

Original Research

Temporal Invariance and Validity of the European Quality of Life 5 Dimensions 3 Level Version Among Taiwanese Older Adults Following Hip Fracture Surgery

Tri Laksono^{1,2}, Yu-Pin Chen^{3,4}, Ka-Ying Yap¹, Yi-Jung Chen¹, Chi Hsien Huang^{5,*}, Mark D. Griffiths⁶, Lien-Chen Wu^{3,4}, Yi-Jie Kuo^{3,4}, Chung-Ying Lin^{1,7,8,9,*}

¹Institute of Allied Health Sciences, College of Medicine, National Cheng Kung University, 701401 Tainan, Taiwan

²Department of Physiotherapy, Faculty of Health Sciences, Universitas 'Aisyiyah Yogyakarta, 55292 Sleman, Indonesia

³Department of Orthopedics, Wan Fang Hospital, Taipei Medical University, 116079 Taipei, Taiwan

⁴Department of Orthopedics, School of Medicine, College of Medicine, Taipei Medical University, 11031 Taipei, Taiwan

⁵Department of Family Medicine and Community Medicine, E-Da Hospital, I-Shou University, 824005 Kaohsiung, Taiwan

⁶Nottingham Institute of Psychology, Nottingham Trent University, NG1 4FQ Nottingham, UK

⁷Biostatistics Consulting Center, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, 701401 Tainan, Taiwan

⁸School of Nursing, College of Nursing, Kaohsiung Medical University, 807378 Kaohsiung, Taiwan

⁹Department of Occupational Therapy, College of Health Sciences, Kaohsiung Medical University, 807378 Kaohsiung, Taiwan

*Correspondence: ed103520@edah.org.tw; evaairgigaa@gmail.com (Chi Hsien Huang); cylin36933@gmail.com; cylin36933@gs.ncku.edu.tw (Chung-Ying Lin)

Academic Editor: Jose Pedro Espada

Submitted: 11 January 2026 Revised: 21 April 2026 Accepted: 28 April 2026 Published: 21 October 2026

Abstract

Background: The EuroQoL five-dimension, 3-level (EQ-5D-3L) instrument is widely used to assess health-related quality of life (HRQoL). However, its longitudinal validity among older adults after hip fracture surgery remains unclear. The present study examined the psychometric properties of EQ-5D-3L, with emphasis on factorial validity, temporal invariance, and score stability. **Methods:** A total of 719 older adults who underwent hip fracture surgery completed the EQ-5D-3L at baseline and 562 completed the EQ-5D-3L at 3 months postoperatively. **Results:** Confirmatory factor analysis (CFA) supported a single-factor structure at both time points (baseline: comparative fit index [CFI] = 0.988, Tucker–Lewis index [TLI] = 0.975; 3 months: CFI = 0.997, TLI = 0.993). Temporal invariance was mostly supported ($\Delta\text{CFI} \leq 0.01$) across configural, metric, and scalar models. However, the correlation between baseline and follow-up utility scores was nonsignificant ($r = 0.072$, $p = 0.052$), indicating varied recovery patterns. **Conclusions:** These findings support the EQ-5D-3L as a valid longitudinal measure of HRQoL among older adults following hip fracture surgery.

Keywords: EQ-5D-3L; psychometric analysis; temporal invariance; confirmatory factor analysis; hip fracture; aged; quality of life

1. Introduction

The global population is aging rapidly, leading to an unprecedented increase in the incidence of hip fractures and their associated health burdens (Harris et al., 2024). According to the Global Burden of Disease Study 2019, the incidence of hip fractures among adults aged ≥ 55 years reached 681.35 per 100,000 population in 2019 (95% uncertainty interval [UI] 508.36–892.27), with a steadily increasing trend over the past three decades (Feng et al., 2023). The incidence of hip fractures among older adults in Taiwan (where the present study was conducted) is 232.1 per 100,000 among women and 100.3 per 100,000 among men (Lin et al., 2018).

Hip fractures among older adults are associated with a one-year mortality rate of 14% to 28%, and the number of global hip fractures is projected to nearly double by 2050 compared with 2018 (Sing et al., 2023). One study reported a five-fold to eight-fold increase in the risk of mortal-

ity within the first three months following a fracture, with this risk continuing to increase over the subsequent years (Haentjens et al., 2010). Although operations are typical among older adults following hip fractures, comorbidities and mortality remain high (Sullivan et al., 2019; Yoon et al., 2021). Therefore, how to elevate care for older adults following hip fracture surgery is crucial among healthcare professionals, including mental healthcare providers. Subsequently, monitoring older adults following hip fracture surgery is important (Min et al., 2021), and health-related quality of life (HRQoL) could be a critical assessment outcome (Chiang et al., 2022).

Hip fractures among older adults (including those who have undergone surgery) often result in a significant loss of independence and a notable decline in HRQoL (Orive et al., 2015), imposing personal and societal costs (Deutschbein et al., 2023). Previous studies indicate that nonoperative groups experience significantly higher mortality rates in both the short-term (one month) and medium-term (up to



one year) compared to surgical groups, particularly among nonoperative patients with more severe baseline conditions, such as significant comorbidities and low functional status (de Jong et al., 2025; Randall et al., 2025; Kim et al., 2020; Yoon et al., 2021). A meta-analysis from the Asia Pacific region indicated that the one-year mortality rate for hip fracture patients undergoing surgery was approximately 15% to 18% in Asian countries (Harvey et al., 2024). Moreover, postoperative HRQoL data have indicated an improvement in both physical and psychological quality of life scores at three months post-surgery when compared to preoperative conditions (Căpăstraru et al., 2025).

The European Quality of Life five-dimension 3 level questionnaire (EQ-5D-3L) is one of the most widely used generic instruments for assessing HRQoL and is valued for its brevity, simplicity, and suitability for economic evaluations (Kaplan and Hays, 2022; Svedbom et al., 2023). Within the population experiencing hip fractures, HRQoL is acknowledged as the paramount outcome measure (Gjertsen et al., 2016). The EQ-5D-3L is the most prevalent instrument and is incorporated into the UK Core Outcome Set for hip fracture research (Parsons et al., 2020). Prospective research has demonstrated a substantial decline in EQ-5D utility scores, from 0.62 (out of 1) prior to the fracture to 0.16 at one-month post-fracture (Amarilla-Donoso et al., 2020), with the most affected domains being self-care, daily activities, and mobility (Amarilla-Donoso et al., 2020). Therefore, using a standardized instrument such as the EQ-5D-3L is crucial for evaluating clinical and rehabilitation outcomes.

Prior research has established the validity and extensive application of the EQ-5D-3L within the population at large through national (Stavem et al., 2018) and cross-national norming (Janssen et al., 2019), as well as validity and feasibility assessments among older adults (Buchholz et al., 2022). However, longitudinal research among postoperative patients, especially those with hip fractures, has been inconsistent. Specific studies have demonstrated significant responsiveness in the initial months (Parsons et al., 2018; Kang, 2021), while others have indicated plateaus or minimal changes over extended follow-up periods (Amphansap and Sujarekul, 2018). Outcomes are also affected by mortality during follow-up and by potential response shifts (Parsons et al., 2018; Svedbom et al., 2023).

