping test from a participant in the experimental group before (gray) and after (black) intervention. The dotted gray and solid black lined triangles represent the end position of the foot before and after the intervention, respectively. The circles represent the reflective markers. It can be seen that after training, the participant was able to place the unaffected big toe closer to the target (B), but not the affected big toe (A).

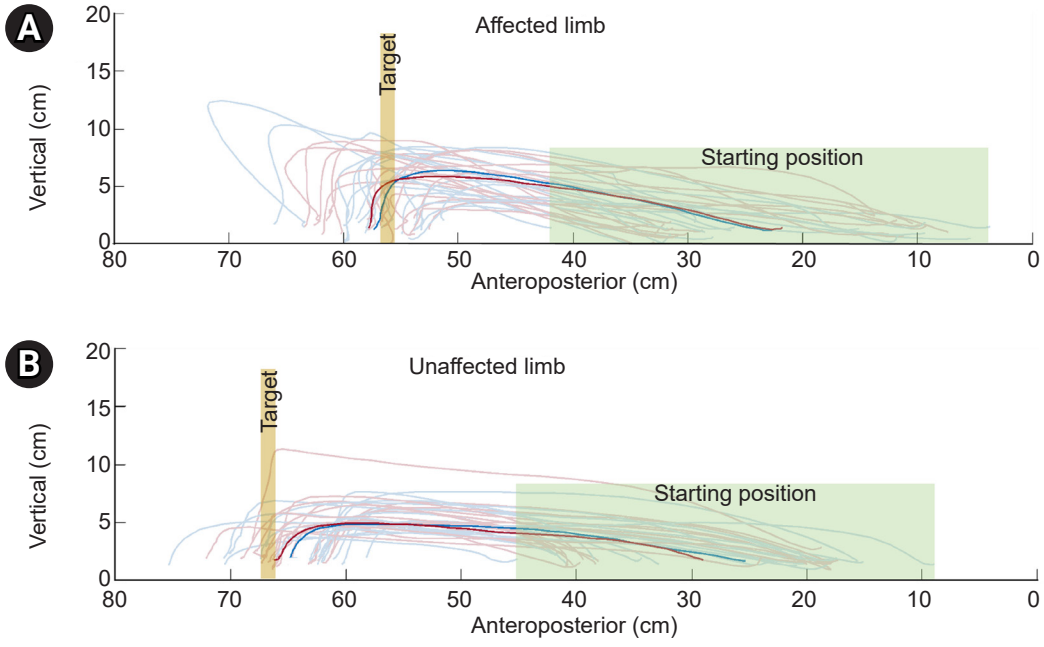

Fig. 5. Trajectories of the affected (A) and unaffected (B) foot motion during targeted stepping of individual subject in experimental group before (light blue) and after (light red) the intervention. The group averages before (dark blue) and after (dark red) the intervention show that the error distance decreased in the unaffected limb but not the affected limb.

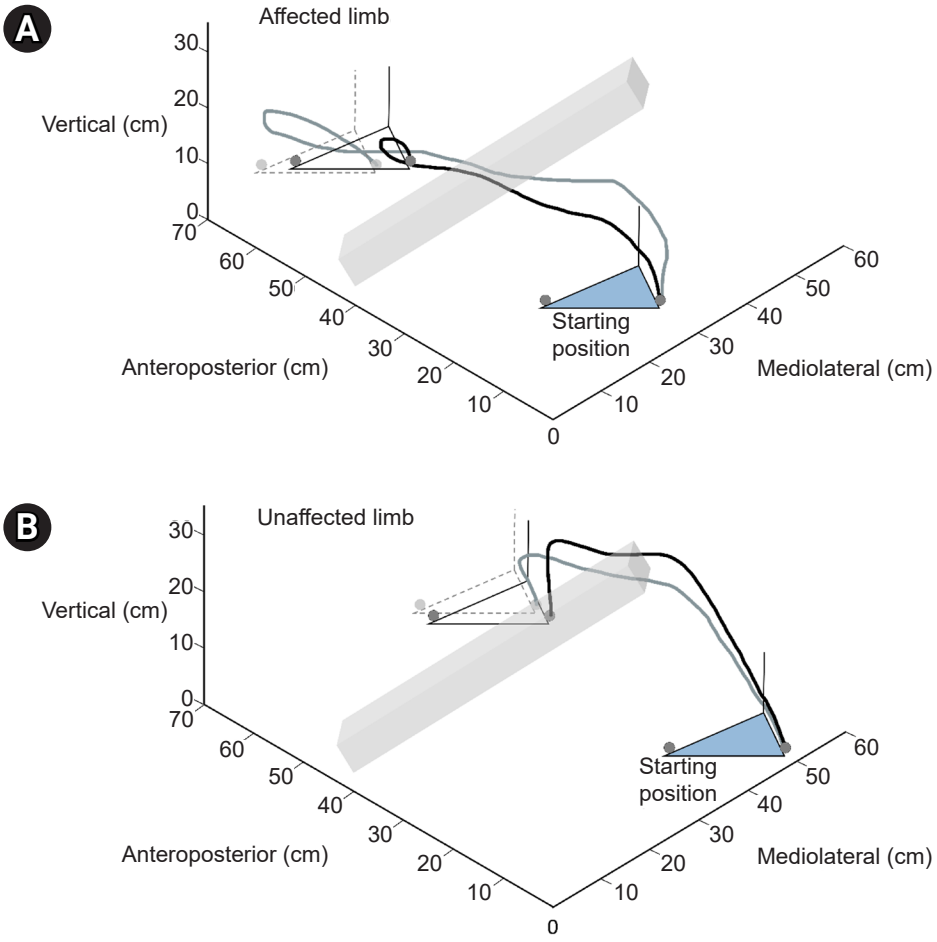

Fig. 6. Representative trajectories for the unaffected limb leading obstacle crossing stepping from a participant in experimental group before (gray) and after (black) intervention. It can be seen that after training, both the affected (A) and unaffected (B) heels were placed closer to the obstacle, indicating greater accuracy.

in post-test (t=4.127, p=0.001), but not for CON (t=0.846, p=0.409). In this condition, changes in landing distance on the unaffected limb were significantly correlated with changes on the affected limb in either group (EXP r=0.693, p=0.01; CON r=0.608, p=0.007). Changes in toe clearance on the unaffected limb after the intervention were significantly correlated with changes on the affected limb in CON (r=0.531, p=0.023), but were nonsignificant in EXP (r=0.254, p=0.294). When the affected leg was the leading limb, multivariate RMANOVA for the leading limb showed significant time main effect (F=4.077, p=0.026) and nonsignificant group main effect (F=0.461, p=0.634) or interaction (F=2.284, p=0.117). Univariate analysis showed significant differences in landing distance (F=4.302, p=0.045) but not toe clearance (F=2.242, p=0.143). In this condition, multivariate RMANOVA for the trailing limb (i.e., unaffected limb) showed nonsignificant time (F=0.686, p=0.51) or group (F=0.416, p=0.663) main effect or interaction (F=3.149, p=0.055). Fig. 7 shows the trajectories of the affected (Fig. 7A) and unaffected (Fig. 7B) foot motion during unaffected limb leading obstacle crossing of individual subject, as well as averages, in EXP before (light blue) and after (light red) the intervention. In this condition, changes in landing distance on the unaffected limb were significantly correlated with changes on the affected limb in either group (EXP r=0.631; CON r=0.465, both p<0.05). Changes in toe clearance on the unaffected limb after the intervention were not significantly correlated with changes on the affected limb in either group (r=0.150 and r=0.373, both p>0.05).

DISCUSSION

Stepping onto a proper location in the environment is import-

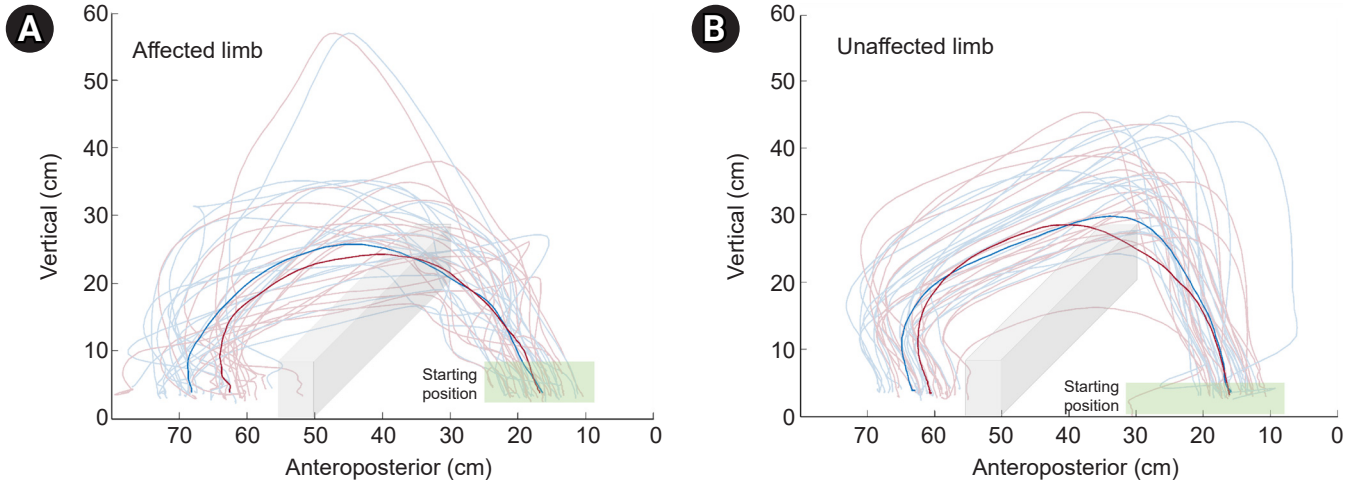

Fig. 7. Trajectories of the affected (A) and unaffected (B) foot motion during the unaffected limb leading obstacle crossing stepping of individual subject in experimental group before (light blue) and after (light red) the intervention. The group averages before (dark blue) and after (dark red) the intervention show that the landing distance decreased in both the affected and unaffected limbs, indicating better accuracy.

ant for meeting the task and environmental demands during daily life. This ability could be affected after stroke. This study used a VR environment to manipulate the on-line visual cues during targeted stepping training. It was found that, compared to providing full on-line visual feedback about the target and the foot in a VR environment, gradually removing this type of feedback led to greater improvement in the accuracy of foot placement of the unaffected limb during targeted stepping and both limbs during unaffected limb leading obstacle crossing. These findings supported the hypothesis that training with gradual removal of on-line visual cues would be superior to similar training with full on-line visual cues for people with chronic stroke. In spite of the positive effects, the removal of visual cues should be used cautiously to avoid loss of balance or falling during stroke rehabilitation.

To place the foot properly in the environment, visual, proprioceptive and exteroceptive inputs are integrated to contribute to the perception of the body in the environment to allow for the planning of motor commands to move the limb directly to the target location [31]. For people with stroke, the control of the upper limb movement and posture has been noted to tend to rely more heavily on on-line visual cues [32,33]. This tendency may interfere with the performance of the tasks that need to be performed without on-line visual cues about the foot or the target, such as targeted stepping during which the gaze typically does not remain on the target or the feet [34]. For the EXP in this study, the participants would be forced to reduce the reliance on visual inputs for perception and movement planning. This type of training could lead to more accurate prediction of movement errors and allow for feedforward updating of motor commands to generate more accurate actions without feedback [35]. Possibly as a result, the EXP showed some improvement in stepping onto the target after training when the visual cues about the target or the feet were not available. It should be noted that although improvements in balance control might also lead to better stepping or obstacle crossing performance, they are unlikely to contribute to the between-group differences, as the balance challenges during training were similar for both groups.

An early systematic review reported that additional visual cues during bilateral standing training for people with stroke had nonsignificant effects on standing weight symmetry, postural sway, or gait [36]. More recent studies that included walking and gait adaptation training showed positive effects on balance and gait when additional visual cues were provided [18,19]. In contrast, this study gradually removed the visual cues in the VR environment during training for the EXP and found that it had positive effects compared to providing more visual cues. These findings corresponded to a study showing that the deprivation of visual cues during balance training improved balance and cadence in people with stroke [20].

In this study, improved accuracy in targeted stepping was found in the unaffected limb in the EXP after training. Although stroke primarily affects the function of the contralateral limbs, it has been shown that in people with chronic stroke, the

ability to place the unaffected (ipsilateral) foot on a target with eyes closed is poorer, compared to non-stroke controls, but this is not the case with eyes open [37]. These findings suggested that after a stroke, the control of the unaffected limb's targeted movement also relied more heavily on visual cues. In this study, with the removal of on-line visual cues during training, the EXP would be forced to develop better feedforward control of targeted stepping, resulting in a smaller error distance in the unaffected limb. The findings of this study also suggest that the control of the unaffected lower limb movement has the potential for improvement in the chronic stage of stroke and support the use of targeted foot placement training in a VR environment in stroke rehabilitation.

When crossing an obstacle, on-line visual cues have been reported to be used primarily in a feedforward manner to regulate foot placement and gait patterns [38]. In this study, the intervention with the gradual removal of on-line visual cues could lead to better learning of feedforward control and possibly result in better training effect: the EXP had better accuracy (i.e. smaller landing distance) after training during obstacle crossing stepping. It is interesting that positive effects in EXP were observed only when the unaffected limb was leading and not when the affected limb was leading. These findings partly correspond to those of a previous study which showed that after one month of conventional physical therapy, walking obstacle crossing performance improved only when the unaffected limb was leading [39]. It has been proposed that the leading limb plays an important role in the trajectory and foot placement of the trailing limb during obstacle crossing or stepping onto an elevated surface in people with stroke [7,40]. This study also found that the landing distance between the leading and trailing limbs was significantly related in both affected and unaffected limb leading obstacle crossing. Further studies are needed to determine whether there are causal relationships between the two limbs during obstacle crossing.

The training in this study did not show significant effects on gait speed or symmetry in either group. And the changes in walking speed or symmetry after training were also not significantly correlated. It is possible that although stepping is an important subtask of walking, stepping training in sitting and standing was not sufficient to improve walking speed or symmetry. Further studies involving task-oriented practice, such as targeted stepping during level walking, are needed to determine whether better effects on gait speed or symmetry could be achieved.

This study is limited in several ways. The training appeared to be shorter than typically used in stroke rehabilitation. However, it should be noted that the same stepping task was repeatedly practiced for 640 trials. This amount of practice would likely be greater than in typical stroke rehabilitation in which various activities or tasks would be practiced. It is not clear if a longer duration training would yield different results. The outcome variables were measured in a laboratory, and it is not clear if the improvement in stepping would transfer to improved functional performance or social participation in real-world situations.

Conclusions

In conclusion, gradually removing the on-line visual cues during the training of foot placement in sitting and standing in a VR environment led to better accuracy in foot placement. This included smaller error distances in the unaffected foot during targeted stepping and smaller landing distances in either foot during obstacle crossing with the unaffected limb leading. However, this type of visual cue manipulation did not show superior effects on gait speed or symmetry, or other aspects of gait adaptation. Manipulation of visual cues of the foot or target may be considered in stroke rehabilitation to improve gait adaptability. The design of VR training programs for stroke rehabilitation may need to carefully consider how on-line visual cues are presented.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

This study was supported by a research grant from the Ministry of Science and Technology, Taiwan (Grant number: 105-2314-B-006-011-MY3).

AUTHOR CONTRIBUTION

Conceptualization: Lee PY, Tsai MR, Lin SI. Formal analysis: Lee PY, Tsai MR, Chen CH, Tseng HY, Lin SI. Funding acquisition: Lin SI. Investigation: Lee PY, Chen CH, Tseng HY, Lin SI. Methodology: Lee PY, Wang HC, Lin SI. Project administration: Lee PY, Lin SI. Software: Lee PY, Wang HC. Supervision: Lee PY, Lin SI. Visualization: Wang HC. Writing – original draft: Lee PY, Tsai MR, Wang HC, Chen CH, Tseng HY, Lin SI. Writ-

ing – review & editing: Lee PY. Approval of final manuscript: all authors.

ORCID

Pei-Yun Lee, https://orcid.org/0000-0003-2655-8048
Meng-Ru Tsai, https://orcid.org/0009-0003-1619-5383
Hsien-Chang Wang, https://orcid.org/0000-0003-1557-0455
Chih-Hung Chen, https://orcid.org/0000-0002-8256-9844
Hui-Yu Tseng, https://orcid.org/0009-0003-2915-0897
Sang-I Lin, https://orcid.org/0000-0003-2201-9127

REFERENCES

1. Kim J, Thayabaranathan T, Donnan GA, Howard G, Howard VJ, Rothwell PM, et al. Global stroke statistics 2019. Int J Stroke 2020;15:819-38.
2. Jang SH. The recovery of walking in stroke patients: a review. Int J Rehabil Res 2010;33:285-9.
3. Den Otter AR, Geurts AC, de Haart M, Mulder T, Duysens J. Step characteristics during obstacle avoidance in hemiplegic stroke. Exp Brain Res 2005;161:180-92.
4. Smulders K, van Swigchem R, de Swart BJ, Geurts AC, Weerdesteyn V. Community-dwelling people with chronic stroke need disproportionate attention while walking and negotiating obstacles. Gait Posture 2012;36:127-32.
5. Houdijk H, van Ooijen MW, Kraal JJ, Wiggerts HO, Polomski W, Janssen TW, et al. Assessing gait adaptability in people with a unilateral amputation on an instrumented treadmill with a projected visual context. Phys Ther 2012;92:1452-60.
6. van Swigchem R, van Duijnhoven HJ, den Boer J, Geurts AC, Weerdesteyn V. Deficits in motor response to avoid sudden obstacles during gait in functional walkers poststroke. Neurorehabil Neural Repair 2013;27:230-9.
7. Said CM, Goldie PA, Culham E, Sparrow WA, Patla AE, Morris ME. Control of lead and trail limbs during obstacle crossing following stroke. Phys Ther 2005;85:413-27.
8. van der Veen SM, Hammerbeck U, Hollands KL. Foot-placement accuracy during planned and reactive target stepping during walking in stroke survivors and healthy adults. Gait Posture 2020;81:261-7.
9. Hollands KL, Pelton TA, van der Veen S, Alharbi S, Hollands MA. A novel and simple test of gait adaptability predicts gold standard measures of functional mobility in stroke survivors. Gait Posture 2016;43:170-5.
10. Balasubramanian CK, Neptune RR, Kautz SA. Foot placement in a body reference frame during walking and its relationship to hemiparetic walking performance. Clin Biomech (Bristol) 2010;25:483-90.
11. Nonnekes JH, Talelli P, de Niet M, Reynolds RF, Weerdesteyn V, Day BL. Deficits underlying impaired visually triggered step adjustments in mildly affected stroke patients. Neurorehabil Neural Repair 2010;24:393-400.
12. Schween R, Taube W, Gollhofer A, Leukel C. Online and post-trial feedback differentially affect implicit adaptation to a visuomotor rotation. Exp Brain Res 2014;232:3007-13.
13. Winstein CJ, Pohl PS, Lewthwaite R. Effects of physical guidance and knowledge of results on motor learning: support for the guidance hypothesis. Res Q Exerc Sport 1994;65:316-23.
14. Sidaway B, Bates J, Occhiogrosso B, Schlagenhaufer J, Wilkes D. Interaction of feedback frequency and task difficulty in children's motor skill learning. Phys Ther 2012;92:948-57.
15. Batcho CS, Gagné M, Bouyer LJ, Roy JS, Mercier C. Impact of online visual feedback on motor acquisition and retention when learning to reach in a force field. Neuroscience 2016;337:267-75.
16. Fujii S, Lulic T, Chen JL. More feedback is better than less: learning a novel upper limb joint coordination pattern with augmented auditory feedback. Front Neurosci 2016;10:251.
17. Lam CF, DeRue DS, Karam EP, Hollenbeck JR. The impact of feedback frequency on learning and task performance: challenging the "more is better" assumption. Organ Behav Hum Decis Process 2011;116:217-28.
18. Hollands KL, Pelton TA, Wimperis A, Whitham D, Tan W, Jowett S, et al. Feasibility and preliminary efficacy of visual cue training to improve adaptability of walking after stroke: multi-centre, single-blind randomised control pilot trial. PLoS One 2015;10:e0139261.
19. Upadhyay S, Verma N. Effect of visual cue training on gait and walking velocity in chronic stroke patients- a quasi-experimental study. J Clin Diagn Res 2022;16:YC06-10.
20. Nam SM, Lee DY. Effects of visual cue deprivation balance training with head control on balance and gait function in stroke patients. Medicina (Kaunas) 2022;58:629.
21. Garay-Sánchez A, Suarez-Serrano C, Ferrando-Margelí M, Jimenez-Rejano JJ, Marcén-Román Y. Effects of immersive and non-immersive virtual reality on the static and dynamic balance of stroke patients: a systematic review and meta-analysis. J Clin Med 2021;10:4473.
22. Han SH, Jang HJ, Lee JW, Cheong JW, Kim YD, Nam HS, et al. The effect of virtual reality-based treadmill gait training on functional mobility and balance in chronic stroke patients: a randomized con-

trolled trial. Front Neurol 2025;16:1603233.
23. Wang X, Qiu J, Zhou Y, Liu W, Zhang S, Gong Y, et al. Effects of virtual reality-assisted and overground gait adaptation training on balance and walking ability in stroke patients: a randomized controlled trial. Am J Phys Med Rehabil 2024;103:480-7.
24. Robinson SM, Jameson KA, Syddall HE, Dennison EM, Cooper C, Aihie Sayer A, et al. Clustering of lifestyle risk factors and poor physical function in older adults: the Hertfordshire cohort study. J Am Geriatr Soc 2013;61:1684-91.
25. Blennerhassett JM, Dite W, Ramage ER, Richmond ME. Changes in balance and walking from stroke rehabilitation to the community: a follow-up observational study. Arch Phys Med Rehabil 2012;93:1782-7.
26. Jeong YG, Koo JW. The effects of treadmill walking combined with obstacle-crossing on walking ability in ambulatory patients after stroke: a pilot randomized controlled trial. Top Stroke Rehabil 2016;23:406-12.
27. Yang YR, Tsai MP, Chuang TY, Sung WH, Wang RY. Virtual reality-based training improves community ambulation in individuals with stroke: a randomized controlled trial. Gait Posture 2008;28:201-6.
28. Cho KH, Lee WH. Effect of treadmill training based real-world video recording on balance and gait in chronic stroke patients: a randomized controlled trial. Gait Posture 2014;39:523-8.
29. Cleland B, Madhavan S. Changes in walking speed after high-intensity treadmill training are independent of changes in spatiotemporal symmetry after stroke. Front Neurol 2021;12:647338.
30. Viteckova S, Kutilek P, Svoboda Z, Krupicka R, Kauler J, Szabo Z. Gait symmetry measures: a review of current and prospective methods. Biomed Signal Process Control 2018;42:89-100.
31. Pick HL, Saltzman E. Modes of perceiving and processing information. Psychology Press; 1978. p. 97-224.
32. Bernard-Espina J, Beraneck M, Maier MA, Tagliabue M. Multisensory integration in stroke patients: a theoretical approach to reinterpret upper-limb proprioceptive deficits and visual compensation. Front Neurosci 2021;15:646698.
33. Kang N, Cauraugh JH. Bimanual force variability in chronic stroke: with and without visual information. Neurosci Lett 2015;587:41-5.
34. Hollands MA, Marple-Horvat DE, Henkes S, Rowan AK. Human eye movements during visually guided stepping. J Mot Behav 1995;27:155-63.
35. Flanagan JR, Vetter P, Johansson RS, Wolpert DM. Prediction precedes control in motor learning. Curr Biol 2003;13:146-50.
36. Van Peppen RP, Kortsmit M, Lindeman E, Kwakkel G. Effects of visual feedback therapy on postural control in bilateral standing after stroke: a systematic review. J Rehabil Med 2006;38:3-9.
37. Lee PY, Chen CH, Tseng HY, Lin SI. Ipsilateral lower limb motor performance and its association with gait after stroke. PLoS One 2024;19:e0297074.
38. Patla AE, Vickers JN. Where and when do we look as we approach and step over an obstacle in the travel path? Neuroreport 1997;8:3661-5.
39. Said CM, Galea M, Lythgo N. Obstacle crossing following stroke improves over one month when the unaffected limb leads, but not when the affected limb leads. Gait Posture 2014;39:213-7.
40. Hösl M, Egger M, Bergmann J, Amberger T, Mueller F, Jahn K. Tempo-spatial gait adaptations in stroke patients when approaching and crossing an elevated surface. Gait Posture 2019;73:279-85.

Original Article

Ann Rehabil Med 2026;50(5):376-388
eISSN: 2234-0653
https://doi.org/10.5535/arm.260090

Annals of Rehabilitation Medicine

Association of Admission Sarcopenia With 3-Month Recovery of Trunk Control and Gait in Subacute Stroke Patients With Impaired Balance

Kyungmin Kim, MD*, Hangi Kim, MD*, Jiwon Lee, MD, Su Jin Choi, MD

Department of Physical Medicine and Rehabilitation, Dong-Eui Medical Center, Busan, Korea

Objective: To examine whether sarcopenia at rehabilitation admission is associated with 3-month recovery of balance, trunk control, and gait in subacute stroke patients with impaired balance.

Methods: This retrospective cohort study included 49 subacute stroke patients who received inpatient rehabilitation for at least 3 months. Participants were classified at admission into sarcopenia group (SG, n=18) and non-sarcopenia group (NSG, n=31) according to the Asian Working Group for Sarcopenia 2019 criteria, using unaffected-side grip strength and skeletal muscle index. The Berg Balance Scale (BBS) was the principal outcome, with the Korean-Trunk Impairment Scale (K-TIS), Motricity Index of Lower Extremity (MI-LE), and Functional Ambulation Category (FAC) as secondary outcomes. Changes were compared by analysis of covariance adjusted for baseline, and threshold-based proportions by Fisher's exact test.

Results: Baseline outcome scores were comparable except for a lower FAC in the SG (p=0.02). After 3 months, the NSG showed greater improvements in all four outcomes, with significant baseline-adjusted differences after Bonferroni correction for BBS (5.87; 95% confidence interval [CI], 1.98 to 9.76), K-TIS (3.68; 95% CI, 1.52 to 5.83), and MI-LE (8.75; 95% CI, 4.65 to 12.85), persisting under stepwise covariate adjustment. The ordinal FAC was compared categorically: fewer SG patients improved by ≥1 category (22.2% vs. 58.1%, p=0.02).

