University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Fernando-Canavan, L;Abraham, P;Devlin, N;Tran-Duy, A

Title:
Health-related quality of life in patients with extremity bone sarcoma after surgical treatment: a systematic review

Date:
2023-12-11

Citation:
Fernando-Canavan, L., Abraham, P., Devlin, N. & Tran-Duy, A. (2023). Health-related quality of life in patients with extremity bone sarcoma after surgical treatment: a systematic review. Quality of Life Research, 33 (5), pp.1157-1174. https://doi.org/10.1007/s11136-023-03554-3.

Persistent Link:
https://hdl.handle.net/11343/357012

Quality of Life Research
https://doi.org/10.1007/s11136-023-03554-3

REVIEW

Health-related quality of life in patients with extremity bone sarcoma after surgical treatment: a systematic review

Liam Fernando-Canavan[1] · Patrick Abraham[1] · Nancy Devlin[1] · An Tran-Duy[1]

Accepted: 23 October 2023
© Springer Nature Switzerland AG 2023

Abstract

Purpose We conducted a systematic review of studies reporting on measurement of health-related quality of life (HRQoL), with a special focus on the use of the preference-weighted instruments, in patients with extremity bone sarcoma treated with limb-salvage surgery or amputation.

Methods We searched MedLine, Embase, Cochrane Library and Web of Science for English-language studies reporting on HRQoL of patients with bone sarcoma from inception to 28 August 2023. All records found were independently reviewed by two reviewers. We used the Newcastle–Ottawa Scale (NOS) and the CONSORT 2010 checklist to assess the quality of the cohort and randomised studies, respectively.

Results The search identified 1225 records, of which 16 studies were included for data extraction. Only one study used a preference-weighted instrument for measuring HRQoL in a small sample of patients ($n = 28$). Ten studies used the generic SF-36 questionnaire, but no preference-weighted HRQoL based on SF-6D was derived from the SF-36 scores. Most studies comparing HRQoL between amputation and limb-salvage surgery reported no significant differences. Twelve cohort studies scored six or more out of nine points based on the NOS. The only randomised study scored 54% on the CONSORT 2010 checklist.

Conclusions The approaches used to measure HRQoL were inconsistent and outcome scores varied substantially. Only one study used preference-weighted instruments for HRQoL measurement. Future research into the surgical treatment of extremity bone sarcoma should consider the use of preference-weighted instruments to measure HRQoL, which will therefore enable economic evaluation for the growing orthopaedic armamentarium of novel surgical interventions.

Registration This systematic review was registered with the PROSPERO International prospective register of systematic reviews (CRD42021282380).

Keywords Health-related quality of life · Health utility · Bone sarcoma · Limb salvage · Amputation · Surgery

Introduction

Bone sarcomas are rare cancers, with osteosarcoma, chondrosarcoma, and Ewing sarcoma being the most frequently diagnosed and usually located in the lower limb [1]. Standard treatment of the most common bone sarcoma, osteosarcoma, involves neoadjuvant chemotherapy and surgical resection of the tumour [2–4]. Amputation was once the standard surgical treatment for extremity tumours, but it is now reserved for non-resectable tumours with soft tissue and neuromuscular contamination that make it too difficult to salvage the limb [5]. Advances in imaging techniques and chemotherapy then led to a major shift away from amputation and toward limb-salvage surgery [2], with limb-salvage surgery currently being applicable to approximately 80–90% of patients with osteosarcoma [6–8].

During the last decade, there have been many advances in limb-salvage surgery, such as the use of computer-assisted surgery, 3D-printed patient-specific implants (PSI) and cutting-guides, and the combination of robotic-assisted surgery and a 3D-printed PSI [9–13]. These and other novel surgical approaches seek to improve long-term surgical outcomes and health-related quality of life (HRQoL) of patients. HRQoL is an important outcome measure in surgical oncology and has

✉ An Tran-Duy
an.tran@unimelb.edu.au

[1] Centre for Health Policy, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

been widely used in economic evaluation of new interventions, allowing patients' views about their own health and functioning to be taken into account in assessing treatment success.

