The landscape of functional outcome measures used in patients with bone sarcoma of the lower extremity or pelvis: a systematic review

Tigelaar, L.G.; Blom, K.J.; Bekkering, W.P.; Bramer, J.A.M.; Schreuder, H.W.B.; Furtado, S.; ... ; EuroEwing Consortium

Journal of Surgical Oncology

WILEY

Journal of SURGICAL ONCOLOGY

REVIEW ARTICLE

The landscape of Functional Outcome Measures Used in Patients with Bone Sarcoma of the Lower Extremity or Pelvis: A Systematic Review

Leonie G. Tigelaar[1] | Kiki J. Blom[1,2] | Willem P. Bekkering[1,3] | Jos A. M. Bramer[1,2,4] | Hendrik W. B. Schreuder[1,5] | Sherron Furtado[6] | Craig H. Gerrand[6] | Emanuela Palmerini[7] | Nathalie Gaspar[8] | Sandra J. Strauss[9] | Michiel A. J. Van de Sande[1,10] | Johannes H. M. Merks[1,11] | Lianne M. Haveman[1] | on behalf of the FOSTER Consortium and the EuroEwing Consortium

[1]Department of Solid Tumors, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands | [2]Department of Orthopedic Surgery, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands | [3]Department of Quality of Life, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands | [4]Department of Orthopedic Surgery and Sports Medicine, Amsterdam Movement Sciences, Cancer Center Amsterdam, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands | [5]Department of Orthopedic Surgery, Radboud University Medical Center, Nijmegen, The Netherlands | [6]Department of Orthopedic Oncology, Royal National Orthopedic Hospital NHS Trust, Stanmore, UK | [7]Department of Medical Oncology, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy | [8]Department of Oncology for Child and Adolescent, Gustave Roussy Cancer Campus, Villejuif, France | [9]Department of Oncology, University College London Cancer Institute, London, UK | [10]Department of Orthopedic Surgery, Leiden University Medical Center, Leiden, The Netherlands | [11]Division of Imaging and Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Correspondence: Lianne M. Haveman (l.m.haveman-3@prinsesmaximacentrum.nl)

Received: 30 August 2024 | **Accepted:** 4 September 2024

Funding: The FOSTER consortium is supported by the "Enfants Cancers Sante (ECS)" and the "Societe Française de Lutte contre les Cancers et les Leucemies de L'Enfant et de L'Adolescent (SFCE)." The EuroEwing Consortium is supported by the Bone Cancer Research Trust.

Keywords: functional status | International Classification of Functioning, Disability and Health | neoplasms, bone tissue | outcome assessment, health care

ABSTRACT

This systematic review provides a structured overview of the measurement instruments of functional outcome used in lower extremity and pelvic bone sarcoma patients. We identified 42 unique instruments covering 18 distinct functional outcome constructs with most studies measuring constructs within the activity domain of the International Classification of Functioning, disability, and health. The MusculoSkeletal Tumor Society 1993 and 1987 score, Toronto Extremity Salvage Score, and range of motion instruments were the measurement instruments most commonly used.

1 | Introduction

For patients with bone sarcoma, local therapy (i.e., surgery and/or radiation therapy) is a crucial part of their treatment. However, it may also impose life-changing consequences, particularly affecting the patient's functional abilities [1–3]. To enhance the functional outcome of these patients, it is essential to comprehensively evaluate their functional status using optimal measurement instruments. By thoroughly assessing all relevant aspects of a patient's functional abilities, healthcare professionals, together with patients, can make better informed decisions about their personalized treatment. Moreover, consistently measured inferior functional outcome associated with a particular local therapy approach should prompt a

Leonie G. Tigelaar and Kiki J. Blom contributed equally to this paper.

Institution of study: Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands.

reassessment and potential adjustment of guidelines when alternative approaches are available.

The World Health Organization (WHO) offers a structured approach to assess the various aspects of functional outcome through its International Classification of Functioning, Disability and Health [4]. The framework encompasses three key domains of functioning: (1) body structures and functions, focusing on the physical aspects of the body including anatomical parts and physiological functions; (2) activity, referring to the individual's ability to perform specific tasks or actions; and (3) participation, considering an individual's involvement in life situations such as participation in sports or social activities. By utilizing this framework, healthcare professionals can comprehensively and systematically map a patient's functional outcome.

In daily practice, the challenge lies in adequately measuring these functional outcome domains. Not only are there many instruments available, there is also disparity in preferences regarding the type of measurement instruments to use, whether it is patient-reported outcome measures (PROMs), clinician reported outcome measures (ClinROMs), or performance-based outcome measures (PerBOMs). The diversity in measurement instruments and approaches results in a lack of consensus on the most appropriate instruments complicating the comparison of findings in the literature [5–7]. This is especially troublesome in the field of bone sarcoma, where reported patient cohorts tend to be relatively small, and comparison of outcomes between studies is essential for gaining a comprehensive understanding of functional outcome and associated risk factors.

Previous reviews have intended to provide an overview of functional outcome measurement instruments used in the bone sarcoma population. A review by Pakulis et al. focused on PROMs and ClinROMs for pediatric patients and Furtado et al. looked at PerBOMs of functional outcome [7, 8]. While both reviews provide valuable information when looking for functional outcome measurement instruments for this population, a comprehensive overview of all instrument types for the various ages is lacking. Moreover, both reviews did not use a framework to structure their findings.