Conclusion: Sarcopenia assessed at rehabilitation admission was associated with smaller 3-month gains in trunk control, balance, lower extremity muscle strength, and gait in this selected cohort of subacute stroke inpatients. Prospective studies with pre-stroke or hyperacute muscle assessment are needed to determine whether premorbid sarcopenia independently predicts recovery.

Keywords: Stroke, Sarcopenia, Recovery of function, Postural balance, Gait

Received: May 8, 2026
Revised: August 6, 2026
Accepted: August 21, 2026

Correspondence:
Su Jin Choi
Department of Physical Medicine and Rehabilitation, Dong-Eui Medical Center, 62 Yangjeong-ro, Busanjin-gu, Busan 47227, Korea.
Tel: +82-51-850-8751
Fax: +82-504-233-3328
E-mail: shucomy@gmail.com

*These authors contributed equally to this work.

© 2026 by Korean Academy of Rehabilitation Medicine

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Stroke is a leading cause of severe physical and cognitive disability in adults worldwide, and rehabilitation aimed at minimizing the resulting functional decline plays a critical role in patient prognosis [1,2]. As societies enter an era of population aging, the importance of musculoskeletal health has gained increasing attention, and there is growing interest in sarcopenia accompanying stroke. Sarcopenia is defined as a condition characterized by low muscle strength along with decreased muscle quality and quantity [3,4].

Stroke and sarcopenia share age as a strong common risk factor, and the prevalence of sarcopenia in stroke patients has been reported to be high [5]. Low grip strength and sarcopenia have also been reported as independent risk factors for cardiovascular disease, including stroke [6,7]. Sarcopenia is not merely a disorder of peripheral muscle but has been proposed as an indicator of impaired muscle-nerve interaction that may affect neuroplasticity, a key recovery mechanism after stroke [8,9]. Muscle activity mediates the secretion of myokines, including brain-derived neurotrophic factor and insulin-like growth factor-1, that play regulatory roles in brain-muscle interaction [10]. Low muscle mass and strength may therefore reflect a state of chronic decline in these factors, which has been proposed as a physiological mechanism that impedes neuroplasticity during post-stroke rehabilitation [11].

Consistent with these mechanisms, sarcopenia has already been reported to be associated with poorer functional recovery after stroke [12,13]; evidence on the recovery of trunk control and balance during the subacute phase, however, remains limited. Two studies have reported reduced functional recovery in stroke patients with sarcopenia, but their observation periods were confined to 8 and 3 weeks and their outcomes largely to gait-related variables, so that the recovery of balance function and trunk control beyond these periods has yet to be elucidated [12,13]. Predicting rehabilitation prognosis and establishing tailored strategies for stroke patients with sarcopenia therefore remains difficult.

To address these limitations, the present study assessed the association of sarcopenia with a broader range of functional recovery outcomes. Balance and trunk control, which are essential prerequisites for functional ambulation, were evaluated with the Berg Balance Scale (BBS) and the Korean-Trunk Impairment Scale (K-TIS), lower extremity muscle strength with the Motricity Index of Lower Extremity (MI-LE), and gait with the Functional Ambulation Category (FAC). The follow-up period was also extended to 3 months, beyond the 3- and 8-week observation periods of these studies. This study examined whether sarcopenia assessed at rehabilitation admission is associated with 3-month recovery of balance, trunk control, and gait in subacute stroke inpatients with impaired balance, and aimed to provide fundamental data for developing tailored rehabilitation.

METHODS

Study design and sarcopenia classification

This study was approved by the Dong-Eui Medical Center Institutional Review Board (IRB) (IRB No. DEMC-2025-06). The requirement for informed consent was waived by the IRB, as this study was based exclusively on retrospectively reviewed medical records obtained during routine clinical care.

This retrospective cohort study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement and was based on the medical records of stroke patients who had received inpatient rehabilitation at a single rehabilitation hospital between September 2019 and February 2024. Completion of at least 3 months of inpatient rehabilitation was required for the source population, and 348 patients were identified and screened for eligibility. Patients who did not complete 3 months of inpatient rehabilitation were not captured in the retrieved records, and their number was not available. The predefined exclusion criteria were then applied sequentially, yielding a final cohort of 49 patients (Fig. 1). The patients were classified into a sarcopenia group (SG) and a non-sarcopenia group (NSG). Unaffected-side grip strength and skeletal muscle index (SMI) measured at admission for inpatient rehabilitation were used to define the exposure of the present study, hereafter referred to as sarcopenia assessed at rehabilitation admission. Pre-stroke health-screening records, previous bioelectrical impedance or grip strength measurements, and pre-stroke frailty, functional, or nutritional assessments were not available; the classification therefore reflects muscle status at rehabilitation admission and does not establish a pre-stroke condition. The diagnostic criteria for sarcopenia were established based on the Asian Working Group for Sarcopenia 2019 (AWGS 2019) guidelines, and patients were classified using two measures: unaffected-side grip strength and SMI obtained through bioelectrical impedance analysis (BIA) [14]. Because most patients had hemiplegia precluding reliable grip measurement on the paretic side, low muscle strength was defined as unaffected-side grip strength <28 kg for men and

<18 kg for women, low SMI was defined as <7.0 kg/m^2 for men and <5.7 kg/m^2 for women, and only patients who met both criteria were classified into the SG, whereas all remaining patients were assigned to the NSG [14-16]. Physical performance measures in the AWGS 2019 algorithm, such as usual gait speed, the five-time chair-stand test, and the Short Physical Performance Battery, were not applied. They require independent standing or walking, which was not feasible in many patients with impaired balance, and would largely reflect the hemiparetic motor deficit under investigation, overlapping with the study outcomes, rather than the underlying muscle status. Requiring concurrently low muscle strength and low muscle mass corresponds to sarcopenia as defined by AWGS 2019 and provides a more specific case definition than possible sarcopenia, which is based on low muscle strength or low physical performance alone [14].

Baseline assessments

Grip strength and BIA were performed on the same day within 7 days of admission for inpatient rehabilitation, with BIA obtained immediately after grip testing. Grip strength was measured on the unaffected side using a hand dynamometer (T.K.K. 5101; Takei Scientific Instruments Co., Ltd.), with the upper limb alongside the trunk; the maximum of three trials separated by 1-minute rests was used. SMI was obtained by BIA (InBody 770; InBody), a direct segmental multifrequency device with an eight-point tactile electrode system, and was calculated as appendicular skeletal muscle mass divided by height squared, with height measured directly at the same session. All patients completed the measurement standing on the foot electrodes and grasping the hand electrodes without physical assistance, and none had a cardiac pacemaker or other implanted electronic device. The measurements were obtained during routine clinical care; fasting status and bladder emptying were not standardized, and the presence of edema was not systematically assessed. The severity of neurological impairment at admission for inpatient rehabilitation was assessed using the National Institutes of Health Stroke Scale (NIHSS). Body weight and height from the same session were used to calculate body mass index (BMI), and serum albumin was recorded from the routine admission laboratory panel. Nutritional risk was summarized with the Geriatric Nutritional Risk Index (GNRI), calculated as [14.89×serum albumin (g/dL)]+[41.7×(body weight/ideal body weight)], with the weight ratio set to 1 when body weight exceeded ideal body weight [17]. Ideal body weight was derived from height and a BMI of 22 kg/m^2, rather than from the Lorentz formula used in the original description. Dysphagia at admission was defined as the prescription of a texture-modified diet, thickened fluids, or tube feeding. Comorbidity burden was summarized as the number of the following conditions documented before the index stroke: diabetes mellitus, cardiovascular disease, chronic kidney disease, and malignancy. Cardiovascular disease comprised ischemic heart disease, heart failure or cardiomyopathy, valvular disease requiring follow-up or surgery, atrial fibrillation or flutter requiring long-term treatment, and peripheral arterial disease; hypertension and dyslipidemia were regarded as vascular risk factors and were not counted.

Exclusion criteria

The exclusion criteria were applied sequentially in the following order: not first-ever stroke; a stroke duration exceeding 6 months; coexisting disorders affecting muscle strength other than stroke and sarcopenia, including neurological, musculoskeletal, or severe systemic conditions; severe aphasia precluding reliable communication; severe cognitive impairment, defined as a Mini-Mental State Examination (MMSE) score below 18 [18]; inability to assess sarcopenia, either due to an inability to maintain independent standing for BIA or bilateral lesions precluding unaffected-side grip strength measurement; and a BBS score of 41 or above at admission, indicating functional independence in balance [19].

Outcome measures

Outcome measures were assessed using the BBS, K-TIS, MI-LE, and FAC at baseline and 3 months after treatment initiation [20-23]. All assessments were performed by rehabilitation physicians using standardized institutional forms, and the same physician assessed a given patient at both time points; formal inter-rater reliability testing was not undertaken. The BBS ranges from 0 to 56, the K-TIS from 0 to 23 (static sitting balance, 0–7; dynamic sitting balance, 0–10; coordination, 0–6), the MI-LE from 0 to 100, and the FAC from 0 to 5, with higher scores indicating better function in all four measures. The MI-LE was measured on the paretic lower limb, and the FAC was rated without the use of any assistive device or orthosis. The assessors were not formally blinded to the admission grip strength and BIA results, which were part of the routine clinical record; the classification into the SG and the NSG, however, was made only for the present retrospective analysis. None of the analyzed patients discontinued rehabilitation because of an adverse event during the 3-month follow-up.

Rehabilitation program

Both groups received the same standard 3-month inpatient rehabilitation program routinely provided to patients with stroke at the hospital. Physical and occupational therapy were each provided for 30 minutes twice daily, 5 days a week, amounting to 2 hours per treatment day; physical therapy comprised balance training, gait training, and lower extremity resistance exercises. Speech therapy, 30 minutes once daily on 2 to 3 days a week, was required in only a small proportion of patients with mild aphasia or dysarthria. All therapy was prescribed by a rehabilitation physician and delivered by licensed physical, occupational, and speech therapists according to the institution's standard stroke rehabilitation protocol. Neither group received any specific nutritional intervention, such as protein, amino acid, or vitamin D supplementation; dietary management followed routine clinical practice, with texture-modified diets, thickened fluids, or diabetic diets prescribed only as clinically indicated.

Statistical analysis

All statistical analyses were performed using SPSS 27.0 (IBM Corp.), and the level of statistical significance was set at $p<0.05$. Because the study was retrospective, no primary outcome was pre-specified; the BBS was regarded as the principal outcome, as impaired balance defined the cohort, and the K-TIS, MI-LE, and FAC as secondary outcomes. All analyses were nonetheless exploratory, and the multiplicity correction described below was applied to all four outcomes. Baseline and 3-month outcome data were complete for all eligible patients; no patient was excluded because of missing data, and no imputation was performed. Categorical variables (sex, stroke laterality, and stroke type) were compared between the two groups using the chi-square test or Fisher's exact test, while continuous variables and changes in measured values were analyzed using the Mann–Whitney U-test, and within-group changes before and after treatment were analyzed using the Wilcoxon signed-rank test; these rank-based tests do not require assumptions about the shape of the underlying distributions. The median with the interquartile range is presented together with the mean and standard deviation for the ordinal functional scales and for the right-skewed stroke duration; the means are retained because the published thresholds for clinically meaningful change are expressed in these terms. For the between-group comparisons of the change scores, the effect size was calculated as $r=|Z|/\sqrt{N}$. Because baseline FAC differed significantly between the groups, ANCOVA with the corresponding baseline score as a covariate was performed for each of the four outcomes to compare the 3-month change scores. In each model, the group was entered as a fixed factor and the between-group difference was expressed as the difference in estimated marginal means with its 95% confidence interval (CI) and the partial eta squared. Before interpretation, the assumptions of each model were examined: the homogeneity of the regression slopes with the group-by-baseline interaction term, the normality of the standardized residuals with the Shapiro–Wilk test, and the homogeneity of the error variance with the Levene test. The between-group differences were additionally re-estimated with Huber–White heteroscedasticity-consistent standard errors. This baseline-adjusted model served as the primary adjusted analysis (Model 1); models with further covariates (Models 2–4) were used as sensitivity analyses. The Bonferroni correction was applied to the between-group comparison of the baseline-adjusted mean changes and, at a corrected level of 0.0125, to the within-group changes across the four outcomes; the sensitivity analyses were regarded as supportive and interpreted at the conventional level.

Sensitivity analyses

All four outcomes were re-examined under progressively wider covariate adjustment, with covariates selected a priori on clinical grounds from the baseline variables in Table 1. Model 2 added the stroke duration, to address the possibility that the timing of assessment influenced the classification of sarcopenia; Model 3 added age, sex, and stroke type; and Model 4 added stroke laterality, the MMSE score, and the NIHSS score, so that every baseline demographic and clinical variable was adjusted for simultaneously. Because the number of covariates that can be entered into a single model is constrained in a cohort of 49 patients, the nutritional and comorbidity measures were instead entered individually into the baseline-adjusted model. BMI and the GNRI incorporate body weight, which is itself partly determined by skeletal muscle mass, so the estimates obtained with these two covariates were regarded as conservative. The cohort was also stratified by the median stroke duration (≤21 days vs. >21 days), and between-group differences in 3-month change scores were compared within each subgroup using the Mann–Whitney U-test. Because the FAC is an ordinal six-level scale, the gait outcome was additionally analyzed in categorical terms. The proportions of patients reaching each predefined threshold, and the distribution of the FAC categories, were compared between the groups. All proportions were compared with Fisher's

Table 1. Baseline clinical and functional characteristics of 49 patients

Characteristic	SG (n=18)	NSG (n=31)	p-value
Age (yr)	71.11±9.62	68.58±10.72	0.30
Sex			0.91
Male	9	16	
Female	9	15	
Stroke duration (day)	34.28±24.06	24.23±16.10	0.12
Median (IQR)	23.50 (18.50–44.00)	20.00 (13.00–35.00)	
Stroke laterality			0.22
Left	6	16	
Right	12	15	
Stroke type			0.71
Hemorrhage	4	5	
Infarction	14	26	
MMSE	25.83±3.26	26.65±2.27	0.58
NIHSS	6.00±2.17	5.58±2.11	0.36
BMI (kg/m^2)	22.84±2.44	23.59±3.17	0.69
Serum albumin (g/dL)	3.98±0.37	4.16±0.41	0.15
GNRI	99.67±5.59	102.71±6.36	0.16
Dysphagia	6	6	0.32
Comorbidity			
Diabetes mellitus	10	8	0.04*
Cardiovascular disease	7	4	0.07
Chronic kidney disease	1	1	>0.999
Malignancy	0	0	-
Comorbidity count	1.00±0.84	0.42±0.62	0.01*
Grip strength, unaffected (kg)			
Male	20.98±5.94	33.64±4.81	<0.001***
Female	12.76±3.86	21.17±5.81	<0.001***
SMI (kg/m^2)			
Male	6.47±0.42	8.77±1.81	<0.001***
Female	5.38±0.31	7.22±1.38	<0.001***
Function			
BBS	21.94±10.09	25.90±9.80	0.16
Median (IQR)	23.00 (14.00–28.75)	27.00 (22.00–32.00)	
K-TIS	6.00±5.31	8.42±4.65	0.10
Median (IQR)	4.50 (2.00–10.25)	9.00 (4.00–12.00)	
MI-LE	48.22±16.92	54.90±17.02	0.21
Median (IQR)	48.00 (38.00–58.00)	58.00 (39.00–70.00)	
FAC	1.22±0.43	1.65±0.66	0.02*
Median (IQR)	1.00 (1.00–1.25)	2.00 (1.00–2.00)	

Values are presented as mean±standard deviation for continuous variables and as numbers of patients for categorical variables. The median (interquartile range, IQR) is additionally presented for the stroke duration and the functional scales.
Stroke laterality refers to the side of the brain lesion. Sex, stroke laterality, stroke type, dysphagia, and the individual comorbid conditions were compared using the chi-square test, with Fisher's exact test applied when any expected cell count was less than 5; all other variables were compared using the Mann–Whitney U-test. Dysphagia was defined as the prescription of a texture-modified diet, thickened fluids, or tube feeding at admission. The comorbidity count is the number of the four listed conditions documented before the index stroke; no patient had a malignancy, and this condition was therefore not tested.
SG, sarcopenia group; NSG, non-sarcopenia group; MMSE, Mini-Mental State Examination; NIHSS, National Institutes of Health Stroke Scale; BMI, body mass index; GNRI, Geriatric Nutritional Risk Index; SMI, skeletal muscle index; BBS, Berg Balance Scale; K-TIS, Korean-Trunk Impairment Scale; MI-LE, Motricity Index of Lower Extremity; FAC, Functional Ambulation Category.
*p<0.05 and ***p<0.001.

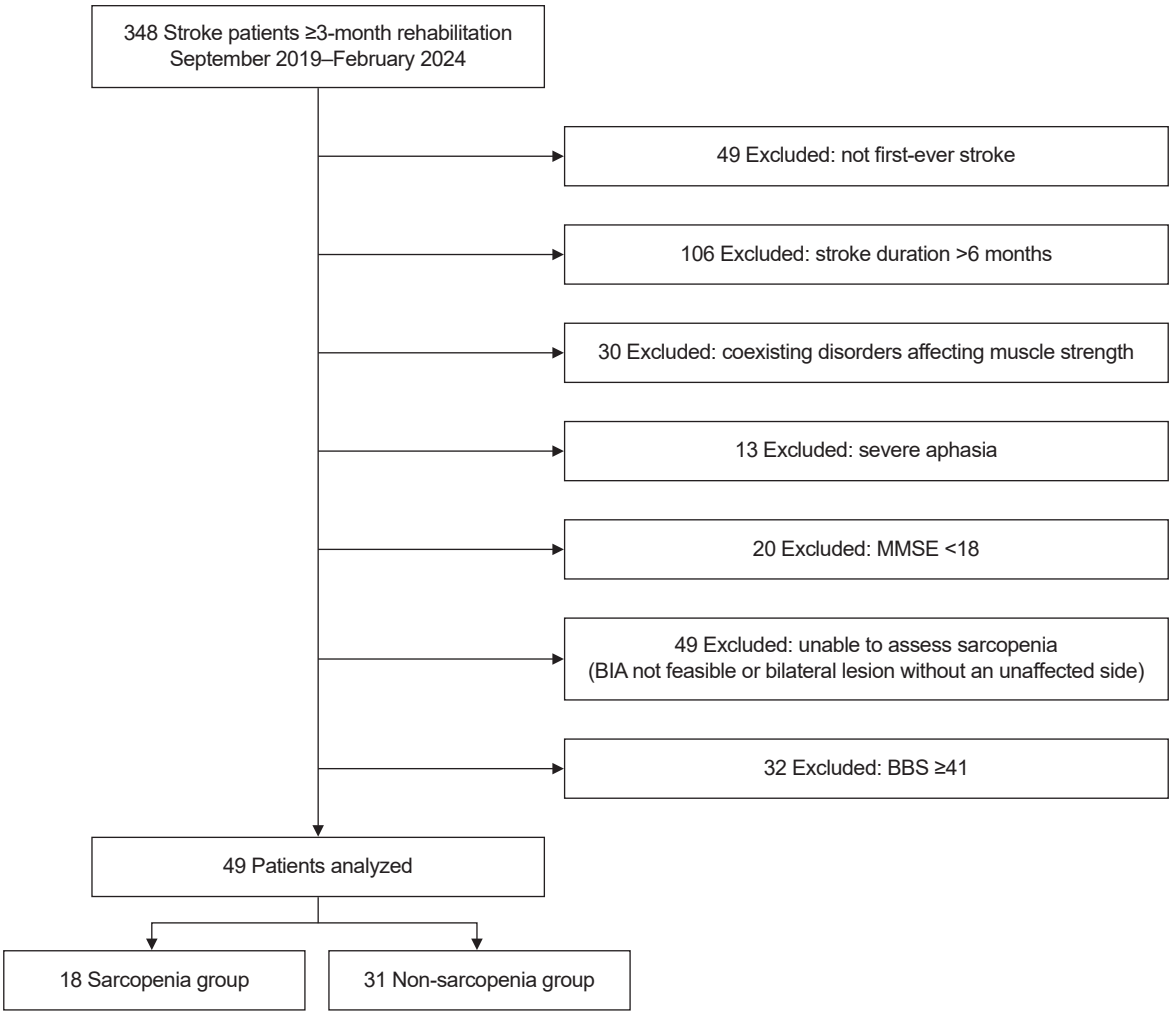

Fig. 1. Flow diagram of patient selection, prepared according to the STROBE recommendations for observational studies. The number of patients excluded at each sequential step and the number analyzed in each group are shown. Patients who had been discharged or had discontinued inpatient rehabilitation before 3 months were not captured in the records retrieved for the present study and are therefore not represented in the diagram. MMSE, Mini-Mental State Examination; BIA, bioelectrical impedance analysis; BBS, Berg Balance Scale.

exact test, and the between-group difference in proportions is reported with its 95% CI calculated by the Newcombe hybrid score method. These category-based comparisons were likewise treated as supportive.

RESULTS

Of the 348 patients screened, 49 met all eligibility criteria (Fig. 1); 18 were classified into the SG and 31 into the NSG. Baseline demographic and clinical characteristics are summarized in Table 1. There were no significant between-group differences in age, sex, stroke laterality, stroke type, stroke duration, MMSE or NIHSS scores, BMI, serum albumin, the GNRI, or the prevalence of dysphagia (all p>0.05). Diabetes mellitus was more prevalent in the SG (p=0.04), whereas cardiovascular disease, chronic kidney disease, and malignancy were comparable, and the comorbidity count was higher in the SG (p=0.01). Baseline BBS, K-TIS, and MI-LE scores did not differ significantly between the groups (all p>0.05), whereas baseline FAC was significantly lower in the SG (p=0.02).

After 3 months of rehabilitation, both groups showed significant within-group improvements in all outcome measures (Fig. 2). In the SG, BBS increased from 21.94±10.09 to 28.83±11.10 (p<0.001), K-TIS from 6.00±5.31 to 9.78±5.19 (p<0.001), MI-

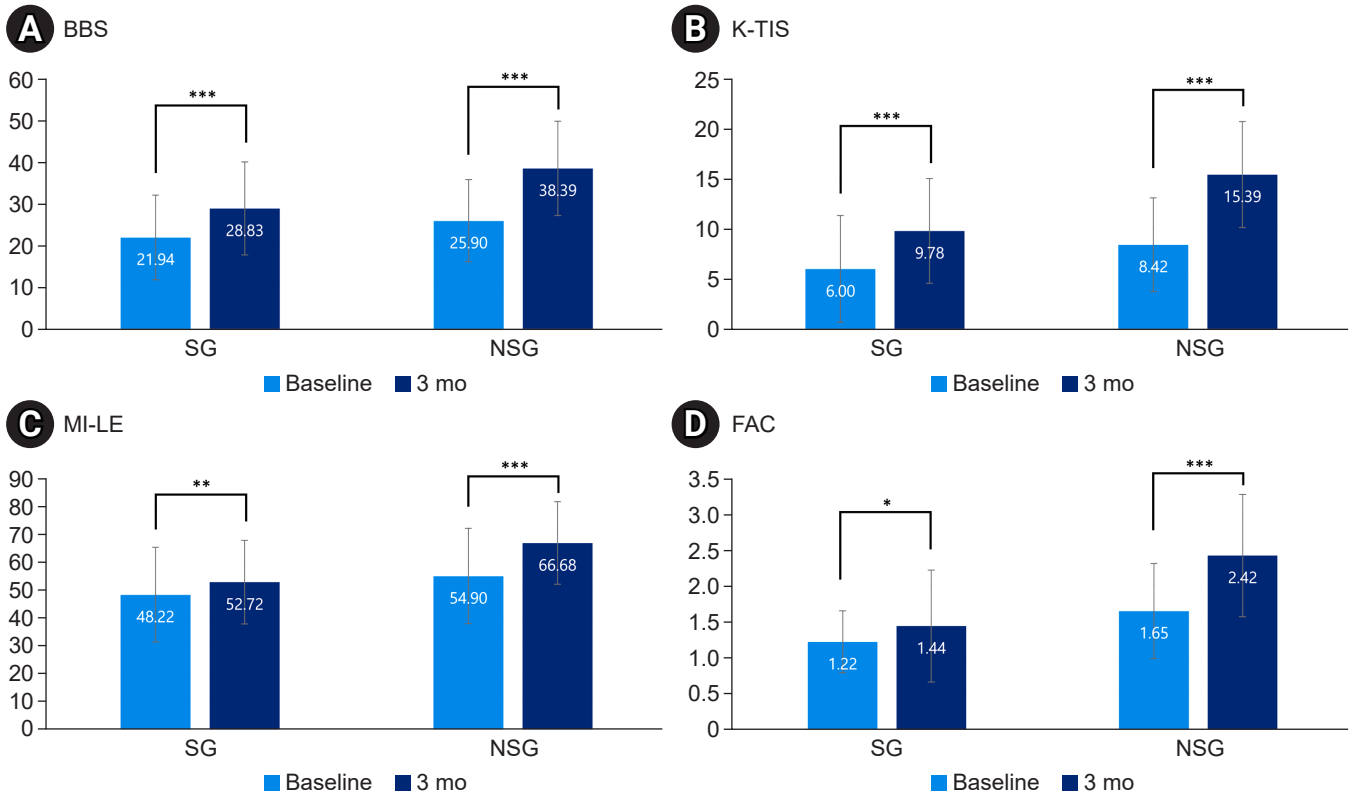

Fig. 2. Comparison of outcome measures at baseline and 3 months after rehabilitation in the sarcopenia group (SG) and non-sarcopenia group (NSG). (A) Berg Balance Scale (BBS), (B) Korean-Trunk Impairment Scale (K-TIS), (C) Motricity Index of Lower Extremity (MI-LE), (D) Functional Ambulation Category (FAC). Data are presented as group means, and error bars indicate standard deviations. Asterisks denote uncorrected within-group significance by Wilcoxon signed-rank test: *p<0.05, **p<0.01, and ***p<0.001.