While previous psychometric evaluations have explored the properties of the EQ-5D-3L among diverse populations (Amphansap and Sujarekul, 2018; Buchholz et al., 2022; Janssen et al., 2019; Parsons et al., 2018; Stavem et al., 2018; Svedbom et al., 2023), no previous study has rigorously examined its temporal measurement invariance specifically among older Taiwanese adults recovering from hip fracture surgery. This particular population, characterized by significant health changes and adaptive processes following major surgery, presents a unique context where temporal measurement invariance may be compromised.

More specifically, phenomena such as ‘response shift’—where an individual’s internal frame of reference for health status can alter over time due to experience or adaptation—could affect the stability of the EQ-5D-3L’s measurement properties between assessment points. Therefore, understanding the temporal stability of the EQ-5D-3L in this critically vulnerable and dynamically changing health state is crucial for accurate interpretation of HRQoL outcomes, representing a key contribution of the present study.

Because subjective HRQoL offers a holistic perspective encompassing physical function, emotional well-being, and social participation, HRQoL at baseline could be a potential predictor for subsequent activities of daily living. More specifically, when an older adult has a better HRQoL at baseline, the older adult should have good physical function (although self-perceived) to perform activities of daily living (ADLs). In this regard, good HRQoL at baseline could predict ADL performance at three-month follow-up. Therefore, the present study used HRQoL at baseline to predict ADLs at three months post-surgery.

Although the inconsistency could be due to the different time points (i.e., HRQoL improved in the short-term and then plateaued in the long-term), the assumption of measurement equivalence across time (*aka* temporal invariance) has never been evaluated. Without the temporal invariance evidence, it is hard to conclude if the findings across time are robust. That is, changes in how patients assess their HRQoL over time, especially in terms of mobility, self-care, and usual activities, might be different because the individuals may interpret the same items in different ways (Svedbom et al., 2023). Moreover, prior findings regarding responsiveness need to be built on the assumptions of supported temporal invariance. Therefore, the present study examined if the EQ-5D-3L was temporally invariant across different assessment timepoints among older adults following hip fracture surgery. More specifically, and based on the aforementioned literature, the present study aimed to: (i) test the factor structure of the EQ-5D-3L using confirmatory factor analysis (CFA) and (ii) verify temporal invariance (including configural, metric, and scalar invariance) across time.

The remainder of the present paper is structured as follows: Section 2 details the research methods, including participant and study design, instruments used, and data analysis. Section 3 presents the key results of the temporal invariance and validity analyses. Section 4 provides an in-depth discussion of these results, their implications, and the study’s limitations. Finally, Section 5 concludes the paper with a summary of the study’s findings and specific clinical recommendations.

2. Materials and Methods

2.1 Participants and Study Design

A longitudinal observational study design was utilized to assess the psychometric characteristics and validity of

the EQ-5D-3L among older individuals post-hip fracture surgery. A total of 719 patients were consecutively recruited at discharge (baseline; 0 months) from Taipei Municipal Wan Fang Hospital, a tertiary medical center in Taiwan, and subsequently re-evaluated at three months post-surgery. The recruitment period spanned from October 2017 to January 2025.

Baseline data were collected at the time of hospital discharge, defined as the point when patients were medically stable for transition from acute care to a home or non-acute care setting (e.g., home-based rehabilitation or a short-term nursing facility). Importantly, the baseline assessment of HRQoL using the EQ-5D-3L was based on patients' retrospective recall of their pre-fracture (i.e., pre-morbid) health status rather than their immediate postoperative condition. This approach was adopted to establish an individualized reference point reflecting patients' perceived health prior to the hip fracture event, thereby enabling a more meaningful evaluation of postoperative changes over time.

Participants were eligible if they had undergone surgical management for hip fracture, were proficient in Mandarin or Taiwanese, and provided written informed consent. Exclusion criteria comprised participants who had a significant cognitive impairment (Mini-Mental State Examination score <18) and/or comorbid conditions precluding reliable follow-up. Patients with postoperative complications were included, provided they were medically stable for discharge and capable of participating in follow-up assessments. However, individuals with severe complications that substantially impaired cognitive or communicative function (e.g., persistent delirium or acute stroke) were excluded to preserve the validity of self-reported outcomes. Participants with minor, clinically manageable postoperative complications were not excluded, thereby maintaining the representativeness of a real-world clinical cohort.

Participants with complete data at baseline and three-month follow-up were included in the primary analysis. To assess potential attrition bias, participants who were lost to follow-up were identified, and baseline characteristics were compared between completers and non-completers. At each time interval, trained research assistants performed structured interviews and examined medical records to gather demographic and clinical information. All methods received approval from the Institutional Review Boards at Taipei Medical University (approval no.: TMU-JIRB N201709053), and signed informed consent was acquired from each participant in accordance with the Declaration of Helsinki.

2.2 Instruments

2.2.1 European Quality of Life 5 Dimensions 3 Level Version (EQ-5D-3L)

The EQ-5D-3L (EuroQol, 1990) is a commonly utilized, general, preference-based instrument for evaluating health-related quality of life. It consists of five in-

dividual dimensions—mobility, self-care, usual activities, pain/discomfort, and anxiety/depression—each evaluated on a three-tier scale (1 = no problems, 2 = moderate problems, 3 = significant problems). This scoring produces 243 potential health states, which were transformed into a single utility score (0 = deceased to 1 = optimal health) by population-derived preference weights, and was considered a continuous scale (Lee et al., 2013). Elevated utility values signify improved self-reported health state, enabling both cross-sectional comparisons and longitudinal monitoring of health outcomes following hip fracture surgery. The EQ-5D-3L has demonstrated acceptable internal consistency (Cronbach's $\alpha \approx 0.60$) and moderate correlation with Short-Form 36, a commonly used HRQoL measure ($r \approx 0.60$), consistent with its multidimensional construct and established use among Taiwanese populations (Liu et al., 2021).

2.2.2 Barthel Index (BI)

The BI consists of 10 components that evaluate an individual's capacity to execute fundamental ADLs, including feeding, bathing, grooming, dressing, bowel control, bladder control, toilet usage, transfers (between bed and chair), mobility (on flat surfaces), and stair climbing. Each item is evaluated on an ordinal scale (often 0, 5, 10, or 15 points), with the aggregate of item scores producing a total score between 0 and 100. A superior total score indicates enhanced functional independence. The BI, initially created to assess disability in stroke rehabilitation, exhibits robust inter-rater reliability (intraclass correlation coefficients >0.90), significant test-retest stability, and commendable concurrent validity with alternative physical function assessments (Collin et al., 1988; Mahoney and Barthel, 1965). In the present study, trained evaluators conducted the BI through a structured interview at baseline and a three-month follow-up. Total score and mobility subscore (i.e., Item 6) were retained for further analysis. The BI exhibited fair-to-moderate inter-rater reliability at the item level and good reliability for the total score, indicating robust assessment of functional independence (Sainsbury et al., 2005).

