Systematic Review

JHSMR Journal of Health Science and Medical Research

Patient-Reported Outcome Measures and Health-Related Quality of Life Instruments in School-Age Children with Upper Limb Fractures: A Scoping Review

Mohamad Zakwan Helmi Abu Bakar, M.D.[1,2], Mai Nurul Ashikin Taib, M.B.B.S., M.Med.(Pediatrics)[3,4], Laila Maisarah A.Rahman, M.B.B.S., M.Orth.[1,2], Mohd Ariff Sharifudin, M.B.B.S., M.Med.(Orthopedics)[1,2,5]

[1]Department of Orthopaedics and Rehabilitation, Faculty of Medicine, Universiti Sultan Zainal Abidin, Jalan Sultan Mahmud, 20400 Kuala Terengganu, Terengganu, Malaysia.

[2]Department of Orthopaedics, Hospital Sultan Zainal Abidin, 21300 Kuala Nerus, Terengganu, Malaysia.

[3]Department of Paediatrics, Faculty of Medicine, Universiti Sultan Zainal Abidin, Jalan Sultan Mahmud, 20400 Kuala Terengganu, Terengganu, Malaysia.

[4]Department of Paediatrics, Hospital Sultan Zainal Abidin, 21300 Kuala Nerus, Terengganu, Malaysia.

[5]Medical and Healthcare Research Cluster, Office of the Deputy Vice-Chancellor (Research and Innovation), Universiti Sultan Zainal Abidin, 21300 Kuala Nerus, Terengganu, Malaysia.

Received 7 April 2026 ● Accepted 6 May 2026 ● Published online 6 November 2026

Abstract:

Objective: Upper limb fractures are common in children, yet the selection of patient-reported outcome measures (PROMs) for research remains inconsistent. This scoping review maps the use of PROMs and health-related quality of life (HRQL) instruments in school-age children with upper limb fractures, evaluates their psychometric properties, and assesses their coverage of pediatric-relevant outcome domains.

Material and Methods: This review followed the Joanna Briggs Institute methodology and is reported in accordance with the PRISMA extension for Scoping Reviews (PRISMA-ScR). Six databases (MEDLINE, EMBASE, Scopus, CENTRAL, CINAHL, and PsycINFO) were searched from 2000 to 2025. Two independent reviewers screened studies, extracted data, and assessed psychometric evidence by using COSMIN guidelines. Instruments were mapped onto the CORE-Kids multidimensional quality-of-life framework.

Contact: Mai Nurul Ashikin Taib, M.B.B.S., M.Med.(Pediatrics)
Department of Paediatrics, Faculty of Medicine, Universiti Sultan Zainal Abidin, Jalan Sultan Mahmud, 20400 Kuala Terengganu, Terengganu, Malaysia.
E-mail: maiashikintaib@unisza.edu.my

J Health Sci Med Res
doi: 10.31584/jhsmr.20261498
www.jhsmr.org

© 2026 JHSMR. Hosted by Prince of Songkla University.
This is an open access article under the CC BY-NC-ND license (http://www.jhsmr.org/index.php/jhsmr/about/editorialPolicies#openAccessPolicy).

Results: From 2,847 records, 68 studies were included, identifying 18 distinct PROMs. The QuickDASH was most frequently used (41% of studies), followed by PROMIS Upper Extremity (18%), PODCI (15%), and PedsQL (13%). Psychometric evidence was strongest for PROOF-UE, but critical gaps existed: 90% of instruments lacked content validity studies, and 80% lacked measurement error data. Domain mapping revealed major gaps in emotional well-being (18% coverage) and social functioning (14%), with critical gaps in school/learning (9%) and overall life satisfaction (5%).

Conclusion: Despite the availability of multiple PROMs, significant gaps in psychometric and content validity persist. Commonly used instruments inadequately assess emotional, social, and school-related outcomes. A multi-measure approach that combines physical function, comprehensive quality-of-life, and psychological outcome tools is recommended for clinical and research practice.

Keywords: child, patient-reported outcome, pediatric fractures, psychometrics, quality of life, upper extremity

Introduction

Upper limb fractures account for more than half of all childhood fractures, with the distal radius being the most commonly affected site[1-3]. While satisfactory healing typically occurs due to the remodeling potential of growing bone, long-term functional outcomes vary considerably based on fracture characteristics, treatment, and patient factors[4,5]. The impact of these injuries extends beyond physical healing, often affecting functional capacity, psychological well-being, school attendance, and social participation during critical developmental periods. Traditional outcome assessments in pediatric fracture research rely heavily on clinical and radiographic measures, which fail to adequately capture the experiences of children and families regarding recovery quality, daily functional limitations, emotional adjustment, and overall quality of life[6-8]. Focusing exclusively on radiographic alignment or clinician-assessed range of motion may not accurately reflect a child's perceived arm function during daily activities, school participation, and play[9,10].

Patient-reported outcome measures (PROMs) and health-related quality of life (HRQL) instruments have transformed outcome assessment in pediatric orthopedics by capturing the subjective experiences of symptom burden, functional capacity, and well-being; outcomes critical for evaluating treatment efficacy[5,11]. Despite this recognition, the pediatric upper limb fracture literature reveals substantial heterogeneity in outcome reporting, with inconsistent application of instruments across studies, limited comparison of psychometric properties, and insufficient evidence regarding which tools best capture outcomes meaningful to children and families[6,12-14]. Fewer than 5% of childhood fracture trials include validated PROMs, and even the most frequently used instruments appear in only 40%–60% of studies[12,14-16]. This heterogeneity is compounded by critical evidence gaps: limited psychometric validation, specifically in pediatric fracture populations[6,16]; an absence of content validity studies[2,10,13,17]; predominant focus on physical function with inadequate coverage of emotional, social, and school-related domains[18,19,20]; and insufficient data on measurement efficiency and responsiveness[18,21]. These limitations prevent meaningful cross-study comparison, hinder evidence synthesis, and impede the development of standardized, patient-centered care protocols.

This scoping review aimed to systematically map the literature in order to address three overarching questions: (1) What PROMs and HRQL instruments have been used to assess outcomes in school-age children (6-17 years) with upper limb fractures, and with what frequency? (2) What are the psychometric properties and methodological quality of evidence supporting these instruments in this population? (3) How do these instruments map against the CORE-Kids multidimensional quality of life framework to

assess coverage of pediatric-relevant outcome domains? Following the established scoping review methodology, this review will clarify the existing tools and their limitations to inform future research and advocate for standardized, child-centered outcome assessment in pediatric upper-limb fracture management.

Material and Methods

Design and methodological framework

This scoping review followed the Joanna Briggs Institute (JBI) methodology[2]. The JBI framework was selected for its suitability in mapping evidence landscapes, identifying knowledge gaps, synthesizing heterogeneous study designs, and informing future systematic reviews and evidence-based practice development. This approach was particularly appropriate given the heterogeneity of PROMs used in pediatric upper limb fracture studies, the variability in instrument development and validation, and the need for a mapping approach to inform future standardization. The review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist[16]. The protocol was not registered but adhered to the principles of transparency and reproducibility in the scoping review methodology. No formal registration of the review was undertaken, as this review aimed to map the existing literature rather than address a specific hypothesis.