LE from 48.22±16.92 to 52.72±14.97 (p=0.003), and FAC from 1.22±0.43 to 1.44±0.78 (p=0.046). In the NSG, BBS increased from 25.90±9.80 to 38.39±11.22, K-TIS from 8.42±4.65 to 15.39±5.27, MI-LE from 54.90±17.02 to 66.68±14.75, and FAC from 1.65±0.66 to 2.42±0.85 (all p<0.001). Under the Bonferroni-corrected threshold (p<0.0125), all improvements in the NSG and the improvements in BBS, K-TIS, and MI-LE in the SG remained significant, whereas the FAC improvement in the SG did not.

The between-group comparison of functional gains revealed significantly greater improvements in the NSG across all outcome measures compared to the SG (Table 2). Because baseline FAC differed significantly between the two groups, the 3-month change scores were compared after adjusting for the corresponding baseline score. The between-group differences remained significant for all four outcomes after adjustment, including ΔFAC (F[1, 46]=8.69, p=0.005), and all four remained significant under the Bonferroni-corrected threshold of 0.0125; the baseline-adjusted mean changes with 95% CIs are presented in Table 3. For the BBS, K-TIS, and MI-LE, the homogeneity of the regression slopes (p=0.30 to 0.62) and the normality of the standardized residuals (Shapiro–Wilk, p=0.21 to 0.56) were sat-

Table 2. Between-group comparison of improvements in outcome measures over 3 months (n=49)

	SG (n=18)	NSG (n=31)	p-value	r
ΔBBS	6.89±3.79	12.48±7.44	0.008**	0.38
Median (IQR)	6.00 (4.75–9.25)	13.00 (7.00–18.00)		
ΔK-TIS	3.78±3.67	6.97±3.58	0.002**	0.44
Median (IQR)	2.00 (1.00–5.75)	7.00 (4.00–10.00)		
ΔMI-LE	4.50±4.20	11.77±9.05	0.008**	0.38
Median (IQR)	5.00 (0.00–6.75)	12.00 (5.00–20.00)		
ΔFAC	0.22±0.43	0.77±0.76	0.009**	0.37
Median (IQR)	0.00 (0.00–0.25)	1.00 (0.00–1.00)		

Values are presented as mean±standard deviation, with the median (interquartile range, IQR) in the following line.
r denotes the effect size calculated as |Z|/√N.
SG, sarcopenia group; NSG, non-sarcopenia group; BBS, Berg Balance Scale; K-TIS, Korean-Trunk Impairment Scale; MI-LE, Motricity Index of Lower Extremity; FAC, Functional Ambulation Category.
**p<0.01 by Mann–Whitney U-test.

Table 3. Baseline-adjusted 3-month changes in functional outcomes with ANCOVA

Outcome	Adjusted mean change (95% CI)		Difference (95% CI)	F(1, 46)	p-value	Partial η^2
	SG (n=18)	NSG (n=31)				
BBS	6.72 (3.64 to 9.79)	12.59 (10.26 to 14.91)	5.87 (1.98 to 9.76)	9.23	0.004**	0.167
K-TIS	3.47 (1.77 to 5.17)	7.15 (5.86 to 8.43)	3.68 (1.52 to 5.83)	11.75	0.001**	0.203
MI-LE	3.57 (0.33 to 6.81)	12.32 (9.86 to 14.77)	8.75 (4.65 to 12.85)	18.43	<0.001***	0.286
FAC	0.18 (-0.14 to 0.51)	0.80 (0.55 to 1.04)	0.61 (0.20 to 1.03)	8.69	0.005**	0.159

Values are estimated marginal means of the 3-month change scores from ANCOVA, with the corresponding baseline score as a covariate (Model 1).
The between-group difference is expressed as non-sarcopenia group (NSG) minus sarcopenia group (SG).
Partial η^2 denotes the proportion of variance in the change score accounted for by the group after adjustment for the baseline score.
CI, confidence interval; BBS, Berg Balance Scale; K-TIS, Korean-Trunk Impairment Scale; MI-LE, Motricity Index of Lower Extremity; FAC, Functional Ambulation Category.
**p<0.01 and ***p<0.001.

Table 4. Sensitivity analyses of the between-group difference in 3-month functional gains under stepwise covariate adjustment

Outcome	Model 2[a]		Model 3[b]		Model 4[c]	
	Difference (95% CI)	p-value	Difference (95% CI)	p-value	Difference (95% CI)	p-value
BBS	5.49 (1.45 to 9.53)	0.009**	5.22 (1.13 to 9.31)	0.01*	4.10 (0.38 to 7.83)	0.03*
K-TIS	3.48 (1.25 to 5.71)	0.003**	3.45 (1.14 to 5.77)	0.004**	3.32 (0.81 to 5.83)	0.01*
MI-LE	9.08 (4.83 to 13.33)	<0.001***	9.22 (4.85 to 13.58)	<0.001***	9.17 (4.57 to 13.77)	<0.001***
FAC	0.58 (0.14 to 1.01)	0.010*	0.58 (0.16 to 1.00)	0.008**	0.50 (0.07 to 0.93)	0.02*

The between-group difference is expressed as non-sarcopenia group (NSG) minus sarcopenia group (SG).
CI, confidence interval; BBS, Berg Balance Scale; K-TIS, Korean-Trunk Impairment Scale; MI-LE, Motricity Index of Lower Extremity; FAC, Functional Ambulation Category.
[a]Model 2: estimated marginal means from analysis of covariance including the corresponding baseline score and the stroke duration as covariates (1 and 45 degrees of freedom).
[b]Model 3: the Model 2 covariates with age, sex, and stroke type added (1 and 42 degrees of freedom).
[c]Model 4: the Model 3 covariates with stroke laterality, the Mini-Mental State Examination score, and the National Institutes of Health Stroke Scale score added (1 and 39 degrees of freedom).
*p<0.05, **p<0.01, and ***p<0.001.

isfied, whereas the error variance was unequal for the BBS and MI-LE (Levene, p=0.008 and p<0.001), with the larger variance in the larger group. For the FAC, all three assumptions were violated (p=0.013, 0.002, and 0.012), as expected for an ordinal six-level scale. Re-estimation with heteroscedasticity-consistent standard errors left the adjusted differences unchanged, with slightly narrower CIs (all p≤0.004). In sensitivity analyses, the smaller gains in the SG relative to the NSG remained significant for all four outcomes under Models 2 to 4 (Table 4). Under Model 4 the estimated differences were somewhat smaller, with direction and significance unchanged. When BMI, serum albumin, the GNRI, dysphagia, diabetes mellitus, or the comorbidity count was entered individually into the baseline-adjusted model, the between-group difference likewise remained significant for every outcome, with adjusted differences ranging from 5.85 to 6.67 for BBS, 3.61 to 4.09 for K-TIS, 8.65 to 9.46 for MI-LE, and 0.48 to 0.63 for FAC (all p≤0.04). When the cohort was divided at the median stroke duration (≤21 days vs. >21 days), the NSG showed larger 3-month gains than the SG for all four outcomes in both subgroups; the between-group difference reached significance for BBS in the early-assessed subgroup and for K-TIS, MI-LE, and FAC in the late-assessed subgroup, with the remaining comparisons limited by small subgroup sizes. In the category-based analyses, the proportions reaching each threshold were consistently higher in the NSG than in the SG (Table 5): improvement by at least one FAC category, an FAC of 3 or above, a BBS of 41 or above, and the reported minimal important change in trunk control all differed significantly (p=0.003 to 0.03), whereas the proportions reaching an FAC of 4 or above and the reported minimal clinically important difference (MCID) for the BBS did not.

DISCUSSION

The present study examined the association between sarcopenia assessed at rehabilitation admission and functional recovery over 3 months in subacute stroke patients with balance impairment. Baseline balance ability, trunk control, and lower extremity muscle strength were statistically comparable between the two groups, whereas the magnitude of functional gains (Δ) after

Table 5. Proportions of patients reaching predefined thresholds of clinically meaningful change at 3 months

	SG (n=18)	NSG (n=31)	Difference, % (95% CI)	p-value
Improvement in FAC by ≥1 category	4 (22.2)	18 (58.1)	35.8 (7.1 to 56.2)	0.02*
FAC ≥3 at 3 months	3 (16.7)	18 (58.1)	41.4 (13.0 to 60.3)	0.007**
FAC ≥4 at 3 months	0 (0)	1 (3.2)	3.2 (-14.6 to 16.2)	>0.999
BBS ≥41 at 3 months	2 (11.1)	14 (45.2)	34.1 (7.1 to 52.9)	0.03*
BBS gain ≥5 points	14 (77.8)	27 (87.1)	9.3 (-11.4 to 33.6)	0.44
K-TIS gain ≥2.5 points	8 (44.4)	27 (87.1)	42.7 (15.6 to 64.0)	0.003**
FAC score distribution				
FAC 1, baseline/3 months	14/13	14/6		
FAC 2, baseline/3 months	4/2	14/7		
FAC 3, baseline/3 months	0/3	3/17		
FAC 4, baseline/3 months	0/0	0/1		

Values are presented as number (%).
All proportions were compared using Fisher's exact test, and the between-group difference in proportions, expressed as non-sarcopenia group (NSG) minus sarcopenia group (SG), is reported with its 95% confidence interval (CI) calculated by the Newcombe hybrid score method. A Functional Ambulation Category (FAC) of 3 or above denotes ambulation with supervision or better, and an FAC of 4 or above independent ambulation on level surfaces. A Berg Balance Scale (BBS) of 41 or above was the threshold used to define independence in balance in the present study [19]; a BBS gain of 5 points and a Korean-Trunk Impairment Scale (K-TIS) gain of 2.5 points correspond to the reported minimal clinically important difference and minimal important change, respectively [24,25]. Validated thresholds have not been established for the Motricity Index of Lower Extremity, which was therefore not included. No patient deteriorated in FAC category; the improvement was confined to one category in the SG, whereas 12 patients in the NSG improved by one category and 6 by two. Baseline and 3-month distributions represent the same 18 SG and 31 NSG patients assessed at each time point rather than a matched transition table.
*p<0.05 and **p<0.01.

3 months of standard rehabilitation was significantly greater in the NSG. Low muscle mass and strength at the beginning of inpatient rehabilitation therefore appear to be associated with smaller functional recovery during the subacute phase after stroke.

Balance and trunk control are essential prerequisites for independent ambulation and activities of daily living after stroke, and the NSG showed greater improvement in both, accompanied by corresponding gains in gait. Given that sarcopenia represents a generalized decline in skeletal muscle mass and strength, the reduced recovery of trunk control observed in the SG may reflect the concomitant involvement of trunk musculature, consequently limiting the recovery of overall motor function. These findings may also reflect the diminished myokine secretion capacity in patients with sarcopenia, which could impair the physiological environment for neuroplasticity and thereby reduce the efficiency of translating neurological recovery into functional gains. The magnitude of the observed gains was further considered in relation to reported thresholds for clinically meaningful change. An improvement of 5 points has been reported as the MCID for the BBS in patients with early subacute stroke requiring assisted walking, based on a receiver operating characteristic analysis anchored to an FAC increase of ≥1 point [24], and for the K-TIS, a minimal important change of about 2.5 points has been described in subacute stroke patients requiring assistance with ambulation [25]. In the present cohort, the mean gains exceeded these thresholds in both groups, but the improvements in the NSG were nearly twice those in the SG (Table 2). The between-group difference itself was of comparable magnitude: the baseline-adjusted difference of 5.87 points for the BBS and 3.68 points for the K-TIS exceeded the reported thresholds of 5 and 2.5 points, although the lower limit of each CI lay below the corresponding threshold, so that the clinical relevance of the difference cannot be regarded as established on this basis alone. For gait, no formal MCID for the FAC has been established, but a ≥1-point increase has been used as the anchor criterion for meaningful improvement in balance outcomes in this population [24]; the mean FAC gain in the SG (0.22) fell well below this level, whereas the gain in the NSG (0.77) approached one category. A validated threshold has not been established for the MI-LE [26], which was therefore interpreted only in relative terms. The same pattern was evident at the level of the individual patient, with fewer patients in the SG reaching each threshold (Table 5). The proportion exceeding the 5-point BBS threshold did not differ between the groups, as this threshold lies well below the mean gains in both. Taken together, these comparisons suggest that sarcopenia at rehabilitation admission was associated with a reduction not only in the statistical magnitude but also in the clinical relevance of functional recovery, most evidently for gait.

The association between sarcopenia and poorer functional recovery after stroke is already established; the present study addresses a more specific question in two respects. First, the observation period was extended to 3 months to capture functional recovery spanning the entire subacute phase. Second, by incorporating the BBS and the K-TIS beyond gait indices, the association of sarcopenia with the fundamental mechanisms of postural control was examined more specifically.

Muscle mass and strength were assessed at rehabilitation admission rather than before stroke onset, and several observations indicate that the between-group difference in these measures was not simply a function of the timing of assessment. The stroke duration did not differ significantly between the groups (p=0.12), indicating that the between-group differences in grip strength and SMI are unlikely to reflect differential durations of post-stroke disuse. Baseline stroke severity, as reflected by the NIHSS, was likewise comparable between the two groups (p=0.36), suggesting that the between-group difference in recovery was independent of the initial neurological deficit. Nutritional indices, including BMI, serum albumin, and the GNRI, likewise did not differ significantly between the groups, whereas diabetes mellitus was more prevalent and the overall comorbidity count was higher in the SG, a pattern consistent with the accumulation of chronic disease that accompanies loss of muscle mass. Furthermore, the smaller gains in the SG persisted across all sensitivity analyses (Table 4) and when the nutritional indices, diabetes mellitus, and the comorbidity count were each taken into account, indicating that the association was independent of these covariates. When the cohort was stratified by stroke duration, the direction of the association was consistent across both the early- and late-assessed subgroups, although statistical power was limited within strata. These findings do not, however, establish that the deficits antedated the stroke, because muscle status before onset was not available; the exposure examined here is therefore best described as sarcopenia at rehabilitation admission.

The present study has several limitations. First, the classification was based on unaffected-side grip strength and SMI obtained at rehabilitation admission. The unaffected side, however, is not entirely spared after stroke, as generalized disuse can also reduce muscle strength beyond the paretic limbs [16], and grip strength on this side may be further influenced by age, bilateral central effects of the stroke, fatigue, comorbid illness, and deconditioning during hospitalization. Handedness was not documented, and grip strength was measured on the non-paretic side irrespective of dominance (right in 27 and left in 22 patients); because grip strength may differ between dominant and non-dominant hands, misclassification cannot be excluded near the sex-specific cut-offs. Because these assessments were conducted at a mean of 24 to 34 days after onset, the measured values inevitably reflect a combination of pre-stroke muscle status and early post-stroke changes, including disuse atrophy, denervation, endocrine dysfunction, and inflammatory processes, and therefore cannot fully isolate the pre-existing component [27,28]. Nevertheless, unaffected-side grip strength correlates with the strength of other ipsilateral muscles in hemiparetic stroke and serves as a simple indicator of overall muscle strength when the paretic side cannot be reliably assessed [29]. It has likewise served as the comparison measure in studies of other indirect markers, such as temporal muscle thickness on cranial imaging [30], although no such measure can substitute for direct assessment before stroke onset. Estimates of muscle mass derived from BIA are also sensitive to body water distribution, and because the measurements were obtained under the routine clinical conditions described above, some misestimation of SMI cannot be excluded. Only one patient in each group had chronic kidney disease, and no patient was receiving dialysis. Because fluid retention tends to increase BIA-derived estimates of muscle mass, any such misestimation, if non-differential, would be expected to attenuate rather than exaggerate the between-group difference in SMI. Although the eligibility window extended to 6 months post-onset, the actual stroke duration was short in most patients (median, 21 days), and the timing of assessment still varied. Future studies should assess sarcopenia at a consistent and earlier time point, ideally in the hyperacute phase, to obtain more precise data on pre-stroke muscle status.

Second, baseline FAC scores differed significantly between the two groups (p=0.02); the group difference in ΔFAC nevertheless remained significant after adjustment for the baseline score, indicating that the association was independent of this imbalance. Third, the eligibility requirements may have narrowed the functional range of the analyzed cohort. Requiring at least 3 months of inpatient rehabilitation excluded patients who were discharged early because they could not tolerate the program or developed medical complications, which may have disproportionately affected those with more severe sarcopenia. The balance-related criteria narrowed the range at both ends: patients with a BBS score of 41 or above were excluded because of functional independence in balance, while those unable to maintain standing balance for BIA, who are likely to have been

the most severely impaired, were also excluded, together confining the sample to an intermediate band of balance function. The analyzed cohort therefore represents a narrowly defined subgroup, which limits the generalizability of the findings; prospective studies with consecutive enrollment and a predefined follow-up window are warranted.

Fourth, the sample size and statistical power were limited. As a retrospective study, the sample size was determined by the number of eligible patients rather than by an a priori power calculation, and the small number of patients in the SG restricted the number of covariates that could be included in a single model. For this reason, the primary analysis was confined to baseline adjustment, with the remaining variables incorporated only in sensitivity analyses. Although the between-group differences remained significant under these adjustments, the present findings should be regarded as preliminary and warrant confirmation in adequately powered prospective studies. In addition, the covariance model for the FAC did not satisfy its distributional assumptions, as expected for an ordinal scale, and the gait findings therefore rest primarily on the category-based analysis. The thresholds applied in that analysis were taken from published estimates of clinically meaningful change; validated minimal detectable change values, which express measurement error, have not been established for the K-TIS or the MI-LE, and the proportions reaching a threshold should therefore be read as descriptive rather than as evidence that the change exceeded measurement error. Fifth, several potential confounders could not be examined at all. Pre-stroke functional status, formal frailty assessment, lesion location and volume, unilateral spatial neglect, ataxia, sensory impairment, spasticity, and depression were not documented in a standardized form and could not be entered into the models. Because each influences the recovery of balance and gait after stroke, residual confounding cannot be excluded, and prospective studies with standardized documentation are needed.

Sixth, individual therapy adherence and nutritional intake were not quantified. Although therapy was delivered according to a uniform institutional protocol, the number of sessions actually delivered, the intensity, progression, and repetitions of the resistance exercises, and daily caloric and protein intake were not recorded because of the retrospective design. Nutritional status could be characterized only by the admission BMI, serum albumin, and GNRI, which do not capture changes in intake over the 3-month program. Residual confounding by differences in rehabilitation dose or nutritional intake therefore cannot be excluded, although the structured inpatient setting is likely to have limited such differences; prospective studies with standardized monitoring are warranted. Seventh, the assessors were not blinded to the admission grip strength and BIA values, and formal inter-rater reliability testing was not undertaken; the risk of differential measurement error is nevertheless likely to be limited, as the group assignment did not exist at the time of assessment.

In conclusion, sarcopenia assessed at rehabilitation admission was associated with smaller 3-month gains in trunk control, balance, lower extremity muscle strength, and gait in this selected cohort of subacute stroke inpatients. Despite both groups receiving the same standard rehabilitation program, the lower functional recovery observed in the SG suggests that conventional rehabilitation protocols alone may be insufficient to fully compensate for the functional deficits associated with sarcopenia. Because no specific intervention was evaluated, tailored resistance exercise and nutritional support are proposed only as directions for future controlled trials rather than as clinical recommendations. Screening for sarcopenia at rehabilitation admission may therefore be useful in early prognostic assessment. Prospective studies with pre-stroke or hyperacute assessment of muscle status will be required to determine whether premorbid sarcopenia independently predicts recovery.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

None.

AUTHOR CONTRIBUTION

Conceptualization: Kim K, Choi SJ. Data curation: Kim K, Kim H, Lee J. Methodology: Kim K, Kim H, Choi SJ. Formal analysis: Kim K, Kim H. Investigation: Kim H, Lee J. Project administration: Kim K. Visualization: Kim K, Kim H. Validation: Kim K, Kim H. Supervision: Choi SJ. Writing – original draft: Kim K. Writing – review and editing: Kim K, Kim H, Choi SJ. Approval of final manuscript: all authors.

ORCID

Kyungmin Kim, https://orcid.org/0000-0003-4507-4022
Hangi Kim, https://orcid.org/0009-0008-8399-216X
Jiwon Lee, https://orcid.org/0009-0000-9128-697X
Su Jin Choi, https://orcid.org/0000-0002-3443-2237

REFERENCES

1. Li S. Stroke recovery is a journey: prediction and potentials of motor recovery after a stroke from a practical perspective. Life (Basel) 2023;13:2061.
2. Yoo DY, Choi JK, Baek CY, Shin JB. Impact of intensive rehabilitation on long-term prognosis after stroke: a Korean nationwide retrospective cohort study. Medicine (Baltimore) 2022;101:e30827.
3. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019;48:16-31.
4. Kirk B, Cawthon PM, Arai H, Ávila-Funes JA, Barazzoni R, Bhasin S, et al.; Global Leadership Initiative in Sarcopenia (GLIS) group. The conceptual definition of sarcopenia: Delphi consensus from the Global Leadership Initiative in Sarcopenia (GLIS). Age Ageing 2024;53:afae052.
5. Su Y, Yuki M, Otsuki M. Prevalence of stroke-related sarcopenia: a systematic review and meta-analysis. J Stroke Cerebrovasc Dis 2020;29:105092.
6. Celis-Morales CA, Welsh P, Lyall DM, Steell L, Petermann F, Anderson J, et al. Associations of grip strength with cardiovascular, respiratory, and cancer outcomes and all cause mortality: prospective cohort study of half a million UK Biobank participants. BMJ 2018;361:k1651.
7. Chin SO, Rhee SY, Chon S, Hwang YC, Jeong IK, Oh S, et al. Sarcopenia is independently associated with cardiovascular disease in older Korean adults: the Korea National Health and Nutrition Examination Survey (KNHANES) from 2009. PLoS One 2013;8:e60119.
8. Moreira-Pais A, Ferreira R, Oliveira PA, Duarte JA. A neuromuscular perspective of sarcopenia pathogenesis: deciphering the signaling pathways involved. Geroscience 2022;44:1199-213.
9. Miao Y, Xie L, Song J, Cai X, Yang J, Ma X, et al. Unraveling the causes of sarcopenia: roles of neuromuscular junction impairment and mitochondrial dysfunction. Physiol Rep 2024;12:e15917.
10. Gao X, Chen Y, Cheng P. Unlocking the potential of exercise: harnessing myokines to delay musculoskeletal aging and improve cognitive health. Front Physiol 2024;15:1338875.
11. Inoue DS, Janini Gomes M. Integrative insights into PNI: low-grade chronic inflammation, skeletal muscle wasting, and brain impairments. Brain Behav Immun Health 2024;40:100838.
12. Kim SY, Cho WS, Park CB, Kim BG. Effect of sarcopenia on functional recovery in acute stroke patients admitted for standard rehabilitation program. Medicina (Kaunas) 2024;60:1716.
13. Park JG, Lee KW, Kim SB, Lee JH, Kim YH. Effect of decreased skeletal muscle index and hand grip strength on functional recovery in subacute ambulatory stroke patients. Ann Rehabil Med 2019;43:535-43.
14. Chen LK, Woo J, Assantachai P, Auyeung TW, Chou MY, Iijima K, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020;21:300-7.e2.
15. Maekawa K, Yoshimura Y, Nagano F, Matsumoto A, Hori K, Shimazu S, et al. Site-specific skeletal muscle mass and functional prognosis in geriatric stroke patients. J Stroke Cerebrovasc Dis 2024;33:108049.
16. Chon J, Soh Y, Shim GY. Stroke-related sarcopenia: pathophysiology and diagnostic tools. Brain Neurorehabil 2024;17:e23.
17. Bouillanne O, Morineau G, Dupont C, Coulombel I, Vincent JP, Nicolis I, et al. Geriatric Nutritional Risk Index: a new index for evaluating at-risk elderly medical patients. Am J Clin Nutr 2005;82:777-83.
18. Monroe T, Carter M. Using the Folstein Mini Mental State Exam (MMSE) to explore methodological issues in cognitive aging research. Eur J Ageing 2012;9:265-74.
19. Joa KL. Outcome measurement in balance problems: Berg Balance Scale. Ann Rehabil Med 2024;48:103-4.
20. Miyata K, Tamura S, Kobayashi S, Takeda R, Iwamoto H. Berg Balance Scale is a valid measure for plan interventions and for assessing changes in postural balance in patients with stroke. J Rehabil Med 2022;54:jrm00359.
21. Ko J, You Y. Reliability and responsiveness of the Korean version of the Trunk Impairment Scale for stroke patients. J Kor Phys Ther 2015;27:175-82.
22. Mehrholz J, Wagner K, Rutte K, Meissner D, Pohl M. Predictive validity and responsiveness of the Functional Ambulation Category in hemiparetic patients after stroke. Arch Phys Med Rehabil 2007;88:1314-9.
23. Yoo YJ, Lim SH. Assessment of lower limb motor function, ambulation, and balance after stroke. Brain Neurorehabil 2022;15:e17.
24. Tamura S, Miyata K, Kobayashi S, Takeda R, Iwamoto H. The minimal clinically important difference in Berg Balance Scale scores among patients with early subacute stroke: a multicenter, retrospec-

tive, observational study. Top Stroke Rehabil 2022;29:423-9.
25. Yamamoto N, Igarashi T, Sato R, Yamaoka H, Maruyama S. Estimating the responsiveness and minimal important change of the Trunk Impairment Scale in inpatients with subacute stroke requiring dependent ambulation. Cureus 2025;17:e82275.
26. Lin C, Arevalo YA, Harvey RL, Prabhakaran S, Martin KD. The minimal clinically important difference of the motricity index score. Top Stroke Rehabil 2023;30:298-303.
27. Ramsay JW, Barrance PJ, Buchanan TS, Higginson JS. Paretic muscle atrophy and non-contractile tissue content in individual muscles of the post-stroke lower extremity. J Biomech 2011;44:2741-6.
28. Kim DH, Cho ES, Park YS, Chang HJ, Park JG, Kim JY, et al. Changes in lower extremity muscle quantity and quality in patients with subacute stroke. Ann Rehabil Med 2023;47:493-501.
29. Takahashi J, Nishiyama T, Matsushima Y. Does grip strength on the unaffected side of patients with hemiparetic stroke reflect the strength of other ipsilateral muscles? J Phys Ther Sci 2017;29:64-6.
30. Park J, Park J, Kim S, Kim DC. Correlation between temporal muscle thickness and grip strength in hemiplegic patients with acute stroke. Front Neurol 2023;14:1252707.