2.3 Data Analysis

2.3.1 Descriptive Statistics, Internal Consistency, and Test of EQ-5D-3L Utility Score Difference Across Time

Means, standard deviations (SDs), frequencies, and percentages of participant characteristics, and instrument scores with item distributions at baseline and follow-up were analyzed. Moreover, internal consistency of the EQ-5D-3L was calculated using McDonald's ω for both time points (i.e., baseline and three-month follow-up). In addition, a paired *t*-test with Cohen's *d* was used to examine if the participants' EQ-5D-3L utility score significantly changed between the two time points.

2.3.2 Confirmatory Factor Analysis (CFA)

The single-factor structure of the EQ-5D-3L was evaluated using CFA with a diagonally weighted least squares (DWLS) estimator to address the ordinal item configuration. The single-factor structure was evaluated at both time points separately (i.e., baseline and three-month follow-up). The model fit was assessed using χ^2 , comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Acceptable criteria were CFI/TLI ≥ 0.95 , RMSEA < 0.08 , and SRMR ≤ 0.08 (Hu and Bentler, 1999).

2.3.3 Assessment of Temporal Invariance

Nested models included configural invariance, metric invariance, and scalar invariance models over time (i.e., baseline vs. three-month follow-up) and were constructed to examine temporal invariance of the EQ-5D-3L. Model comparisons (i.e., metric invariance model vs. configural invariance model, and scalar invariance model vs. metric invariance model) were performed to evaluate whether temporal invariance was supported. Temporal invariance is supported when $\Delta\text{CFI} \leq 0.01$ together with $\Delta\text{RMSEA} \leq 0.015$ and $\Delta\text{SRMR} < 0.03$ (Chen, 2007).

2.3.4 Predictive Validity

Pearson correlations were first used to assess the relationships between EQ-5D utility score and BI scores (including total score and mobility subscore) at both baseline and three-month follow-up. Based on Cohen's definition of effect size, an $|r| > 0.50$ is considered strong; between 0.3 and 0.5 is medium; and between 0.1 and 0.3 is weak (Cohen, 1988). Then, two regression models with forecasted three-month outcomes (BI total score and mobility subscore) as dependent variables were constructed utilizing the following variables: baseline EQ-5D-3L utility score and demographic factors. The assumptions of linearity, homoscedasticity, multicollinearity (variance inflation factor [VIF] < 5), and normality of residuals were met for both regression models.

2.3.5 Software Programs

All psychometric and regression analyses were conducted using *JASP version 0.19.3* (University of Amsterdam, Nieuwe Achtergracht 129B, Amsterdam, The Netherlands). Expectation-maximization for missing data and DWLS estimation were applied using *JASP's* structural equation modeling module as indicated.

3. Results

There were 157 participants lost to follow-up ($n = 157$; i.e., non-completers), and their baseline characteristics were compared with those who were not lost to follow-up (i.e., completers). **Supplementary Table 1** shows that there were no significant differences between completers

and non-completers with regard to age ($p = 0.56$) and gender distribution ($p = 0.87$). However, completers had significantly better EQ-5D-3L utility score (0.87 vs. 0.82; $p = 0.01$), BI total score (88.15 vs. 81.40; $p < 0.001$), and BI mobility subscore (13.43 vs. 11.94; $p < 0.001$) than non-completers.

A total of 719 older adults (28.7% male, 71.3% female) were included (Table 1). Their mean age was 82.1 years ($SD = 8.26$), and mean body mass index (BMI) was 22.5 ($SD = 4.01$). At baseline, participants reported a mean EQ-5D-3L utility score of 0.87 ($SD = 0.18$), which decreased to 0.45 ($SD = 0.26$) at three-month follow-up. Mean BI total score decreased from 88.15 ($SD = 18.26$) at baseline to 69.23 ($SD = 30.13$) at follow-up. The BI mobility subscore decreased from 13.43 ($SD = 3.10$) to 12.25 ($SD = 11.60$) at follow-up. Most participants lived with family (84.8%), and the primary caregiver was a family member in 77.2% of cases. Educational attainment was low: 57.0% had elementary education or below, and only 11.8% had attended college or university. The majority were married (66.8%), with 28.4% widowed. Moreover, the EQ-5D-3L utility scores were significantly different between baseline ($M = 0.87$) and three-month follow-up ($M = 0.45$; $p < 0.001$; Cohen's $d = -0.603$).

Table 2 shows the distribution of responses for each EQ-5D-3L dimension at both time points. At baseline, $\geq 70\%$ of participants reported no problems (level 1) on each item, with $< 6\%$ reporting extreme problems (level 3) across dimensions. By three months, reports of moderate (level 2) and extreme problems (level 3) increased markedly. For example, the proportion reporting some problem self-care rose from 22.1% to 41.3%, and extreme problem self-care from 5.0% to 17.7%, indicating substantial deterioration in perceived health status following hip fracture surgery. Moreover, internal consistency of the EQ-5D-3L was good at both baseline (ω [95% CI] = 0.813 [0.791–0.834]) and at three-month follow-up (ω [95% CI] = 0.851 [0.834–0.868]).

CFAs testing single-factor structure for the EQ-5D-3L at baseline and three months are summarized in Table 3. At baseline, the model demonstrated acceptable fit: $\chi^2(5) = 39.56$, $p < 0.01$; CFI = 0.988; TLI = 0.975; RMSEA = 0.098; SRMR = 0.061. At three months, the fit improved further: $\chi^2(5) = 33.03$, $p < 0.01$; CFI = 0.997; TLI = 0.993; RMSEA = 0.088; SRMR = 0.056. Both timepoint models met or approached conventional thresholds for good fit (CFI/TLI ≥ 0.95 ; RMSEA/SRMR ≤ 0.08), supporting the unidimensional factorial structure of the EQ-5D-3L across assessments.

Table 3 additionally reports nested model comparisons testing temporal invariance. Configural invariance (M1) yielded $\chi^2(10) = 72.92$, CFI = 0.994; RMSEA = 0.094; SRMR = 0.059. Metric invariance (M2) produced no change in CFI ($\Delta\text{CFI} = 0.000$), and improved RMSEA by -0.015 , although SRMR increased slightly by 0.014.

Table 1. Descriptive characteristics of participants (N = 719).

Variable	Category	N (%)	Mean (SD)
Sex	Male	206 (28.65%)	
	Female	513 (71.35%)	
Age		719 (100.00%)	82.1 (8.26)
Body mass index (kg/m ²)		719 (100.00%)	22.5 (4.01)
EQ-5D-3L			
Utility score	Baseline	719 (100.00%)	0.87 (0.18)
	3-month follow-up	719 (100.00%)	0.45 (0.26)*
Barthel Index			
Total score	Baseline	719 (100.00%)	88.15 (18.26)
	3 months follow-up	719 (100.00%)	69.23 (30.13)
Barthel Index mobility subscore	Baseline	719 (100.00%)	13.43 (3.10)
	3-month follow-up	719 (100.00%)	12.25 (11.60)
Living arrangements	With family	610 (84.84%)	
	Living alone	62 (8.62%)	
	Nursing home	47 (6.54%)	
Primary caregiver	Family	555 (77.19%)	
	Nursing home	41 (5.70%)	
	Foreign caregiver	88 (12.24%)	
	Friends/neighbor	22 (3.06%)	
	Nursing	13 (1.81%)	
Education level	Elementary/below	410 (57.02%)	
	Junior high school	92 (12.80%)	
	Senior high school	125 (17.39%)	
	College/university	85 (11.82%)	
	Postgraduate	7 (0.97%)	
Marital status	Married	480 (66.76%)	
	Widower/widow	204 (28.37%)	
	Unmarried	21 (2.92%)	
	Divorced	14 (1.95%)	

EQ-5D-3L, European Quality of Life 5 Dimensions 3 Level Version. * The paired *t*-test for EQ-5D-3L utility scores (Baseline vs. three-month follow-up) was $p < 0.001$ ($t [718] = -16.18$, Cohen's $d = -0.603$).