With emerging interventions becoming available to orthopaedic oncologists, access to new technology, provided it is safe and effective, will depend on robust evidence that it is cost-effective. Payers in many countries such as the UK and Australia use Health Technology Assessment (HTA) processes to judge cost-effectiveness and ensure that increased costs are justified by increased benefits to patients. The most common measure of effectiveness in HTA is quality-adjusted life years (QALYs), a metric that combines both the quality and quantity of life lived into a single value and then allows for comparisons to be made across different disease areas [14]. To estimate QALYs, HRQoL is usually measured by asking patients to complete concise, generic health questionnaires which are accompanied by preference weights. These preference-weighted questionnaires represent multi-attribute health status classification systems in which each health state is associated with a pre-scored weight based on population's preference for that health state. HRQoL can also be measured in other ways, including the use of condition-specific HRQoL questionnaires either accompanied by or mapped to preference weights, or by directly seeking preferences for health state vignettes using time-consuming techniques such as standard gamble and time trade-off [15]. Well-known preference-weighted instruments for measuring HRQoL in a manner that allows QALY estimation include EQ-5D [16], Health Utility Index Mark 2 (HUI2) and Mark 3 (HUI3) [17], and SF-6D [18]. Different from the other instruments, the SF-6D is not a stand-alone questionnaire; it is derived from the SF-36 questionnaire [19], a widely used generic tool to measure health outcomes in eight dimensions (physical functioning, role limitations due to physical health problems, role limitations due to emotional problems, vitality (or energy/fatigue), mental health (or emotional well-being), social functioning, bodily pain, and general health perceptions), with higher scores indicating a more favourable health state in the respective dimension.

Little has been done to assess the extent to which preference-weighted instruments have been used to measure HRQoL in extremity bone sarcoma. The only systematic review of HRQoL studies in extremity bone sarcoma by Tang et al. [20] did not focus on the use of preference-weighted instruments. In addition, this study [20] was published over 10 years ago and used suboptimal search strategies based on only two key words ("sarcoma" and "quality of life"), which might not have captured studies that used only other variants of "sarcoma" such as "bone cancer", or studies that used terms for preference-weighted HRQoL instruments such as "HUI" [17] or "health utility" without mentioning "quality of life". Currently, it is unclear which instruments and preference weights or mapping methods have been used to capture HRQoL for surgical management of extremity bone sarcoma. In this study, we conducted a systematic review of studies reporting on measurement of HRQoL, with a special focus on the use of preference-weighted instruments, in patients with bone sarcoma in the extremities who underwent limb-salvage surgery or amputation.

Methods

The methods reported in this paper were based on the Cochrane Handbook for Systematic Reviews of Interventions [21]. The review was conducted and reported in accordance with the Preferred Items for Reporting in Systematic Reviews and Meta-Analyses (PRISMA) statement [22].

Eligibility criteria

We searched MedLine, Embase, Cochrane Library and Web of Science for English-language studies that reported on HRQoL of patients with bone sarcoma and published in peer-reviewed journals from inception to 28 August 2023. We included randomised controlled trials, observational studies, and case series that involved patients aged 15 or older and who were diagnosed with extremity bone sarcoma and underwent limb-salvage or amputation. We chose this cut-off point for age because patients with bone sarcoma aged 15 and older are usually treated using common guidelines for adults [23, 24]. By focusing on patients aged 15 or older, we reduced the heterogeneity of the studies which reported the same type of surgery, but might employ different treatment pathways which would influence HRQoL. We excluded case reports, studies where full text was not available, studies with cohorts less than ten participants, and studies involving soft tissue sarcoma, non-extremity sarcoma, or involving a mixed cohort with many types of cancer.

Search strategy

We used search terms related to surgical type (e.g., limb-salvage and amputation), the disease (e.g., sarcoma) and instruments used to measure HRQoL (e.g., SF-6D) to build the search queries. We also used terms related to instruments for functional outcome measurement because studies that report on functional outcomes may also include HRQoL measurement. Our search strategies were reviewed and validated by a senior librarian at the University of Melbourne. Detailed search terms and queries and a glossary of the abbreviated terms are provided in the Supplementary Tables S1–S4.

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Study selection and data extraction

All records found from the search were imported to Covidence [25] and independently reviewed in two stages by two reviewers (LFC and PA). In the first stage, titles and abstracts were screened using the eligibility criteria. Uncertainty was discussed between the two reviewers and disagreements referred to a third reviewer (AT-D or ND) for a resolution. In the second stage, full text of the studies included in the first stage were screened using the same method as in the first stage. The following information was independently extracted by two reviewers (LFC and PA) from the studies included after the full-text review: *year of publication, journal, country, study design, study cohort, sample size, sex, mean age at time of study, duration between surgery and HRQoL measurement, time frame of treatment, instrument for HRQoL measurement, HRQoL comparison, type of surgery, summary of findings* and *HRQoL scores in different dimensions*. Disagreements in extracted data were referred to a third reviewer (AT-D or ND) for resolution.