Original Article

Ann Rehabil Med 2026;50(5):389-400
eISSN: 2234-0653
https://doi.org/10.5535/arm.260097

arm
Annals of Rehabilitation Medicine

Reliability and Concurrent Validity of the Motor Evaluation Scale for Upper Extremity in People With Chronic Stroke

Shamay S.M. Ng, PhD[1,2], Peiming Chen, PhD[1], Siyue Li, PhD[1], Shiu Hang Chan, BSc[1], Yin Tung Chan, BSc[1], Man Chi Cheng, BSc[1], Chung Him Cheung, BSc[1], Yau Ming Cheung, BSc[1], Mimi M.Y. Tse, PhD[3], Cynthia Y.Y. Lai, PhD[1]

[1]*Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China*
[2]*Research Centre for Chinese Medicine Innovation, The Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR, China*
[3]*School of Nursing and Health Sciences, Hong Kong Metropolitan University, Ho Man Tin, Hong Kong SAR, China*

Objective: To evaluate the reliability and concurrent validity of the Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES) and establish a cut-off score distinguishing people with chronic stroke from healthy older adults.

Methods: This cross-sectional study included 64 individuals with chronic stroke and 31 healthy older adults at a university-based rehabilitation center. Assessments included MESUPES, Fugl-Meyer Assessment for Upper Extremity (FMA-UE), muscle strength and stiffness, wrist range of motion, Action Research Arm Test (ARAT), Quick Disabilities of the Arm, Shoulder, and Hand (QuickDASH), and Oxford Participation and Activities Questionnaire.

Results: MESUPES total scores demonstrated excellent inter-rater (intraclass correlation coefficient [ICC]=0.976; 95% confidence interval [CI]: 0.960–0.985) and test–retest reliability (ICC=0.965; 95% CI: 0.943–0.978). A ceiling effect, but no floor effect, was observed; sensitivity analyses confirmed that excellent reliability was not solely attributable to ceiling effects. MESUPES scores correlated with FMA-UE (r=0.872–0.922), ARAT (r=0.821–0.885), affected elbow strength (r=0.585–0.661), wrist range of motion (dorsal–palmar: r=0.727–0.751; ulnar–radial: r=0.750–0.817), handgrip strength (r=0.674–0.783), and QuickDASH (r=-0.433 to -0.521; all p<0.001). A cut-off score of 57.5 differentiated participants with stroke from healthy older adults, with 93.5% sensitivity and 82.8% specificity.

Conclusion: MESUPES demonstrated excellent reliability and concurrent validity for assessing upper-limb movement quality in chronic stroke. The cut-off score of 57.5 discriminated between participants with stroke and healthy older adults. Associations with established impairment- and activity-level measures support its use as a complementary assessment in chronic stroke rehabilitation.

Keywords: Reliability and validity, Outcome assessment, Upper extremity

Received: April 23, 2026
Revised: September 9, 2026
Accepted: October 13, 2026

Correspondence:
Shamay S.M. Ng
Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, QT515, 5/F, Core S, Hung Hom, Hong Kong SAR, China.
Tel: +852 2766 4889
E-mail: shamay.ng@polyu.edu.hk

© 2026 by Korean Academy of Rehabilitation Medicine

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Stroke is the leading cause of long-term disability among adults, with over 94 million stroke survivors worldwide [1,2]. Stroke incidence increases one-fold each decade after 55 years of age [3]. Paresis is the most common impairment after stroke [4], with more than 70% of the people with stroke suffering from upper-limb paresis, most commonly presenting unilaterally as motor weakness [5]. It generally limits functional movement including reaching, grasping, and manipulating objects using the affected upper limb [6]. Post-stroke upper-limb dysfunctions have long-term effects on activities of daily living, occupation, and leisure activities, adversely affecting their quality of life. Effective assessment of upper-limb function is essential for establishing rehabilitation goals and tracking recovery.

While systematic reviews and consensus statements highly recommend tools such as the Fugl-Meyer Assessment for Upper Extremity (FMA-UE) and the Action Research Arm Test (ARAT) for evaluating motor impairment and functional capacity, these measures often prioritize completion time or task success over the quality of movement. This is a clinically important gap, as two patients may achieve the same task score using markedly different movement strategies, with one using efficient, well-coordinated motor patterns, and another relying on compensatory trunk displacement or abnormal muscle synergies. As a result, a distinction that capacity-based tools cannot capture. The Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES) was first developed to address the need for a comprehensive tool that assesses both the quality of active and passive upper-limb movements in stroke patients [7]. Compared to the FMA-UE and the Wolf Motor Function Test (WMFT), the MESUPES requires less time to complete, typically taking between 5 to 15 minutes. Additionally, while the ARAT and WMFT require standardized equipment, the MESUPES uses more readily accessible equipment, enhancing its feasibility for use in routine clinical and community settings.

The scale has demonstrated high reliability, with excellent relative reliability (intraclass correlation coefficient [ICC]=0.98), absolute reliability (standard error of measurement=2.68), and internal consistency (Cronbach's alpha ranging from 0.97 to 0.99). Its internal construct validity was evaluated in the Rasch framework by Van de Winckel et al. [8], who suggested removing five items from the original scale and separating the two subscales for arm and hand. The scale has shown excellent construct validity through strong correlations with arm scores on the Modified Motor Assessment Scale (Spearman's rho=0.87) [7], as well as excellent convergent validity when compared with the Stroke Upper Limb Capacity Scale (correlation coefficients ranging from 0.87 to 0.91) [9]. The minimal detectable change at the 95% confidence interval was determined to be 7.43 points, indicating that an improvement of at least 8 points is needed to reflect a true functional change [7].

However, most prior psychometric evaluations of the MESUPES were conducted in patients with subacute stroke [7,8], a period during which spontaneous neurological recovery and rapid changes in motor capacity are still occurring and may confound the interpretation of reliability and validity findings. As chronic stroke survivors represent the majority of the global stroke population and are increasingly managed in community-based rehabilitation settings, a validated, quality-focused outcome measure with established reliability and concurrent validity in this population is needed to guide long-term goal-setting, monitor subtle functional changes, and support decision-making regarding continued therapy provision. Building on this rationale, despite the unique ability of the MESUPES to measure movement quality in patients with stroke, further evaluation of its psychometric properties is necessary to strengthen its evidence-based implementation. Specifically, its test-retest reliability and the correlation of the scale's scores with recommended post-stroke outcome measures and stroke-specific impairments have not been examined in patients with chronic stroke. Establishing these correlations is particularly important because movement quality, as captured by the MESUPES, is theoretically influenced by multiple underlying impairment-level deficits commonly observed after stroke, including reduced muscle strength, altered muscle stiffness/spasticity, and restricted joint range of motion (ROM) [10]. Understanding the strength and pattern of these relationships would clarify whether the MESUPES captures a construct distinct from, yet related to, these impairments, thereby supporting its interpretability alongside standard impairment-based measures in clinical practice. Furthermore, given that residual spasticity, weakness, and joint restrictions often persist into the chronic phase and continue to influence functional movement patterns [11,12], examining these associations specifically in chronic stroke patients would help clinicians determine whether the MESUPES can serve as a sensitive indicator of ongoing impairment-related changes in movement quality, beyond what is captured by capacity-based measures such as the FMA-UE or ARAT.

In addition, our study focuses on more chronic stroke pa-

tients whose stroke onset was more than 1 year prior, and who were community-dwelling when compared to previous studies [7,8]. Taken together, this study extends the psychometric evidence for the MESUPES to the chronic stroke population, addressing a critical gap not covered by prior subacute validation work. Specifically, we seek to: (1) determine the test-retest and inter-rater reliabilities of the MESUPES, (2) examine the ceiling and floor effects, (3) investigate concurrent validity by evaluating the correlations between the scale's scores and established outcome measures, and (4) establish a cut-off score to differentiate individuals with stroke from healthy older adults.

METHODS

Study design

This was a psychometric validation study employing a longitudinal design with repeated measurements (Assessment 1 and Assessment 2, seven days apart) conducted in the Wellness and Exercise Laboratory of The Hong Kong Polytechnic University. The study was approved by the Ethics Committee of The Hong Kong Polytechnic University (approval no. HSEARS20230807001) and was conducted in accordance with the Declaration of Helsinki. All subjects were informed of the objectives and procedures of the study, and written consent was obtained from them on the first day of the study.

Sample size

Sample size calculations were tailored to the specific psychometric aims of the study. The total score of the MESUPES had an ICC of 0.98 for test-retest reliability in a previous study with people with subacute stroke [7]. For test-retest reliability, calculations were based on the desired confidence interval width for ICCs. Assuming a minimum acceptable ICC of 0.85 and an expected ICC of 0.95, a minimum of 40 subjects was required [13]. To establish concurrent validity, assuming an expected moderate-to-strong correlation magnitude (ρ=0.4) between the scale and established measures, an alpha error probability of 0.05 (two tails), and a power of 0.8, the minimum sample size required was 64. The sample size was calculated by G*Power (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf).

Setting and participants

Sixty-four community-dwelling individuals with stroke were recruited from the Hong Kong Stroke Association, and 31 healthy older adults were recruited from the local community through invitation. Individuals with stroke were included if they (1) were aged 55–80 years, (2) were diagnosed with stroke for more than 1 year, (3) had at least minimal antigravity upper-limb motor function (defined as the ability to actively achieve at least 15 degrees of shoulder flexion or abduction against gravity); only patients presenting with unilateral upper-limb motor impairment (hemiparesis) were included; individuals with bilateral upper-limb involvement (quadriparesis) were excluded, (4) had normal cognitive function with an Abbreviated Mental Test (AMT) score >7, and (5) could understand and follow simple instructions, as determined by the participant's ability to correctly follow standardized one- and two-step verbal commands administered by the assessor during the AMT and prior to formal assessment (e.g., "raise your arm," "open your hand and then close it"). Participants with significant receptive aphasia or other communication impairments that precluded comprehension of test instructions were excluded. Subjects were excluded if they had any pre-existing (1) neurological disorders other than stroke (e.g., Parkinson's disease, multiple sclerosis, peripheral neuropathy, traumatic brain injury), (2) cardiopulmonary conditions that would preclude safe participation in physical assessment (e.g., unstable angina, uncontrolled hypertension, chronic obstructive pulmonary disease with exertional limitation), or (3) musculoskeletal disorders affecting the upper limb (e.g., pre-stroke shoulder pathology such as rotator cuff tear or frozen shoulder, upper-limb fracture, or rheumatoid arthritis) that would affect the assessment results.

Main measures

The MESUPES is the first outcome measure to assess the quality of upper-limb functional movements in patients with stroke [7]. Based on previous Rasch analysis recommendations, the scale is divided into two subscales, the arm test and the hand test, comprising eight items (40 points) and nine items (18 points), respectively (total score range: 0–58). Items 1–4 are performed in a supine position, while the others are performed in a supported sitting position. For each task, scoring is based on muscle tone, muscle contraction, and proper coordination of joints to finish the task. Patients are graded on a 6-level ordinal scale (0–5) for the arm test and a 3-level ordinal scale (0–2) for the hand test. For the arm test, patients are scored based on passive, assisted, and active execution of actions. For the hand test, subjects were required to actively execute hand movements without external help. A maximum of 3 trials was allowed for each test component. A higher score indicates better quality of

upper-limb movement. The scale score has demonstrated excellent inter-rater reliability (ICC=0.95–0.97) [9]. The validated Hong Kong-Chinese version of the scale was used. The scoring sheet is provided in Supplementary Material S1.

According to the International Classification of Functioning, Disability and Health model, the body function and body structure was assessed by the FMA-UE (score range: 0–66) [14], upper limb muscle strength, upper limb muscle stiffness, and wrist ROM. The activity level was assessed by the ARAT (score range: 0–57) [15], and Quick Disabilities of the Arm, Shoulder, and Hand (QuickDASH) (score range: 0–100) [16]. The participation level was assessed by Oxford Participation and Activities Questionnaire (OxPAQ) (score range: 0–100) [17]. All of these outcome measures showed good to excellent test-retest reliability [17-23]. Validated Hong Kong-Chinese language versions were used for all self-reported questionnaires to ensure appropriate psychometric properties for this population. The detailed instruction can be found in the Supplementary Material S2.

The FMA-UE is designed to assess upper-limb motor control in people with stroke [14]. It assesses movements of the shoulder, elbow, wrist, and hand and the coordination of the upper limb in a sitting position. Patients are graded on a 3-level ordinal scale (0–2), yielding a total score range of 0–66, with a higher score indicating better motor functions. The assessment has excellent inter-rater and intra-rater reliabilities (ICC=0.98) [18].

The isometric muscle strengths of the biceps brachii and triceps brachii were assessed using a Nicholas handheld dynamometer (Model 01165A; Lafayette Instrument Company) (Fig. 1A). The average of three maximum isometric voluntary contraction strengths was recorded. Isometric muscle strength has demonstrated excellent inter-rater reliability (ICC=0.90–0.99) and test-retest reliability (ICC=0.89–0.99) [19].

Handgrip strength was assessed using a digital spring handgrip dynamometer (Model EH101; Zhongshan Camry Electronic Co. Ltd.) (Fig. 1B). Subjects were instructed to squeeze the handheld dynamometer with maximum effort for 5 seconds. The average handgrip strength across three trials was recorded. The Camry EH101 electronic hand dynamometer has an excellent reliability (ICC=0.815–0.854) [20].

Muscle stiffness of the biceps brachii and triceps brachii was measured using a handheld MyotonPRO device (Myoton AS) (Fig. 1C). Stiffness is defined as the resistance of muscle tissue

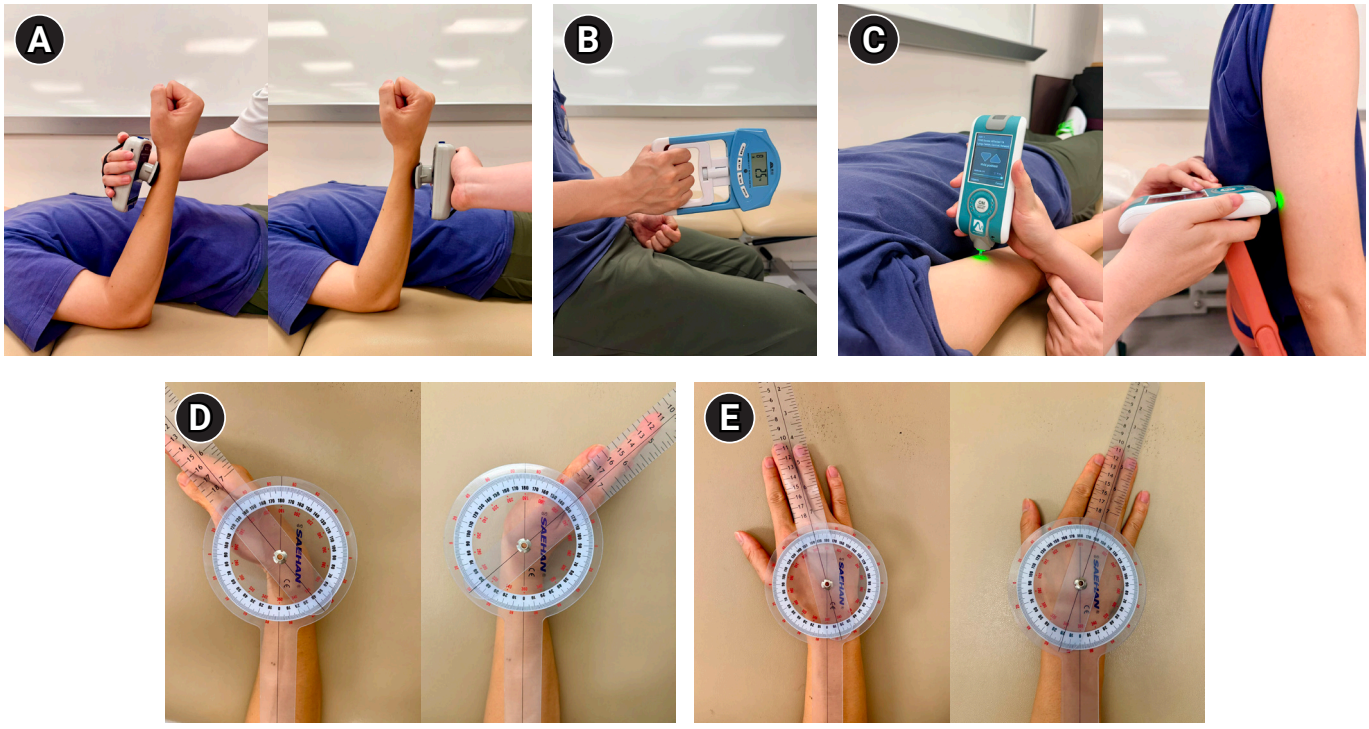

Fig. 1. Measurement techniques for upper extremity evaluation. (A) Measurement of biceps brachii and triceps brachii muscle strength. (B) Measurement of handgrip strength. (C) Measurement of biceps brachii and triceps brachii muscle stiffness. (D) Measurement of wrist flexion and extension range of motion. (E) Measurement of wrist radial and ulnar deviation range of motion.

to passive deformation by an external force. The probe of the device was placed perpendicular to the skin surface directly above the muscle to be measured with a device-set pre-pressure of 0.18 N to slightly compress the subcutaneous tissues. The Myoton demonstrated very good inter-rater reliability (ICC=0.85–0.86) and test-retest reliability (ICC=0.82–0.88) in measuring upper-limb muscle stiffness in stroke survivors [21].

Wrist ROM was measured using a universal goniometer. A dual-arm goniometer with a 360° span in 1° increments was used. The measured wrist flexion and extension ranges were summed to obtain the wrist dorsal–palmar ROM (Fig. 1D), whereas the ROM of wrist ulnar deviation and radial deviation were summed to obtain the ulnar–radial ROM (Fig. 1E).

The ARAT is an observational test of upper-limb function [15]. The test consists of 19 items classified into four subtests, namely grasp, grip, pinch, and gross arm movement. All tasks are conducted in a sitting position. Each item is rated on a 4-level ordinal scale (0–3), yielding a total score range of 0–57, with a higher score demonstrating better performance. The test has shown high inter-rater and test-retest reliabilities (ICC>0.98) [22].

The QuickDASH is a self-reported questionnaire measuring physical functions and symptoms of the upper limb [16]. It consists of 11 questions, with eight related to pain and three related to functions. Patients rate their difficulty in completing daily tasks and interference with daily life on a 5-level ordinal scale (1–5), with a lower score indicating better ability in completing daily tasks. The questionnaire demonstrated excellent test-retest reliability (ICC=0.90) [23]. The validated Hong Kong-Chinese version was used in this study to ensure cultural and linguistic appropriateness.

The OxPAQ is a patient-reported questionnaire assessing the level of participation and activities [17]. The 23 items in the questionnaire are split into three domains: Routine Activities, Emotional Well-being, and Social Engagement. Each item is measured on a 5-level ordinal scale, with a higher score reflecting greater problems in activities and participation. The questionnaire has demonstrated excellent external reliability in all three domains (ICC=0.83–0.96) [17]. The validated Hong Kong-Chinese version was utilized.

Procedures

Inter-rater reliability of the scale for stroke participants was assessed on Assessment 1 by two independent raters (A and B), each with over three years of neurorehabilitation experience. Both had attended two 2-hour training sessions led by a senior specialist with 10+ years of experience. They scored participants simultaneously under identical conditions, without communication. To assess test-retest reliability, Rater A repeated the scale approximately one week later (Assessment 2), with the two assessments intended to be 7 days apart; the exact interval varied slightly for some participants to accommodate scheduling needs. The same setup and instructions were used for both assessments, and this interval minimized fatigue, learning effects, and recovery-related changes. Rater A also conducted all additional clinical and functional assessments for the stroke group on Assessment 1, immediately after the initial assessment. Participants received 5–10 minutes of rest between tasks to avoid fatigue.

Healthy older adults were assessed by one randomly assigned physiotherapist from five trained raters. Procedures mirrored those for stroke participants, ensuring consistency in test order, rest intervals, and environmental conditions. All assessments were conducted in a quiet, climate-controlled lab with standardized lighting and equipment setup. Testing materials and instructions remained consistent across sessions. Assessment schedules for both groups are summarized in Supplementary Table S1.

Statistical analysis

Data analysis was conducted using the SPSS software, version 29.0.2.0 (IBM Corp.). Data obtained from Assessment 1 were used for all cross-sectional comparisons, validity correlation analyses, and inter-rater reliability calculations. Data from Assessment 2 were exclusively used alongside Assessment 1 data to compute test-retest reliability.

Descriptive statistics were used to encapsulate the demographic data of the subjects. The Kolmogorov–Smirnov test was applied to test the normality of the data. For between-group comparisons of categorical demographic variables (e.g., sex, affected side, stroke type) between stroke individuals and healthy controls, the chi-square test was used. For between-group comparisons of MESUPES arm, hand, and total scores between stroke individuals and healthy controls, independent t-tests were used for normally distributed data, and the Mann–Whitney U-test was used for non-normally distributed data. For within-group comparisons between Assessment Day 1 and Assessment Day 2 MESUPES scores in stroke individuals (used to inform test-retest reliability calculations), paired t-tests were used for normally distributed data, and the Wilcoxon signed-rank test was used for non-normally distributed data.

Because the scale scores were based on individual rater as-

sessments rather than averaged ratings across multiple raters, both inter-rater and test-retest reliabilities were estimated using a two-way random-effects model with absolute agreement (ICC $_{2,1}$) [24]. Youden's index was used to determine the optimal cut-off point based on a trade-off between sensitivity and specificity of each cut-off point by plotting the receiver-operating characteristic (ROC) curve [25]. Overall accuracy was determined by the area under the ROC curve to compare test performance between two groups of subjects. Ceiling and floor effects were deemed present if more than 15% of patients scored either the maximum or minimum score, respectively [26]. To investigate whether ceiling effects artificially inflated reliability estimates, a sensitivity analysis was planned to recalculate the ICCs after excluding participants who achieved the maximum score on either the arm or hand subscales. To quantify the correlation of the scale scores with other outcome measures, Pearson's r was used for normally distributed variables (elbow flexion strength, elbow extension strength, and QuickDASH scores), and Spearman's rho was used for non-normally distributed variables (FMA-UE, handgrip strength, muscle stiffness of biceps and triceps brachii, wrist ROM, ARAT, and OxPAQ subscale scores), consistent with the results of the Kolmogorov–Smirnov normality testing.

RESULTS

A total of 64 community-dwelling individuals with stroke (34 males, 30 females; mean age=66.83±6.42 years and mean years since stroke=9.13±5.01 years) and 31 healthy older adults (17 males, 14 females; mean age=64.84±7.85 years) were recruited. No significant differences were found in baseline demographics between the two subject groups (p>0.05) (Table 1). Table 2 summarizes the scale scores of participants. Compared with individuals with stroke, healthy older adults achieved significantly higher scale arm, hand, and total scores (p<0.001) in MESUPES.

A ceiling effect was found in the scale, as 48.4% and 20.3% of stroke survivors achieved the maximum score in arm and hand components, respectively, and 17.2% scored the maximum total score. Regarding the floor effect, none scored zero in the arm component and total score, but 20.3% scored zero in the hand component. Thus, no floor effect was noted in the scale's total scores. The performance profiles from other outcome measures are summarized in Table 3.

The scale exhibited excellent inter-rater and test-retest reliability. The ICCs ranged from 0.944 to 0.984 for inter-rater reliability and from 0.907 to 0.983 for test-retest reliability (Table 4). Subsequent sensitivity analyses, performed after removing participants who achieved the maximum scores on the respective subscales, confirmed that the ICCs remained robustly high (inter-rater ICCs>0.90; test-retest ICCs>0.80). This indicates that the excellent reliability was a true reflection of the scale's stability rather than merely an artifact of the ceiling effect. ROC analysis determined an optimal scale total score cut-off of 57.5 (area under the curve=0.897, sensitivity=93.5%, specificity=82.8%, p<0.001) for differentiating stroke participants from healthy older adults (Fig. 2). Because the scale's items are scored in whole numbers, a clinical cut-off score of 57 or 58 may be more practically applied in clinical settings to indicate near-normal functional quality.

Table 1. Demographic and clinical characteristics of people with stroke and healthy older adults

Characteristic	Stroke (n=64)	Healthy (n=31)	p-value
Age (yr)	66.83±6.42	64.84±7.85	0.090
Sex			0.875
Male	34 (53.13)	17 (54.84)	
Female	30 (46.87)	14 (45.16)	
Height (cm)	162.30±7.99	164.93±8.79	0.153
Weight (kg)	63.00±10.26	64.63±11.70	0.557
Body mass index (kg/m^2)	23.94±3.31	23.77±4.09	0.833
Years since stroke	9.13±5.01	N/A	N/A
Affected side		N/A	N/A
Left	31 (48.44)		
Right	33 (51.56)		
Type of stroke		N/A	N/A
Ischemic	39 (60.94)		
Hemorrhagic	24 (37.50)		
Both	1 (1.56)		

Values are presented as mean±standard deviation or number (%).
N/A, not applicable.

Table 2. Performance of the Motor Evaluation Scale for Upper Extremity in stroke individuals and healthy older adults

Parameter	Stroke individual (n=64)	Healthy control (n=31)	p-value
MESUPES Assessment 1			
Arm score	34.30±8.22	40.00±0.00	<0.001
Hand score	9.80±7.41	17.81±0.91	<0.001
Total score	44.09±14.64	57.81±0.91	<0.001
MESUPES Assessment 2		N/A	N/A
Arm score	33.61±8.85		
Hand score	9.69±7.51		
Total score	43.30±15.04		

Values are presented as mean±standard deviation.
MESUPES, Motor Evaluation Scale for Upper Extremity in Stroke Patients; N/A, not applicable.