Population, concept, and context (PCC) framework

The PCC framework guided the review scope as follows:

Population (P): School-age children and adolescents aged 6–17 years with upper limb fractures (clavicle, humerus, forearm bones, carpal bones, metacarpals, and phalanges). Studies including adolescents up to age 17 years were included to maintain developmental continuity, as cognitive and emotional development continues through late adolescence. Mixed-age studies were included if >80% of participants were within this range or subgroup data were reported separately. Studies were excluded if pediatric participants represented <25% of the sample without separate data, the mean age fell outside the 6–17 years range, or the sample included prenatal or early childhood populations. Studies focusing on non-fracture upper limb conditions (e.g., cerebral palsy, arthritis, soft-tissue injuries) were excluded. Studies of lower limb, spinal, or trunk fractures were also excluded unless they examined upper limb fractures as the primary focus.

Concept (C): PROMs and HRQL instruments capturing children's or parents' subjective assessment of outcomes. Included measures encompassed condition-specific instruments (e.g., Disabilities of the Arm, Shoulder and Hand (DASH/QuickDASH), Patient-Rated Wrist Evaluation (PRWE)), pediatric-specific generic HRQL instruments (e.g., Pediatric Quality of Life Inventory (PedsQL), PROMIS Pediatric measures), generic HRQL tools adapted for pediatrics (e.g., EuroQol Five-Dimension Youth (EQ-5D-Y)), and functional measures with subjective components (e.g., Canadian Occupational Performance Measure [COPM])[9,22]. Excluded measures were purely objective clinician-assessed tools (e.g., dynamometry, goniometry), radiographic parameters, and performance-based assessments without subjective components.

Context (C): Clinical studies (randomized controlled trials (RCTs), cohort studies, case series, cross-sectional, and qualitative studies) published in English-language peer-reviewed journals, conducted in any geographical or healthcare setting. Editorials, commentaries, expert opinions without primary data, grey literature, conference abstracts, and dissertations were excluded.

Study selection criteria

Studies were included if published in English between January 1, 2000, and December 31, 2025, involved participants aged 6–17 years (≥80% of the sample) with

documented upper limb fractures, reported at least one PROM or HRQL instrument, employed a clinical study design with primary data collection, and were available as full-text peer-reviewed articles. Exclusion criteria comprised adult-only populations (mean age >17 years) without pediatric subgroup data, chronic upper limb conditions without acute fracture (e.g., cerebral palsy, arthritis, and neuromuscular disease), reported only clinician-measured outcomes, non-English publications, grey literature, case reports (<3 participants), and studies focused exclusively on instrument development without application in fracture populations.

Information sources and search strategy

Electronic databases searched

Systematic searches were conducted across six databases: MEDLINE (PubMed), EMBASE, Scopus, Cochrane Central Register of Controlled Trials (CENTRAL), Cumulative Index to Nursing and Allied Health Literature (CINAHL), and PsycINFO. These databases were selected to capture biomedical literature (MEDLINE), European and drug/device-related research (EMBASE), multidisciplinary citation indexing (Scopus), registered trials and systematic reviews (CENTRAL), nursing and rehabilitation literature (CINAHL), and psychological and behavioral outcomes (PsycINFO).

Search strategy

The search strategy combined controlled vocabulary (e.g., MeSH, Emtree) and title/abstract field searches, with database-specific adaptations to accommodate each platform's syntax requirements. The search employed three core components connected by Boolean operators, as detailed in Table 1. Multi-word phrases were enclosed in quotation marks to ensure proximity searching. All three core components were combined by using AND to form the final search strategy. The MEDLINE (PubMed) example in Table 1 illustrates the application of title/abstract field

Table 1 Core search components and MEDLINE (PubMed) search strategy

Component	Search Terms
Fracture population	pediatric OR pediatric OR child* OR children OR childhood OR minor* OR adolescent OR teenager* OR youth OR school age OR school-age OR student* AND (fracture* OR trauma* OR injury*)
Upper limb location	upper limb* OR upper extremity OR upper arm OR arm OR shoulder OR clavicle OR collarbone OR humerus OR humeral OR supracondylar OR distal humerus OR radius OR radial OR radioulnar OR ulna OR ulnar OR forearm OR distal forearm OR wrist OR carpal OR carpometacarpal OR metacarp* OR phalang* OR distal radius OR distal ulna OR both bones forearm OR Monteggia OR Galeazzi OR epiphyseal
Outcome measurement/ PROM Terms	patient-reported OR patient reported OR self-report* OR PROM* OR HRQoL OR health-related quality of life OR HRQL OR health related quality of life OR outcome* OR outcome measure* OR quality of life OR health outcome OR treatment outcome OR clinical outcome* OR functional status OR functional outcome* OR disability OR disabilities OR impairment* OR questionnaire* OR assessment* OR evaluation* OR survey OR instrument*
MEDLINE (PubMed) example	(pediatric[Title/Abstract] OR child[Title/Abstract] OR adolescent[Title/Abstract]) AND (fracture[Title/Abstract] OR trauma[Title/Abstract] OR injur[Title/Abstract]) AND ("upper limb"[Title/Abstract] OR "upper extremity"[Title/Abstract] OR shoulder[Title/Abstract] OR humerus[Title/Abstract] OR forearm[Title/Abstract] OR radius[Title/Abstract] OR ulna[Title/Abstract] OR wrist[Title/Abstract] OR hand[Title/Abstract]) AND ("patient-reported"[Title/Abstract] OR "patient reported"[Title/Abstract] OR PROM[Title/Abstract] OR "outcome measure"[Title/Abstract] OR "quality of life"[Title/Abstract] OR "functional outcome"[Title/Abstract] OR questionnaire[Title/Abstract])

Truncation/wildcard symbol used in database search syntax (e.g., child retrieves child, children, childhood). Truncation and phrase syntax were adapted to each database's requirements.

tags for this database; similar adaptations were made for other databases according to their respective syntax requirements.

Secondary searching included citation tracking of the included studies and relevant systematic reviews, as well as manual reviews of reference lists.

Study selection and screening process

Retrieved records were exported to Covidence (Veritas Health Innovation, Melbourne, Australia) for systematic review management and screening. Two independent reviewers (M.N.A.T. and L.M.A.) screened titles and abstracts against eligibility criteria. Studies with insufficient information from the title/abstract were automatically advanced to full-text review. Both reviewers jointly screened a 20% sample to establish screening consistency before independently screening the remaining records. Inter-rater agreement was quantified using Cohen's kappa coefficient (kappa ≥0.60 considered acceptable). Potentially relevant studies underwent full-text review by the two reviewers, with disagreements resolved through discussion or a third reviewer (M.A.S.). Reasons for exclusion were documented, and a PRISMA-ScR flow diagram was completed.