This study aims to identify the existing measurement instruments used to assess functional outcome after treatment for lower extremity or pelvic bone sarcoma. The objectives of this review are to: (1) provide an overview of the measurement instruments utilized structured by the ICF domains, (2) identify the most commonly used measurement instruments, (3) describe the basic characteristics per instrument such as assessed construct, type of report, number of items and language translations, and (4) describe characteristics of studies assessing functional outcome in lower extremity and pelvic bone sarcoma patients. The results will inform clinicians and researchers in selecting instruments for standardized measurement sets and upcoming European bone sarcoma trials. Importantly, results are to be implemented in the upcoming European trial within the FOSTER (Fight Osteosarcoma Through European Research) consortium and could also be considered for use in future EuroEwing Consortium (EEC) trials.

2 | Methods

The study protocol was registered in PROSPERO (CRD42022353668).

2.1 | Search Strategy

The study was conducted by performing systematic searches in the electronic databases MEDLINE (via PubMed) and Embase. The search was performed on October 31, 2023. The search was centered around three primary themes: bone sarcoma, lower extremity or pelvic localization, and functional outcome. The complete search strategy can be found in Supporting Infomation S1: Table S1. Articles were considered eligible for inclusion if they met the following criteria: (1) a minimum of 20 bone sarcoma patients with a tumor located in the lower extremity or pelvis to ensure sufficient representation of the population of interest; (2) data collected after local therapy (i.e., surgery or radiation therapy); (3) measurement instrument assessed functional outcome as main concept of interest; (4) original research study published in peer-reviewed journal; and (5) article published in English. In case of multiple articles covering the same cohort including the same measurement instruments, the largest study was included. Studies were excluded if they involved animal studies, case reports, reviews or trial protocols, or if the study evaluated measurement properties of a functional outcome instrument. Additionally, studies were excluded if the assessment of functional outcome was conducted through items that were study-specific and did not contribute to the formation of a broadly applicable instrument.

2.2 | Selection Process

The results of the database searches were merged and duplicate entries were removed. To assess eligibility, two independent reviewers (K.J.B. and L.G.T.) screened the articles. When the title and abstract suggested that an article might meet the inclusion criteria, full text was obtained. Any disagreements regarding study eligibility were resolved through discussion until a consensus was reached. In cases of persistent disagreement, a third reviewer (L.M.H.) decided.

2.3 | Data Extraction

The data extraction process was conducted by two reviewers (K.J.B. and L.G.T.). Extracted data were compared to ensure accuracy and reliability. Data from studies meeting the inclusion criteria were systematically tabulated using Microsoft Excel. Information was collected on title, author, year of publication, country, number of study subjects, number of bone sarcoma patients, age at evaluation (mean, median, range), time since local therapy (mean, median, range), and used measurement instrument(s). Additionally, age at evaluation was scored as a categorical variable distinguishing between patients < 18 years and ≥ 18 years old. If age range was not explicitly reported, we made an effort to categorize patients by estimating

their age at evaluation based on age at diagnosis or surgery and the duration of the follow-up.

We adopted the World Health Organization's (WHO) International Classification of Functioning, Disability, and Health (ICF) as a foundational structure for organizing the various measurement instruments under investigation. Measurement instruments were classified according to the ICF based on the construct for which the instrument was originally developed. It is important to note that some instruments are used to assess a broader range of constructs beyond their initial development scope. For example, the 6-Minute Walk Test was developed to assess endurance (ICF body functions domain), but is also regularly used to assess walking capabilities (ICF activity domain). Subsequently, a systematic categorization process was employed, resulting in the creation of seven distinct categories that corresponded to the ICF domains and the type of instrument: (1) "Body functions and structures—PerBOMs," (2) "Body functions and structures—ClinROMs," (3) "Body functions and structures—PROMs," (4) "Activities—PerBOMs," (5) "Activities—ClinROMs," (6) "Activities—PROMs," and (7) "Participation—PROMs."

Some of the measurement instruments were grouped because of assessment of the same construct using similar techniques, albeith with different instrument brands. For example, gait analyses were executed by devices of five different brands, but they were all grouped as "gait analysis." The same was performed for range of motion devices, isokinetic dynamometers, energy expenditure systems, and wearables (Supporting Information S1: Table S2).

Data visualization was accomplished using the programming tool R [9].

2.4 | Basic Measurement Instrument Characteristics

Basic characteristics of the measurement instruments, such as the assessed construct, number of items, unit of measurement, age specifications, and language translations, were collected from a variety of resources, including development articles, translation articles, manuals, measurement instrument websites, and measurement instrument databases. Translations were documented if an article was identified describing the translation process with or without validation of the translation.

3 | Results

Our search strategy initially yielded 4986 publications after removal of duplicates (Figure 1). After exclusion based on title and abstract, 840 articles were assessed for full text screening. Ultimately, we extracted data from a total of 347 studies for this review.