Study quality assessment

As all observational studies included for data extraction were cohort studies, we used the Newcastle–Ottawa Scale (NOS) designed for cohort studies to assess the quality of the these studies as recommended by the Cochrane Handbook for Systematic Reviews of Interventions [21]. When using the NOS for cohort studies, a study is judged by giving points ('stars' in the NOS terminology) to items belonging to three categories: selection of the study groups, the comparability of the cohorts on the basis of the design or analysis, and the outcome of interest [26]. For quality assessment of included studies reporting on randomised trials, we used the CONSORT 2010 checklist [27]. The CONSORT 2010 checklist consists of 37 items, and if an item was reported in an included article, a score of 1 was given for that item, while a score of 0 was given if an item was not reported. Overall adherence to the CONSORT 2010 checklist was scored as the percentage of the total number of included items. LFC and PA independently reviewed each included study using the NOS and CONSORT 2010 checklist, with disagreements referred to a third reviewer (AT-D or ND) for resolution. Details of the NOS criteria and CONSORT 2010 checklist used for quality assessment are provided in Supplementary Tables S5 and S6.

Results

Studies and patient characteristics

Figure 1 shows the PRISMA flow diagram of the study selection. The search identified 1225 records, of which 876 studies were excluded after title and abstract screening, and 333 studies excluded in the full-text screening. As a result, 16 studies [28–38] were included in the final review and data extraction. Table 1 summarises characteristics of included studies and patients.

Instruments used for HRQoL measurement

Only one study [37] used a preference-weighted instrument for measuring HRQoL (HUI2 and HUI3) [17, 37]. The most commonly used instrument across all studies was the generic SF-36 questionnaire [19] (ten studies [28–30, 35, 37–42]) (Table 1). While SF-36 could be used to derive the preference-weighted SF-6D, no studies reported such a derivation. One study [36] used SF-12, a shortened version of the SF-36 [43]. Five studies [33, 34, 37, 38, 44] used the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30) [45, 46]. In addition to using SF-36, one Dutch study [29] used the Bone tumour short-form DUCATQOL (Bt-DUX)[47] and TNO-AZL Quality of Life (TAAQOL) [48] questionnaires [29], while another Dutch study used SF-36 and TAAQOL [39] – both of these studies then compared their mean scores amongst each of the HRQoL questionnaires. Lastly, one study [31] used the Quality of Life Questionnaire (QLQ) [49] only, and another study [32] used the Quality of Life-Cancer Specific Scale (QOL-CSS) [50] questionnaire only.

Types of surgery

Eleven of the 16 studies involved both limb-salvage surgery and amputation (Table 2) [28–35, 37, 39, 42]. Four of these 11 studies also included individuals treated with rotationplasty, with three studies defining rotationplasty as a form of amputation [32, 35, 37], and one defining rotationplasty as a form of limb-salvage surgery [28]. One of the 16 studies involved rotationplasty only, defined by the authors as "limb preserving surgery" [41]. Four of the 16 studies involved limb-salvage surgery only [36, 38, 40, 44].

Study quality assessment

Of the 16 included studies, 15 were cohort studies and assessed using the NOS criteria with a maximum of nine points for each study. Among these cohort studies, one study [36] achieved eight points, six studies [30–32, 35, 39, 42] achieved seven points, and five studies [28, 33, 37, 38, 41, 44] scored six points. Of the remaining three cohort studies, two studies [29, 34] scored five points, and one study [40] scored three points (Table 3). Extremity bone sarcoma is a rare disease typically managed with a centre-based approach and therefore specialist oncology centres provide services to patients across large geographical