Data extraction and management

A standardized data extraction form was pilot-tested on 10 included studies. Extracted data included study characteristics (first authors, year, country, design, setting), participant characteristics (sample size, age, sex), fracture characteristics (type, mechanism, laterality, associated complications), intervention details, PROM details (name, type, items, administration, timing), reported psychometric properties, and outcome results. A 30% random sample was independently extracted by both reviewers to assess agreement; discrepancies were resolved through discussion. Data were entered into a secure database (REDCap [Research Electronic Data Capture]) for management and analysis.

PROM categorization and coding

Identified PROMs were categorized into five types: (1) upper extremity/region-specific measures (e.g., QuickDASH); (2) generic pediatric HRQL instruments (e.g., PedsQL, PROMIS Pediatric); (3) generic HRQL instruments adapted for pediatric use (e.g., EQ-5D-Y); (4) pediatric-specific functional measures (e.g., PODCI); and (5) condition-specific pediatric fracture measures (e.g., PROOF-UE questionnaire).

Psychometric evidence appraisal

Psychometric evidence was appraised using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) Risk of Bias Checklist. Two independent reviewers (L.M.A. and M.A.S.) assessed each measurement property across 10 domains (content validity, structural validity, internal consistency, reliability, measurement error, criterion validity, construct validity, responsiveness, cross-cultural validity, interpretability). Each property was rated on a four-point scale (Adequate, Doubtful, Inadequate, Not Applicable). Overall quality ratings (High, Moderate, Low, Very Low) were assigned in accordance with the COSMIN guidelines. Evidence was synthesized narratively, with the overall strength of the evidence rated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria.

Mapping to the CORE-Kids framework

Identified instruments were mapped to the Core Outcomes in Pediatric Research (CORE-Kids) multidimensional quality-of-life framework, which comprises six domains: physical function and independence, emotional well-being, social functioning, school and learning, pain and discomfort, and overall life satisfaction. Two independent

reviewers (M.N.A.T. and M.A.S) assessed each instrument's domain coverage as comprehensive (≥3 items), moderate (1–2 items), or limited/no coverage (0 items). Discrepancies were resolved by consensus.

Data analysis, knowledge synthesis, and gap identification

Descriptive statistics were calculated for study and instrument characteristics, including frequencies by publication year, country, study design, fracture type, and PROM application. Narrative synthesis was employed to describe instrument characteristics, summarize psychometric evidence by measurement property with quality ratings, identify evidence gaps, and evaluate CORE-Kids domain coverage. Subgroup analyses were conducted by fracture type, age group (6–9, 10–13, 14–17 years), study design, and primary outcome type.

The review synthesized knowledge to identify methodological gaps (measurement properties lacking published evidence), application gaps (fracture types or age groups with limited evidence), coverage gaps (quality-of-life domains insufficiently addressed), and evidence-quality gaps (properties evaluated with insufficient quality per COSMIN standards).

Results

Study identification and selection

Database searches yielded 2,847 unique records after duplicate removal. Citation tracking added 23 records, bringing the total to 2,870 records screened. Title and abstract screening by the two independent reviewers (κ=0.78; 92% agreement) identified 387 potentially relevant records. Full-text review (κ=0.81; 94% agreement) resulted in 68 studies meeting the inclusion criteria. Common exclusion reasons were no upper limb fracture population (27.3%), no PROM or HRQL reported (28.8%), adult-only population without pediatric subgroup data (20.4%), and chronic upper limb conditions without acute fracture (12.9%). A PRISMA-ScR flow diagram is presented in Figure 1.

Characteristics of included studies

Included studies were published between 2005 and 2025, with increasing frequency over time: 2005–2009 (7.4%), 2010–2014 (16.2%), 2015–2019 (33.8%), and 2020–2025 (42.6%). Studies originated from 28 countries across five continents: Europe (42.6%), North America (26.5%), Asia-Pacific (17.6%), South America (7.4%), and the Middle East (5.9%).

Study designs comprised retrospective case series or cohort studies (45.6%), prospective cohort studies (27.9%), RCTs (16.2%), cross-sectional surveys (7.4%), and qualitative studies (2.9%). A total of 5,247 children across all studies, with mean ages ranging from 5.2 to 17.8 years. Sex distribution was 54.3% male and 45.7% female. Ethnicity data were reported in only 23.5% of studies, with predominantly White/Caucasian populations.

Fracture types studied included distal radius (35.3% of studies), supracondylar humerus (26.5%), proximal humerus (11.8%), distal humerus (10.3%), forearm shaft (13.2%), medial epicondyle (8.8%), hand (5.9%), wrist (4.4%), and clavicle (2.9%). Treatment modalities included non-operative management only (33.8%), operative management only (39.7%), and mixed cohorts (26.5%).

PROMs identified

The review identified 18 distinct PROMs applied across 68 studies, with 186 total applications (range 1–4 PROMs per study). Characteristics and frequency of the commonly used instruments are summarized in Table 2.

QuickDASH was the most commonly applied instrument (41% of studies), reflecting its efficient 11-item format, applicability across multiple fracture types, and multilingual availability. PROMIS Upper Extremity (17.6% of studies) offers the advantages of computerized