Table 5 and Fig. 2 show correlations between the scale scores and stroke-related impairments. The MESUPES scores were significantly and moderately to very strongly correlated with the FMA-UE (r=0.872–0.922, p<0.001), ARAT (r=0.821–0.885, p<0.001), affected elbow strength (r=0.585–0.661, p<0.001), affected wrist ROM (dorsal–palmar: r=0.727–0.751; ulnar–radial: r=0.750–0.817; both p<0.001), handgrip strength (r=0.674–0.783, p<0.001), and QuickDASH (r=-0.433 to -0.521, p<0.001). A weak but significant correlation was also found with the Routine Activities domain of the OxPAQ sub-scores (r=-0.271, p=0.030).

DISCUSSION

This study is the first to comprehensively examine the test-retest reliability, ceiling and floor effects, concurrent validity, and severity cut-offs of the MESUPES in people with chronic stroke. The scale showed excellent inter-rater and test-retest reliability, with a ceiling effect but no floor effect. The scale's scores strongly correlated with recommended capacity measures such as the FMA-UE and the ARAT, as well as affected elbow strength, affected wrist ROM, handgrip strength and the QuickDASH. An optimal cut-off score was established to differentiate the performance between individuals with stroke from healthy older adults.

As predicted, stroke survivors had poorer upper-limb performance than healthy controls. Damage to the descending upper motor neuron tracts typically results in contralateral weakness and spasticity [27]. In patients with spasticity, an exaggerated stretch reflex and prolonged muscle contraction give rise to increased muscle tone and hyperreflexia [28]. Hand spasticity is a common disorder after stroke and often limits mobility, making it difficult for patients to grasp, transfer, and manipulate objects. The moderate correlation between MESUPES scores and muscle stiffness measures likely reflects the complex interplay between spasticity and voluntary motor control in chronic stroke. It is worth clarifying that spasticity and muscle weakness, although related, represent distinct post-stroke impairments [29]. Rather than spasticity directly causing weakness, increased

Table 3. Performance of Upper Extremity Outcomes in people with stroke

Parameter	Stroke individuals (n=64), affected
FMA-UE	45.72±16.97
Strength (kg)	
Elbow flexion	8.65±4.02
Elbow extension	7.19±3.34
Handgrip	12.41±8.43
Muscle stiffness (N/m)	
Biceps brachii	184.03±58.53
Triceps brachii	179.94±43.76
Wrist Range of motion (°)	
Dorsal–palmar	60.75±51.73
Ulnar–radial	35.90±32.88
ARAT	33.06±23.98
QuickDASH	27.58±9.64
OxPAQ	
Routine Activities	19.22±19.21
Social Engagement	11.52±17.30
Emotional Well-being	18.36±20.26

Values are presented as mean±standard deviation.
FMA-UE, Fugl-Meyer Assessment for Upper Extremity; ARAT, Action Research Arm Test; QuickDASH, Quick Disabilities of the Arm, Shoulder, and Hand; OxPAQ, Oxford Participation and Activities Questionnaire.

Table 4. Inter-rater and test-retest reliabilities of the MESUPES subscores and total scores

Reliability type	Subscale	ICC	95% CI	p-value
Inter-rater reliability	Arm	0.944	0.910–0.966	<0.001
	Hand	0.984	0.974–0.990	<0.001
	Total	0.976	0.960–0.985	<0.001
Test-retest reliability	Arm	0.907	0.852–0.943	<0.001
	Hand	0.983	0.971–0.989	<0.001
	Total	0.965	0.943–0.978	<0.001

MESUPES, Motor Evaluation Scale for Upper Extremity in Stroke Patients; ICC, intraclass correlation coefficient; CI, confidence interval.

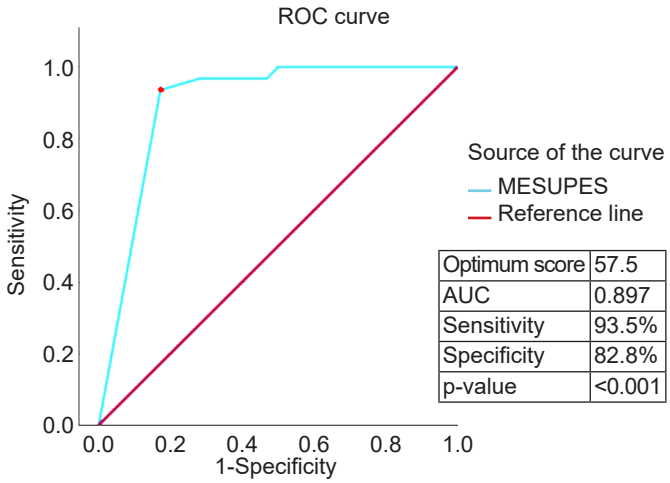

Fig. 2. Receiver-operating characteristic (ROC) curve of the Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES) for differentiating the performance of individuals with stroke from healthy older adults. The red dot indicates the optimal cut-off score of 57.5, which yields an area under the curve (AUC) of 0.897, a sensitivity of 93.5%, and a specificity of 82.8% (p<0.001).

Table 5. Correlations between MESUPES total scores and stroke-specific impairments (n=64)

Parameter	Total		Arm		Hand	
	r-value	p-value	r-value	p-value	r-value	p-value
FMA-UE[a]	0.922	<0.001	0.872	<0.001	0.893	<0.001
Affected-upper-limb muscle strength						
Elbow flexion[b]	0.661	<0.001	0.602	<0.001	0.585	<0.001
Elbow extension[b]	0.539	<0.001	0.494	<0.001	0.445	<0.001
Handgrip[a]	0.783	<0.001	0.674	<0.001	0.765	<0.001
Affected-upper-limb muscle stiffness						
Biceps brachii[a]	-0.318	0.010	-0.305	0.014	-0.327	0.008
Triceps brachii[a]	-0.072	0.574	-0.072	0.570	-0.095	0.454
Affected-wrist range of motion						
Dorsal–palmar[a]	0.727	<0.001	0.751	<0.001	0.742	<0.001
Ulnar–radial[a]	0.808	<0.001	0.750	<0.001	0.817	<0.001
ARAT[a]	0.877	<0.001	0.821	<0.001	0.885	<0.001
QuickDASH[b]	-0.498	<0.001	-0.521	<0.001	-0.433	<0.001
OxPAQ						
Routine Activities[a]	-0.271	0.030	-0.304	0.015	-0.203	0.108
Social Engagement[a]	-0.167	0.186	-0.202	0.110	-0.136	0.284
Emotional Well-being[a]	-0.039	0.760	-0.107	0.398	-0.028	0.824

MESUPES, Motor Evaluation Scale for Upper Extremity in Stroke Patients; FMA-UE, Fugl-Meyer Assessment for Upper Extremity; ARAT, Action Research Arm Test; QuickDASH, Quick Disabilities of the Arm, Shoulder, and Hand; OxPAQ, Oxford Participation and Activities Questionnaire.
[a]Spearman's rho.
[b]Pearson's r.

muscle stiffness may coexist with reduced voluntary strength due to shared underlying mechanisms, including abnormal co-contraction of agonist–antagonist muscle pairs, impaired selective motor control, and disrupted descending corticospinal drive. As a result, patients with greater spasticity may show lower MESUPES scores not because spasticity itself weakens the muscle, but because it interferes with the ability to generate coordinated, high-quality movement, which is the very construct the MESUPES is designed to capture. Some stroke survivors may also experience paresis, loss of dexterity, and uncoordinated movement, all of which can further diminish movement quality and help explain their poorer performance on the scale. Taken together, the difference in MESUPES scores between groups in this study likely stems from these stroke-specific motor impairments rather than a single isolated deficit. Regarding measurement properties, the scale's total scores demonstrated excellent inter-rater and test-retest reliabilities. Both arm and hand components also showed similar results. These findings were in line with previous studies indicating good-to-very good inter-rater reliability of the scale (ICC=0.95–0.97) [7,8]. In our study, the raters were briefed on the standardized assessment procedures and scoring criteria of the scale to ensure all assessors had the same understanding and were able to provide clear instructions for the patients. This indicates that, with adequate rater training and standardization, the instrument can produce highly consistent scores across sessions and assessors. Crucially, our sensitivity analysis confirmed that the excellent reliability coefficients were not merely an artifact of the observed ceiling effect, demonstrating the true stability of the scale. However, a ceiling effect was identified in subscales and the total score, suggesting that the assessment may be less sensitive for detecting improvement in participants with good upper-limb function. This limitation should be considered when applying the scale in rehabilitation or research contexts targeting higher-functioning individuals.

Systematic reviews and consensus statements highly recommend the FMA-UE and the ARAT as robust, well-validated measures for evaluating post-stroke upper limb capacity. However, these tools primarily focus on task completion and quantitative capacity rather than the qualitative nuances of movement. In this study, the scale's scores had significant and strong positive correlations with the FMA-UE and the ARAT scores. The FMA-UE contains all basic shoulder, arm, wrist, and hand movements performed in the scale, such as putting a hand on the stomach, shoulder abduction, moving the hand from the knee to a table and to the mouth, putting the hand on top of the head, wrist extension range, gripping a plastic bottle, performing a pinch grip, and performing a hook grasp. Similarly, the ARAT

evaluates one's ability in grasping, gripping, pinching, and gross motor movement, which are common components of the scale (e.g., using cylindrical and cubic objects). The strong concurrent validity demonstrated between the scale and these established measures suggests that it accurately captures motor impairment while providing unique, complementary insights into movement quality that standard capacity measures may overlook.

Stroke survivors often present with impaired muscle properties, including reduced strength [29], increased stiffness, and limited joint ROM, all of which significantly influence upper-limb function. The scale's scores showed significant positive correlations with biceps and triceps strength, as well as handgrip strength, with markedly weaker strength observed on the affected side. Muscle weakness in stroke may result from altered activation patterns, including reduced agonist activation and increased antagonist co-contraction. Additionally, motoneuron loss can reduce firing rates [30,31], while disuse and compensatory motor patterns contribute to muscle atrophy [32]. As expected, elbow flexion and extension strength correlated more strongly with the MESUPES arm subscale than the hand subscale, while handgrip strength more correlated with the hand subscale than arm subscale. This finding suggested that MESUPES subscores have a higher resolution to measure something distinct and specific, rather than just measuring one general score to quantify the strength level. Weak elbow muscles can limit the ability to perform antigravity movements such as reaching or lifting during the scale's tasks, while poor handgrip strength may impair the ability to grip objects like the test bottle [33]. Increased biceps-brachii stiffness, which limits elbow extension and requires greater force to overcome [34], was negatively correlated with the scale's scores. This stiffness reduces movement efficiency and quality, particularly in tasks involving gross motor movements such as bringing the hand to the mouth or lifting it overhead. Wrist ROM, including dorsal–palmar flexion and ulnar–radial deviation, also showed significant positive correlations with the scale's scores. Ulnar–radial wrist ROM correlated more strongly with the hand subscale than the arm subscale, which is consistent with expectations, as ulnar–radial mobility is directly required for hand-positioning and grasp-orientation items scored under the hand subscale. Dorsal–palmar ROM showed a more comparable pattern across subscales, this finding suggests that dorsal–palmar mobility functions as a shared biomechanical prerequisite across both subscales rather than being uniquely associated with either. Functional tasks typically require 70% of the maximum wrist ROM [35], and limited mobility can compromise fine motor tasks such as reaching, lifting, and placing the hand in specific positions—core components of the assessment. Prior studies [36] have shown that upper-limb function worsens under wrist-restricted conditions, highlighting the importance of adequate wrist flexibility. Together, these findings suggest that muscle weakness, increased stiffness, and reduced wrist ROM on the affected side are interrelated impairments that contribute to reduced functional performance in stroke survivors. These findings suggest that the MESUPES demonstrates meaningful associations with established impairment and activity-level measures, supporting its concurrent validity. However, further longitudinal research is needed to confirm its responsiveness to change in chronic stroke populations.

The scale's total scores showed a very strong positive correlation with the ARAT scores. The ARAT evaluates one's ability in grasping, gripping, pinching, and gross motor movement, which are also common components of the scale. Specifically, tasks such as placing the hand on top of the head and moving the hand to the mouth are present in both the ARAT and the scale. Gripping tasks are also included in both tests, wherein both cylindrical and cubic objects are used for testing, while the ARAT also involves objects of different sizes. These similarities between the two tests explain the significant correlation between the scale and the ARAT scores.

Additionally, the scale's total scores showed a significant negative correlation with the QuickDASH and the Routine Activities domain of the OxPAQ scores. Both questionnaires assess the ability to perform daily living tasks, such as getting dressed and doing household chores, which require sufficient ability of upper-limb movement. Higher scores on these self-reported questionnaires indicate greater difficulty in completing daily tasks, whereas higher scores on the scale indicate better quality of upper-limb movement. Because the scale reflects the quality of gross and fine upper-limb movements, its scores naturally align with the patient's subjective experience of functional activities tested in these questionnaires. However, negligible correlations were found between the scale's total scores and the Emotional Well-being and Social Engagement sub-scores of the OxPAQ, as the latter two items focus on interviewees' emotions and social engagements, respectively, which are not assessed in the scale.

While stroke patients are identified through clinical diagnosis rather than functional assessment, establishing an optimal cut-off score differentiate individuals with stroke from healthy older adults. The cut-off score of 57.5 for the scale's total scores

demonstrated a very good ability to differentiate the performance of patients with stroke from that of healthy older adults. The high AUC value, sensitivity, and specificity implied that the scale is a discriminative scale that can effectively identify functional deficits. However, because the test items are scored in whole numbers, a precise cut-off of 57.5 has limited direct applicability; therefore, a clinical cut-off of 57 or 58 is more practical for clinicians to define near-normal qualitative motor performance and tailor rehabilitation goals accordingly. The cut-off score of 57.5 provides clinicians with a practical benchmark for identifying residual upper-limb motor impairment among community-dwelling chronic stroke survivors, many of whom may appear to have regained substantial function on capacity-based measures alone. This threshold may be particularly useful in community and outpatient settings to screen for subtle, quality-related motor deficits that might otherwise go undetected, thereby supporting decisions regarding continued rehabilitation referral, goal-setting focused on movement quality rather than task completion alone, and long-term monitoring of upper-limb function in the chronic stroke population.

Limitation

Several limitations merit consideration. First, the sample comprised primarily chronic, community-dwelling stroke survivors, with relatively mild upper-limb impairment (given that inclusion required only minimal antigravity motor function, ≥15° shoulder flexion/abduction), which limits generalizability to individuals in acute or subacute stages of recovery, as well as to those with more moderate-to-severe impairment. This may also explain the observed ceiling effect, which could reflect the relatively high baseline function and frequent upper-limb use typical of community-dwelling chronic survivors, rather than an inherent limitation of the scale itself. Future studies with larger samples spanning a broader range of impairment severity and recovery stages are needed to clarify this. Second, although the scale demonstrated concurrent validity with existing outcome measures in this study, future research is recommended to further examine its relationship with established measures of manual dexterity, such as the Box and Block Test, the Nine-Hole Peg Test, or the Purdue Pegboard Test, to confirm its validity in assessing dexterous hand function in stroke patients. Third, although a seven-day test-retest interval was intended, minor scheduling variations occurred for some participants, which may have slightly influenced the reliability estimates. Fourth, assessments were administered in a fixed order to standardize procedures, which may have introduced fatigue effects affecting performance on later-administered tests and potentially the concurrent validity correlations; future studies should consider counterbalancing or randomizing test order to address this. Finally, while this study establishes concurrent validity, sensitivity to change requires longitudinal evidence showing that the scale's scores vary appropriately with clinical improvement over time. Future research should investigate the longitudinal responsiveness of the scale to fully establish its sensitivity to motor recovery over time, as well as its predictive ability for independent living and returning to work.

Conclusions

The MESUPES is a highly reliable and valid outcome measure for assessing upper-limb movement quality in chronic stroke survivors. It demonstrates strong concurrent validity with standard assessments (e.g., FMA-UE, ARAT) and provides severity cut-offs that help clinicians accurately classify deficits and tailor rehabilitation strategies. Future research should validate the scale in larger, acute populations and evaluate its longitudinal sensitivity to motor recovery, as well as its predictive value for independent living.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

Grant from the General Research Fund (Reference Number: 15126224), the Research Grants Council of the Hong Kong Special Administrative Region, China, awarded to Professor Shamay S.M. Ng and her team.

AUTHOR CONTRIBUTION

Conceptualization: Ng SSM, Chen P. Methodology: Chen P, Li S, Chan SH, Chan YT, Cheng MC, Cheung CH, Cheung YM. Data curation: Chen P, Li S, Chan SH, Chan YT, Cheng MC, Cheung CH, Cheung YM. Formal analysis: Chen P, Li S, Chan SH, Chan YT, Cheng MC, Cheung CH, Cheung YM. Funding acquisition: Ng SSM. Project administration: Ng SSM, Chen P. Visualization: Tse MMY, Lai CYY. Writing – original draft: Chen P, Li S, Chan SH, Chan YT, Cheng MC, Cheung CH, Cheung YM.

Writing – review and editing: Tse MMY, Lai CYY. Approval of final manuscript: all authors.

DATA AVAILABILITY STATEMENT

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

SUPPLEMENTARY MATERIALS

Supplementary materials can be found via https://doi.org/10.5535/arm.260097.

ORCID

Shamay S.M. Ng, https://orcid.org/0000-0003-1660-0548
Peiming Chen, https://orcid.org/0000-0002-2575-367X
Siyue Li, https://orcid.org/0009-0009-9024-2225
Shiu Hang Chan, https://orcid.org/0009-0003-0788-6617
Yin Tung Chan, https://orcid.org/0009-0009-6512-2782
Man Chi Cheng, https://orcid.org/0000-0002-7406-5101
Chung Him Cheung, https://orcid.org/0009-0000-8008-6681
Yau Ming Cheung, https://orcid.org/0009-0003-4096-7988
Mimi M.Y. Tse, https://orcid.org/0000-0001-9428-8719
Cynthia Y.Y. Lai, https://orcid.org/0000-0001-6032-7089

REFERENCES

1. Katan M, Luft A. Global burden of stroke. Semin Neurol 2018;38:208-11.
2. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. Int J Stroke 2022;17:18-29.
3. Yousufuddin M, Young N. Aging and ischemic stroke. Aging (Albany NY) 2019;11:2542-4.
4. Krakauer JW. Arm function after stroke: from physiology to recovery. Semin Neurol 2005;25:384-95.
5. Dutta D, Sen S, Aruchamy S, Mandal S. Prevalence of post-stroke upper extremity paresis in developing countries and significance of m-Health for rehabilitation after stroke - a review. Smart Health 2022;23:100264.
6. Shelton FN, Reding MJ. Effect of lesion location on upper limb motor recovery after stroke. Stroke 2001;32:107-12.
7. Johansson GM, Häger CK. Measurement properties of the Motor Evaluation Scale for Upper Extremity in Stroke patients (MESUPES). Disabil Rehabil 2012;34:288-94.
8. Van de Winckel A, Feys H, van der Knaap S, Messerli R, Baronti F, Lehmann R, et al. Can quality of movement be measured? Rasch analysis and inter-rater reliability of the Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES). Clin Rehabil 2006;20:871-84.
9. Branco JP, Oliveira S, Páscoa Pinheiro J, L Ferreira P. Assessing upper limb function: transcultural adaptation and validation of the Portuguese version of the Stroke Upper Limb Capacity Scale. BMC Sports Sci Med Rehabil 2017;9:15.
10. Pandyan AD, Gregoric M, Barnes MP, Wood D, Van Wijck F, Burridge J, et al. Spasticity: clinical perceptions, neurological realities and meaningful measurement. Disabil Rehabil 2005;27:2-6.
11. Zackowski KM, Dromerick AW, Sahrmann SA, Thach WT, Bastian AJ. How do strength, sensation, spasticity and joint individuation relate to the reaching deficits of people with chronic hemiparesis? Brain 2004;127(Pt 5):1035-46.
12. Ada L, Vattanasilp W, O'Dwyer NJ, Crosbie J. Does spasticity contribute to walking dysfunction after stroke? J Neurol Neurosurg Psychiatry 1998;64:628-35.
13. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998;17:101-10.
14. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehabil Med 1975;7:13-31.
15. Lyle RC. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. Int J Rehabil Res 1981;4:483-92.
16. Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. BMC Musculoskelet Disord 2006;7:44.
17. Morley D, Dummett S, Kelly L, Dawson J, Fitzpatrick R, Jenkinson C. Validation of the Oxford Participation and Activities Questionnaire. Patient Relat Outcome Meas 2016;7:73-80.
18. See J, Dodakian L, Chou C, Chan V, McKenzie A, Reinkensmeyer DJ, et al. A standardized approach to the Fugl-Meyer assessment and its implications for clinical trials. Neurorehabil Neural Repair 2013;27:732-41.
19. Morin M, Hébert LJ, Perron M, Petitclerc É, Lake SR, Duchesne E. Psychometric properties of a standardized protocol of muscle strength assessment by hand-held dynamometry in healthy adults: a reliability study. BMC Musculoskelet Disord 2023;24:294.
20. Huang L, Liu Y, Lin T, Hou L, Song Q, Ge N, et al. Reliability and validity of two hand dynamometers when used by communi-

ty-dwelling adults aged over 50 years. BMC Geriatr 2022;22:580.
21. Davidson MJ, Bryant AL, Bower WF, Frawley HC. Myotonometry reliably measures muscle stiffness in the thenar and perineal muscles. Physiother Can 2017;69:104-12.
22. Van der Lee JH, De Groot V, Beckerman H, Wagenaar RC, Lankhorst GJ, Bouter LM. The intra- and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. Arch Phys Med Rehabil 2001;82:14-9.
23. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. J Shoulder Elbow Surg 2009;18:920-6.
24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-63.
25. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J 2005;47:458-72.
26. Monticone M, Galeoto G, Berardi A, Tofani M. Psychometric properties of assessment tools. In: Galeoto G, Berardi A, Tofani M, Auxiliadora Marquez M, editors. Measuring spinal cord injury. Springer; 2021. p. 7-15.
27. Emos MC, Agarwal S. Neuroanatomy, upper motor neuron lesion. StatPearls Publishing; 2023.
28. Burke D, Wissel J, Donnan GA. Pathophysiology of spasticity in stroke. Neurology 2013;80(3 Suppl 2):S20-6.
29. Abdollahi I, Taghizadeh A, Shakeri H, Eivazi M, Jaberzadeh S. The relationship between isokinetic muscle strength and spasticity in the lower limbs of stroke patients. J Bodyw Mov Ther 2015;19:284-90.
30. Chang SH, Francisco GE, Zhou P, Rymer WZ, Li S. Spasticity, weakness, force variability, and sustained spontaneous motor unit discharges of resting spastic-paretic biceps brachii muscles in chronic stroke. Muscle Nerve 2013;48:85-92.
31. Li X, Holobar A, Gazzoni M, Merletti R, Rymer WZ, Zhou P. Examination of poststroke alteration in motor unit firing behavior using high-density surface EMG decomposition. IEEE Trans Biomed Eng 2015;62:1242-52.
32. Hunnicutt JL, Gregory CM. Skeletal muscle changes following stroke: a systematic review and comparison to healthy individuals. Top Stroke Rehabil 2017;24:463-71.
33. Boissy P, Bourbonnais D, Carlotti MM, Gravel D, Arsenault BA. Maximal grip force in chronic stroke subjects and its relationship to global upper extremity function. Clin Rehabil 1999;13:354-62.
34. Chuang LL, Lin KC, Wu CY, Chang CW, Chen HC, Yin HP, et al. Relative and absolute reliabilities of the myotonometric measurements of hemiparetic arms in patients with stroke. Arch Phys Med Rehabil 2013;94:459-66.
35. Ryu JY, Cooney WP 3rd, Askew LJ, An KN, Chao EY. Functional ranges of motion of the wrist joint. J Hand Surg Am 1991;16:409-19.
36. Adams BD, Grosland NM, Murphy DM, McCullough M. Impact of impaired wrist motion on hand and upper-extremity performance. J Hand Surg Am 2003;28:898-903.

Original Article

Ann Rehabil Med 2026;50(5):401-411
eISSN: 2234-0653
https://doi.org/10.5535/arm.260110

arm
Annals of Rehabilitation Medicine

Longitudinal Mean HbA1c and Electrodiagnostic Severity of Distal Symmetric Sensorimotor Polyneuropathy in Type 2 Diabetes: Comparison With Index HbA1c

Songi Han, MD, Jongkyu Kim, MD, Sujeong Choi, MD, Eun Sang Yoon, MD, Ho Kyung Yoo, MD

Department of Physical Medicine and Rehabilitation, Seoul Medical Center, Seoul, Korea

Objective: To evaluate the association of longitudinal mean glycosylated hemoglobin (HbA1c) with electrodiagnostic severity of distal symmetric sensorimotor polyneuropathy (DSPN), using index HbA1c as a comparator.

Methods: We retrospectively reviewed patients with type 2 diabetes undergoing nerve conduction studies (NCS) during 2022–2024. Eligibility required ≥4 HbA1c measurements spanning ≥2 calendar years within a capped 4-year look-back window including the NCS month. Mean HbA1c averaged these measurements; index HbA1c was the measurement closest to NCS within ±3 months. The primary outcome was ordinal Baba stage (0–4). Separate ordinal models used parsimonious and full covariate adjustment. Paired bootstrap resampling and joint models directly compared associations. Exploratory paired receiver operating characteristic analyses assessed advanced electrodiagnostic DSPN (Baba stages 3–4). Penalized analyses assessed stability; additional analyses evaluated time-weighted means and HbA1c coefficient of variation (CV).

Results: Of 263 screened patients, 73 were included. Mean HbA1c correlated with Baba stage (ρ=0.320, p=0.006). Fully adjusted odds ratios were 1.51 for mean HbA1c (95% confidence interval [CI], 1.03–2.22) and 1.31 for index HbA1c (95% CI, 0.98–1.76). The paired bootstrap 95% CI for the difference in log-odds ratios included zero (-0.152 to 0.650). Mean HbA1c remained associated with advanced DSPN; areas under the curve were 0.701 and 0.618, respectively (DeLong p=0.153), with an exploratory threshold of 7.0%. Ordinal associations were attenuated with time-weighted means or CV adjustment.

Conclusion: Mean HbA1c showed more consistent associations than index HbA1c across separate models, but superiority was not established. Sensitivity analyses indicated uncertainty in the ordinal association.