Scalar invariance (M3) demonstrated a small decline in CFI ($\Delta\text{CFI} = -0.007$) and marginal increases in RMSEA and decreases in SRMR ($\Delta\text{RMSEA} = 0.023$; $\Delta\text{SRMR} = -0.003$). All ΔCFI values remained within the ≤ 0.01 criterion, indicating support for full temporal invariance of factor loadings and item thresholds over the three-month period.

Zero-order correlations in Table 4 indicate strong cross-sectional associations between EQ-5D-3L utility and functional measures. At baseline, utility correlated with BI total ($r = 0.800$, $p < 0.001$) and mobility subscore ($r = 0.652$, $p < 0.001$). At follow-up, correlations attenuated but remained significant for BI total ($r = 0.269$, $p < 0.001$) and were negligible for mobility ($r = 0.009$, $p = 0.817$), reflecting differential predictive strength across domains. Utility scores showed a modest negative relationship with age (baseline $r = -0.217$, $p < 0.001$; follow-up $r = -0.062$, $p = 0.095$) but no meaningful associations with BMI.

Two multiple linear regressions were conducted to understand how baseline EQ-5D-3L was associated with BI (including its total score and mobility subscore) at three months (Table 5). Both models were significant at the model level: $F(16, 702) = 15.982$, $p < 0.001$, $R^2 = 0.267$ for BI total score; $F(16, 702) = 4.46$, $p < 0.001$, $R^2 = 0.092$ for BI mobility subscore. When examining the predictors used in the regression models, baseline EQ-5D-3L performed as a strong factor explaining BI total score ($\beta = 0.392$; $p < 0.001$) and BI mobility subscore ($\beta = 0.147$; $p < 0.001$).

4. Discussion

The present study examined the factor structure of the EQ-5D-3L through CFAs and confirmed its single-factor structure at both time points. This single-factor structure was supported for its temporal invariance over time and within pertinent groups among a population of older individuals following hip fracture surgery. Measurements us-

Table 2. Distribution of the item scores among European Quality of Life 5 Dimensions 3 Level Version (EQ-5D-3L) at baseline and at three-month follow-up (N = 719).

n%			
EQ-5D-3L	Score 1	Score 2	Score 3
Baseline			
Item 1	508 (70.7)	202 (28.0)	9 (1.3)
Item 2	524 (72.9)	159 (22.1)	36 (5.0)
Item 3	510 (71.0)	171 (23.7)	38 (5.3)
Item 4	587 (81.6)	127 (17.7)	5 (0.7)
Item 5	586 (81.5)	124 (17.2)	9 (1.3)
Three-month follow-up			
Item 1	235 (32.7)	425 (59.1)	59 (8.2)
Item 2	295 (41.0)	297 (41.3)	127 (17.7)
Item 3	309 (43.0)	310 (43.1)	100 (13.9)
Item 4	597 (83.0)	115 (16.0)	7 (1.0)
Item 5	570 (79.3)	138 (19.2)	11 (1.5)

ing the EQ-5D-3L were performed at baseline and three months after surgery. To the best of the authors' knowledge, the present study is the first psychometric investigation of the EQ-5D-3L that has been conducted among a population of older individuals with hip fractures in Taiwan.

The present study provides real-world data on the validity and temporal invariance of the EQ-5D-3L among older individuals with hip fractures, which supplements the psychometric evidence of EQ-5D-3L in the literature (Buchholz et al., 2018; Buchholz et al., 2022). Given that various disorders may influence HRQoL in distinct ways and yield diverse health profiles and utilities (Buchholz et al., 2018), it is important to comprehend how the selection of the EQ-5D-3L may impact the estimated utility of patients with hip fractures in Taiwan.

As aforementioned, the CFA results showed that the single-factor structure of the EQ-5D-3L was confirmed at both measurement times (i.e., baseline and three months after surgery). Although RMSEA has a relatively higher value (i.e., 0.098 at baseline and 0.088 at three-month follow-up), all other fit indices were satisfactory (i.e., CFI and TLI >0.95 together with SRMR <0.08). Therefore, the single-factor structure consistently represents the EQ-5D-3L construct across two time points. The slightly high value of RMSEA in the borderline range (≈ 0.09 – 0.10) can be justified by the very simple factor structure of the EQ-5D-3L. More specifically, the single-factor structure of the EQ-5D-3L only had five degrees of freedom ($df = 5$). Therefore, the CFI/TLI and SRMR provide more valuable insights in this context (Kenny et al., 2015). Additionally, assessments of temporal invariance confirmed that configural, metric, and scalar invariance were achieved ($\Delta CFI = 0.000$ for metric; $\Delta CFI = -0.007$ for scalar), adhering to the $\Delta CFI \leq 0.01$ standard for validating the equality of factor loadings and item thresholds. Moreover, the corresponding $\Delta RMSEA/\Delta SRMR$ metrics also remained within the widely ac-

cepted range. This confirms that the comparison of latent means and latent changes across time is free from measurement bias (Chen, 2007; Cheung and Rensvold, 2002).

The present study employed consecutive sampling from a single tertiary center (Taipei Municipal Wan Fang Hospital), which, while ensuring internal consistency, may constrain external validity. Patients managed in alternative care settings (e.g., regional hospitals or community facilities), as well as those with severe complications leading to prolonged acute care or early mortality, were likely underrepresented, thereby limiting generalizability across the full clinical spectrum of hip fracture populations in Taiwan. Therefore, the observed temporal invariance and validity of the EQ-5D-3L are most appropriately interpreted within the context of older adults who are medically stable for discharge and capable of participating in a structured follow-up, rather than the broader population of all post-hip fracture patients.

The findings also indicated a non-significant correlation between baseline and three-month postoperative measurements ($r = 0.072$; $p = 0.052$), suggesting that pre-fracture HRQoL, assessed via recall, has limited predictive value for short-term recovery outcomes. This likely reflects the inherently heterogeneous and dynamic nature of post-hip fracture recovery, where trajectories are shaped by interacting factors such as postoperative complications, rehabilitation intensity, evolving comorbidities, and psychosocial adaptation (Bellova et al., 2024; de Haan et al., 2024, 2025; Yoon et al., 2021). Consequently, baseline HRQoL alone appears insufficient as a prognostic indicator of subsequent health status.