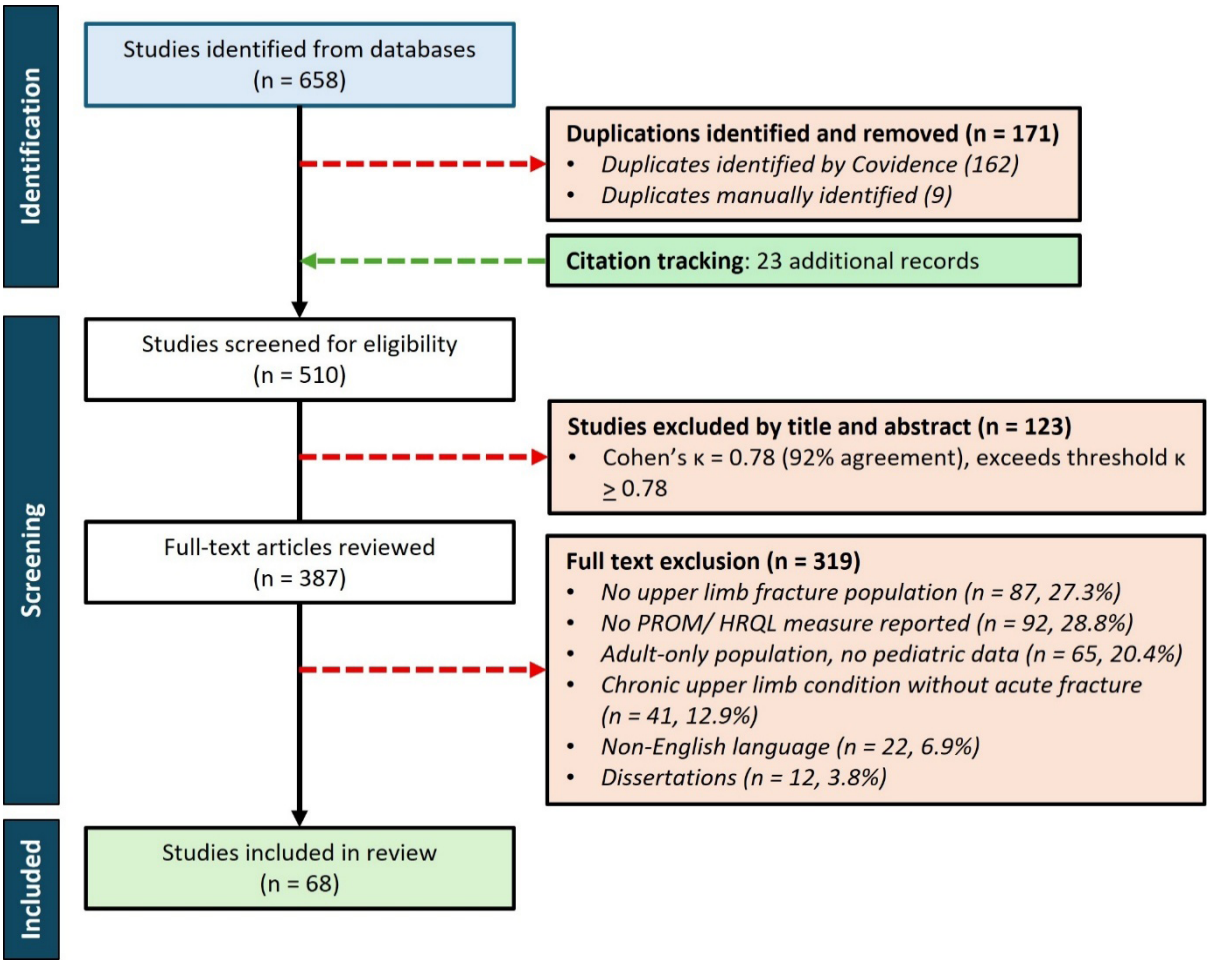

Figure 1 PRISMA-ScR flow chart

adaptive testing and integration into contemporary pediatric orthopedic registries[18]. PODCI (14.7% of studies) provides a comprehensive multidimensional assessment but requires longer administration time. Detailed feasibility and administration characteristics for all instruments are presented in Supplemental Table S1.

Identified PROMs were categorized into five types based on their intended use and target population. Upper extremity or region-specific instruments formed the largest group, comprising nine instruments (50.0%). Generic pediatric health-related quality of life (HRQL) instruments accounted for three instruments (16.7%), while pediatric-specific functional instruments represented two instruments (11.1%). Generic HRQL instruments adapted for pediatric use comprised three instruments (16.7%). The PROOF-UE questionnaire was classified as a condition-specific pediatric fracture measure, representing one instrument (5.6%). A detailed breakdown of instrument categorization is provided in Supplemental Table S2.

Psychometric evidence summary

Validity evidence

Construct validity: QuickDASH demonstrated acceptable convergent validity with PROMIS Upper Extremity

Table 2 Characteristics and frequency of patient-reported outcome measures (PROMs) used in pediatric upper limb fracture studies

Instrument	Type	Items	Domains	Studies using n (%)
QuickDASH	Region-specific	11	Physical function, disability, pain	28 (41.2)
PROMIS upper extremity	Generic pediatric HRQL	Variable (CAT)*	Upper extremity function	12 (17.6)
PODCI	Pediatric-specific functional	36	UE function, mobility, sports, pain, happiness, global	10 (14.7)
PedsQL 4.0	Generic pediatric HRQL	23	Physical, emotional, social, and school	9 (13.2)
PROMIS pediatric (other domains)	Generic pediatric HRQL	Variable (4–12)**	Pain Interference, Mobility, etc.	8 (11.8)
Patient-rated wrist evaluation (PRWE)	Region-specific	15	Wrist pain, function	7 (10.3)
PROOF-upper extremity (UE)	Condition-specific pediatric fracture	21	Appearance, pain/feeling, function, healing	4 (5.9)
EQ-5D-Y	Generic HRQL adapted	5 + VAS	Mobility, self-care, activities, pain, anxiety	3 (4.4)
Jebsen-Taylor hand function test	Region-specific	7	Simulated ADLs	3 (4.4)
ABILHAND-Kids	Region-specific	22	Manual ability (bimanual)	2 (2.9)
ASK/ASK-P	Region-specific	15	Physical function, ADLs	1 (1.5)
Canadian occupational Performance measure (COPM)	Region-specific	Variable	Self-identified performance, satisfaction	1 (1.5)
Modified ashworth scale	Pediatric-specific functional	Variable (4–6)***	Muscle spasticity	1 (1.5)
Oxford shoulder score	Region-specific	12	Shoulder pain, function	1 (1.5)
Oxford elbow score	Region-specific	12	Elbow pain, function	1 (1.5)
Short form 36 (SF-36)	Generic HRQL adapted	36	Physical and mental health domains	1 (1.5)
Visual analogue scale (VAS) pain	Generic HRQL adapted	1	Pain intensity	1 (1.5)
Mayo Elbow performance score (MEPS)	Region-specific	4	Pain, motion, stability, function	1 (1.5)

ADLs=activities of daily living, CAT=computerized adaptive testing, VAS=visual analogue scale
Percentages sum to >100% due to multiple PROMs per study.
*Computerized Adaptive Testing (CAT); number of items varies (typically 4–12) based on stopping rules.
**Number of items depends on domain and format (short form or CAT).
***Number of items varies depending on the number of muscle groups assessed, typically, 4–6 muscle groups in upper limb evaluation.

(r=0.61–0.79)[23], though structural validity limitations were noted (RMSEA=0.089, CFI=0.891). PROMIS Upper Extremity showed strong convergent validity with PROMIS Mobility (r=0.78), PedsQL Physical Function (r=0.82), and PODCI Upper Extremity subscale (r=0.52), with known-groups validity supported by significantly different T-scores across fracture severity levels[18,24]. PODCI correlated appropriately with PROMIS Upper Extremity (r=0.52; p-value<0.001)[24,25]. PROOF-UE demonstrated excellent known-groups validity, with complicated fractures scoring significantly worse (p-value<0.05)[11].

Content validity: Formal content validity studies in pediatric upper limb fracture populations were absent for all identified instruments. Only PROOF-UE documented concept elicitation interviews with children and parents during development. This represents a critical evidence gap.

Reliability and internal consistency

Most instruments demonstrated good to excellent internal consistency (Cronbach's α ≥0.75). PROOF-UE showed slightly lower domain-specific α (0.62–0.88) but excellent total score consistency (α>0.90). EQ-5D-Y showed lower internal consistency (α=0.69–0.78), consistent with its brief item structure. Test-retest reliability was excellent for PROOF-UE (ICC up to 0.98), ABILHAND-Kids (0.98), and PODCI (≥0.92), and good for QuickDASH, PRWE, and PedsQL. Notably, PROMIS-UE demonstrated direction-dependent reliability, with high reliability in the clinically worse direction but lower reliability (ICC<0.70) in normal functioning ranges[18]. Detailed reliability data are provided in Table 3.