3.1 | Study Characteristics

Study cohorts were generally small, with 80% (277/347) of studies including less than 100 subjects. Studies were characterized by heterogeneity (Table 1). Most study cohorts (70%, 244/347) included either mixed type of tumors (i.e., both bone sarcoma and non-bone sarcoma patients), mixed localization of tumors, or both. Non-bone sarcoma patients most often consisted of soft-tissue sarcoma and/or benign bone tumor patients. The age range of patients at functional outcome evaluation was available in only 12% (41/347) of the included studies. However, we were able to categorically score age at evaluation for 71% (248/347) of the studies, with 4 studies including only pediatric

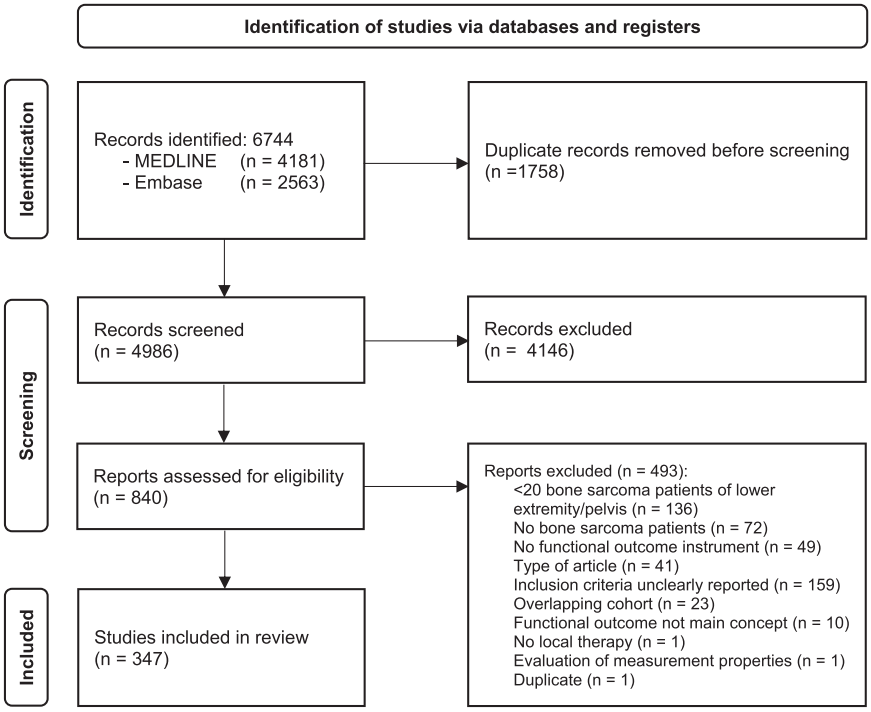

FIGURE 1 | PRISMA flow diagram.

TABLE 1 | Descriptives of the included studies.

	Median (Q1–Q3)	Range	$N_{studies}$ (%)
Subjects per study (*N*)	52 (33–90)	20–1667	347 (100)
Bone sarcoma patients per study (*N*)	41 (28–69)	20–618	333 (96)
Age at evaluation (years)	42 (19–60)	12–76	41 (12)
Follow-up (years)	12 (8–18)	1–62	263 (76)
Measurement instruments per study (*N*)	1 (1–2)	1–8	347 (100)

Note: $N_{studies}$ = number of studies for which data were available; *N* = number.

patients, 52 including only adult patients, and 192 studies combining pediatric and adult patients in their cohort. Additionally, for 64% (168/263) of studies patient follow-up ranged 10 years or more. Due to high inconsistency in presentation of age at evaluation and follow-up, we did not pool this data.

3.2 | Measurement Instruments Used

Included studies measured functional outcome 517 times using 42 unique instruments (Figure 2). The MusculoSkeletal Tumor Society Score 1993 (MSTS93) was the measurement instrument most frequently used (52%, 268/517). Other instruments frequently found were the Toronto Extremity Salvage Score (TESS) (11%, 59/517), range of motion (ROM) instruments (9%, 46/517), MusculoSkeletal Tumor Society Score 1987 (MSTS87) (5%, 28/517), and wearables (2%, 8/517). Studies measured functional outcome mostly within the activity domain (81%, 421/517), followed by body function and structures (18%, 94/517) and participation (0.4%, 2/517) domains. With regard to the type of instruments, ClinROMs (66%, 341/517) were used most often, followed by PerBOMs (20%, 105/517) and PROMs (14%, 71/517).

3.3 | Measurement Instrument Characteristics

Twenty measurement instruments assessed functional outcome within the body functions and structures domain, 21 within the activity domain, and 1 within the participation domain. Across all ICF domains, 18 distinct constructs of functional outcome were identified.

Among the measurement instruments identified, only the MSTS87, MSTS93, TESS, and Functional Mobility Assessment (FMA) are bone sarcoma specific. The remaining instruments were developed either for the general population or other disease groups. Regarding age applicability, PerBOMs often lack a clearly defined age range. Their suitability for children and elderly will mostly depend on individual capabilities. In contrast, PROMs and ClinROMs are typically tailored to specific age groups, with most designed for adults.

PROMs consist of a median of 14 items, ranging from 1 to 30 items (Table 2). While predominantly developed in English, some have undergone extensive translation, reaching up to 16 available languages for the TESS, 26 for the Western Ontario McMaster Universities Osteoarthritis Index, and 31 for the Oxford Knee Score. ClinROMs were typically shorter with a median of six items (range 1–14) and generally translated in less languages than PROMs with a maximum of eight languages for the Harris Hip Score and MSTS93.