Longitudinal studies have demonstrated that functional recovery after hip fracture is highly heterogeneous, with patients exhibiting very different trajectories, even when they have similar demographic and clinical characteristics (Amarilla-Donoso et al., 2020; Young et al., 2010). Some patients experience rapid recovery with dramatic improvements in function within the first few weeks, whereas others show slower recovery or even decline in function over time. The study by Young et al. (2010) identified specific outliers that exhibited recovery patterns that deviated significantly from the group average. For example, some patients who initially showed good recovery at two months subsequently experienced functional deterioration at one-year follow-up, and other patients with poor function at one month continued to improve progressively. In addition to recovery heterogeneity, variability in postoperative pain trajectories further complicates patient outcomes. Other studies have identified variations in pain experience that not only impact patient comfort but also affect early mobilization, participation in rehabilitation, and ultimately mid-term functional outcomes, contributing to inconsistencies between baseline status and outcomes three months later (Goto et al., 2020; Morrison et al., 2003).

Table 3. Temporal invariance findings for the European Quality of Life 5 Dimensions 3 Level Version (EQ-5D-3L) (N = 719).

Model											
Model number	χ^2	df	CFI	RMSEA	SRMR	Comparison	$\Delta\chi^2$	Δdf	ΔCFI	$\Delta RMSEA$	$\Delta SRMR$
EQ-5D-3L											
M1	72.924*	10	0.994	0.094	0.059	-	-	-	-	-	-
M2	76.419*	14	0.994	0.079	0.073	M2-M1	3.495	4	0.000	-0.015	0.014
M3	153.536*	18	0.987	0.102	0.070	M3-M2	77.117*	4	-0.007	0.023	-0.003

* p -value < 0.01; χ^2 , chi-square; df, degrees of freedom; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; $\Delta\chi^2$, delta chi-square; Δdf , delta degree of freedom; ΔCFI , delta comparative fit index; $\Delta RMSEA$, delta root mean square error of approximation; $\Delta SRMR$, delta standardized root mean square residual; M1, configural invariance; M2, metric invariance; M3, scalar invariance.

Table 4. Correlation of utility score the European Quality of Life 5 Dimensions 3 Level Version (EQ-5D-3L) with relevant criteria (N = 719).

r (p -value)/95% confidence interval of r		
	Baseline EQ-5D-3L	Follow-up three months EQ-5D-3L
Age	-0.217 (<0.001)/(-0.285, -0.146)	-0.062 (0.10)/(-0.135, 0.011)
Body mass index	0.023 (0.54)/(-0.050, 0.096)	0.003 (0.94)/(-0.070, 0.076)
Baseline EQ-5D-3L	-	0.072 (0.052)/(-7.304 $\times 10^{-4}$, 0.145)
Follow-up three months EQ-5D-3L	0.072 (0.052)/(-7.304 $\times 10^{-4}$, 0.145)	-
Baseline Barthel Index (total score)	0.800 (<0.001)/(0.772, 0.825)	0.133 (<0.001)/(0.060, 0.204)
Follow-up three months Barthel Index (total score)	0.462 (<0.001)/(0.403, 0.518)	0.269 (<0.001)/(0.200, 0.336)
Baseline Barthel Index (mobility subscore)	0.652 (<0.001)/(0.608, 0.692)	0.109 (0.003)/(0.037, 0.181)
Follow-up three months Barthel Index (mobility subscore)	0.213 (<0.001)/(0.142, 0.281)	0.009 (0.82)/(-0.065, 0.082)

Within these constraints, the confirmation of factorial validity and longitudinal invariance (i.e., configural, metric, and scalar) from baseline to three months supports the EQ-5D-3L as a practical instrument for monitoring HRQoL during early post-hip-fracture rehabilitation. Observed changes in utility scores over this interval can be interpreted as likely reflecting changes in perceived health, rather than measurement instability. The weak, non-significant correlation between baseline and follow-up utility scores underscores substantial inter-individual variability in early recovery. The establishment of full scalar invariance indicates that observed changes in scores over time can be interpreted as true changes in the underlying construct rather than artifacts of measurement. Had scalar invariance not been supported, any observed longitudinal differences could reflect shifts in item intercepts rather than genuine clinical change, thereby compromising the validity of change score interpretation.

The lack of a significant correlation between the BI mobility subscore and EQ-5D-3L utility score at three months likely reflected both conceptual divergence and temporal misalignment between the measures. The BI assesses objective functional independence, whereas the EQ-5D-3L assesses multidimensional, preference-based HRQoL. Consequently, early gains in mobility may not translate into improved perceived health, particularly in the presence of persistent pain, fear of falling, and psychological distress (Lee et al., 2013). This pattern suggests a tran-

sient decoupling between physical recovery and subjective health perception during the subacute phase, further compounded by differential sensitivity and potential ceiling effects of the EQ-5D-3L (Garratt and Stavem, 2022).

Limitations

The present study had several limitations. First, although the sample of 719 older adults provided adequate power for CFA and longitudinal invariance testing, recruitment from a single national context limited external validity. Replications in diverse languages, different health systems, and different care settings are warranted, ideally by comparing country-specific value sets (tariffs) to evaluate score comparability. Second, the three-month follow-up may not adequately reflect long-term stability or sensitivity to change during post-hip-fracture recovery. A prolonged follow-up period (e.g., 6–12 months) is necessary to evaluate temporal invariance and responsiveness in the subsequent phases of rehabilitation.

Third, the EQ-5D-3L relies on self-report and a three-level ordinal response structure, rendering it less sensitive to transient health fluctuations, cognitive variability, and emotional states. Moreover, its limited scaling is prone to ceiling effects (baseline mean utility = 0.87) that may mask subtle yet clinically meaningful changes. The dependence on self-reported measures further introduces methodological constraints among individuals in this cohort. Even without overt cognitive impairment, mild cognitive decline

Table 5. Linear regression analysis of baseline European Quality of Life 5 Dimensions 3 Level Version (EQ-5D-3L) utility scores predicting total Barthel Index score and Barthel Index mobility subscore at three-month follow-up.