Keywords: Diabetic neuropathies, Glycated hemoglobin, Nerve conduction studies, Electrodiagnosis, Diabetes mellitus, type 2

Received: June 1, 2026
Revised: September 22, 2026
Accepted: October 13, 2026

Correspondence:
Jongkyu Kim
Department of Physical Medicine and Rehabilitation, Seoul Medical Center, 156 Sinnae-ro, Jungnang-gu, Seoul 02053, Korea.
Tel: +82-2-2276-7474
Fax: +82-2-2276-8504
E-mail: jongkyukim@seoulmc.or.kr

© 2026 by Korean Academy of Rehabilitation Medicine

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Diabetes mellitus (DM) is a chronic metabolic disorder associated with microvascular complications including neuropathy [1,2]. Although glycosylated hemoglobin (HbA1c) is widely used to assess recent glycemic exposure, a single contemporaneous value may not adequately reflect the cumulative metabolic burden associated with chronic peripheral nerve injury.

Distal symmetric sensorimotor polyneuropathy (DSPN) is the most common neuropathic complication of DM [1-5]. It is a chronic, length-dependent, progressive disorder that may remain clinically silent despite electrophysiological involvement [4-6]. As DSPN progresses, it increases the risk of foot ulceration, amputation, falls, and functional decline [1,4-6]. An ordinal electrodiagnostic staging framework may provide more nuanced information on disease burden than a simple binary definition and may therefore have greater clinical and research utility [7-9]. Nerve conduction studies (NCS) are considered a comprehensive and objective approach for evaluating large-fiber DSPN [2,4,9]. However, DSPN severity is often assessed using neurological findings or symptom-based measures rather than structured electrodiagnostic staging [7-9]. A practical NCS-based ordinal staging system (stages 0–4) may be useful in referral-based rehabilitation practice [7,8]. This framework provides a basis for examining whether longitudinal glycemic exposure, reflected by mean HbA1c, is more closely associated with graded electrodiagnostic severity than a single contemporaneous HbA1c value is.

Previous studies examined the association between glycemic control and electrodiagnostic neuropathy severity but did not directly compare longitudinal mean HbA1c with index HbA1c using an ordinal electrodiagnostic framework [10,11]. Long-term HbA1c variability has also been associated with electrophysiological neuropathy severity, supporting its consideration alongside mean HbA1c [12].

In this study, we aimed to evaluate the association of longitudinal mean HbA1c with electrodiagnostic DSPN severity in patients with type 2 DM (T2DM), using index HbA1c as a comparator. We also explored a candidate mean HbA1c threshold for advanced electrodiagnostic DSPN. Sensitivity analyses using time-weighted mean HbA1c addressed irregular measurement intervals. Exploratory analyses evaluated the contribution of HbA1c coefficient of variation (CV) alongside mean HbA1c.

METHODS

Study design

This retrospective observational study was conducted at a single general hospital. We reviewed the medical records of all patients with T2DM who underwent NCS between January 1, 2022, and December 31, 2024. All included patients had a documented diagnosis of T2DM and were receiving oral glucose-lowering medications, insulin, or both at the time of NCS.

Patients were eligible if they had symptoms clinically consistent with typical DSPN and at least four HbA1c measurements spanning ≥2 calendar years within a capped 4-year look-back window that included the month of NCS. This window was selected to capture cumulative glycemic exposure relevant to chronic nerve injury while applying a consistent upper limit to the observation period. When longer HbA1c records were available, only values within the 4-year window were utilized. Patients with focal neuropathies, asymmetric patterns of neuropathy inconsistent with typical DSPN, or other neurological disorders that could account for the clinical or nerve conduction findings were excluded.

An a priori sample size calculation was performed during protocol development using G*Power version 3.1.9.7 (Heinrich Heine University Düsseldorf) for a two-sided comparison of two independent means. Based on HbA1c values of 8.14±1.63% and 6.69±1.01% reported in a previous study [10], an effect size of approximately 1.07 was assumed. With a significance level of 0.05, 95% power, and an allocation ratio of 0.222, the required sample size was 80 patients. A target sample size of 96 patients was set after adding a 20% allowance for exclusions or incomplete records. The final analysis included 73 eligible patients and therefore did not meet the calculated minimum sample size. This calculation was based on the originally planned exploratory comparison of HbA1c between patients with Baba stages ≥2 and 0–1 and was not designed to determine the sample size required for the ordinal analyses or direct comparisons between mean and index HbA1c.

The study protocol was approved by the Institutional Review Board (IRB) of Seoul Medical Center (SEOUL IRB 2026-04-003). The requirement for informed consent was waived by the IRB.

The Strengthening the Reporting of Observational Studies in Epidemiology checklist is provided in Supplementary Material S1.

Electrodiagnostic assessment and severity grading

In patients with T2DM and clinically suspected DSPN, NCS findings were assessed for a distal symmetric, length-dependent pattern of polyneuropathy involving the lower extremities. Patients with focal entrapment neuropathy, lumbosacral radiculopathy, plexopathy, or other alternative neurological disorders that better explained the findings were excluded [2-4].

NCS were performed by experienced medical technologists using a Sierra Wave electrodiagnostic system (Cadwell Industries, Inc.). Tibial motor responses were recorded over the abductor hallucis following stimulation at the ankle and knee, and sural sensory responses were recorded antidromically at the lateral malleolus following calf stimulation. Distances were measured individually. Bandpass filters were set at 10 Hz–10 kHz for motor studies and 20 Hz–2 kHz for sensory studies. F-wave studies were performed using at least 20 stimuli, and the minimum F-wave latency was used for assessment. Lower-limb skin temperature was maintained at ≥32°C.

DSPN severity was graded using the Baba classification, an NCS-based lower-extremity staging system, ranging from stage 0 (normal) to stage 4 (severe) (Fig. 1) [7,8]. Stage 0 denotes normal findings under the Baba criteria and does not exclude clinically suspected DSPN. Prolonged tibial F-wave latency was incorporated into stage 1 [13].

The primary outcome was ordinal Baba stage (0–4). For exploratory clinical translation, advanced electrodiagnostic DSPN was defined as Baba stage 3–4 versus 0–2 because stage 3 indicates progression beyond isolated sural sensory abnormality to additional motor involvement, reflected by reduced tibial compound muscle action potential. Lower-extremity NCS parameters were interpreted based on established laboratory standards with reference to published normative data from healthy Korean adults [14].

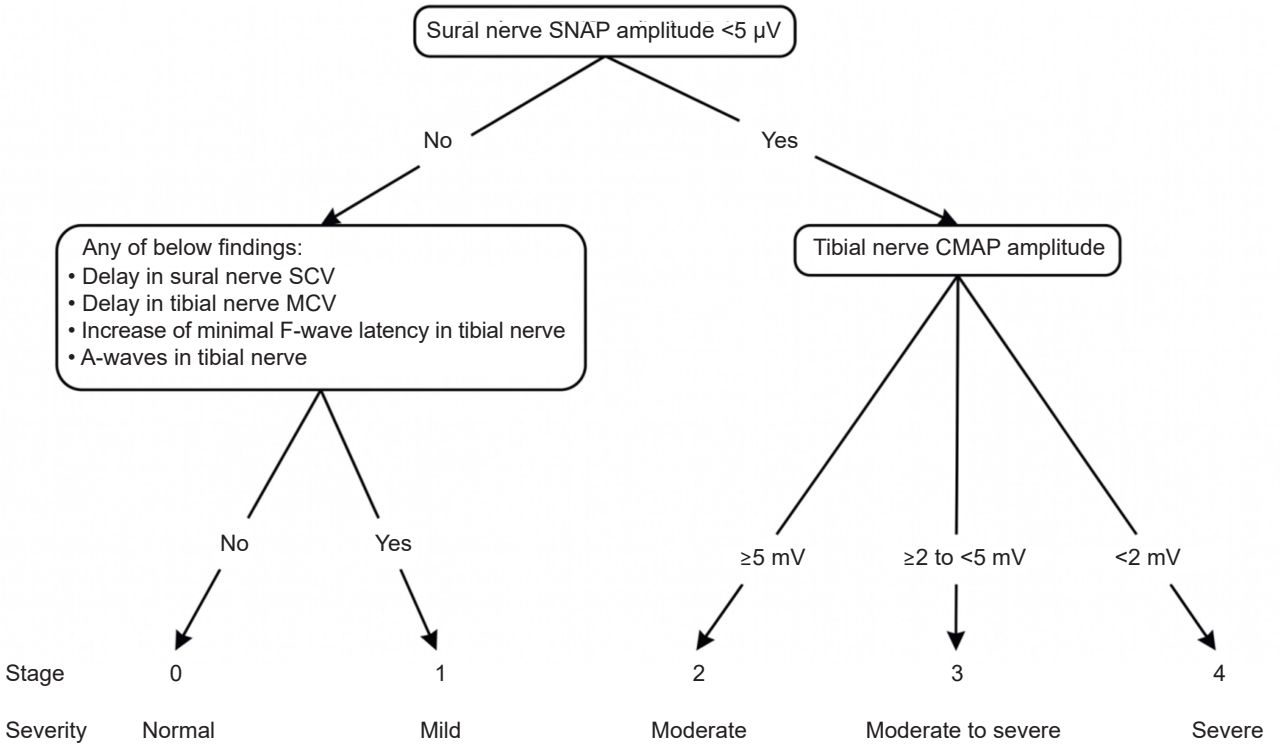

Fig. 1. Schematic diagram of the Baba classification system indicating the electrodiagnostic severity of distal symmetric sensorimotor polyneuropathy. The Baba classification grades the electrodiagnostic severity from stage 0 (normal) to 4 (severe) using lower-extremity nerve conduction study parameters. Stage 1 is defined by a slowing of tibial MCV or sural SCV, prolonged tibial minimum F-wave latency, or the presence of tibial A-waves. Stage 2 is defined as sural SNAP amplitude <5 μV with tibial CMAP amplitude ≥5 mV. Stage 3 is defined as sural SNAP <5 μV with tibial CMAP ≥2 and <5 mV. Stage 4 is defined as sural SNAP <5 μV with tibial CMAP <2 mV. SNAP, sensory nerve action potential; SCV, sensory conduction velocity; MCV, motor conduction velocity; CMAP, compound muscle action potential. Based on Baba et al. [7]; adapted from Himeno et al. (J Diabetes Investig 2020;11:5-16) [8].

Glycemic exposure variables

Longitudinal mean HbA1c (hereafter, mean HbA1c) was the primary exposure, defined as the arithmetic mean of all available HbA1c values within the qualifying capped 4-year look-back window, including the month of NCS. Index HbA1c was defined as the HbA1c value closest to the NCS date within ±3 months and was examined as a secondary comparator. HbA1c measurements were obtained as part of routine clinical care, rather than at protocol-defined intervals. Time-weighted mean HbA1c was calculated within the same qualifying window using trapezoidal integration between successive measurements, divided by the interval between the first and last measurements. Time intervals were based on recorded calendar months, without extrapolation beyond the observed period. Measurements recorded within the same month were averaged for this calculation. HbA1c CV was calculated as the within-patient standard deviation divided by the arithmetic mean, multiplied by 100, using the same measurements as the primary exposure.

Statistical analysis

The analysis proceeded in two steps: a primary ordinal analysis using Baba stage 0–4 and an exploratory binary analysis comparing stages 3–4 versus 0–2. Categorical variables were presented as number (percentage); and continuous variables as mean±standard deviation or median [interquartile range], as appropriate. One-way ANOVA or Welch's t-test was used for group comparisons of continuous variables. All 73 patients had complete data for the regression variables; no imputation was performed.

The association between Baba stage and mean HbA1c was assessed using Spearman's correlation. Separate ordered logistic regression models were constructed for mean and index HbA1c using identical covariate sets. The parsimonious model (M1) included age, sex, diabetes duration, and estimated glomerular filtration rate (eGFR). The fully adjusted model (M3) additionally included retinopathy and atherosclerotic cardiovascular disease. M1 was used to limit model complexity while accounting for demographic factors, diabetes duration, and renal function. Covariates were not selected based on statistical significance. HbA1c odds ratios (ORs) were expressed per 1 percentage-point increase. For unpenalized regression coefficients, 95% confidence intervals (CIs) and p-values were calculated using Wald methods.

Sensitivity analyses replaced arithmetic mean HbA1c with time-weighted mean HbA1c in M1 and M3. Exploratory ordinal analyses evaluated CV alone and jointly with arithmetic mean HbA1c using the M1 covariate set. Pearson correlation and variance inflation factors (VIFs) were also assessed to evaluate collinearity between mean HbA1c and CV. A likelihood-ratio test assessed the additional contribution of CV to the mean HbA1c model. CV ORs were expressed per 1 percentage-point increase.

To directly compare the associations estimated from the separate ordinal models, paired patient-level bootstrap resampling was performed with 3,000 iterations to obtain percentile 95% CIs for the difference in log-ORs, calculated as mean minus index HbA1c. Both models were fitted to the same resampled patients in each iteration. Pearson correlation between the two HbA1c measures and VIFs were assessed to evaluate collinearity in joint ordinal models. At each adjustment level, likelihood-ratio tests evaluated the additional contribution of each measure, and a Wald contrast tested equality of the two coefficients in the joint model. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were also compared between the separate mean and index HbA1c models at each adjustment level, with lower values indicating better relative model fit.

Separate binary logistic regression models for mean and index HbA1c were fitted for advanced electrodiagnostic DSPN using the same parsimonious and full covariate sets as the ordinal models. Receiver operating characteristic (ROC) analyses evaluated mean and index HbA1c values alone, and the Youden index was used to identify an exploratory mean HbA1c threshold. Paired ROC curves were compared using the DeLong test [15]. A paired, outcome-stratified bootstrap with 3,000 iterations was used to obtain percentile 95% CIs for each area under the curve (AUC) and for the difference in AUC between the two measures. ROC comparisons were additionally performed for Baba stage ≥2 as a sensitivity analysis.

The separate-exposure ordinal models contained five and seven slope parameters under parsimonious and full adjustment, respectively, in addition to four ordinal thresholds. The joint mean HbA1c–CV model contained six slope parameters and four ordinal thresholds. For the exploratory binary outcome, 22 patients had Baba stage ≥3 and 51 had Baba stage 0–2, corresponding to 4.4 events per slope parameter in M1 and 3.14 in M3. These counts describe the binary analysis rather than a single events-per-variable measure for the ordinal outcome. The proportional odds assumption was assessed using the Brant test in R [16] and a supplementary Rao score test implemented in Python. Global Brant test results were considered interpretable

only when the estimated covariance matrix of the coefficient contrasts was positive definite. Separate cumulative binary logistic regression models were also fitted across Baba stage cut-points (≥1 vs. 0, ≥2 vs. 0–1, ≥3 vs. 0–2, and ≥4 vs. 0–3). These proportional odds diagnostics and cumulative binary analyses were also performed for the time-weighted mean and CV models.

Model stability was examined using ridge-penalized ordinal regression and leave-one-out analyses. Ridge penalties of 0.5, 1, and 2 were applied to standardized slope coefficients, with ordinal thresholds left unpenalized. These penalties were examined as sensitivity choices rather than selected by optimization. For the parsimonious ridge model with a penalty of 1, a percentile 95% CI for the mean HbA1c OR was obtained using 3,000 bootstrap resamples. Firth penalized logistic regression was additionally performed for the exploratory binary outcome, with profile-likelihood CIs and tests [17]. After excluding three patients receiving hemodialysis, the ordinal models and fully adjusted binary model containing mean HbA1c were refitted using the same covariate sets.

Statistical analyses were conducted using Python 3.12 with NumPy, SciPy, pandas, and scikit-learn. Custom estimation routines were used for ordinal regression, penalized regression, and the Rao score test. Unpenalized ordinal and cumulative binary models, including the additional time-weighted mean and CV models, were cross-checked in R 4.6.0 (R Foundation) using the MASS and stats packages; Brant tests were performed using the brant package. All tests were two-sided, with $p<0.05$ considered statistically significant.

RESULTS

Study population and baseline characteristics

Of 263 patients with T2DM who underwent NCS during the study period, 140 were excluded because of neurological disorders other than typical DSPN, including 117 with radiculopathy. Of the remaining 123 patients, 50 did not meet the longitudinal HbA1c eligibility criteria, leaving 73 patients for the final analysis (Fig. 2). Baba stage distribution was stage 0, n=29; stage 1, n=4; stage 2, n=18; stage 3, n=12; and stage 4, n=10. Baseline characteristics are summarized in Table 1.

Mean HbA1c across the full Baba severity spectrum

Mean HbA1c was positively associated with Baba stage (Spearman ρ=0.320, p=0.006), and one-way ANOVA indicated differences across stages (p=0.018). However, the relationship was not strictly monotonic: the median mean HbA1c was higher at stage 3 than at stage 4 (8.30% vs. 7.83%; Fig. 3).

Ordered logistic regression for ordinal Baba stage

In the parsimonious ordered logistic model (M1), mean HbA1c was associated with higher Baba stage (OR, 1.49; 95% CI, 1.03–2.16; p=0.035), while the estimate for index HbA1c was 1.28 (95% CI, 0.96–1.69; p=0.092). In the fully adjusted models (M3), cor-

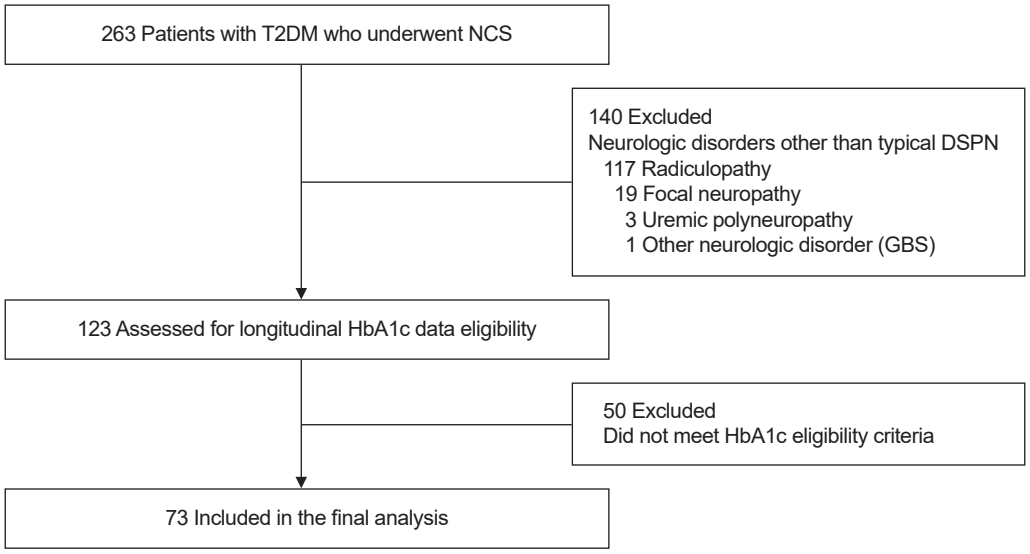

Fig. 2. Flow diagram of study participant selection. T2DM, type 2 diabetes mellitus; NCS, nerve conduction study; DSPN, distal symmetric sensorimotor polyneuropathy; GBS, Guillain–Barré syndrome; HbA1c, glycosylated hemoglobin.

responding estimates were 1.51 (95% CI, 1.03–2.22; p=0.034) and 1.31 (95% CI, 0.98–1.76; p=0.065), respectively (Table 2). Full covariate estimates are provided in Supplementary Table S1.

Mean and index HbA1c were correlated (Pearson r=0.730), and the maximum VIF in the joint models was 2.33. Paired bootstrap 95% CIs for the difference in separate-model log-ORs included zero in both M1 (-0.139 to 0.586) and M3 (-0.152 to 0.650). Adding mean HbA1c to models containing index HbA1c did not provide a statistically significant additional contribution (likelihood-ratio p=0.183 for M1 and p=0.221 for M3); corresponding p-values for adding index HbA1c were 0.806 and 0.650. Joint-model coefficient contrasts also did not demonstrate a statistically significant difference between the measures (p=0.489 for M1 and p=0.596 for M3; Supplementary Table S2). AIC and BIC were slightly lower for mean HbA1c models than for the corresponding index HbA1c models, with index-minus-mean differences of 1.71 in M1 and 1.29 in M3 for both criteria (Supplementary Table S2).

For the arithmetic mean HbA1c models, global Brant test

Table 1. Baseline characteristics of the study population

Variable	Value (n=73)
Age (yr)	66.41±9.94
Sex, male	53 (72.6)
DM duration (yr)	10.00 [7.00–13.00]
eGFR (mL/min/1.73 m^2)	76.27±26.47
Index HbA1c (%)	7.29±1.50
Mean HbA1c (%)	7.46±1.25
HbA1c CV (%)	8.43 [5.98–13.15]
Insulin user	24 (32.9)
ESRD on HD	3 (4.1)
DM retinopathy	21 (28.8)
DM foot disease	13 (17.8)
PAD	11 (15.1)
ASCVD (MI or stroke)	13 (17.8)

Values are presented as mean±standard deviation (SD), number (%), or median [interquartile range].
Visit-to-visit HbA1c CV was calculated as within-patient SD/mean×100.
DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; HbA1c, glycosylated hemoglobin; CV, coefficient of variation; ESRD, end-stage renal disease; HD, hemodialysis; PAD, peripheral artery disease; ASCVD, atherosclerotic cardiovascular disease; MI, myocardial infarction.

Table 2. Adjusted associations of mean and index HbA1c with higher Baba stage (0–4): ordered logistic regression analysis

Adjustment	HbA1c measure (%)	OR (95% CI)	p-value
M1	Mean HbA1c	1.49 (1.03–2.16)	0.035*
M1	Index HbA1c	1.28 (0.96–1.69)	0.092
M3	Mean HbA1c	1.51 (1.03–2.22)	0.034*
M3	Index HbA1c	1.31 (0.98–1.76)	0.065

Separate ordered logistic models were fitted for each HbA1c measure (n=73).
M1 adjusted for age, sex, diabetes duration, and eGFR; M3 additionally adjusted for diabetic retinopathy and ASCVD (myocardial infarction or stroke).
ORs are expressed per 1 percentage-point increase in HbA1c, with values >1 indicating higher odds of a higher Baba stage under the proportional odds model.
Full covariate estimates are provided in Supplementary Table S1.
HbA1c, glycosylated hemoglobin; OR, odds ratio; CI, confidence interval; eGFR, estimated glomerular filtration rate; ASCVD, atherosclerotic cardiovascular disease.
*p<0.05.

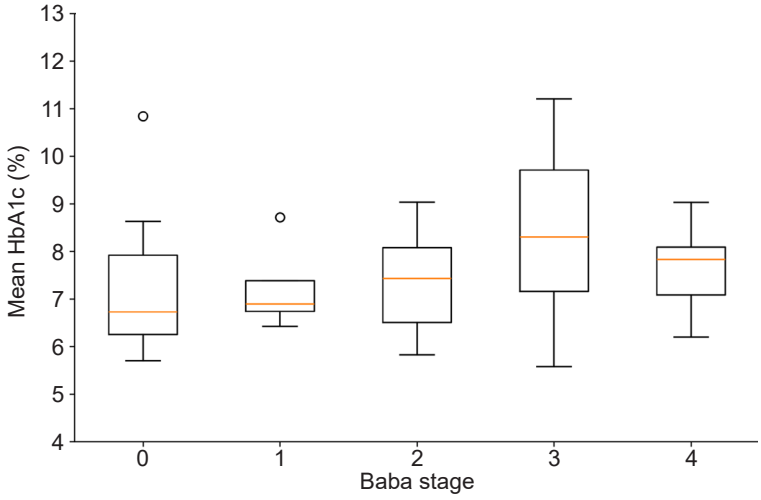

Fig. 3. Distribution of mean glycosylated hemoglobin (HbA1c) across Baba stages 0–4. Boxplots show the distribution of the longitudinal mean HbA1c across Baba stages 0–4. The central line indicates the median; the box represents the interquartile range; and whiskers indicate the range excluding the outliers.

results were not interpretable because the estimated covariance matrices of the coefficient contrasts were not positive definite. Supplementary Rao score tests did not reject the proportional odds assumption (M1, p=0.323; M3, p=0.191). However, sparse outcome–predictor combinations warranted cautious interpretation. In cumulative binary analyses, the mean HbA1c association was most pronounced for Baba stage ≥3 versus 0–2, whereas associations at lower cut-points and stage 4 versus 0–3 were less clear. Estimates varied across cut-points and were imprecise, particularly for stage 4 versus 0–3 (Supplementary Table S3).

In the ridge-penalized M1 ordinal model with a penalty of 1, the OR for mean HbA1c was 1.46 (bootstrap 95% CI, 1.09–2.37). In the unpenalized M1 model, leave-one-out ORs ranged from 1.43 to 1.72, with p-values ranging from 0.008 to 0.062. The direction of association was maintained, although statistical significance was sensitive to individual observations. After excluding the three patients receiving hemodialysis, the M1 ordinal estimate for mean HbA1c remained similar (OR, 1.50; 95% CI, 1.03–2.17; n=70).

Arithmetic and time-weighted mean HbA1c were highly correlated (Pearson r=0.993). When time-weighted mean HbA1c replaced the arithmetic mean, ordinal estimates were attenuated in both M1 (OR, 1.41; 95% CI, 0.98–2.03; p=0.066) and M3 (OR, 1.43; 95% CI, 0.98–2.08; p=0.061). Point estimates remained positive, but both CIs included 1 (Supplementary Table S4).

Mean HbA1c and CV were correlated (Pearson r=0.368), and the maximum VIF in their joint model was 1.43. In the model containing CV and the M1 covariates, CV was associated with higher Baba stage (OR per 1 percentage-point increase, 1.055; 95% CI, 1.001–1.112; p=0.047). When mean HbA1c and CV were included together, the mean HbA1c estimate was attenuated (OR, 1.39; 95% CI, 0.94–2.04; p=0.100), and the CV estimate was 1.043 (95% CI, 0.988–1.101; p=0.131). Adding CV to the mean HbA1c model did not provide a statistically significant additional contribution (likelihood-ratio p=0.124; Supplementary Table S4).

For the time-weighted mean models, global Brant tests were not interpretable because the covariance matrices were not positive definite; Rao score tests did not reject proportional odds (M1, p=0.398; M3, p=0.229). For the CV and joint mean HbA1c–CV models, global Brant tests indicated departures from proportional odds (both p<0.001), whereas Rao score tests did not (p=0.860 and p=0.527, respectively). These discordant diagnostics warrant cautious interpretation of the exploratory ordinal estimates (Supplementary Table S4).