4 | Discussion

This systematic review provides a structured overview of measurement instruments used in patients with lower extremity or pelvic bone sarcoma aiming to facilitate clinicians and researchers in the selection of appropriate instruments. By far the most commonly used measurement instrument is the MSTS93. This 6-item ClinROM is an easy to use instrument without imposing a burden on the patient. There is, however, a growing preference for using PROMs in assessing outcomes, as is illustrated by the incorporation of patient-reported outcome endpoints in drug approval processes [10, 11]. This preference aligns with the evolving shift towards patient-centered care and research, and the observed discrepancies between patient and clinician reports [12, 13]. The most frequently used PROM we found is the TESS: a 30-item questionnaire designed specifically to assess physical disability in patients with bone and soft tissue sarcomas.

Range of motion instruments were also commonly used which can be explained by the notable impact of local therapy on joint mobility. Although muscle strength is also frequently affected, its measurement was less frequent. This discrepancy may stem from the uncertainty regarding its overall impact on functioning, given patients' adaptive use of compensation mechanisms. Among the other PerBOMs identified are wearables, gait analysis instruments, isokinetic dynamometers, and energy expenditure systems. These are relatively complex instruments compared to PerBOMs like the MRC scale, 6-Minute Walk Test, Timed Up and Down Stairs and Timed Up and Go. Because of their complexity, particularly in terms of interpretability, availability, and costs, they may be less suitable for inclusion in a standardized measurement set and often remain reserved for in-depth analysis.

The ICF's activity domain of functional outcome was most frequently measured, mainly by using PROMs and ClinROMs. This is probably due to wide availability and familiarity of certain PROMs and ClinROMs, such as the MSTS93 and TESS, their low costs, and clear link of the individual items to everyday functioning. Measurement of the ICF's participation domain of functional outcome, however, was uncommon with only one instrument found in two studies. Whiteneck et al. attribute this shortage of participation measurement instruments to the