Predictor	Outcome: total Barthel Index score at three-month follow-up ($R^2 = 0.267$; adj. $R^2 = 0.250$; F [df] = 15.982 [16, 702]; $p < 0.001$; $\eta p^2 = 0.267$)				
	B	SE	β	t	p
Baseline EQ-5D-3L	64.334	5.740	0.392	11.208	<0.001
Age	−0.503	0.130	−0.138	−3.861	<0.001
Sex (Ref: male)	0.184	2.288		0.080	0.94
Educational level (Ref: elementary/below)					
Junior high school	−0.870	3.096		−0.281	0.78
Senior high school	−4.400	2.770		−1.588	0.11
College/university	6.467	3.315		1.951	0.051
Postgraduate	5.011	10.042		0.499	0.62
Marital status (Ref: married)					
Widower/widow	−3.678	2.356		−1.561	0.12
Unmarried	−5.232	6.166		−0.848	0.40
Divorced	−4.297	7.475		−0.575	0.57
Living arrangements (Ref: with family)					
Living alone	−2.221	4.324		−0.514	0.61
Nursing home	−27.060	10.135		−2.670	0.008
Primary caregiver (Ref: family)					
Nursing home	18.896	10.617		1.780	0.08
Foreign caregiver	−7.214	3.148		−2.292	0.02
Friends/neighbor	0.667	7.054		0.095	0.93
Nursing	4.388	8.233		0.534	0.59
Predictor	Outcome: Barthel Index mobility subscore at three-month follow-up ($R^2 = 0.092$; adj. $R^2 = 0.072$; F [df] = 4.463 [16, 702]; $p < 0.001$; $\eta p^2 = 0.092$)				
	B	SE	β	t	p
Baseline EQ-5D-3L	9.273	2.458	0.147	3.772	<0.001
Age	−0.153	0.056	−0.109	−2.737	0.006
Sex (Ref: male)	0.192	0.980		0.196	0.85
Educational level (Ref: elementary/below)					
Junior high school	0.211	1.326		0.159	0.87
Senior high school	0.994	1.187		0.838	0.40
College/university	2.282	1.420		1.607	0.11
Postgraduate	−0.675	4.301		−0.157	0.88
Marital status (Ref: married)					
Widower/widow	−1.850	1.009		−1.834	0.07
Unmarried	−0.612	2.641		−0.232	0.82
Divorced	2.502	3.201		0.782	0.44
Living arrangements (Ref: family)					
Living alone	2.837	1.852		1.532	0.13
Nursing home	−6.278	4.341		−1.446	0.15
Primary caregiver (Ref: family)					
Nursing home	4.040	4.547		0.889	0.38
Foreign caregiver	−1.312	1.348		−0.973	0.33
Friends/neighbour	2.265	3.021		0.750	0.45
Nursing	−2.254	3.522		−0.643	0.52

B, unstandardized coefficient; SE, standard error; β , standardized coefficient; adj. R^2 , adjusted R^2 ; F (df) indicates F-value and degrees of freedom for the model with the main predictor; t , t -statistics; ηp^2 , partial eta-squared.

may compromise comprehension, recall, and response consistency, increasing susceptibility to measurement bias and potentially attenuating the accuracy of HRQoL estimates and their associations with functional outcomes. Fourth, the present study has the potential for survival bias. The exclusion of individuals who developed severe conditions or experienced mortality prior to the follow-up assessment restricts the generalizability of the findings because the study cohort may not fully represent the complete spectrum of outcomes within the broader population of older Taiwanese adults recovering from hip fracture surgery.

Fifth, attrition bias may have been present in the study because differences were observed between participants who completed the follow-up and those lost to follow-up at baseline. More specifically, variations in key characteristics such as baseline HRQoL and functional status suggest that participants who remained in the study may not be fully representative of the initial sample. This differential loss to follow-up could have potentially influenced the estimated outcomes and led to biased interpretations. Therefore, the findings of the study should be interpreted with caution. Finally, a further limitation arises from the use of retrospective recall to assess baseline HRQoL because participants completed the EQ-5D-3L at discharge by reflecting on their pre-fracture health status. Such an approach is susceptible to recall bias, whereby current postoperative experiences may influence or distort recollection of prior health, potentially leading to systematic overestimation of baseline utility values. Such bias may have inflated the apparent magnitude of decline at follow-up and affected the accuracy of longitudinal comparisons.

Future research should extend follow-up to 6–12 months to evaluate responsiveness, establish minimal important change thresholds, and assess predictive validity for clinically relevant outcomes, including readmission and mortality. External validity could be strengthened through multi-site and cross-national studies examining measurement invariance across languages, cultures, and healthcare systems, alongside comparisons of country-specific value sets. In parallel, adequately powered, subgroup-specific analyses are needed to evaluate temporal invariance across clinically relevant strata, including fracture type, sex, and cognitive status. Such targeted approaches would refine the interpretability of the EQ-5D-3L and enhance its application in personalized rehabilitation and outcome assessment among older adults following hip fracture.

The observed sensitivity of the EQ-5D-3L in the mobility, self-care, and usual activities dimensions, reflected by an increased proportion of patients reporting moderate-to-severe problems in these dimensions at three months, underscores the need for intensified, individualized postoperative functional ability program. This should include routine and standardized program tailored to the needs of older adults to mitigate persistent symptoms and support functional recovery. Moreover, the lower HRQoL scores

among patients residing in nursing homes highlight the necessity of optimizing rehabilitation care in these settings. Targeted strategies should encompass increased frequency and intensity of supervised rehabilitation, enhancement of mobility-supportive environments, and structured training of nursing staff in postoperative hip fracture management to improve continuity and quality of care.

In routine practice, brief, repeated EQ-5D-3L assessments may facilitate early identification of patients at risk of delayed or suboptimal recovery, enabling timely adjustment of intervention strategies, including modulation of therapy intensity, optimization of pain and mood management, and reinforcement of environmental and social support. This finding further highlights the limited prognostic value of baseline measures alone and supports a shift toward longitudinal, adaptive care models. Continuous monitoring of HRQoL and functional status is therefore essential to capture evolving recovery trajectories and inform individualized clinical decision-making. From a research standpoint, these results underscore the need for predictive frameworks that integrate dynamic, multidimensional factors rather than relying solely on static baseline indicators. In interdisciplinary teams, utilizing EQ-5D-3L's utility score can help in discussing progress, promoting collaborative decision-making, and yielding standardized outcomes for assessing program efficacy. Moreover, implementing country-specific value sets would likely improve sensitivity and interpretability.

5. Conclusions

The findings of the present study demonstrate that the EQ-5D-3L possesses a stable and temporally valid single-factor structure for assessing HRQoL among older adults following hip fracture surgery. CFA conducted at two time points showed a satisfactory model fit. The results of temporal invariance testing confirmed full configural, metric, and scalar invariance, indicating consistency of the measurement structure over time. The correlation of EQ-5D-3L's utility score between baseline and three-month follow-up was low and non-significant ($r = 0.072$; $p = 0.052$). This result likely indicates clinical variation and individual subjective perceptions of health recovery, rather than deficiencies in the instrument. The results support the validity of the EQ-5D-3L as an instrument for longitudinal studies and the evaluation of rehabilitation outcomes among older adult patients with post-hip fracture surgery, illustrating its robust ability to detect authentic changes in health status over time. These findings hold implications for mental health professionals and interdisciplinary teams by identifying patients at risk of poor or delayed recovery and yielding standardized outcomes for assessing program efficacy.

Availability of Data and Materials

Due to patient privacy and institutional review board (IRB) restrictions, de-identified individual-level data can-

not be posted publicly. Qualified researchers may request access to the de-identified dataset and analysis code for reproducible research purposes by submitting a brief proposal and a signed data use agreement. Requests should be directed to the corresponding author (Chung-Ying Lin, cylin36933@gmail.com) and will be reviewed by the study team and ethics committee. A data dictionary, variable list, and statistical analysis scripts will be shared upon approval.