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

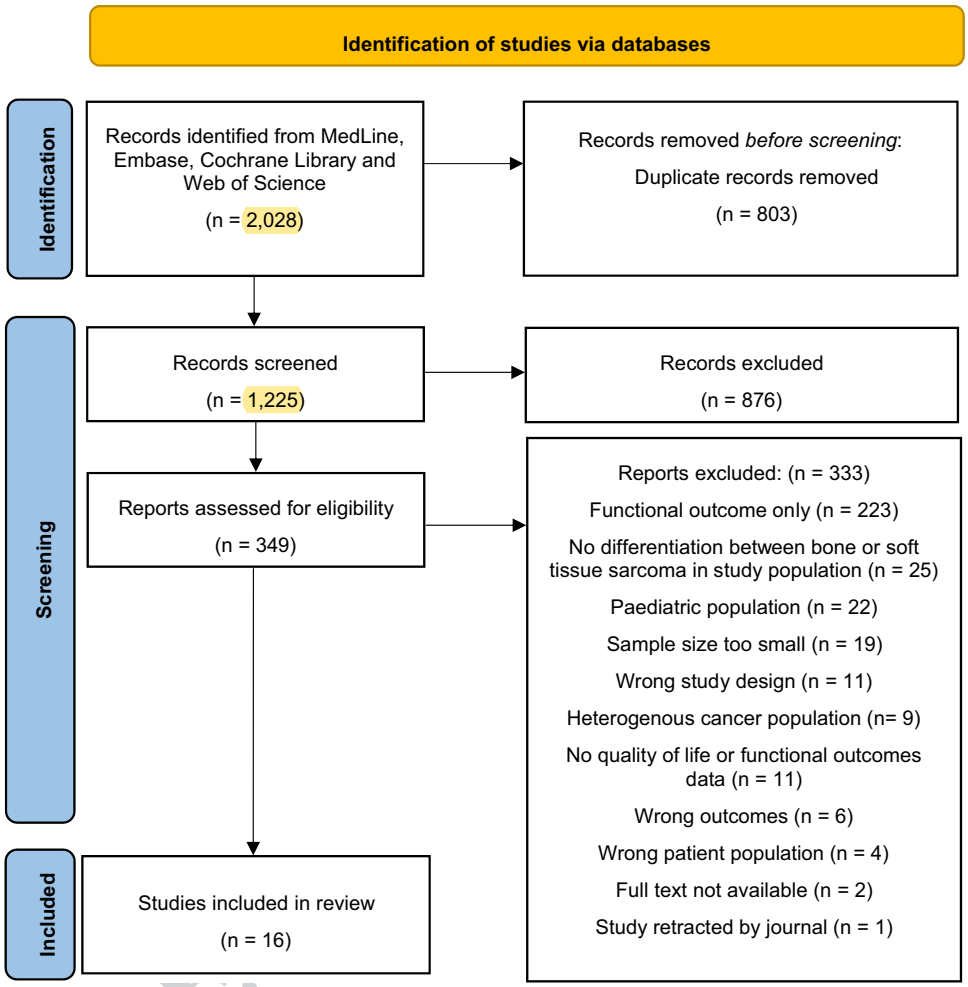


Fig. 1 PRISMA flow diagram of the study selection

areas [51]. In all studies except one [40], participants were reportedly recruited from dedicated specialist oncology centres and therefore they were most likely representative of the population of patients with bone sarcoma. In addition, the selected patients were confirmed as having their diagnosis and surgical management obtained via secure hospital records. In studies that compared patients undergoing surgical treatment with a healthy cohort [39–42, 44], both cohorts were chosen from the same regions. Except for one study [36] which obtained a baseline HRQoL measurement at the time of the surgical intervention, all cohort studies were conducted several years after the surgery and did not obtain a baseline HRQoL measurement. Eleven studies [28, 30–33, 35–38, 41, 42, 44] conducted comparisons between different surgical treatments, such as limb-salvage surgery versus amputation. However, none of these studies that involved comparisons between different surgical treatments adjusted the differences in HRQoL for important confounders such as age, sex, socioeconomic status, and comorbidities. In summary, 12 of the 16 studies achieved a score of six or more out of nine using the NOS criteria (Table 3).

Only one of the included studies used a randomised trial design [38], with this study adhering to 20 out of 37 items (54%) on the CONSORT 2010 checklist (Table 4).