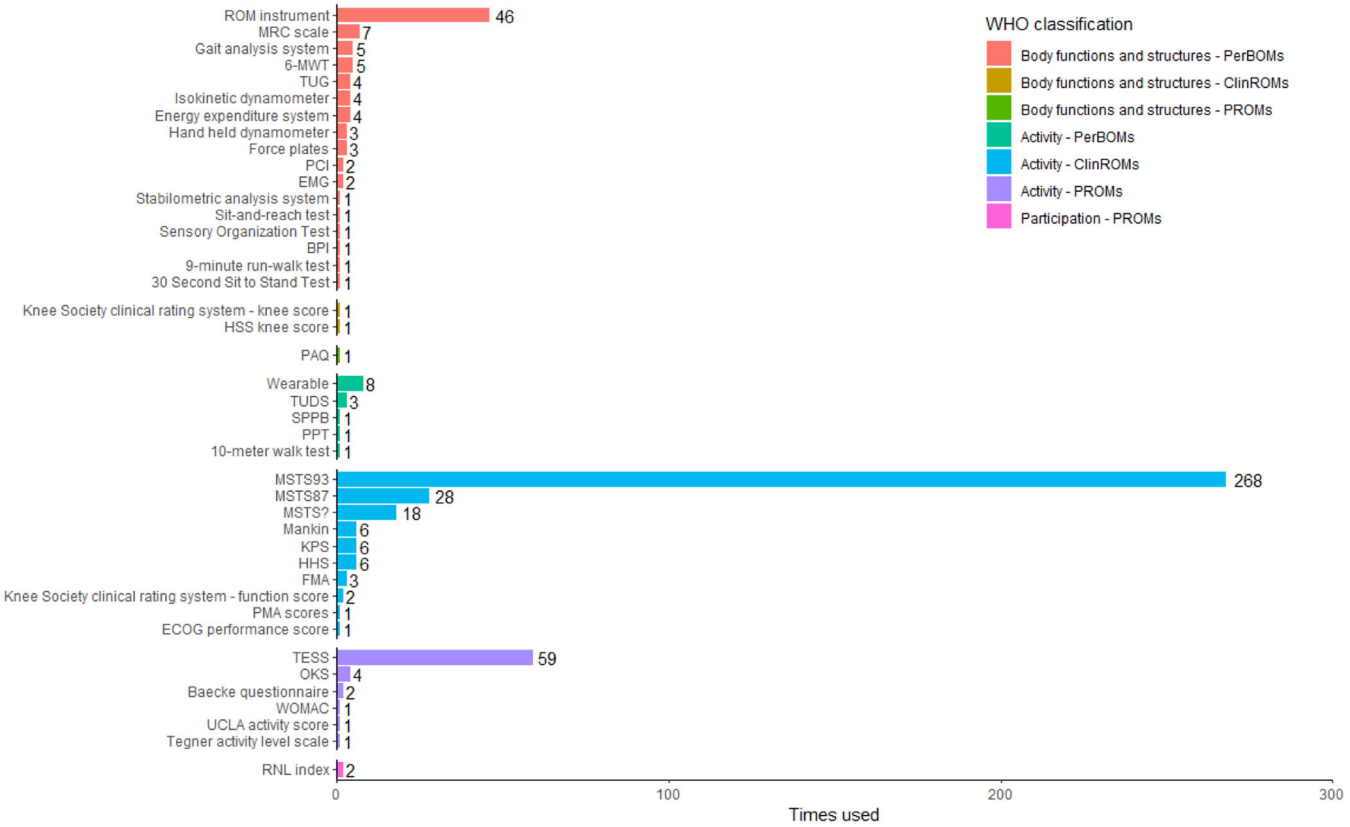

FIGURE 2 | Measurement instruments assessing functional outcome in bone sarcoma literature, including frequency of use. 6-MWT = 6-Minute Walk Test; BPI = Blackburne-Peel Index; ECOG performance score = Eastern Cooperative Oncology Group performance score; EMG = electromyography; FMA = Functional Mobility Assessment; HHS = Harris Hip Score; KPS = Karnofsky Performance Scale; Mankin = Mankin score; MRC = Medical Research Council scale; MSTS? = MusculoSkeletal Tumor Society Score, version unknown; MSTS87 = MusculoSkeletal Tumor Society Score 1987; MSTS93 = MusculoSkeletal Tumor Society Score 1993; OKS = Oxford Knee Score; PAQ = Physical Activity Questionnaire; PCI = Physiological Cost Index; PMA scores = Postel-Merle d'Aubigné scores; PPT = Physical Performance Test; RNL = Reintegration to Normal Living index; ROM = range of motion; SPPB = Short Physical Performance Battery; TESS = Toronto Extremity Salvage Score; TUDS = Timed Up and Down Stairs; TUG = Timed Up and Go; UCLA sports activity score = University of California Los Angeles sports activity score; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

indistinctness in defining and explaining the concept of participation by the ICF [14]. As a result, the content of a measurement instrument for this domain remains unclear, leading to limited acceptance of the few instruments available and little progress in developing new ones. This is unfortunate since we believe that the participation domain is probably the most crucial domain for a patient's experience of functioning in daily living and therefore plays a pivotal role in decision making processes in various healthcare fields including orthopedic care, rehabilitation, and physical therapy. Despite the scarcity of participation measurement instruments we found in our review, there are some options available, for example, assessment of life habits (LIFE-H), Participation Measure for PostAcute Care (PM-PAC), or the Utrecht Scale for Evaluation of Rehabilitation—Participation (USER-P) [15, 16].

Functional outcome is a broad concept as is shown by the many distinct constructs that are assessed by the measurement instruments identified. When selecting a measurement instrument for a standardized outcome set or trial, it is crucial to consider which constructs are relevant to the population and outcomes of interest to ensure collection of meaningful data.

Exploring the basic population characteristics of the included studies, cohorts were heterogeneous with mixed tumor types and wide ranges of age at evaluation and duration of follow-up. This is partially understandable since large bone sarcoma cohorts are relatively difficult to assemble. However, it underlines the importance of adequate description of cohort characteristics and standardized measurement of outcomes to be able to compare outcomes across the cohorts.

One of the limitations of this review is the exclusive focus on measurement instruments currently used within the lower extremity and pelvic bone sarcoma population. Therefore, the list presented may not encompass all the instruments suitable for this patient group. There may be other measurement instruments, yet to be employed in this population, that could hold value. Another limitation encompasses the exclusion of instruments that contain a subscale measuring functional outcome within a broader concept, for example, quality of life. These instruments can also be of value depending on the (research) question to be answered.

We want to emphasize that researchers and clinicians should always consider measurement properties such as content

TABLE 2 | Properties of identified measurement instruments of functional outcome.

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
Body function & structures	LorAN Engineering Srl- EPS-R1 system board	Balance	General	PerBOM	—	Speed of center mass (mm/s)	Children and adults[b]	—	—
	Sensory organization test	Balance	General	PerBOM	—	Equilibrium score (%) Strategy score (%)	≥ 3	—	—
	TUG	Dynamic balance	Elderly	PerBOM	—	s	≥ 3	—	—
	Fourteen-channel electromyographic amplifier (Fritz Schwarzer)	Electrical muscle activity	General	PerBOM	—	mV	All	—	—
	MA-300 EMG (Motion Lab Systems)	Electrical muscle activity	General	PerBOM	—	mV	All	—	—
	6-MWT	Endurance	Chronic respiratory disease	PerBOM	—	m	≥ 3	—	—
	9-min walk/run test	Endurance	General	PerBOM	—	m	Children and adults[b]	—	—
	Cosmed K4b^2 system	Energy expenditure	General	PerBOM	—	vO_2, vCO_2, mL air_{exp}	Children and adults[b]	—	—
	Douglas bag technique	Energy expenditure	General	PerBOM	—	vO_2, vCO_2, mL air_{exp}	Children and adults[b]	—	—
	Oxycon-4 (OX-4)	Energy expenditure	General	PerBOM	—	vO_2, vCO_2, mL air_{exp}	Children and adults[b]	—	—
	PAQ	Energy expenditure	General	PROM	16	METs	12–150	English	—
	PCI	Energy expenditure	General	PerBOM	—	Heart beats per meter	≥ 5	—	—
	Sit-and-reach test	Flexibility	General	PerBOM	—	cm	≥ 5	—	—

(Continues)

TABLE 2 | (Continued)