Author Contributions

Conception & design: TL, Y-PC, CHH, C-YL; Data acquisition: Y-PC, L-CW, Y-JK; Data analysis/interpretation: TL, K-YY, Y-JC, MDG, C-YL; Drafting of the manuscript: TL, C-YL; Critical revision for important intellectual content: TL, Y-PC, K-YY, Y-JC, CHH, MDG, L-CW, Y-JK, C-YL; All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The Declaration of Helsinki was strictly followed and adhered to during the entire study procedure, and the study was approved by the Institutional Review Boards at Taipei Medical University (approval no.: TMU-JIRB N201709053) on October 23, 2017. Respondents gave written consent for review and signature before starting interviews.

Acknowledgments

We would like to thank the older adults who participated in the study and the multidisciplinary teams in the orthopedic, geriatrics, and rehabilitation units for their support in recruitment and data collection. We are also grateful to the research assistants for fieldwork and data management, and to our clinical collaborators for valuable feedback on study procedures. We also appreciate the constructive comments from colleagues who reviewed earlier versions of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest. Chung-Ying Lin is serving as one of the Editorial Board members of this journal. We declare that Chung-Ying Lin had no involvement in the peer review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Jose Pedro Espada.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used ChatGpt-3.5 and QuillBot for spelling and grammar checks. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/BP49899>.

References

- Amarilla-Donoso FJ, López-Espuela F, Roncero-Martín R, Leal-Hernandez O, Puerto-Parejo LM, Aliaga-Vera I, et al. Quality of life in elderly people after a hip fracture: a prospective study. *Health and Quality of Life Outcomes*. 2020; 18: 71. <https://doi.org/10.1186/s12955-020-01314-2>
- Amphansap T, Sujarekul P. Quality of life and factors that affect osteoporotic hip fracture patients in Thailand. *Osteoporosis and Sarcopenia*. 2018; 4: 140–144. <https://doi.org/10.1016/j.afos.2018.11.082>
- Bellova P, Linne M, Postler AE, Günther KP, Stiehler M, Goronzy J. Girdlestone resection arthroplasty for femoral neck fractures has poorer outcomes than hemiarthroplasty in frail patients with increased risk for arthroplasty-related complications: a retrospective case study of 21 patients. *Acta Orthopaedica*. 2024; 95: 61–66. <https://doi.org/10.2340/17453674.2024.34901>
- Buchholz I, Janssen MF, Kohlmann T, Feng YS. A Systematic Review of Studies Comparing the Measurement Properties of the Three-Level and Five-Level Versions of the EQ-5D. *PharmacoEconomics*. 2018; 36: 645–661. <https://doi.org/10.1007/s40273-018-0642-5>
- Buchholz I, Marten O, Janssen MF. Feasibility and validity of the EQ-5D-3L in the elderly Europeans: a secondary data analysis using SHARE(d) data. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*. 2022; 31: 3267–3282. <https://doi.org/10.1007/s11136-022-03158-3>
- Căpăstraru BF, Levai CM, Mederle OA, Velimirovici MD, Folescu R, Bogdan H, et al. Comparative Assessment of Quality of Life in Hip Fracture Patients Before and After Surgery: A Prospective Longitudinal Observational Study. *Healthcare (Basel, Switzerland)*. 2025; 13: 1126. <https://doi.org/10.3390/healthcare13101126>
- Chen FF. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: a Multidisciplinary Journal*. 2007; 14: 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: a Multidisciplinary Journal*. 2002; 9: 233–255. https://doi.org/10.1207/s15328007sem0902_5

- Chiang MH, Huang YY, Kuo YJ, Huang SW, Jang YC, Chu FL, et al. Prognostic Factors for Mortality, Activity of Daily Living, and Quality of Life in Taiwanese Older Patients within 1 Year Following Hip Fracture Surgery. *Journal of Personalized Medicine*. 2022; 12: 102. <https://doi.org/10.3390/jpm12010102>
- Cohen J. Statistical power analysis for the behavioral sciences. 2nd edn. Lawrence Erlbaum Associates: Hillsdale, NJ. 1988. <https://doi.org/10.4324/9780203771587>
- Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. *International Disability Studies*. 1988; 10: 61–63. <https://doi.org/10.3109/09638288809164103>
- de Haan E, de Ronde A, de Jong L, van Rijckevorsel VAJIM, DHFR, Roukema GR. Preoperative identification of low-risk patients for major complications after acute hip fracture surgery in geriatric patients: an observational cohort study. *BMC Musculoskeletal Disorders*. 2025; 26: 590. <https://doi.org/10.1186/s12891-025-08720-x>
- de Haan E, Roukema GR, van Rijckevorsel VAJIM, Kuijper TM, de Jong L, Dutch Hip Fracture RegistryCollaboration. Risk Factors for 30-Days Mortality After Proximal Femoral Fracture Surgery, a Cohort Study. *Clinical Interventions in Aging*. 2024; 19: 539–549. <https://doi.org/10.2147/CIA.S441280>
- de Jong L, de Haan E, van Rijckevorsel VAJIM, Kuijper TM, Roukema GR. Multidimensional Approach for Predicting 30-Day Mortality in Patients with a Hip Fracture: Development and External Validation of the Rotterdam Hip Fracture Mortality Prediction-30 Days (RHMP-30). *The Journal of Bone and Joint Surgery. American Volume*. 2025; 107: 459–468. <https://doi.org/10.2106/JBJS.23.01397>
- Deutschbein J, Lindner T, Möckel M, Pigorsch M, Gilles G, Stöckle U, et al. Health-related quality of life and associated factors after hip fracture. Results from a six-month prospective cohort study. *PeerJ*. 2023; 11: e14671. <https://doi.org/10.7717/peerj.14671>
- EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy (Amsterdam, Netherlands)*. 1990; 16: 199–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
- Feng JN, Zhang CG, Li BH, Zhan SY, Wang SF, Song CL. Global burden of hip fracture: The Global Burden of Disease Study. *Osteoporosis International*. 2023; 35: 41–52. <https://doi.org/10.1007/s00198-023-06907-3>
- Garratt A, Stavem K. Comparison of EQ-5D-3L and 5L versions following operative fixation of closed ankle fractures. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*. 2022; 31: 2071–2082. <https://doi.org/10.1007/s11136-022-03105-2>
- Gjertsen JE, Baste V, Fevang JM, Furnes O, Engesaeter LB. Quality of life following hip fractures: results from the Norwegian hip fracture register. *BMC Musculoskeletal Disorders*. 2016; 17: 265. <https://doi.org/10.1186/s12891-016-1111-y>
- Goto K, Kataoka H, Honda A, Yamashita J, Morita K, Hirase T, et al. Factors Affecting Persistent Postoperative Pain in Patients with Hip Fractures. *Pain Research & Management*. 2020; 2020: 8814290. <https://doi.org/10.1155/2020/8814290>
- Haentjens P, Magaziner J, Colón-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B, et al. Meta-analysis: excess mortality after hip fracture among older women and men. *Annals of Internal Medicine*. 2010; 152: 380–390. <https://doi.org/10.7326/0003-4819-152-6-201003160-00008>
- Harris E, Clement N, MacLulich A, Farrow L. The impact of an ageing population on future increases in hip fracture burden. *The Bone & Joint Journal*. 2024; 106-B: 62–68. <https://doi.org/10.1302/0301-620X.106B1.BJJ-2023-0740.R1>
- Harvey LA, Payne NL, Tan A, Zhang J, Lai YC, Taylor ME, et al. Variation in mortality following hip fracture across the Asia Pacific region: Systematic review and proportional meta-analysis. *Archives of Gerontology and Geriatrics*. 2024; 126: 105519. <https://doi.org/10.1016/j.archger.2024.105519>
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*. 1999; 6: 1–55. <https://doi.org/10.1080/10705519909540118>
- Janssen MF, Szende A, Cabases J, Ramos-Goñi JM, Vilagut G, König HH. Population norms for the EQ-5D-3L: a cross-country analysis of population surveys for 20 countries. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care*. 2019; 20: 205–216. <https://doi.org/10.1007/s10198-018-0955-5>
- Kang S. Assessing responsiveness of the EQ-5D-3L, the Oxford Hip Score, and the Oxford Knee Score in the NHS patient-reported outcome measures. *Journal of Orthopaedic Surgery and Research*. 2021; 16: 18. <https://doi.org/10.1186/s13018-020-02126-2>
- Kaplan RM, Hays RD. Health-Related Quality of Life Measurement in Public Health. *Annual Review of Public Health*. 2022; 43: 355–373. <https://doi.org/10.1146/annurev-publhealth-052120-012811>
- Kenny DA, Kaniskan B, McCoach DB. The Performance of RMSEA in Models With Small Degrees of Freedom. *Sociological Methods & Research*. 2015; 44: 486–507. <https://doi.org/10.1177/0049124114543236>
- Kim SJ, Park HS, Lee DW. Outcome of nonoperative treatment for hip fractures in elderly patients: A systematic review of recent literature. *Journal of Orthopaedic Surgery (Hong Kong)*. 2020; 28: 2309499020936848. <https://doi.org/10.1177/2309499020936848>
- Lin KB, Yang NP, Lee YH, Chan CL, Wu CH, Chen HC, et al. The incidence and factors of hip fractures and subsequent morbidity in Taiwan: An 11-year population-based cohort study. *PloS One*. 2018; 13: e0192388. <https://doi.org/10.1371/journal.pone.0192388>
- Liu TH, Ho AD, Hsu YT, Hsu CC. Validation of the EQ-5D in Taiwan using item response theory. *BMC Public Health*. 2021; 21: 2305. <https://doi.org/10.1186/s12889-021-12334-y>
- Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Maryland State Medical Journal*. 1965; 14: 61–65.