Findings on health-related quality of life

Most of the studies comparing HRQoL between amputation and limb-salvage surgery reported no significant differences between the two modalities (Table 2) [28, 30, 32–34, 37, 39]. The study [37] that used preference-weighted measures of HRQoL found that amputees had a significantly higher emotion score (using HUI2 or HUI3) and higher fatigue score (using the EORTC QLQ-C30) compared to limb-salvage surgery. However, this study [37] found no other differences in any other dimensions of the HUI2, HUI3 or EORTC QLQ-C30, and no significant differences between amputation and limb-salvage surgery using SF-36. One study found that limb-salvage surgery had significantly better overall HRQoL than amputation using the QLQ [31], while another using SF-36 reported the dimension of social functioning was higher for limb-salvage surgery when compared to amputation, but no

UNCORRECTED

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 1 Summary of study characteristics

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Aksnes et al. [28]	2008	Norway and Sweden	Cross-sectional	Norwegian and Swedish patients treated for non-metastatic osteosarcoma or Ewing's sarcoma of the limbs with amputation or limb-salvage surgery	118	55%	31	13	1982–2000	SF-36	TESS and MSTS
Avasthi et al. [36]	2023	India	Cross-sectional	Two groups of Indian patients; one having GCT of distal end femur and the other group with GCT of proximal end tibia, both treated with limb-salvage surgery	21	61.9%	33.67	N/A	2015–2019	SF-12	MSTS
Barrera et al. [37]	2012	Canada	Cross-sectional	Canadian patients with malignant osteosarcoma or Ewing sarcoma of the lower limbs treated with limb-salvage surgery or amputation	28	50%	25.1	Not stated	Not stated	SF-36, EORTC QLQ-C30, HUI2, HUI3	TESS

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 1 (continued)

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Bekkering et al. [29]	2013	UK	Cross-sectional	Patients from the UK who underwent limb-salvage surgery or amputation for a malignant osteosarcoma or Ewing sarcoma of the leg	17	59%	19.9	3.7	2004–2009	SF-36, Bt-DUX and TAAQOL	TESS
Bekkering et al. [39]	2010	Netherlands	Cross-sectional with controls	Dutch patients with malignant osteosarcoma or Ewing sarcoma located around the knee treated with limb-salvage surgery or amputation. The patients were compared to a healthy population of Dutch people	48	48%	19.7	3.1	2003–2006	SF-36, TAAQOL and qualitative data from patient interviews	No

UNCORRECTED PROOF

Table 1 (continued)

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Fernandes et al. [44]	2021	Denmark	Cross-sectional with controls	Danish patients diagnosed with bone sarcoma or giant cell tumour of bone located in the hip or knee and treated with resection and reconstruction with a prosthesis. Patients with amputation were excluded. A control group of 30 healthy Danish patients matched by age ± 4 years and gender on a 1:1 ratio were included	60	53%	51	7	2006–2016	EORTC QLQ-C30	TESS

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 1 (continued)

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Fowler et al. [40]	2021	USA	Cross-sectional with controls	American patients post-proximal tibial tumour resection for malignant bone tumour compared to a control group ($n=9$) of similar age and sex distribution on a 1:1 ratio	18	44%	Cases 31.3; Control 31.6	14.1	Not stated	SF-36	MSTS
Gradl et al. [41]	2015	Germany	Cross-sectional with controls	German patients who underwent rotationplasty for management of bone sarcoma of the distal or proximal femur compared to a representative healthy German sample cohort ($n=2043$)	12	58%	32	14	1991–2001	SF-36	MSTS
Jing et al. [38]	2022	China	Randomised controlled trial	Chinese patients diagnosed with osteosarcoma and requiring limb-salvage surgery	90	Control = 60% Study = 55.5%	Control = 29.9 Study = 29.9	0.5	2016–2018	SF-36, EORTC QLQ-C30	MSTS

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 1 (continued)

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Malek et al. [30]	2012	USA	Cross-sectional	American patients treated for malignant bone sarcoma around the knee with limb-salvage surgery or above-knee amputation	20	60%	34	4.7	1994–2004	SF-36	TESS, RNL and PCI
Mason et al. [31]	2013	USA	Cross-sectional	American patients with lower extremity bone sarcoma who underwent limb-salvage surgery or amputation	82	52%	Amputation = 27.6; LSS = 25.6	Amputation = 8.1; LSS = 5	Not stated	QLQ	No
Robert et al. [32]	2010	USA	Cross-sectional	American patients diagnosed with osteosarcoma of the upper and lower limb and treated with limb-salvage surgery or amputation	57	35%	33.8	Amputation = 21.8; LSS = 13.6	Not stated	QOL-CSS	TESS
Silva et al. [33]	2020	Brazil	Cross-sectional	Brazilians with lower limb bone sarcoma treated with amputation or limb-salvage surgery	45	58.6%	Amputation = 32.2; LSS = 27	Amputation = 10.6; LSS = 7.1	Not stated	EORTC QLQ-C30	MSTS