Responsiveness to change

Responsiveness evidence was identified for 8 of the 10 most frequently used instruments. PROOF-UE demonstrated the strongest responsiveness, with standardized response means exceeding 1.0 across all domains[11]. QuickDASH (SRM=0.42–1.45) and PRWE (SRM=0.55–1.28) showed excellent responsiveness in the appropriate fracture populations[26]. PROMIS-UE showed moderate responsiveness (SRM=0.38–0.98) with variability by baseline severity. EQ-5D-Y demonstrated modest responsiveness (d=0.28–0.65), suitable for group-level comparisons but limited for detecting individual patient change[8]. ABILHAND-Kids, Mayo MEPS, and Oxford Shoulder/Elbow Scores lacked sufficient responsiveness data

Table 3 COSMIN quality ratings, reliability, and floor/ceiling effects summary for the most frequently used PROMs

PROMs	Overall quality (Q) and content validity (C)	Internal consistency (α)	Test-retest (ICC)	Floor effect (%)	Ceiling effect (%)	Key limits
PROOF-UE	Q: Excellent C: Adequate	0.62–0.88 (domain); >0.90 (total)	0.67–0.98 (child); 0.57–0.84 (parent)	0	5	Condition-specific only
QuickDASH	Q: Good C: Doubtful	0.75–0.82	0.80–0.89	0–2	0–3	Adult-derived; limited pediatric content validity
PODCI	Q: Good C: Doubtful	0.90–0.96	0.82–0.94	<1	<5	Length; limited fracture-specific validation
PRWE	Q: Good C: Doubtful	0.85–0.92	0.84–0.89	<1	2–5	Limited pediatric evidence
PROMIS-UE	Q: Moderate C: Doubtful	>0.80	0.84 (clinically worse)	<1	13	Ceiling effects; limited responsiveness
PedsQL 4.0	Q: Moderate C: Doubtful	0.77–0.89	0.78–0.91	0	18–22	Generic; may lack fracture-specific sensitivity
ABIL HAND-Kids	Q: Good C: Adequate	0.94	0.98	NR	NR	Bimanual assessment only; interview format required
ASK/ ASK-P	Q: Good C: Doubtful	>0.85	0.86–0.91	NR	NR	Limited fracture-specific validation
EQ-5D-Y	Q: Low C: Inadequate	0.69–0.78	0.75–0.85	<1	12–18	Limited pediatric fracture validation
Mayo MEPS	Q: Moderate C: Doubtful	Not tested	0.88–0.94	<1	25–35	Substantial ceiling effects; limited pediatric evidence
Oxford shoulder score	Q: Moderate C: Doubtful	NR	NR	3–5	8–12	Limited pediatric validation; ceiling effects
Oxford Elbow Score	Q: Moderate C: Doubtful	NR	NR	2–4	10–15	Limited pediatric validation; ceiling effects

PROMs=patient-reported outcome measures, NR=not reported. Floor effect ≥10% or ceiling effect ≥15% considered problematic

in pediatric fracture populations. Detailed responsiveness data are provided in Table 4.

Minimum clinically important differences (MCID): Limited pediatric-specific MCID data exist. QuickDASH MCID is estimated at 10–14 points (based on adult data). PROMIS-UE T-score changes of 3–5 points may represent a clinically meaningful change. PROOF-UE domain score changes of 10–15 points or total score changes of 15–20 points are suggested as clinically meaningful.

Floor and ceiling effects

Most instruments demonstrated acceptable floor and ceiling effects (<10% floor, <15% ceiling). Notable exceptions included Mayo Elbow Performance Score (25%–35% ceiling), PedsQL Physical Functioning (18%–22% ceiling), and PROMIS-UE (13% ceiling), which limited the ability to detect improvement in well-functioning children. QuickDASH, PROOF-UE, and PODCI showed excellent floor/ceiling effect profiles. Detailed floor and ceiling effect data for all instruments are provided in Table 3.

COSMIN quality assessment results

COSMIN quality assessment was completed for the 10 most frequently used instruments. PROOF-UE demonstrated the highest overall quality, with adequate content validity and good-to-excellent psychometric properties across domains[11]. QuickDASH, PODCI, and PRWE showed good reliability and construct validity but doubtful content validity in pediatric populations[27]. PROMIS-UE and PedsQL 4.0 demonstrated moderate-quality evidence, with PROMIS-UE showing ceiling effects and PedsQL lacking fracture-specific sensitivity. EQ-5D-Y demonstrated the lowest quality, with inadequate content validity and limited psychometric testing in pediatric fracture

Table 4 Responsiveness to change for PROMs

PROMs	Effect size (Cohen's d)	Standardized response mean (SRM) range	Responsiveness classification	Clinical contexts tested	MCID estimate
PROOF-UE	1.10–1.85	1.02–1.63	Excellent	Upper extremity fracture healing	Domain: 10–15 points; Total: 15–20 points
QuickDASH	0.65–1.82	0.42–1.45	Good to Excellent	Fracture healing, surgical intervention	10–14 points (adult-derived)
PRWE	0.72–1.45	0.55–1.28	Good to Excellent	Wrist fracture treatment	Not established in pediatrics
PODCI	0.58–1.34	0.40–1.12	Good to Excellent	Orthopedic intervention, rehabilitation	Not established in pediatrics
ASK/ASK-P	0.82–1.18	0.65–0.95	Good to Excellent	Pediatric functional intervention	NE
PROMIS-UE	0.45–1.23	0.38–0.98	Good	Operative and non-operative treatment	3–5 T-score points (adult-derived)
PedsQL 4.0	0.35–0.92	0.25–0.85	Moderate to Good	General pediatric populations	NE
EQ-5D-Y	0.28–0.65	NR	Moderate	Generic health assessment	NE
ABIL HAND-Kids	NR	NR	Insufficient data	Bimanual hand function	NE
Mayo MEPS	NR	NR	Insufficient data	Elbow function	NE
Oxford shoulder score	NR	NR	Insufficient data	Shoulder function	NE
Oxford Elbow Score	NR	NR	Insufficient data	Elbow function	NE

PROMs=patient-reported outcome measures, MCID=minimal clinically important difference, NE=Not established, NR=Not reported, SRM=standardized response mean (change score divided by standard deviation of change)
Effect size interpretation: <0.2=trivial, 0.2–0.5=small, 0.5–0.8=moderate, >0.8=large

populations[8,18]. Detailed COSMIN quality ratings are provided in Table 3. Critical evidence gaps were identified across multiple measurement properties,[14] as summarized in Table 5.

CORE-Kids domain mapping analysis

All 18 identified instruments were systematically mapped to the CORE-Kids framework, assessing coverage of six pediatric quality-of-life domains. Mapping was performed independently by two reviewers with excellent inter-rater agreement (κ=0.89). Results are summarized in Table 6. Detailed domain coverage by instrument category is provided in Supplemental Table S3.