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
	BPI	Joint anatomy (patellar height)	Dislocation and chondromalacia of the patella	PerBOM	—	cm	Adolescents and adults	—	—
	HSS knee score	Joint function	Disease-specific: Total knee arthroplasty	ClinROM	7	Total score 0–100	Adults	English	Chinese, Turkish
	Knee Society clinical rating system—Knee score	Joint function	Disease-specific: Total knee arthroplasty	ClinROM	7	Total score 0–100	Adults	English	—[c]
	Goniometer	Joint mobility	General	PerBOM	—	Degrees	All	—	—
	NORATRACK 360 electrogoniometer	Joint mobility	General	PerBOM	—	Degrees	All	—	—
	Force plates AMTI	Gait pattern	General	PerBOM	—	Ground reaction force (N)	Children and adults[b]	—	—
	Force plates Hentschel system	Gait pattern	General	PerBOM	—	Ground reaction force (N)	Children and adults[b]	—	—
	Force plates Kistler	Gait pattern	General	PerBOM	—	Ground reaction force (N)	Children and adults[b]	—	—
	Falcon CCD cameras (Motion Analysis Corporation)	Gait pattern	General	PerBOM	—	Amongst others[d]: velocity (m/s), step width (cm), step length (cm), gait cycle time (s)	Children and adults[b]	—	—
	GaitMatTMII	Gait pattern	General	PerBOM	—	Amongst others[d]: velocity (m/s), step width (cm), step length (cm), gait cycle time (s)	Children and adults[b]	—	—
	Passive marker system with four infrared cameras (Motion Analysis Corporation)	Gait pattern	General	PerBOM	—	Amongst others[d]: velocity (m/s), step width (cm), step length (cm), gait cycle time (s)	Children and adults[b]	—	—

(Continues)

TABLE 2 | (Continued)

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
	VA-Rancho Footswitch Stride Analyzer	Gait pattern	General	PerBOM	—	Amongst others[d]: velocity (m/s), step width (cm), step length (cm), gait cycle time (s)	Children and adults[b]	—	—
	Vicon 612	Gait pattern	General	PerBOM	—	Amongst others[d]: velocity (m/s), step width (cm), step length (cm), gait cycle time (s)	Children and adults[b]	—	—
	30 s Sit to Stand Test	Muscle strength & endurance	Elderly	PerBOM	—	No.	Children and adults[b]	—	—
	Biodex System 3.0	Muscle strength	General	PerBOM	—	Nm	Children and adults[b]	—	—
	Biodex System 4 Pro	Muscle strength	General	PerBOM	—	Nm	Children and adults[b]	—	—
	Cybex 6000	Muscle strength	General	PerBOM	—	Nm	Children and adults[b]	—	—
	Cybex II	Muscle strength	General	PerBOM	—	Nm	Children and adults[b]	—	—
	Hand held dynamometer	Muscle strength	General	PerBOM	—	N	Children and adults[b]	—	—
	MRC scale	Muscle strength	Nerve injury	PerBOM	—	Score 0–5	All	—	—
Activity	ECOG performance score	ADL ability	Cancer	ClinROM	1	Grade 0–5	Adults	English	Turkish
	KPS	ADL ability	Cancer	ClinROM	1	Score 0–10	Children and adults[b]	English	Turkish
	Knee Society clinical rating system—Function score	Functional ability	Total knee arthroplasty	ClinROM	3	Total score 0–100	Adults	English	—[c]

(Continues)

TABLE 2 | (Continued)

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
	Mankin Score	Functional ability	Bone tumor	ClinROM	1	4-point ordinal scale (Excellent–Good–Fair–Failure)	Children and adults[b]	English	—
	MSTS 1987 score	Functional ability	Bone and soft tissue sarcoma	ClinROM	7	Total score 0–100	Children and adults[b]	English	Greek
	MSTS 1993 score	Functional ability	Bone and soft tissue sarcoma	ClinROM	6	Total score 0–100	Children and adults[b]	English	Brazilian Portuguese, Chinese, Danish, Finnish, French, German, Italian, Japanese, Turkish, Romanian
	PMA scores	Functional ability	Hip arthroplasty	ClinROM	3	Total score 0–18	Adults	French	—
	WOMAC	Functional ability	Hip and knee osteoarthritis	PROM	24	Total score 0–96	Adults	English	Arabic, Bengali, Brazilian Portuguese, Chinese (simplified), Dutch, Finnish, French Canadian, German, Greek, Gujarati, Hebrew, Hindi, Indonesian, Italian, Korean, Malay, Marathi, Moroccan Arabic, Nepali, Persian, Spanish, Swedish, Thai, Turkish, Yoruba[e]

(Continues)

TABLE 2 | (Continued)

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
	OKS	Functional disability	Total knee replacement	PROM	12	Total score 0–48	Adults	English	Arabic, Chinese, Danish, Dutch, Estonian, Finnish, French, German, Greek, Gujarati, Hindi, Hungarian, Indonesian, Italian, Japanese, Korean, Malay, Marathi, Persian, Polish, Portuguese, Punjabi, Russian, Serbian, Spanish, Swedish, Tamil, Thai, Turkish, Welsh
	10-m walk test	Functional mobility	General	PerBOM	—	s	≥ 2	—	—
	FMA	Functional mobility	Lower extremity sarcoma	ClinROM	14	Total score 0–70	Children and adults[b]	English	—
	HHS	Functional mobility	Total hip athroplasty	ClinROM	10	Total score 0–100	Children and adults[b]	English	Arabic, Brazilian Portuguese, French, Italian, Persian, Slovenian, Turkish
	PPT	Functional mobility	Elderly	PerBOM	9	Total score 0–36	Adults	—	—
	SPPB	Functional mobility	Elderly	PerBOM	5	Total score 0–12	Adults	—	—

(Continues)

TABLE 2 | (Continued)