Exploratory analysis of advanced electrodiagnostic DSPN (Baba stages 3–4)

Mean HbA1c was higher in the advanced group than in the non-advanced group (8.13±1.45% vs. 7.18±1.05%; p=0.009; Table 3). In binary logistic regression, mean HbA1c was associated with advanced electrodiagnostic DSPN in both M1 (OR, 2.02; 95% CI, 1.20–3.41; p=0.008) and M3 (OR, 2.05; 95% CI, 1.21–3.46; p=0.008; Table 4). Index HbA1c was not significantly associated with advanced electrodiagnostic DSPN in either M1 (OR, 1.30; 95% CI, 0.90–1.90; p=0.165) or M3 (OR, 1.30; 95% CI, 0.89–1.91; p=0.176). The Firth M1 model yielded a mean HbA1c OR of 1.85 (profile-likelihood 95% CI, 1.19–3.18; p=0.005). Lower eGFR was also associated with advanced electrodiagnostic DSPN in the M3 mean HbA1c model (OR per 1 mL/min/1.73 m^2 increase, 0.970; 95% CI, 0.943–0.998; p=0.033).

After excluding the three patients receiving hemodialysis, mean HbA1c remained associated with advanced electrodiagnostic DSPN in M3 (OR, 2.01; 95% CI, 1.19–3.40; p=0.010). The inverse association with eGFR was attenuated and no longer statistically significant (OR, 0.976; 95% CI, 0.945–1.008; p=0.143; n=70; Supplementary Table S3).

ROC analysis showed an AUC of 0.701 for mean HbA1c (bootstrap 95% CI, 0.567–0.822), with an exploratory Youden-derived threshold of 7.0% (sensitivity, 81.8%; specificity, 54.9%; Fig. 4). The corresponding AUC for index HbA1c was 0.618 (bootstrap 95% CI, 0.469–0.754). The AUC difference, calculated as mean minus index HbA1c, was 0.083 (DeLong 95% CI, -0.031 to 0.198; p=0.153), and the paired bootstrap 95% CI also included zero (-0.027 to 0.203; Supplementary Table S2). At the Baba ≥2 sensitivity cut-point, the AUC difference was 0.047 (DeLong 95% CI, -0.055 to 0.149; p=0.366).

Table 3. Unadjusted comparison of mean HbA1c between non-advanced and advanced electrodiagnostic DSPN

Variable	Non-advanced (n=51)	Advanced (n=22)	p-value
Mean HbA1c (%)	7.18±1.05	8.13±1.45	0.009*

Values are presented as mean±standard deviation.
Non-advanced was defined as Baba stage 0–2 and advanced as Baba stage 3–4.
Groups were compared using Welch's t-test.
HbA1c, glycosylated hemoglobin; DSPN, distal symmetric sensorimotor polyneuropathy.
*p<0.05.

Table 4. Adjusted associations of mean and index HbA1c with advanced electrodiagnostic DSPN

Variable	Mean HbA1c model OR (95% CI)	p-value	Index HbA1c model OR (95% CI)	p-value
M1 (parsimonious)				
HbA1c (%)	2.02 (1.20–3.41)	0.008*	1.30 (0.90–1.90)	0.165
Age (yr)	1.03 (0.96–1.10)	0.394	1.02 (0.96–1.08)	0.566
Sex, male (vs. female)	1.51 (0.33–6.88)	0.597	1.92 (0.47–7.89)	0.367
DM duration (yr)	1.01 (0.94–1.09)	0.788	1.03 (0.96–1.10)	0.484
eGFR (mL/min/1.73 m^2)	0.973 (0.949–0.997)	0.029*	0.974 (0.952–0.997)	0.028*
M3 (fully adjusted)				
HbA1c (%)	2.05 (1.21–3.46)	0.008*	1.30 (0.89–1.91)	0.176
Age (yr)	1.03 (0.96–1.11)	0.387	1.02 (0.95–1.09)	0.539
Sex, male (vs. female)	1.61 (0.36–7.34)	0.535	2.02 (0.48–8.47)	0.338
DM duration (yr)	1.00 (0.92–1.09)	0.930	1.02 (0.94–1.10)	0.614
eGFR (mL/min/1.73 m^2)	0.970 (0.943–0.998)	0.033*	0.973 (0.948–0.998)	0.036*
DM retinopathy	1.11 (0.25–4.80)	0.894	1.16 (0.29–4.65)	0.832
ASCVD (MI or stroke)	0.50 (0.10–2.48)	0.400	0.58 (0.13–2.61)	0.476

Separate binary logistic models were fitted for each HbA1c measure (n=73), comparing advanced electrodiagnostic DSPN (Baba stages 3–4; n=22) with non-advanced DSPN (stages 0–2; n=51).
M1 adjusted for age, sex, diabetes duration, and eGFR; M3 additionally adjusted for diabetic retinopathy and ASCVD (myocardial infarction or stroke).
Maximum-likelihood estimates are presented with Wald 95% CIs and two-sided Wald p-values.
HbA1c ORs are per 1 percentage-point increase; other continuous-variable ORs are per 1-unit increase. Binary-variable ORs compare presence with absence unless otherwise stated.
HbA1c, glycosylated hemoglobin; DSPN, distal symmetric sensorimotor polyneuropathy; OR, odds ratio; CI, confidence interval; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; ASCVD, atherosclerotic cardiovascular disease; MI, myocardial infarction.
*p<0.05.

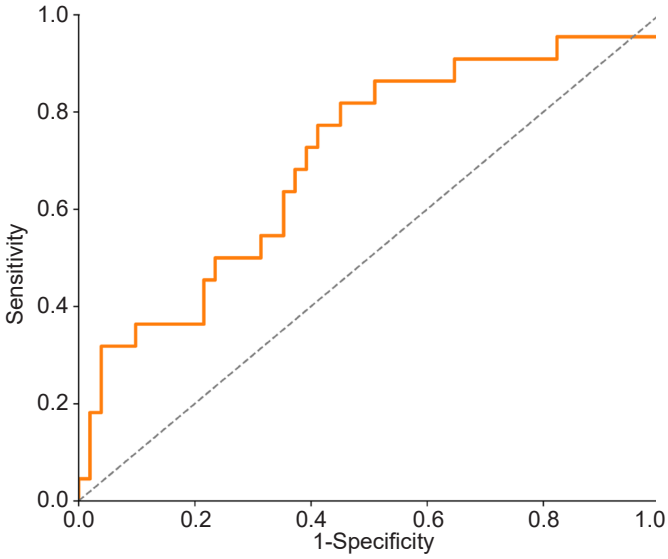

Fig. 4. Receiver operating characteristic curve of mean glycosylated hemoglobin (HbA1c) for identifying advanced electrodiagnostic distal symmetric sensorimotor polyneuropathy (Baba stages 3–4 vs. 0–2). The area under the curve is 0.701 (bootstrap 95% confidence interval, 0.567–0.822). The exploratory Youden-derived mean HbA1c threshold was 7.0% (sensitivity, 81.8%; specificity, 54.9%).

DISCUSSION

In this retrospective study of patients with T2DM referred for electrodiagnostic evaluation, mean HbA1c showed a more consistent pattern of association with electrodiagnostic DSPN severity than index HbA1c did across the separate ordinal and exploratory binary models. However, direct comparisons did not demonstrate a statistically significant difference between the two associations or their AUCs. Thus, the observed pattern does not establish that mean HbA1c better reflects electrodiagnostic severity than index HbA1c, and the nonsignificant comparisons should not be interpreted as evidence of equivalence.

As nerve injury in DSPN develops gradually over years, progressing from early conduction slowing to sequential sensory and motor axonal loss, electrodiagnostic abnormalities reflect accumulated dysfunction of large, myelinated fibers [1,2,5]. From this perspective, the observed association between mean HbA1c and electrodiagnostic neuropathy severity is pathophysiologically plausible. However, mean HbA1c was calculated from four or more measurements and may therefore be less affected by random measurement error than a single index value. Its numerically larger association may partly reflect reduced regression dilution rather than a biological effect of cumulative

metabolic burden [18]. The present study cannot distinguish between these explanations.

Previous research has linked long-term HbA1c variability with electrophysiological neuropathy severity [12]. In our exploratory analyses, CV was associated with higher Baba stage when evaluated with the M1 covariates. However, when mean HbA1c and CV were included together, neither association reached statistical significance, and adding CV did not significantly improve model fit. The mean HbA1c estimate was attenuated after CV adjustment, but this does not establish that variability explains its association with severity. Limited precision and proportional odds concerns warrant cautious interpretation.

This study advances the literature by directly comparing the mean HbA1c with index HbA1c within an ordinal electrodiagnostic severity framework. Previous studies have demonstrated a correlation between contemporaneous HbA1c levels and electrodiagnostic neuropathy severity but did not evaluate mean HbA1c [10,11]. Using Baba stage 0–4 as the primary endpoint, this study examined the association between mean HbA1c and electrodiagnostic severity across the full Baba stage spectrum, rather than limiting the outcome to the presence or absence of DSPN [7-9]. Recent studies have emphasized the importance of structured yet clinically applicable electrodiagnostic algorithms for DSPN, such as the electrodiagnostic pathway proposed by Lee et al. [4] and related editorial perspectives [9]. In line with these efforts, the Baba classification offers an easily applicable NCS-based staging approach for characterizing severity in patients already referred for electrodiagnostic evaluation [7,8].

For exploratory clinical translation, we additionally examined advanced electrodiagnostic DSPN, defined as Baba stage 3–4 versus 0–2, and found that the mean HbA1c remained independently associated with this outcome. We selected this threshold because Baba stage 3 reflects more advanced large-fiber involvement, including motor amplitude reduction and sural sensory amplitude loss [7,8]. The ROC-derived threshold of 7.0% should nevertheless be interpreted cautiously. Discrimination was modest and imprecisely estimated (AUC, 0.701; bootstrap 95% CI, 0.567–0.822), with specificity of only 54.9%. Because the threshold was selected and evaluated in the same cohort, its apparent performance may be optimistic. Furthermore, 7.0% coincides numerically with the commonly used American Diabetes Association HbA1c target of <7% for many nonpregnant adults, although treatment goals are individualized and a longitudinal mean is not equivalent to a contemporaneous treatment-monitoring value [19]. Without external validation and evidence of incremental clinical utility, this exploratory threshold offers limited information beyond existing glycemic targets. The four-year look-back window and minimum measurement requirements were research definitions and should not be interpreted as clinical monitoring recommendations. This study does not establish that a 4-year HbA1c history is necessary or determine the optimal number or spacing of measurements for neuropathy assessment.

Lower eGFR was associated with advanced electrodiagnostic DSPN after adjustment. This finding may reflect shared diabetic microvascular burden, a contribution of renal dysfunction to peripheral nerve injury, or residual confounding. After excluding three patients receiving hemodialysis, the inverse association with eGFR was attenuated and no longer statistically significant, whereas the association with mean HbA1c remained similar. These findings warrant cautious interpretation of the renal association and do not establish a causal effect of renal dysfunction. Renal function also warrants consideration when interpreting HbA1c, because its accuracy may be reduced in advanced kidney disease, particularly among patients receiving dialysis [20].

This study has certain limitations. First, its retrospective single-center design limits causal inference and generalizability. Second, the cohort was restricted to patients with clinically suspected DSPN referred for NCS, which may have favored symptomatic or more severe presentations. Requiring repeated HbA1c measurements over at least two calendar years may also have preferentially selected patients with sustained follow-up or more frequent monitoring because of difficult glycemic control. These selection processes may have altered the distributions of glycemic exposure and neuropathy severity, although their effects on the observed association could not be determined. Differences in follow-up patterns between included and excluded patients were not assessed. Radiculopathy accounted for the exclusion of 117 of 263 screened patients (44.5%). Excluding radiculopathy and other neurological disorders may have reduced non-DSPN contributions to abnormal NCS findings but limits applicability to patients with coexisting DSPN and these disorders, including older adults with T2DM. The direction and magnitude of the resulting selection bias remain uncertain, limiting generalizability to the broader T2DM population. Third, HbA1c measurements were obtained in routine clinical practice rather than at uniform intervals, introducing heterogeneity into the assessment of longitudinal glycemic exposure. Time-weight-

ed analyses yielded smaller positive ordinal estimates with CIs including 1, indicating that the statistical evidence for the association was sensitive to how longitudinal exposure was summarized. These analyses used calendar-month intervals and linear interpolation between observed measurements and could not reconstruct unmeasured changes in HbA1c. Fourth, the final sample of 73 patients fell below the protocol-based minimum of 80, which was calculated for the originally planned exploratory binary comparison. The small sample and sparse Baba stage categories increased the risk of overfitting and limited precision, particularly for direct comparisons and exploratory analyses. Parsimonious and penalized analyses addressed some concerns regarding model stability but could not overcome the limited information available. Some leave-one-out analyses crossed the 0.05 significance threshold. Global Brant tests were not interpretable for the arithmetic and time-weighted mean HbA1c models because the estimated covariance matrices were not positive definite. Contrarily, interpretable Brant tests indicated departures from proportional odds in the CV and joint mean HbA1c–CV models, whereas supplementary Rao score tests did not. These discordant diagnostics and the heterogeneous, imprecise cut-point-specific estimates limit confidence in the proportional odds assumption. Accordingly, the common ordinal OR should be interpreted cautiously, and the nonsignificant direct comparisons should be considered in light of the limited precision. Fifth, the Baba classification reflects large-fiber dysfunction and does not fully capture small-fiber neuropathy [1,2,5,6]. Finally, several clinically relevant variables were not systematically assessed. Detailed classes and durations of oral glucose-lowering medications were unavailable, limiting assessment of factors such as possible vitamin B12 deficiency associated with long-term metformin use [21]. Although mean HbA1c associations were similar after excluding patients receiving hemodialysis in both ordinal and binary analyses, residual confounding by renal disease and limitations in HbA1c measurement remain possible. The use of Korean population-specific NCS reference standards may also limit generalizability. Metabolic risk factors, including obesity and dyslipidemia [21,22], as well as clinical symptom scores and sensory examination findings, were not fully assessed or adjusted for. Therefore, residual confounding cannot be excluded.

Despite these limitations, this study has the following strengths: It used an electrodiagnostically grounded ordinal severity framework, explicitly compared the mean HbA1c with index HbA1c, and explored a candidate HbA1c threshold for advanced electrodiagnostic DSPN. This framework may also be particularly useful in referral-based rehabilitation practice, where severity stratification can complement a simple presence-or-absence diagnosis. Future studies should evaluate whether continuous glucose monitoring-derived metrics, including time in range and glycemic variability indices, provide information beyond the mean HbA1c in explaining electrodiagnostic DSPN severity [23,24].

Overall, mean HbA1c showed a more consistent pattern of association with electrodiagnostic DSPN severity than index HbA1c across the separate ordinal and exploratory binary models examined. However, direct comparisons did not establish a statistically significant difference between the two measures in this cohort. Attenuation of the ordinal estimates with time weighting or CV adjustment further indicates uncertainty in the mean HbA1c association. Larger independent studies are needed to evaluate the comparative value of longitudinal mean versus index HbA1c and externally validate the exploratory threshold.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING INFORMATION

None.

AUTHOR CONTRIBUTION

Conceptualization: Han S. Data curation: Han S. Formal analysis: Han S. Investigation: Han S. Methodology: Han S. Resources: Han S. Project administration: Kim J. Visualization: Han S, Kim J. Writing – original draft: Han S. Writing – review and editing: Kim J, Han S, Choi S, Yoon ES, Yoo HK. Supervision: Kim J. Approval of final manuscript: all authors.

DATA AVAILABILITY STATEMENT

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

SUPPLEMENTARY MATERIALS

Supplementary materials can be found via https://doi.org/10.5535/arm.260110.

ORCID

Songi Han, https://orcid.org/0000-0002-5416-8576
Jongkyu Kim, https://orcid.org/0000-0003-1441-9227
Sujeong Choi, https://orcid.org/0009-0006-3829-4974
Eun Sang Yoon, https://orcid.org/0000-0002-3320-9316
Ho Kyung Yoo, https://orcid.org/0009-0008-9547-7093

REFERENCES

1. Pop-Busui R, Boulton AJ, Feldman EL, Bril V, Freeman R, Malik RA, et al. Diabetic neuropathy: a position statement by the American Diabetes Association. Diabetes Care 2017;40:136-54.
2. Tesfaye S, Boulton AJ, Dyck PJ, Freeman R, Horowitz M, Kempler P, et al. Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments. Diabetes Care 2010;33:2285-93.
3. England JD, Gronseth GS, Franklin G, Miller RG, Asbury AK, Carter GT, et al. Distal symmetrical polyneuropathy: definition for clinical research. Muscle Nerve 2005;31:113-23.
4. Lee HJ, Kim DH, Kim KH. Diabetic distal symmetric sensorimotor polyneuropathy: a proposal of new electrodiagnostic evaluation. Ann Rehabil Med 2023;47:231-3.
5. Sloan G, Selvarajah D, Tesfaye S. Pathogenesis, diagnosis and clinical management of diabetic sensorimotor peripheral neuropathy. Nat Rev Endocrinol 2021;17:400-20.
6. Tankisi H, Pugdahl K, Beniczky S, Andersen H, Fuglsang-Frederiksen A. Evidence-based recommendations for examination and diagnostic strategies of polyneuropathy electrodiagnosis. Clin Neurophysiol Pract 2019;4:214-22.
7. Baba M, Suzuki C, Ogawa Y. [Severity grading system of diabetic neuropathy in type-2 diabetes by nerve conduction study: five-year prospective study on occurrence of diabetic foot, macroangiopathic events, and eventual death]. Jpn J Clin Neurophysiol 2018;46:71-7. Japanese.
8. Himeno T, Kamiya H, Nakamura J. Lumos for the long trail: strategies for clinical diagnosis and severity staging for diabetic polyneuropathy and future directions. J Diabetes Investig 2020;11:5-16.
9. Oh BM. Refined diagnostic protocol for diabetic polyneuropathy: paving the way for timely detection. Ann Rehabil Med 2023;47:234-6.
10. Lee WJ, Jang S, Lee SH, Lee HS. Correlation between the severity of diabetic peripheral polyneuropathy and glycosylated hemoglobin levels: a quantitative study. Ann Rehabil Med 2016;40:263-70.
11. Tkac I, Bril V. Glycemic control is related to the electrophysiologic severity of diabetic peripheral sensorimotor polyneuropathy. Diabetes Care 1998;21:1749-52.
12. Lai YR, Chiu WC, Huang CC, Tsai NW, Wang HC, Lin WC, et al. HbA1C variability is strongly associated with the severity of peripheral neuropathy in patients with type 2 diabetes. Front Neurosci 2019;13:90.
13. Andersen H, Stålberg E, Falck B. F-wave latency, the most sensitive nerve conduction parameter in patients with diabetes mellitus. Muscle Nerve 1997;20:1296-302.
14. Kim JY, Kim E, Shim HS, Lee JH, Lee GJ, Kim K, et al. Reference standards for nerve conduction studies of individual nerves of lower extremity with expanded uncertainty in healthy Korean adults. Ann Rehabil Med 2022;46:9-23.
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.
16. Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 1990;46:1171-8.
17. Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80:27-38.
18. Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. Am J Epidemiol 1999;150:341-53.
19. American Diabetes Association Professional Practice Committee. 6. Glycemic goals and hypoglycemia: Standards of Care in Diabetes—2024. Diabetes Care 2024;47(Suppl 1):S111-25.
20. de Boer IH, Khunti K, Sadusky T, Tuttle KR, Neumiller JJ, Rhee CM, et al. Diabetes management in chronic kidney disease: a consensus report by the American Diabetes Association (ADA) and Kidney Disease: Improving Global Outcomes (KDIGO). Diabetes Care 2022;45:3075-90.
21. American Diabetes Association Professional Practice Committee. 12. Retinopathy, neuropathy, and foot care: Standards of Care in Diabetes—2024. Diabetes Care 2024;47(Suppl 1):S231-43.
22. Callaghan BC, Gallagher G, Fridman V, Feldman EL. Diabetic neuropathy: what does the future hold? Diabetologia 2020;63:891-7.
23. Gouveri E, Papanas N. The emerging role of continuous glucose monitoring in the management of diabetic peripheral neuropathy: a narrative review. Diabetes Ther 2022;13:931-52.
24. Kim MY, Kim G, Park JY, Choi MS, Jun JE, Lee YB, et al. The association between continuous glucose monitoring-derived metrics and cardiovascular autonomic neuropathy in outpatients with type 2 diabetes. Diabetes Technol Ther 2021;23:434-42.

Instructions for authors

Enacted in May, 2011
Revised on January 31, 2013
Revised on August 1, 2018
Revised on March 11, 2019
Revised on March 16, 2020
Revised on June 9, 2023
Revised on January 19, 2024
Revised on April 16, 2025
Revised on February 10, 2026

arm
Annals of Rehabilitation Medicine

1. AIMS & SCOPE

Annals of Rehabilitation Medicine (ARM) is the official journal of the Korean Academy of Rehabilitation Medicine. It is an international, peer-reviewed open access journal, which aims to be a global leader in sharing up-to-date knowledge dedicated to the advancement of care and enhancing the function and quality of life of persons with various disabilities and chronic illnesses. As the official journal of one of the largest societies of rehabilitation medicine in Asia and Oceania, more than 3,500 physiatrists regularly receive this journal's table of contents via email, and the number is continually increasing. This journal is endorsed by the International Society of Physical and Rehabilitation Medicine (ISPRM) and the Asia-Oceanian Society of Physical and Rehabilitation Medicine (AOSPRM). International members comprise approximately half the editorial board, conducting peer-reviews of submitted manuscripts or acting as handling editors by assigning peer-review tasks to other reviewers.

The journal encompasses all aspects of physical medicine and rehabilitation, including basic, translational, clinical, and policy research, as well as rehabilitation education. Research areas covered by this journal include rehabilitation for brain disorders, spinal cord injuries, neuromuscular diseases, and dysphagia; cancer, cardiopulmonary, geriatric, pain & musculoskeletal, and pediatric rehabilitation; electrodiagnosis, orthosis & prosthesis, physical therapy, and sports medicine; and other emerging fields in rehabilitation medicine.

2. COPYRIGHTS AND CREATIVE COMMONS ATTRIBUTION LICENSE

The Korean Academy of Rehabilitation Medicine is the owner of all copyright to papers published in *ARM*, and has the right to publish, reproduce, distribute, and print the contents in other types of media. Authors of accepted papers must complete the Copyright Transfer Form. A letter of permission is required for any and all material that has been published previously. It is the responsibility of the author to request permission from the publisher for any material that is being reproduced. This requirement applies to text, illustrations, and tables. This is identical to the Creative Commons (Attribution-Noncommercial) license available at http://creativecommons.org/.

3. GENERAL GUIDELINES

The manuscript guidelines for *ARM* are based on the "Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals" published by the International Committee of Medical Journal Editors (ICMJE, http://www.icmje.org/), and instructions which are not mentioned in the present guidelines are referred to the guidelines stated in the ICMJE Recommendations.

1) ARTICLE TYPES

Manuscripts include original articles, review articles, brief reports, case reports, images in this issue, and letters to the editor.

(1) Original articles

This form of publication represents original research articles reporting the results of basic and clinical investigations that are sufficiently well documented to be acceptable to critical readers.

(2) Review articles

The Editorial Board welcomes state-of-the-art review articles. The *ARM* strongly prefers systematic reviews of the literature. Invited review articles provide a comprehensive review of a subject of importance to clinicians and researchers and are commissioned by the editorial board to an invited expert in the field.

(3) Brief reports

These manuscripts are short but important reports to provide preliminary communications with less complete data sets than would be appropriate for original contributions that present novel and impactful clinical and basic research of a more preliminary nature.

(4) Case reports

Case reports are considered for publication when at least one of

the following criteria is met: (a) a rare condition is reported, (b) atypical symptoms and signs are observed, (c) new diagnostic or therapeutic methods are introduced, (d) atypical clinical and laboratory findings for populations residing in Asia and the Pacific Rim. Descriptions of clinical cases (individual or a series) should be unique, should deal with clinical cases of exceptional interest or innovation and should preferably be a first-time report.

(5) Images in this issue

This form of publication represents images (e.g.,radiographs, CT, MRI, electrodiagnostic tracings, pathology, physical examination findings, photos of a patient or medical device) that are interesting and unique.

(6) Letters to the editor

Critical comments are welcomed for providing alternative interpretations or views about articles published in *ARM*. Letters should be directly related to the published article on which it comments. Letters being considered for publication ordinarily will be sent to the authors, who will be given the opportunity to reply. Letters will be published at the discretion of the editors and are subject to abridgement and editing for style and content.

2) LANGUAGE OF MANUSCRIPT

All manuscripts must be written in clearly understandable English. Authors whose first language is not English are requested to have their manuscripts checked for grammatical and linguistic correctness before submission. Correct medical terminology should be used, and jargon should be avoided. Use of abbreviations should be minimized and restricted to those that are generally recognized. When using an abbreviated word, it should be spelled out in full on first usage in the manuscript followed by the abbreviation in parentheses. Numbers should be written in Arabic numerals, but must be spelled out when placed in the beginning of a sentence. Measurements should be reported using the metric system, and hematologic and biochemical markers should be reported in International System (SI) of Units. All units must be preceded by one space except percentage (%), temperature (°C), and degree (°).

4. RESEARCH AND PUBLICATION ETHICS

All manuscripts should be written with strict adherence to the research and publication ethics guidelines recommended by Council of Science Editors (http://www.councilscienceeditors.org/), International Committee of Medical Journal Editors (ICMJE, http://www.icmje.org/), World Association of Medical Editors (WAME, http://www.wame.org/), and the Korean Association of Medical Journal Editors (KAMJE, https://www.kamje.or.kr/en/main_en). For all studies involving human subjects, the principles embodied in the Declaration of Helsinki (https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/) should be upheld, informed consent must be obtained from all participants, and must be approved by a recognized Institutional Review Board (IRB) or research ethics committee. The editor may request submission of copies of informed consents from human subjects in clinical studies or IRB approval documents. Experiments involving animals should comply with the NIH guidelines for the use of laboratory animals (https://www.nlm.nih.gov/services/research_report_guide.html) and/or be reviewed by an appropriate committee (e.g., Institutional Animal Care and Use Committee, IACUC) to ensure the ethical treatment of animals in research. Also, studies with pathogens requiring a high degree of biosafety should pass review of a relevant committee (e.g., Institutional Biosafety Committee, IBC). *ARM* will follow the guidelines by the Committee on Publication Ethics (COPE, http://publicationethics.org/) for settlement of any misconduct.