Physical Function and Pain domains demonstrated acceptable to moderate coverage. However, major gaps were identified in Emotional Well-being (18% comprehensive) and Social Functioning (14% comprehensive)[12,19]. Critical gaps were observed in School/Learning (9% comprehensive) and Overall Life Satisfaction (5% comprehensive). The most frequently used instruments (QuickDASH, PROMIS-UE, PRWE, PODCI-UE subscale) do not assess emotional adjustment, anxiety, depression, peer relationships, school attendance, or global satisfaction with recovery[14,28], all of which are domains central to pediatric quality of life.

Subgroup analyses

Instrument selection patterns varied by fracture type, age group, and study design. Key findings are summarized below; detailed data are provided in Supplemental Tables S4–S6.

By Fracture Type: QuickDASH was the most frequently used instrument across all fracture types. PRWE was more commonly applied to distal radius fractures (33% of studies). PROMIS-UE was most frequently used in supracondylar humerus fractures (39%). PODCI was preferred for studies involving mixed fracture types (57%).

By Age Group: For children aged 5–8 years, parent-proxy measures were used in 78% of studies, with PODCI

Table 5 Critical evidence gaps and recommendations

Measurement property	Instruments with evidence (%)	Gap severity (%)	Priority	Recommended action
Content validity	10 (1 of 10)	90	Critical	Conduct concept elicitation studies with children and families to identify outcomes that matter most
Measurement error (SEM/MDC)	20 (2 of 10)	80	Critical	Establish the standard error of measurement and the minimal detectable change for each instrument
Interpretability (MCID)	30 (3 of 10)	70	High	Determine pediatric-specific minimal clinically important differences using anchor-based and distribution-based methods
Criterion validity	60 (6 of 10)	40	Moderate	Correlate PROMs with objective functional measures and clinician-assessed outcomes
Structural validity	70 (7 of 10)	30	Moderate	Conduct confirmatory factor analyses to validate instrument structure in pediatric fracture populations
Responsiveness	70 (7 of 10)	30	Moderate	Evaluate the ability to detect clinically important change across fracture types and treatment modalities
Construct validity	80 (8 of 10)	20	Low	Further validate convergent and divergent validity with reference measures
Test-retest reliability	90 (9 of 10)	10	Low	Confirm stability over time in stable fracture populations
Floor/ceiling effects	90 (9 of 10)	10	Low	Evaluate instrument performance across the full range of function, particularly in well-functioning children

SEM=standard error of measurement, MDC=minimal detectable change, MCID=minimal clinically important difference

Table 6 CORE-Kids domain coverage across all instruments (n=22)

Domain	Comprehensive coverage (≥3 items)	Comprehensive (%)	Gap severity
Physical function/independence	11	50	Acceptable
Pain/discomfort	8	36	Moderate
Emotional well-being	4	18	Major
Social functioning	3	14	Major
School/learning	2	9	Critical
Overall life satisfaction	1	5	Critical

and PedsQL being the most common. For children aged 9–12 years, child self-report increased, with QuickDASH use rising to 48%. For children aged 13–16 years, self-report predominated (74%), with QuickDASH use highest (58%).

By study design: RCTs showed the highest QuickDASH use (64%). Prospective cohort studies demonstrated more diverse instrument use, with PROOF-UE adoption in 11% of studies. Retrospective case series showed the most heterogeneous selection. Cross-sectional surveys preferred PROMIS-UE (60%), reflecting computerized adaptive testing efficiency.

Knowledge synthesis and gap identification

This scoping review identified four overarching evidence gaps:

Methodological gaps: Content validity studies are absent for 90% of instruments; measurement error and interpretability data are lacking for most instruments, limiting clinical utility.

Application gaps: Fracture types such as clavicle, hand, and elbow fractures have limited outcome measurement evidence; younger children (5–8 years) are underrepresented in self-report studies.

Coverage gaps: Critical gaps exist in the assessment of school functioning and overall life satisfaction; major gaps exist in emotional well-being and social functioning domains.

Evidence quality gaps: Most instruments have at most moderate-quality evidence for pediatric fracture populations, with few meeting COSMIN criteria for high-quality psychometric validation.

Discussion

This scoping review mapped the landscape of PROMs and HRQL instruments used in school-age children with upper limb fractures, synthesizing evidence on instrument characteristics, psychometric properties, and coverage of pediatric-relevant outcome domains. The review identified 22 distinct PROMs across 68 studies (2005–2025), with QuickDASH being the most frequently applied (41% of studies), followed by PROMIS Upper Extremity (18%), PODCI (15%), and PedsQL (13%). While this heterogeneity reflects diverse fracture types and clinical contexts, it also demonstrates the absence of consensus on optimal outcome measurement. The dominance of QuickDASH, an adult-derived instrument lacking pediatric content validity, highlights a critical tension between measurement efficiency and child-centeredness[23]. In contrast, PROOF-UE, the only fracture-specific instrument developed with pediatric patient and family input, appeared in only 5.9% of studies, suggesting limited adoption despite superior content validity and psychometric properties[11].

Psychometric properties and methodological quality

Using COSMIN methodology, the authors found that most instruments demonstrated good to excellent

reliability and internal consistency. However, critical evidence gaps emerged. Content validity, the foundational property establishing that instruments measure what matters to patients[21], was documented for only one instrument (PROOF-UE). This gap is concerning because instruments without content validity risk measuring constructs that clinicians deem important rather than those meaningful to children and families. This finding aligns with the previous reviews of upper- and lower-extremity PROMs, which similarly reported that while reliability data were available, content validity remained a major gap[13,17,29]. Similarly, measurement error (SEM/MDC) data were absent for 80% of instruments, and pediatric-specific minimal clinically important differences were unavailable for 70%, limiting clinicians' ability to distinguish true change from measurement variation. EQ-5D-Y demonstrated the lowest overall quality, with inadequate content validity and limited validation for pediatric fractures[8,17,18,27].

CORE-Kids domain coverage

Mapping to the CORE-Kids framework revealed substantial gaps in the coverage of pediatric quality of life[12]. While physical function (50% comprehensive) and pain (36%) were adequately assessed, major gaps existed in emotional well-being (18% comprehensive) and social functioning (14%). Critical gaps were identified in school/learning (9%) and overall life satisfaction (5%). Most frequently used instruments (QuickDASH, PROMIS-UE, PRWE) do not assess emotional adjustment, peer relationships, school attendance, or global satisfaction, domains central to pediatric recovery. This gap is particularly significant given evidence that anxiety predicts worse functional outcomes in children with fractures[28], and that social deprivation influences patient-reported outcomes[25,30].