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
	TUDS	Functional mobility	Children with cerebral palsy	PerBOM	—	s	≥ 6	—	—
	ActiLog V3.0	Physical activity	General	PerBOM	—	Accelerations (m/s^2) > can be translated into a GPA score[f] (amongst others)	Children and adults[b]	—	—
	Baecke questionnaire	Physical activity	General	PROM	16	Total score 3–15	Adults	Dutch	English, Japanese, Persian, Brazilian Portuguese
	Dynaport ADL monitor	Physical activity	General	PerBOM	—	Accelerations (m/s^2) > can be translated to time spent in different activity categories, movement intensity, and a PA[g] I.	Children and adults[b]	—	—
	Omron Health Counter HJ-5	Physical activity	General	PerBOM	—	No. of steps	Children and adults[b]	—	—
	Step Activity Monitor (Cyma)	Physical activity	General	PerBOM	—	No. of steps	≥ 5	—	—
	Step-Watch Activity Monitor (Orthocare Innovations)	Physical activity	General	PerBOM	—	No. of steps	≥ 5	—	—
	Tegner activity level scale	Physical activity	Knee ligament injury	PROM	1	Score 0–10	Children and adults[b]	Swedish	Arabic, Chinese (simplified), Danish, Dutch, English, German, Greek, Indonesian, Italian, Persian, Thai

(Continues)

TABLE 2 | (Continued)

WHO ICF domain	Measurement instrument	Construct	Development target population	Type	No. of items	Measurement unit	Age (years)	Development language	Translations[a]
	UCLA activity score	Physical activity	Joint arthroplasty	ClinROM	1	Score 1–10	Children and adults[b]	English	Chinese (simplified), Danish, Italian, Persian
	TESS – Lower extremity	Physical disability	Bone and soft tissue sarcoma	PROM	30	Total score 0–100	≥ 12	English	Arabic, Brazilian Portuguese, Chinese, Danish, Dutch, Finnish, French, German, Greek, Italian, Japanese, Korean, Mexican Spanish, Norwegian, Swedish
Participation	RNL index	Reintegration to normal living	Incapacitating illness or severe trauma	PROM	11	Total score 0–100	Adults	English & French	Chinese, German, Igbo, Korean, Spanish, Turkish

Abbreviations: 6-MWT = 6-Minute Walk Test, ADL = Activities of Daily Living, BPI = Blackburne-Peel Index, ClinROM = Clinician-Reported Outcome Measure, cm = centimeter, ECOG = Eastern Cooperative Oncology Group, FMA = Functional Mobility Assessment, GPA = General Physical Activity, HHS = Harris Hip Score, HSS = Hospital for Special Surgery, KPS = Karnofsky Performance Scale, MET = Metabolic Equivalent of Task, mL air_{exp} = volume of expired air, MRC = Medical Research Council, MSTS = MusculoSkeletal Tumor Society, mV = millivolt, m/s = meter per second, N = Newton, Nm = Newton meter, No. = number, OKS = Oxford Knee Score, PAI = Physical Activity Index, PAQ = Physical Activity Questionnaire, PCI = Physiological Cost Index, PerBOM = Performance-Based Outcome Measure, PMA = Postel-Merle d'Aubigné, PPT = Physical Performance Test, PROM = Patient-Reported Outcome Measure, RNL = Reintegration to Normal Living, s = seconds, SPPB = Short Physical Performance Battery, TESS = Toronto Extremity Salvage Score, TUDS = Timed Up and Down Stairs, TUG = Timed Up and Go, UCLA = University of California Los Angeles, vCO_2 = difference between the volumes of expired and inspired carbon dioxide, vO_2 = difference between the volumes of inspired and expired oxygen, WHO ICF = World Health Organization International Classification of Functioning, disability and health, WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

[a]Translations are included when an article was found dedicated to translate (with or without validation of) the measurement instrument.

[b]No clear age range defined for usage of this measurement instrument.

[c]A new version of the Knee Society clinical rating system ("The 2011 Knee Society Knee Scoring System") was developed in 2011. This version has been translated into several languages (amongst others: Brazilian Portuguese, Dutch, French, German, Japanese, Korean, Persian, Spanish, and Turkish).

[d]Gait analysis systems often provide an extensive range of possible outcome. Please refer to the company of the gait analysis system for details on the possibilities.

[e]According to the website https://www.womac.com/womac/index.php, over 100 translations of the WOMAC are available. Only the languages for which an article was found on translating (with or without validating) the measurement instrument are included in the table.

[f]Average number of accelerations per 5 min period throughout the day.

[g]Index ranging from 1 to 6 with lying, sitting and standing being rated as 0, 1 and 2, respectively.

validity, construct validity, and reliability in their selection process, rather than solely relying on the frequency of use and basic measurement instrument properties. This assessment is essential to make an informed selection of the most suitable measurement instruments. Therefore, as a next step, we will evaluate the quality of the PROMs and ClinROMs identified in this review following the COSMIN (COnsensus-based Standards for the selection of health Measurement Instruments) guidelines [17, 18]. These guidelines offer a systematic approach to evaluating measurement properties. The results of this evaluation will be reported separately.

5 | Conclusion

A variety of measurement instruments assessing functional outcome in patients with a bone sarcoma of the lower extremity or pelvis were identified, with the MSTS93 score and TESS most commonly used. The provided structured overview of the measurement instruments including their basic properties shows the variety of constructs that can be assessed across the ICF domains, all within the scope of functional outcome. Together with the different types of report, distinct applicability for specific populations (e.g., children), available language translations and measurement properties, it is crucial to carefully weigh which measurement instrument to select for use in clinical practice or trials.