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 1 (continued)

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Solooki et al. [34]	2018	Iran	Cross-sectional	Iranians diagnosed with osteosarcoma and treated with either limb-salvage surgery or amputation	48	Not stated	Amputation = 34.06; LSS = 25.81	Not stated	2006–2016	EORTC QLQ-C30	No
Suresh Nathan et al. [42]	2021	Singapore	Cross-sectional with controls	Singaporeans diagnosed with primary bone cancers classified into categories of treatment of the joints and segmental loss and the material used. A control group of 17 patients with acute lymphocytic leukaemia were included as limbs with cancer experience controls	162	56.8	38.7	9.1	1978–2008	SF-36	TESS and MSTS

UNCORRECTED PROOF

Table 1 (continued)

Authors	Year	Country	Study design	Study cohort	Sample size	Male (%)	Mean age at time of study (years)	Mean duration between surgery and HRQoL measurement (years)	Time-frame of surgery	Instrument for HRQoL measurement	Functional outcome measure
Yonemoto et al. [35]	2007	Japan	Cross-sectional	Japanese patients diagnosed with osteosarcoma of the extremities and treated with either limb-salvage surgery or amputation	33	39.3%	33.5	Not stated	1976–2000	SF-36	MSTS

Bt-DUX Bone tumour short form of DUCATQOL questionnaire; *EORTC QLQ-C30* European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire; *GCT* giant cell tumour; *HRQoL* health-related quality of life; *HUI2* Health Utilities Index Mark 2; *HUI3* Health Utilities Index Mark 3; *LSS* limb-salvage surgery; *MSTS* Musculoskeletal Tumour Society Score; *PCI* Physiological Cost Index; *RNL* Reintegration to Normal Living Index; *SF-36* Short Form survey-36; *TESS* Toronto Extremity Salvage Score; *TAAQOL* TNO-AZL Adult Quality Of Life; *QLQ* Quality of Life Questionnaire; *QOL-CSS* Quality of Life Cancer Specific Scale

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 2 Summary of study findings

Authors, year	HRQoL comparison description	Type(s) of surgery included in the study	Study aim	Findings
Aksnes et al., 2008 [28]	Long-term HRQoL outcomes between amputees and those with limb-salvage surgery	Limb-salvage surgery (including rotationplasty) and amputation	Evaluated the long-term functional outcome in 118 patients treated for osteosarcoma or Ewing's sarcoma in the extremities a minimum of five years after treatment. Also examined if impaired function influenced HRQoL and ability to work	There were no significant differences in HRQoL between amputees and those with limb-salvage surgery except in physical functioning ($p < 0.01$). Amputees had a significantly lower MSTS score than those with limb-salvage surgery ($p < 0.001$), but there was no difference for the TESS. 11% (13) of patients did not work or study
Avasthi et al., 2023 [36]	Comparing the SF-12 mental and physical scores preoperatively and at two years	Limb-salvage surgery only	To compare functional and HRQoL outcomes before and after surgery for patients with GCT of the distal femur and proximal tibia	SF-12 mental and physical scores preoperatively and at two years showed statistically significant improvement in both the distal femur and proximal tibia groups. Intergroup analysis also showed significantly better SF-12 mental and physical score in the distal femur group compared to the proximal tibia group during the preoperative and two-year follow-up. Patients with GCT of the proximal tibia as compared to the distal femur also have poorer functional scores in the pre-operative and post-operative follow-up periods
Barrera et al., 2012 [37]	Measurement of HRQOL of malignant bone sarcoma survivors using SF- 36, HUI, and EORTC-QLQ-C30	Limb-salvage surgery and amputation (including rotationplasty defined as amputation)	Measurement of HRQOL of malignant bone sarcoma survivors using several general and specific measures of HRQOL	Survivors of malignant bone sarcoma reported HRQOL equivalent to the general population, with the exception of physical functioning. LSS reported poorer HRQoL than amputation for emotional functioning and fatigue. Males reported better HRQOL compared with females for overall HRQOL, general health, physical functioning, and social functioning. Finally, younger participants generally reported better HRQOL than older participants for overall HRQOL and physical functioning