1) REDUNDANT PUBLICATION AND PLAGIARISM

All submitted manuscripts should be original and should not be considered by other scientific journals for publication at the same time. No part of the accepted manuscript should be duplicated in any other scientific journal without the permission of the editorial board. If plagiarism or duplicate publication related to the papers of this journal is detected, the manuscripts may be rejected, the authors will be announced in the journal, and their institutes will be informed. There will also be penalties for the authors.

2) AUTHORSHIP

ARM follows the recommendations by International Committee of Medical Journal Editors (ICMJE,http://www.icmje.org/) and and the Korean Association of Medical Journal Editors (KAMJE, https://www.kamje.or.kr/en/main_en). Authorship is credited to those who have direct involvement in the study and have made significant contributions to (a) substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (b) drafting the work or revising it critically for important intellectual content; AND (c) final approval of the version to be published; AND (d) agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved as recommended by ICMJE. The primary investigator is designated the first author of

the study, unless contested by the other authors. The corresponding author is directly responsible for communication and revision of the submitted manuscript. Authors are required to include a statement of responsibility in the manuscript that specifies the contribution of every author at the end of the manuscript, in a section entitled "Author contribution". All persons who have made substantial contribution, but who are not eligible as authors should be named in the acknowledgments. In the case of change of authorship, a written explanation must be submitted. Change in either the first author or the corresponding author requires approval by the editorial board, and any changes in the other authors require approval by the editor-in-chief.

3) CONFLICT OF INTEREST

The corresponding author of an article is asked to inform the editor of the authors' potential conflicts of interest possibly influencing their interpretation of data. A potential conflict of interest must be disclosed during the online submission process on the appropriate web page. Such conflicts may be financial support or private connections to pharmaceutical companies, political pressure from interest groups, or academic problems based on the "ICMJE Uniform Disclosure Form for Potential Conflicts of Interest" (http://www.icmje.org/coi_disclosure.pdf). The editor will decide whether the information on the conflict should be included in the published paper. Before publishing such information, the editor will consult with the corresponding author. In particular, all sources of funding for a study should be explicitly stated.

4) REGISTRATION OF CLINICAL TRIAL

Clinical trial defined as "any research project that prospectively assigns human subjects to intervention and comparison groups to study the cause-and-effect relationship between a medical intervention and a health outcome" is recommended to be registered to the primary registry to be prior publication. *ARM* recommend, as a condition of consideration for publication, registration in a public trials registry. *ARM* accepts the registration in any of the primary registries that participate in the WHO International Clinical Trials Portal (http://www.who.int/ictrp/en/), NIH ClinicalTrials.gov (http://www.clinicaltrials.gov/), ISRCTN Resister (www.isrctn.org), ANZCTR (https://www.anzctr.org.au/), EudraCT Database (https://eudract.ema.europa.eu/), Clinical Trials Information System (https://euclinicaltrials.eu/), University Hospital Medical Information Network (www.umin.ac.jp/ctr/index/htm), EU Clinical Trials Register (https://www.clinicaltrialsregister.eu/) or The Clinical Research Information Service (http://cris.nih.go.kr/). The clinical trial registration number will be published at the end of the abstract.

5) DATA AVAILABILITY STATEMENT

ARM follows the ICMJE Recommendations regarding data sharing statements (http://icmje.org/icmje-recommendations.pdf). All manuscripts reporting clinical trial results should include a data sharing statement in accordance with the ICMJE guidelines.

ARM requires authors of all manuscripts reporting results derived from research data to include a Data Availability Statement. The provision of a Data Availability Statement will be verified as a condition of publication. Data Availability Statements should include information on where the data supporting the results reported in the article can be found, including, where applicable, hyperlinks or persistent identifiers (e.g., DOI/URL) to publicly archived datasets analyzed or generated during the study.

Where research data are not publicly available, this must be stated in the manuscript along with the reason(s) (e.g., participant confidentiality, ethical or legal restrictions, institutional policies, or third-party ownership/licensing) and any conditions for accessing the data.

6) PROCESS FOR MANAGING RESEARCH AND PUBLICATION MISCONDUCT

When the journal faces suspected cases of research and publication misconduct such as redundant (duplicate) publication, plagiarism, fraudulent or fabricated data, changes in authorship, an undisclosed conflict of interest, ethical problems with a submitted manuscript, a reviewer who has appropriated an author's idea or data, complaints against editors, and so on, the resolution process will follow the flowchart provided by the Committee on Publication Ethics (COPE, https://publicationethics.org/guidance/Flowcharts). The discussion and decision on the suspected cases are carried out by the Editorial Board.

7) PROCESS FOR HANDLING CASES REQUIRING CORRECTIONS, RETRACTIONS, AND EDITORIAL EXPRESSIONS OF CONCERN

Cases that require editorial expressions of concern or retraction shall follow the COPE flowcharts available from: https://publicationethics.org/guidance/Flowcharts. If correction needs, it will follow the ICMJE Recommendation for Corrections, Retractions, Republications and Version Control available from: http://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/corrections-and-version-control.html.

8) DISCLOSURE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

If generative AI-based technologies (e.g., large language models, chatbots, or image generators) are used at any stage of manuscript

preparation (e.g., writing assistance, data analysis, figure generation), authors must transparently disclose their use. This disclosure should be included in the manuscript (e.g., in the Methods or Acknowledgments) and is also recommended in the cover letter. It is important to note that generative AI tools cannot be listed as authors, as they are not capable of accountability or authorship. Authors must take full responsibility for all AI-generated content, including inaccurate, incomplete, biased, or plagiarized material, and must ensure its appropriateness, integrity, and adherence to research and publication ethics.

5. MANUSCRIPT SUBMISSION

All submissions are made online at the journal's online manuscript submission site (https://www.editorialmanager.com/arm/default.aspx) by the corresponding author. Submitted manuscripts are initially examined for format, and then appointed a submission number. For nonbiased peer review, authors' names and institutional affiliations should not be mentioned in the text. The revised manuscript should be submitted through the same web system under the same identification numbers. The date of final review for the manuscript will be the date of acceptance for publication. If you have any questions about the online submission process, contact the Editorial Office by e-mail at edit@e-arm.org.

6. PEER REVIEW PROCESS

1) EDITORIAL REJECT POLICY

Conformity of the submitted manuscript to the submission instructions is examined upon submission. The Editorial Board may reject the manuscript or request the author to resubmit in the following cases: 1) Topic clearly out of scope / insufficient perceptual content 2) Work clearly does not meet sufficient standards of novelty or quality 3) Manuscript incomplete or incorrectly formatted 4) Suspected plagiarism in the manuscript.

2) PEER REVIEW PROCESS

Submitted manuscripts will be reviewed by two or more peer reviewers selected from the board's database of expert reviewers. In addition, if deemed necessary, a review of statistics may be requested. Following review, the editorial board will decide whether the manuscript will be 1) accepted for publication, 2) subject to minor revision, 3) subject to major revision, or 4) rejected for publication. For manuscripts which are either subject to minor revision or subject to major revision, the corresponding author must resubmit the revised manuscript online. The revised manuscript should clearly indicate all changes by highlighting the modified text or by using a different color for the altered sections. In addition, the corresponding author must reply to both reviewers' comments point by point, and explain in detail what changes were made in the manuscript. When considered necessary, the editorial board may make changes to the structure and phrases of the manuscript without compromising the integrity of the original paper. After completion of the peer review process, the editorial board will determine acceptance for publication and notify the corresponding author by e-mail. Manuscripts which do not comply with the present guidelines will be notified for correction or withheld from publication.

When a manuscript is not resubmitted within 2 months of notification, it will be considered that the authors have withdrawn the manuscript from submission. Manuscripts accepted for publication are generally published in order of submission, depending on the category of the manuscript and the date of acceptance for publication.

7. PREPARATION OF THE MANUSCRIPT

All materials must be written in proper and clear English. Use Microsoft Office Word (versions after 2003) and ensure correct spelling and grammar. Setup the MS Word document for 1-inch margins on letter or A4-sized paper. The manuscript must be written in 12-point font and the sentences must be double-spaced, including tables and figure legends. Each page should be numbered in the middle of the lower margin, and all sentences must be numbered sequentially throughout the entirety of the manuscript, starting with the title page.

All papers must be accompanied by a title page. The title page should contain the title of the manuscript, a short running title, the authors' names, academic degrees, respective affiliations, and ORCID. The corresponding author must be identified, and his or her contact information (postal address, e-mail, telephone and fax numbers) should be listed. The title should clearly describe the objective of the study and contain less than 20 words. The first letter of each word of the title should be in capital letters except for prepositions, articles, and conjunctions. Provide a short running title containing less than 10 words. In cases in which the authors belong to multiple affiliations, the affiliations during the study being reported should be matched to the authors' names using a superscript of Arabic numerals. Conflicts of interest, funding information, author contribution and acknowledgements (when applicable) should also be located in the title page.

1) ORIGINAL ARTICLES

Original papers should be structured in the following order: Ab-

stract, Introduction, Methods, Results, Discussion, Conflict of interests, Funding information, Author contribution, Acknowledgments (when applicable), References, Tables, Figure legends, and Figures. Maximum word count is limited to 5,000 words.

(1) Abstract

A structured abstract with the headings of Objective, Methods, Results, and Conclusion must succinctly describe the paper in 250 words or less. Use complete sentences and do not number the results. At the end of the abstract, list up to 5 relevant keywords which are in accordance to the Medical Subject Headings (MeSH) in the Index Medicus (http://www.nlm.nih.gov/mesh).

(2) Introduction

Clearly present the objective of the study and its relationship to earlier work in the field. A brief background to inform the readers of the relevance of the study may be necessary. However, avoid extensive review of the literature.

(3) Methods

Describe the participants or research materials of the study, and explain in detail the inclusion and exclusion criteria for both the experimental and control groups. Describe the experimental methods in a logical and systematic manner so that they can be reproducible by another investigator. Experimental drugs should be stated in the generic name. When proprietary brands are used, include the brand name and the name of the manufacturer in parentheses after the first mention of the generic name. When using experimental devices or other products, state the brand name then follow with the name of the manufacturer, in parentheses, e.g., Flow Cytometer (Coulter Electronic Inc.). To ensure anonymity during the peer review process, the authors' affiliations or institutional setting of the study should not be revealed. Statistical analysis and criteria for determining significance should be described in enough detail to allow the knowledgeable reader with access to the original data to verify the reported results. An ethics statement should be placed here when the studies are performed using clinical samples or data, and animals.

Ensure correct use of the terms sex (when reporting biological factors) and gender (identity, psychosocial or cultural factors), and, unless inappropriate, report the sex and/or gender of study participants, the sex of animals or cells, and describe the methods used to determine sex and gender. If the study was done involving an exclusive population, for example in only one sex, authors should justify why, except in obvious cases (e.g., prostate cancer). Authors should define how they determined race or ethnicity and justify their relevance.

Authors who used AI technology to conduct their study should describe its use in the Methods section with sufficient detail to enable replication of the approach. This should include the name of the tool, version number, and the prompts used, where applicable.

(4) Results

Summarize and describe logically the significant findings and trends observed in the results using text, figures and tables. Avoid extensive repetition of contents of the tables and figures in the text.

In statistical expression, mean and standard deviation should be described as mean ± SD, and mean and standard error as mean ± SE. In general, p-values larger than 0.01 should be reported to two decimal places, those between 0.01 and 0.001 to three decimal places; p-values smaller than 0.001 should be reported as $p < 0.001$.

(5) Discussion

Interpret the results in respect to the objective of the study, and describe differences with previous studies and significant findings which lead to the deduction of the conclusion. Refrain from excessive review of historic studies, textbook facts, or irrelevant references. Accentuate newly obtained observations from the study, and include significant limitations of the study.

(6) Conflicts of interest

Any potential conflicts of interest relevant to the manuscript should be described. If there are no conflicts of interest, authors should state that none exists.

(7) Funding information

All sources of funding applicable to the study should be stated here explicitly. All original articles, editorials, reviews, and new technology articles must state funding sources for the study.

(8) Data Availability Statements

Data Availability Statements should take one of the following forms (or a combination if multiple types of data apply):

- The datasets generated and/or analyzed during the current study are available in the [NAME] repository, [PERSISTENT LINK/DOI].
- The datasets generated and/or analyzed during the current study are included in this published article [and its supplementary information files].
- The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

- The datasets generated and/or analyzed during the current study are not publicly available due to [REASON] but are available from the corresponding author on reasonable request [and/or with approval/permission of …].
- Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

(9) Author contribution

The individual contributions of the authors to the manuscript should be specified in this section. Each author's contribution must be clearly described using the CRediT (Contributor Roles Taxonomy) system. To qualify for authorship, individuals must have participated in at least one of the core 12 CRediT roles (e.g., conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization) and at least one of the writing contributions (preparation of the original draft and/or critical review and revision of the manuscript). These contributions will be published in the final article, and all authors must review and approve their statements before submission.

(10) Acknowledgments

Persons who have made contributions to the study, but who are not eligible for authorship can be named in this section. Their contribution must be specified, such as data collection, financial support, statistical analysis, or experimentation. The corresponding author must inform the named contributor of the acknowledgment, and acquire consent before manuscript submission.

(11) References

- Cite only references which are quoted in the text. Limit the number of references 40.
- When quoting a reference in the text, refrain from stating the author's name, and identify references with Arabic numerals in brackets such as [1], [2-4], and [5,7,9].
- The references should be listed in order of citation in the text.
- List all authors when there are 6 or fewer; when there are 7 or more, list the first 6, followed by "et al."
- Journal names should be abbreviated according to the format listed in the Index Medicus. If the journal is not listed in the Index Medicus, refer to the list of title word abbreviations by the ISSN network (http://www.issn.org/2-22660-LTWA.php).
- For more on references, refer to the NLM Style Guide for Authors, Editors, and Publishers (http://www.nlm.nih.gov/citingmedicine).

Journals

1. Jeon JH, Jung YJ, Lee JY, Choi JS, Mun JH, Park WY, et al. The effect of extracorporeal shock wave therapy on myofascial pain syndrome. Ann Rehabil Med 2012;36:665-74.
2. Geraud G, Spierings EL, Keywood C. Tolerability and safety of frovatriptan with short- and long-term use for treatment of migraine and in comparison with sumatripan. Headache 2002;42 Suppl:S93-9.

Book & Chapter of book

3. Frontera W, Silver JK, Rizzo TD. Essentials of physical medicine and rehabilitation. 2nd ed. Saunders; 2008. p. 579-82.
4. Esquenazi A. Upper limb amputee rehabilitation and prosthetic restoration. In: Braddon RL, editor. Physical medicine and rehabilitation. 2nd ed. Saunders; 2000. p. 263-78.

Proceedings of academic conference

5. Harnden P, Joffe JK, Jones WG. Germ cell tumours V. Proceedings of the 5th Germ Cell Tumour Conference; 2001 Sep 13-15; Leeds, UK. New York: Springer; 2002.

Thesis (Dissertation)

6. Borkowski MM. Infant sleep and feeding: a telephone survey of Hispanic Americans [dissertation]. Mount Pleasant, MI: Central Michigan University; 2002.

(11) Tables

Tables should be submitted separately from the text, and each table should be created in MS Word on separate pages, using double space throughout. They should be simple, self-explanatory, and not redundant with the text or the figures. Limit 5 tables per manuscript. The title of the tables should be written in phrases, and capitalized the first letter of the first word. The title should be placed above the table, and abbreviations and footnotes should be placed under the table. Number the tables in order of appearance in the text (e.g., Table 1, Table 2). All abbreviations used in the table must be spelled-out in full under the table in the following order: abbreviation, comma, full word (e.g., RM, rehabilitation medicine;). Table footnotes should be indicated in superscripts in the following order: [a], [b], [c]… but p-values should be indicated by asterisk (e.g., *$p < 0.05$, **$p < 0.01$. ***$p < 0.001$).

(12) Figure legends

Legends should be submitted separately from the text, and each legend should be typed on separate pages. They should be written in full sentences to describe the content of the figure, and only the first letter of the legend should be capitalized. For lengthy legends

continuing beyond one line, the left margin of the following lines should start at the same point as the first line. Any symbols, marks or abbreviations made in the figure must be explained in the legend. Figures containing histologic slides should be accompanied by legends explaining tissue origin, stain method, and microscopic amplification.

(13) Figures

Figures should be uploaded online as separate files and numbered in order of appearance in the text (e.g., Fig. 1). When a single numbered figure contains 2 or more figures, the figure should be numbered with an alphabet letter following the number (e.g., Fig. 1A, Fig. 1B). Indicate focus points in the figures with markers such as arrows and arrowheads, etc. Image files must be of resolutions higher than 300 dpi, and less than 3 MB, in JPEG, GIF, TIFF, or Microsoft PowerPoint format. A single numbered figure containing more 2 or more figures such as Fig. 1A and Fig. 1B should be uploaded as a single file.

(14) Graphical Abstract (Optional)

Authors submitting Original Articles, Review Articles, Brief Reports, or Case Reports are encouraged to include a graphical abstract summarizing the key message of the study. A template and further instructions are available on the journal and submission website. The graphical abstract should be uploaded as a separate file during manuscript submission.

2) REVIEW ARTICLES

The abstract should contain no more than 250 words and 5 keywords. The text is structured in the order of Introduction, Main text, Conclusion, Conflict of interests, Funding information, Author contribution, Acknowledgments (when applicable), References, Tables, Figure legends, and Figures.

3) BRIEF REPORTS

General guidelines are the same as for the original article. The manuscript is structured in the order of Abstract, Main text, Conflict of interests, Funding information, Author contribution, Acknowledgments (when applicable), References, Tables, Figure legends, and Figures. A structured abstract is required and limited to 150 words, with no more than 3 keywords attached. Manuscripts should be limited to 1,500 words of text including references and figure legends (not including abstract, tables, and figures), and no more than 10 references. The total number of figures and/or tables is limited to 3.

4) CASE REPORTS

General guidelines and order of manuscript preparation are the same as for the original article. Case reports are considered for publication only if they report rare conditions, atypical symptoms and signs, novel diagnostic or therapeutic approaches, or describe atypical findings for populations residing in Asia and the Pacific Rim. The editorial board will determine whether the case report fulfills the above criteria for acceptance of publication. The manuscript is structured in the order of Abstract, Introduction, Case report, Discussion, Conflict of interests, Funding information, Author contribution, Acknowledgments (when applicable), References, Tables, Figure legends, and Figures. The abstract should be nonstructured and limited to 150 words, with no more than 3 keywords attached. The introduction should briefly state the background and significance of the case. The actual case report should describe the clinical presentation and the diagnostic and therapeutic measures taken. The discussion should focus on the uniqueness of the case and should not contain extensive review of the disease or disorder. The combined number of tables and figures is limited to 5, and the number of references is limited to 10. Maximum word count is limited to 1,500 words including references and figure legends.

5) IMAGES IN THIS ISSUE

All images should be accompanied by a short description of the image and a brief and concise clinical review of the specific patient or clinical issue of no more than 500 words (excluding references) with references limited to 5. Image files must be of resolutions higher than 300 dpi for photographs, and 900 dpi for line art, waveforms, and graphs, in JPEG, GIF, TIFF, or Microsoft PowerPoint format. Images should make up a single figure, although they may contain more than one frame. The manuscript does not have an abstract.

6) LETTERS TO THE EDITOR

Letters should not have an abstract, tables, figures, and data supplements. Letters must be limited to roughly 500 words of text and no more than 5 references, 1 of which should be to the recent *ARM* article. Letters may have no more than 3 authors.

7) REPORTING GUIDELINES FOR SPECIFIC STUDY DESIGNS

Authors are encouraged to adhere to internationally accepted reporting guidelines for their study type, such as CONSORT (randomized controlled trials), PRISMA (systematic reviews), STROBE (observational studies), and CARE (case reports). These can be found at the EQUATOR Network (https://www.equator-network.org).

Initiative	Type of study	Source
CONSORT	Randomized controlled trials	http://www.consort-statement.org
STARD	Studies of diagnostic accuracy	http://www.stard-statement.org
PRISMA	Preferred reporting items of systematic reviews and metaanalyses	http://www.prisma-statement.org
STROBE	Observational studies	http://www.strobe-statement.org
MOOSE	Meta-analyses of observational studies in epidemiology	http://www.consort-Statement.org/resources/downloads/other-instruments/moose-statement-2000pdf

8. SUBMISSION APPLICATION & COPYRIGHT TRANSFER

All submitted manuscripts must be accompanied by the official Submission Application & Copyright Transfer Form of the Korean Academy of Rehabilitation Medicine. The Submission Application & Copyright Transfer Form must contain the title of the manuscript, date of submission, names of all authors, authors' affiliations, and written signatures. Note the corresponding author and provide his/her affiliation, e-mail, telephone and fax numbers, and mailing address.

For the copyrights of the contributions published in *ARM*, see Creative Commons (Attribution-Noncommercial) at http://creativecommons.org.

9. MANUSCRIPTS AFTER ACCEPTANCE

ARM is published in English bi-monthly on the last days of February, April, June, August, October, and December.

1) FINAL VERSION UPLOAD

When accepted for publication, the authors' institutional affiliations should be inserted into the text of the final revised manuscript and uploaded to the online submission system. Files containing figures should be of the highest resolution (at least 300 dpi for color figures, and 900 dpi for line art and graphs) should be also be uploaded in JPEG, GIF, or TIFF format, and must be named according to the figure number (e.g., Fig. 1.jpg).

2) GALLEY PROOF

Galley proofs will be sent to the corresponding author for final corrections. Corrections should be kept to a minimum, must be returned within 2 days, otherwise publication may be delayed. Any fault found after the publication is the responsibility of the authors. We urge our contributors to proofread their accepted manuscripts very carefully. After the publication, if there are critical errors, they should be corrected as Corrigendum or Erratum.

3) PUBLICATION

The editorial board retains the right to request minor stylistic and major alterations that might influence the scientific content of the paper. The final manuscript will be published following final approval by the editor-in-chief.

4) ARTICLE PROCESSING CHARGES

Annals of Rehabilitation Medicine is an open access journal. To publish in *Annals of Rehabilitation Medicine*, authors are asked to pay an article processing charge (APC) on acceptance of their research paper. The APC for all published papers is as follows, plus VAT or local taxes where applicable. The currency KRW will be applied to the submissions from South Korea.

- Original article / Brief Reports / Unsolicited review article: 600 USD or 600,000 KRW
- Solicited article: Free
- Case report: 300 USD or 300,000 KRW
- Image in this issue: 170 USD or 170,000 KRW
- Correspondence / erratum: 100 USD or 100,000 KRW

†Note: As categorized by the World Bank, low-income countries will be eligible for a 20% discount on the updated APCs for original articles/brief reports/review articles/case reports. To find out if your country qualifies, please refer to the World Bank's classification available at https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html.

Checklist for Authors

Annals of Rehabilitation Medicine

General

☐ All elements of the manuscript are printed in English and double-spaced with 1-inch margins at top, bottom, and sides. Right margins are unjustified.

☐ All pages are numbered in the following order: title page, structured or standard abstract, body of the text, conflict of interests, Funding information, Author contribution, Acknowledgments (when applicable), references, legends, and tables.

☐ The text is consecutively line numbered.

☐ The Submission Application & Copyright Transfer Form is signed by the guarantor at original submission.

Abstract (applied to original articles, review articles, brief reports, and case reports)

☐ A structured abstract with the headings of Objective, Methods, Results, and Conclusion (A nonstructured abstract for case reports) must succinctly describe the paper.

☐ At the end of the abstract, relevant keywords are listed.

References

☐ All references have been checked for accuracy and completeness.

☐ Cite only references which are quoted in the text. Limit the number of references 40 for original articles, 10 for brief reports and case reports, and 5 for images in this issue and letters to the editor.

☐ Are numbered consecutively in the order they are cited in the text; all listed references have been cited in the text.

☐ Do NOT parenthesize the superscript numerals, and hyphenate (-) when citing 3 or more references in consecutive order.

☐ The format prescribed by the "Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals (ICMJE Recommendations)" has been followed. Examples provided under Instructions for Authors have been uploaded.

☐ List all authors when there are 6 or fewer; when there are 7 or more, list the first 6, followed by "et al."

☐ Journal names should be abbreviated according to the format listed in the Index Medicus. If the journal is not listed in the Index Medicus, refer to the list of title word abbreviations by the ISSN network (http://www.issn.org/2-22660-LTWA.php).

Figure Legends

☐ Figure legends are provided for each figure.

☐ Figure legends are numbered and presented together in numeric order following reference page(s).

Tables

☐ Each table is headed by a title and numbered in Arabic numerals on a separate page.

☐ The title of the tables should be written in phrases, and capitalized the first letter of the first word.

☐ There are less than 5 tables in the text of original articles.

☐ Tables are cited in numeric sequence in the text.

Figures

☐ Each is numbered with an Arabic numeral and cited in numeric sequence in the text.

☐ Photographs of recognizable persons should be accompanied by a signed release from the patient or legal guardian authorizing publication. Masking eyes to hide identity is not sufficient.

Credits & Permissions

☐ In addition to the notice of informed consent and releases to publish photographs of recognizable persons, submit with the manuscript written permission to use non-original material (quotations exceeding 100 words, any table or illustration) from copyright holder of the original. No article will be accepted as a submission to Archives without all required permissions. Credit the source in a text or table footnote or in a legend.

A more complete description of each item that must be checked is provided under the appropriate heading in the Instructions for Authors.

I have reviewed this Checklist and have complied with its requirements.

Every author took a certain role and made contribution to the study and the manuscript. In case of publication, I agree to transfer all copyright ownership of the manuscript to the Korean Academy of Rehabilitation Medicine to use, reproduce, or distribute the article.

☐ Agree ☐ Disagree