This review serves as an initial step to guide the adequate selection of functional outcome instruments for the lower extremity and pelvic bone sarcoma population. As a next step, assessment of the quality of the individual instruments needs to be conducted to complement the findings of this review and provide definite guidance on instrument selection.

Acknowledgments

We acknowledge Esperanza Perez as the FOSTER consortium project manager. The FOSTER consortium is supported by the "Enfants Cancers Sante (ECS)" and the "Societe Française de Lutte contre les Cancers et les Leucemies de L'Enfant et de L'Adolescent (SFCE)." The EuroEwing Consortium is supported by the Bone Cancer Research Trust.

Data Availability Statement

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

References

1. M. W. Bishop, K. K. Ness, C. Li, et al., "Cumulative Burden of Chronic Health Conditions in Adult Survivors of Osteosarcoma and Ewing Sarcoma: A Report From the St. Jude Lifetime Cohort Study," *Cancer Epidemiology, Biomarkers & Prevention* 29 (2020): 1627–1638.

2. A. R. Cooper, B. Wilke, M. Scarborough, C. P. Gibbs, and A. Spiguel, "Pediatric Sarcoma Patients With Worse Physical Function but Better Peer Relationships and Depressive Symptoms Than U.S. Pediatric Population as Measured by PROMIS®," *Cureus* 12 (2020): e7040.

3. L. Fernandes, C. E. Holm, A. Villadsen, M. S. Sørensen, M. K. Zebis, and M. M. Petersen, "Clinically Important Reductions in Physical Function and Quality of Life in Adults With Tumor Prostheses in the Hip and Knee: A Cross-Sectional Study," *Clinical Orthopaedics & Related Research* 479 (2021): 2306–2319.

4. World Health Organization, *International Classification of Functioning, Disability and Health: ICF* (Geneva, Switzerland: World Health Organization, 2001), https://iris.who.int/handle/10665/42407.

5. W. P. Bekkering, T. P. M. Vliet Vlieland, M. Fiocco, et al., "Quality of Life, Functional Ability and Physical Activity After Different Surgical Interventions for Bone Cancer of the Leg: A Systematic Review," *Surgical Oncology* 21 (2012): e39–e47.

6. J. Stokke, L. Sung, A. Gupta, A. Lindberg, and A. R. Rosenberg, "Systematic Review and Meta-Analysis of Objective and Subjective Quality of Life Among Pediatric, Adolescent, and Young Adult Bone Tumor Survivors," *Pediatric Blood & Cancer* 62 (2015): 1616–1629.

7. S. Furtado, L. Errington, A. Godfrey, L. Rochester, and C. Gerrand, "Objective Clinical Measurement of Physical Functioning After Treatment for Lower Extremity Sarcoma—A Systematic Review," *European Journal of Surgical Oncology* 43 (2017): 968–993.

8. P. J. Pakulis, N. L. Young, and A. M. Davis, "Evaluating Physical Function in an Adolescent Bone Tumor Population," *Pediatric Blood & Cancer* 45 (2005): 635–643.

9. R Core Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2023), https://www.R-project.org/.

10. EMA, *Appendix 2 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man: The Use of Patient-Reported Outcome Measures in Oncology Studies* (European Medicines Agency, 2016).

11. FDA, *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims* (Food and Drug Administration, 2009).

12. C. Xiao, R. Polomano, and D. W. Bruner, "Comparison Between Patient-Reported and Clinician-Observed Symptoms in Oncology," *Cancer Nursing* 36 (2013): E1–E16.

13. E. K. Fromme, K. M. Eilers, M. Mori, Y. C. Hsieh, and T. M. Beer, "How Accurate Is Clinician Reporting of Chemotherapy Adverse Effects? A Comparison With Patient-Reported Symptoms From the Quality-Of-Life Questionnaire C30," *Journal of Clinical Oncology* 22 (2004): 3485–3490.

14. G. Whiteneck and M. P. Dijkers, "Difficult to Measure Constructs: Conceptual and Methodological Issues Concerning Participation and Environmental Factors," *Archives of Physical Medicine and Rehabilitation* 90 (2009): S22–S35.

15. L. Resnik and M. A. Plow, "Measuring Participation As Defined By the International Classification of Functioning, Disability and Health: An Evaluation of Existing Measures," *Archives of Physical Medicine and Rehabilitation* 90 (2009): 856–866.

16. M. W. M. Post, C. H. van der Zee, J. Hennink, C. G. Schafrat, J. M. A. Visser-Meily, and S. B. van Berlekom, "Validity of the Utrecht Scale for Evaluation of Rehabilitation-Participation," *Disability and Rehabilitation* 34 (2012): 478–485.

17. C. A. C. Prinsen, L. B. Mokkink, L. M. Bouter, et al., "COSMIN Guideline for Systematic Reviews of Patient-Reported Outcome Measures," *Quality of Life Research* 27 (2018): 1147–1157.

18. C. B. Terwee, C. A. C. Prinsen, A. Chiarotto, et al., "COSMIN Methodology for Evaluating the Content Validity of Patient-Reported Outcome Measures: A Delphi Study," *Quality of Life Research* 27 (2018): 1159–1170.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.