Implications for research and practice

For research, priority should be given to conducting content validity studies for existing instruments through concept elicitation with children and families, establishing measurement error and pediatric-specific MCID values, developing or adapting instruments to address emotional well-being, school functioning, and life satisfaction, and developing a core outcome set following COMET methodology to standardize measurement across trials. For practice, no single instrument captures all meaningful outcomes; therefore, a multi-measure approach is recommended[31]. PROOF-UE is preferred for physical function assessment[11], with QuickDASH serving as an acceptable alternative[7,23], while PODCI or PedsQL should be used for comprehensive quality-of-life evaluation across emotional, social, and school domains[5,12,25]. PROMIS Anxiety and Depressive Symptoms are recommended for psychological assessment in at-risk children[28]. Age-appropriate administration is essential: parent-proxy reporting is preferred for children aged 6–9 years, transition to self-report is feasible for aged 10–13 years, and self-report predominates for adolescents aged 14–17 years.

Limitations

This review has limitations. Inclusion of only English-language publications may have excluded relevant studies from non-English-speaking countries. The focus on school-age children (5–16 years) excludes younger children and older adolescents, for whom different instruments may be appropriate. Heterogeneity in PROM application and fracture types limited quantitative synthesis. Additionally, the rapid development of new instruments may have outpaced the evidence captured through 2025.

Conclusion

A total of 22 PROMs were identified across 68 studies of school-age children with upper limb fractures, revealing that QuickDASH is most frequently used despite lacking pediatric content validity. In contrast, PROOF-UE, the only instrument developed with child and family input, remains underutilized. Critical evidence gaps persist:

90% of instruments lack content validity studies, 80% lack measurement error data, and 70% lack pediatric-specific MCID values. Most importantly, commonly used instruments inadequately assess emotional well-being, social functioning, school outcomes, and overall life satisfaction—domains central to pediatric quality of life. Until these gaps are addressed, clinicians should adopt a multi-measure approach that combines physical function assessment with comprehensive quality-of-life and psychological outcome measures, ensuring that outcome assessment captures what matters most to children and families recovering from upper limb fractures.

References

1. Badhe N, Busby C, See A, Deacon C, Altell T, Ollivere BJ, et al. Waterproof casts for the management of upper limb fractures in children: a systematic review and meta-analysis. Bone Joint J 2025;107-B:587-94.
2. Cordray H, Fiandeiro M, Banala M, Vaile JR, Struble SL, Pehnke M, et al. Gamified assessments of pediatric upper extremity function: a systematic review and appraisal. J Hand Ther 2025;38:715–25.
3. Higgins MJ, Gomez RW, Storino M, Wang N, Strouse A, Zook Z, et al. 3D-printed casts improve patient-reported function and satisfaction in children with stable fracture patterns. J Pediatr Orthop 2025;45:546–51.
4. Musters L, Roth KC, Diederix LW, Edomskis PP, Benner JL, Reijman M, et al. Does early conversion to below-elbow casting for pediatric diaphyseal both-bone forearm fractures adversely affect patient-reported outcomes and ROM? Clin Orthop Relat Res 2024;482:1873-81.
5. Taib MNA, Kamalanathan CK, Sharifudin MA, A.Rahman LM, Abu Bakar MZH, Awang MS. Temporal dynamics of quality of life and functional recovery in pediatric supracondylar humerus fractures: a prospective study. Mal J Med Health Sci 2024;20:157–68.
6. Alberghina F, Andreacchio A, Pavone V, Mansour M, Dimeglio A, Canavese F. Review of pediatric functional outcomes measures used to evaluate surgical management in pediatric patients with an upper extremity fracture. J Pediatr Orthop B 2022;31:260–9.
7. Ernat JJ, Wimberly RL, Ho CA, Riccio AI. Vascular examination predicts functional outcomes in supracondylar humerus fractures: a prospective study. J Child Orthop 2020;14:495-501.
8. Nowak LL, Davis AM, Mamdani M, Beaton D, Kennedy C, Schemitsch EH. A systematic review and standardized comparison of available evidence for outcome measures used to evaluate proximal humerus fracture patients. J Orthop Trauma 2019;33:e256-62.
9. Fiandeiro M, Cordray H, Vaile JR, Struble SL, Banala M, Pehnke M, et al. Selecting patient-reported outcome measures for pediatric upper extremity function: a systematic review. J Pediatr Orthop 2025;45:508–18.
10. Fiandeiro M, Cordray H, Vaile JR, Struble SL, Banala M, Pehnke M, et al. Measuring what matters: a guide to selecting the optimal patient-reported outcome measure for pediatric shoulder and elbow function. J Shoulder Elbow Surg 2026;35:e136–46.
11. Narayanan U, Di Nola L, Ho E, Borkhoff CM, Camp M, Murphy A, et al. Validation of the patient/parent-reported outcome measure of fracture healing (PROOF-UE) questionnaire for upper extremity fractures in children. Orthop Procs 2025;107-B (Suppl 11):24.
12. Morris EJ, Gray K, Gibbons PJ, Grayson J, Sullivan J, Amorim AB, et al. Evaluating the use of PROMs in pediatric orthopaedic registries. Children (Basel) 2023;10:1552.
13. Abbot S, Proudman S, Sim YP, Williams N. Psychometric properties of patient-reported outcomes measures used to assess upper limb pathology: a systematic review. ANZ J Surg 2022;92:3170-5.
14. Nuttall J, Rolls C. A survey of practice on the use of condition-specific patient reported outcome measures with patients who have distal radius fractures. Hand Ther 2025;30:63-71.
15. Arguelles GR, Shin M, Lebrun DG, Kocher MS, Baldwin KD, Patel NM. The majority of patient-reported outcome measures in pediatric orthopaedic research are used without validation. J Pediatr Orthop 2021;41:e74-9.
16. Marson BA, Craxford S, Deshmukh SR, Grindlay DJC, Manning JC, Ollivere BJ. Quality of patient-reported outcomes used for quality of life, physical function, and functional capacity in trials of childhood fractures. Bone Joint J 2020;102-B:1599-607.
17. Kalle JPR, Saris TFF, Sierevelt IN, Eygendaal D, van Bergen CJA. Quality of patient- and proxy-reported outcomes for children with impairment of the upper extremity: a systematic

review using the COSMIN methodology. J Patient Rep Outcomes 2022;6:58.