- Min K, Beom J, Kim BR, Lee SY, Lee GJ, Lee JH, et al. Clinical Practice Guideline for Postoperative Rehabilitation in Older Patients With Hip Fractures. *Annals of Rehabilitation Medicine*. 2021; 45: 225–259. <https://doi.org/10.5535/arm.21110>
- Morrison SR, Magaziner J, McLaughlin MA, Orosz G, Silberzweig SB, Koval KJ, et al. The impact of post-operative pain on outcomes following hip fracture. *Pain*. 2003; 103: 303–311. [https://doi.org/10.1016/S0304-3959\(02\)00458-X](https://doi.org/10.1016/S0304-3959(02)00458-X)
- Orive M, Aguirre U, García-Gutiérrez S, Las Hayas C, Bilbao A, González N, et al. Changes in health-related quality of life and activities of daily living after hip fracture because of a fall in elderly patients: a prospective cohort study. *International Journal of Clinical Practice*. 2015; 69: 491–500. <https://doi.org/10.1111/ijcp.12527>
- Parsons N, Griffin XL, Achten J, Chesser TJ, Lamb SE, Costa ML. Modelling and estimation of health-related quality of life after hip fracture: A re-analysis of data from a prospective cohort study. *Bone & Joint Research*. 2018; 7: 1–5. <https://doi.org/10.1302/2046-3758.71.BJR-2017-0199>
- Parsons NR, Costa ML, Achten J, Griffin XL. Baseline quality of life in people with hip fracture: results from the multicentre WHiTE cohort study. *Bone & Joint Research*. 2020; 9: 468–476. <https://doi.org/10.1302/2046-3758.98.BJR-2019-0242.R1>
- Randall ZD, Barber HF, Buesser KE, Obey MR, Wilson JL, McAndrew CM, et al. The impact of biological age and age acceleration on 1-year mortality rates in elderly hip fracture patients: a prospective cohort study. *Archives of Osteoporosis*. 2025; 20: 84. <https://doi.org/10.1007/s11657-025-01572-x>
- Sing CW, Lin TC, Bartholomew S, Bell JS, Bennett C, Beyene K, et al. Global Epidemiology of Hip Fractures: Secular Trends in Incidence Rate, Post-Fracture Treatment, and All-Cause Mortality. *Journal of Bone and Mineral Research: the Official Journal of the American Society for Bone and Mineral Research*. 2023; 38: 1064–1075. <https://doi.org/10.1002/jbmr.4821>
- Stavem K, Augestad LA, Kristiansen IS, Rand K. General population norms for the EQ-5D-3 L in Norway: comparison of postal and web surveys. *Health and Quality of Life Outcomes*. 2018; 16: 204. <https://doi.org/10.1186/s12955-018-1029-1>
- Sullivan NM, Blake LE, George M, Mears SC. Palliative Care in the Hip Fracture Patient. *Geriatric Orthopaedic Surgery & Rehabilitation*. 2019; 10: 2151459319849801. <https://doi.org/10.1177/2151459319849801>
- Svedbom A, Borgstöm F, Hernlund E, Alekna V, Bianchi ML, Clark P, et al. An experience- and preference-based EQ-5D-3L value set derived using 18 months of longitudinal data in patients who sustained a fracture: results from the ICUROS. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*. 2023; 32: 1199–1208. <https://doi.org/10.1007/s11136-022-03303-y>
- Yoon SH, Kim BR, Lee SY, Beom J, Choi JH, Lim JY. Influence of comorbidities on functional outcomes in patients with surgically treated fragility hip fractures: a retrospective cohort study. *BMC Geriatrics*. 2021; 21: 283. <https://doi.org/10.1186/s12877-021-02227-5>
- Young Y, Xiong K, Pruzek RM, Brant LJ. Examining heterogeneity of functional recovery among older adults with hip fractures. *Journal of the American Medical Directors Association*. 2010; 11: 132–139. <https://doi.org/10.1016/j.jamda.2009.11.007>
- Lee HY, Hung MC, Hu FC, Chang YY, Hsieh CL, Wang JD. Estimating quality weights for EQ-5D (EuroQol-5 dimensions) health states with the time trade-off method in Taiwan. *Journal of the Formosan Medical Association*. 2013; 112: 699–706. <https://doi.org/10.1016/j.jfma.2012.12.015>
- Sainsbury A, Seebass G, Bansal A, Young JB. Reliability of the Barthel Index when used with older people. *Age and Ageing*. 2005; 34: 228–232. <https://doi.org/10.1093/ageing/afi063>