18. Korteling DL, Ketelaar M, Limmen S, Terwee CB, Bloemen MAT, Rameckers EAA, et al. Psychometric properties and reference values of the patient-reported outcomes measurement information system (PROMIS®) pediatric item banks mobility, upper extremity, and pain interference in the Dutch population. J Clin Epidemiol 2025;184:111855.
19. Dorich JM, Cornwall R, Uhl T. ICF linking of patient-reported therapy goals for children with acquired upper extremity impairment. J Hand Ther 2023;36:74-84.
20. Quax MLJ, Krijnen P, Schipper IB, Termaat MF. Managing patient expectations about recovery after a distal radius fracture based on patient reported outcomes. J Hand Ther 2023;36:903-12.
21. Morales-Dubois LM, Julian KR, Maruyama M, Kamal RN, Shapiro LM. Have commonly utilised patient-reported outcome measures been validated for the Japanese language and culture amongst patients with distal radius fractures? a systematic review. J Hand Surg Asian Pac 2026;31:137-46.
22. Cordray H, Fiandeiro M, Struble SL, Vaile JR, Banala M, Pehnke M, et al. Choosing the most appropriate patient-reported outcome measures for hand function: a guide for pediatric hand surgeons. Hand (N Y) 2026;21:1206-16.
23. Kämppä N, Hulkkonen S, Grahn P, Laaksonen T, Repo J. The construct validity and internal consistency of QuickDASH in pediatric patients with upper extremity fractures. Acta Orthop 2024;95:192-9.
24. Nguyen HB, Miller PE, Sullivan N, Mahan ST, Spencer SA, Kasser JR, et al. Psychometric evaluation of patient-reported outcome measurement information system (PROMIS) measures in children and adolescents with lower limb orthopaedic conditions: construct validity, reliability, and responsiveness. Bone Joint J 2026;108-B:259-66.
25. Gates PE, Campbell SR. Effects of age, sex, and comorbidities on the pediatric outcomes data collection instrument (PODCI) in the general population. J Pediatr Orthop 2015;35:203-9.
26. Yaakobi D, Ron I, Agbaria H, Yelton MJ, Peskin B, Ghrayeb N, et al. Postoperative grip strength and range of motion correlated with better patient-reported outcomes following surgical management of distal radius fractures. J Wrist Surg 2024;15:141-6.
27. Strydom JP, Pretorius HS, Burge MC. Functional and patient-reported outcomes of terrible triad elbow fracture-dislocations. S Afr Orthop J 2025;24:169-75.
28. Puckett HD, Borgida JS, Doxey SA, Kleinsmith RM, Bohn DC, Husband JB, et al. Anxiety and the attainment of clinically significant outcomes in operatively treated distal radius fracture patients: a retrospective study of 78 patients. Hand (N Y) 2026:15589447261415641. doi: 10.1177/15589447261415641.
29. Saris TF, Kalle JR, Sierevelt IN, Eygendaal D, van Bergen CJ. Quality of patient- and proxy-reported outcomes for children with impairment of the lower extremity: a systematic review using the consensus-based standards for selection of health measurement instruments methodology. J Child Orthop 2023;18:96-110.
30. Evans S, Okoroafor UC, Calfee RP. Is social deprivation associated with PROMIS outcomes after upper extremity fractures in children? Clin Orthop Relat Res 2021;479:826-34.
31. Uppal H, Garcia D, Abdelmalek G, Attia P, Sahai N, Sinha K, et al. Evaluating the readability of patient-reported outcome measures in pediatric orthopaedic registries: alignment with National Institutes of Health and American Medical Association Standards. J Pediatr Orthop 2025;45:599-606.

Supplemental Table S1 Feasibility and administration characteristics of patient-reported outcome measures (PROMs) used in pediatric upper limb fracture studies

Instrument	Admin time (minutes)	Items	Literacy level	Completion rate (%)
QuickDASH	2–3	11	Grade 6	95–98
PROMIS-UE CAT	2–5	Variable (8–12)	Grade 5	94–96
PODCI	10–15	36	Grade 7	88–92
PedsQL 4.0	5–10	23	Grade 6	93–95
PRWE	5–10	15	Grade 7	94–96
PROOF-UE	5–7	21	Grade 6	97–99
EQ-5D-Y	2–3	5 + VAS	Grade 5	98–99
ASK/ASK-P	5–10	15	Grade 6	92–94
ABILHAND-Kids	10–15	22	Grade 7	88–91

Supplemental Table S2 PROM instruments by category

Category	Instruments	Count	Percentage (%)
Upper extremity/region-specific	QuickDASH, PRWE, Oxford Shoulder Score, Oxford Elbow Score, ABILHAND-Kids, COPM, ASK/ASK-P, Jebsen-Taylor, Mayo MEPS	9	50.0
Generic Paediatric HRQL	PedsQL, PROMIS Upper Extremity, PROMIS Pediatric (other domains)	3	16.7
Paediatric-specific functional	PODCI, Modified Ashworth Scale	2	11.1
Generic HRQL adapted for Paediatric use	EQ-5D-Y, SF-36, VAS pain	3	16.7
Condition-specific	PROOF-UE	1	5.6
Total		18	100.0

Supplemental Table S3 CORE-Kids domain coverage by instrument category

Domain	Upper extremity region-specific (n=11) (%)	Generic paediatric HRQL (n=4)(%)	Paediatric-specific functional (n=3)(%)	Generic HRQL adapted (n=3)(%)	Fracture-specific (n=1)(%)
Physical function	100	100	100	67	100
Pain/discomfort	91	50	100	67	100
Emotional well-being	9	100	33	67	0
Social functioning	9	100	67	33	0
School/learning	0	75	33	0	100
Overall life satisfaction	0	25	33	33	100

Supplemental Table S4 Instrument application by fracture type (18 instruments)

Fracture type	Most frequent PROM (% studies)	2nd most frequent (% studies)	3rd most frequent (% studies)	Key pattern
Distal radius (n=24)	QuickDASH (75%)	PRWE (33%)	PODCI (25%)	PRWE is more common in distal radial fractures
Supracondylar humerus (n=18)	QuickDASH (67%)	PROMIS-UE (39%)	PODCI (28%)	Highest PROMIS-UE use
Proximal humerus (n=8)	QuickDASH (75%)	PROMIS-UE (38%)	PODCI (25%)	Limited shoulder-specific measures
Forearm shaft (n=9)	QuickDASH (67%)	PROMIS-UE (44%)	PODCI (33%)	Supination/pronation underassessed
Hand (n=4)	QuickDASH (50%)	ABILHAND-Kids (50%)	PedsQL (25%)	Bimanual function importance
Mixed fractures (n=7)	PODCI (57%)	QuickDASH (43%)	PedsQL (43%)	PODCI preferred for multifracture

Percentages reflect the proportion of studies within each fracture type using the specified PROM

Supplemental Table S5 Instrument application by age group (18 instruments)

Age Group	QuickDASH use (%)	PROMIS-UE use (%)	PODCI use (%)	PedsQL use (%)	Parent-proxy use (%)
6-9 (or 5-8) years (n=18)	22	28	39	44	78
10-13 (or 9-12) years (n=31)	48	19	16	19	52
14-17 (or 13-16) years (n=19)	58	18	11	16	26

Supplemental Table S6 Instrument application by study design (18 instruments)

Study design	QuickDASH (%)	PROMIS-UE (%)	PODCI (%)	PedsQL (%)	PROOF-UE (%)	Other/varied (%)
RCT (n=11)	64	27	18	18	0	27
Prospective cohort (n=19)	42	21	16	16	11	32
Retrospective case series (n=31)	35	16	16	10	0	35
Cross-sectional survey (n=5)	20	60	40	0	0	20

Percentages reflect the proportion of studies within each design type using the specified PROM.