UNCORRECTED PROOF

Table 2 (continued)

Authors, year	HRQoL comparison description	Type(s) of surgery included in the study	Study aim	Findings
Bekkering et al., 2013 [29]	Bt-DUX, TAAQOL, TESS and SF-36 scores for the entire cohort were compared to determine the validity of the Bt-DUX questionnaire	Limb-salvage surgery and amputation	To translate the Dutch DUX questionnaire for lower extremity bone tumour patients (Bt-DUX), a disease-specific HRQoL instrument, into the English language and preliminary validate the English version in patients who were treated for lower-extremity bone tumours	Results were pooled across types of surgery. The median scores of the SF-36 were 11.7 (range: 0–64) for the PCS and 48.2 (range: 0–64) for the MCS. The median TAAQOL domain scores varied between 41.7 and 100. English Bt-DUX translation is a valid disease-specific instrument for evaluating QoL of adolescents with lower extremity bone cancer. The mean scores of the TESS, TAAQOL, and SF-36 as found in the present study were all slightly lower than those found in other cross-sectional studies, suggesting that the study population may not be representative for all adolescent patients after surgery for a malignant bone tumour of the leg
Bekkering et al., 2010 [39]	TAAQOL and SF-36 of patients compared to healthy Dutch population (reference sample); and differences between limb-salvage surgery vs. amputation, and male vs. female gender	Limb-salvage surgery and amputation	Compare the HRQoL of children and adolescents after malignant bone tumour surgery of the leg with healthy controls. Also examined the impact of gender and type of surgical procedure on the HRQoL outcome in this patient population	In comparison with healthy controls, patients had significantly ($p < 0.05$) poorer HRQoL within the domains gross motor function, cognitive functioning, pain, daily activities and of the TAAQOL. Significantly lower scores in physical functioning, physical role, general health, and the physical and mental component summary scales of the SF-36 were observed. No differences were observed between surgery types. No significant difference between gender except female patients had significantly higher SF-36 role physical scores than male patients

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 2 (continued)

Authors, year	HRQoL comparison description	Type(s) of surgery included in the study	Study aim	Findings
Fernandes et al., 2021 [44]	Differences between patients and age-matched controls for body functions (range of motion, muscle strength, pain), activity, and participation (walking, getting up from a chair, daily tasks), and HRQoL (EORTC QLQ-C30); and within the patient group, does body function, activity and participation outcome scores correlate with HRQoL	Limb-salvage surgery only	To discern any differences in body functions and HRQoL for patients 2–12 years after a bone resection and reconstruction surgery of the hip and knee following resection of a bone sarcoma and without walking limitations and do body functions, activity, and participation outcome scores correlate with HRQoL	Compared to healthy controls, bone tumour patients showed significantly different HRQoL ($p < 0.001$) in global health status, and physical, role and emotional functions. Comparing results with population based references, patients' scores for global health were at equal levels and physical functioning were at levels of 70- to 79-year-old patients, despite the cohort having a mean age of 51 ± 18 years. Patients showed poorer walking and chair-stand capacity and had lower TESS scores than patients in the control group. Higher pain scores correlated to lower physical functioning of the EORTC QLQ-C30. Less muscle strength in knee extension, knee flexion, and hip abduction correlated to lower physical functioning of the EORTC QLQ-C30
Fowler et al., 2021 [40]	The proximal tibial tumour resection group was compared to the control group for gait outcomes, however, for the SF-36 data, patients were compared to US national norms	Limb-salvage surgery only	The impact of proximal tibial tumour resection on gait and HRQoL	The surgical limb exhibited significantly poorer gait. Norm-based proximal tibial reconstruction group SF-36 component scores were within normal values (53.4 PCS, 56.5 MCS). The mean SF-36 MCS and PCS scores exceeded the US national average of 50. Despite gait deficits, walking speed and HRQoL were relatively normal in patients undergoing proximal tibial tumour resection
Gradl et al., 2015 [41]	SF-36 data for the patient cohort was compared to a healthy German sample cohort	Rotationplasty only	To assess long-term results regarding HRQoL and functionality after rotationplasty and to compare the obtained results to a representative healthy German sample cohort	The comparison of the presented results to a representative healthy German sample cohort revealed significantly higher scores for the patients regarding the SF-36 dimensions vitality ($p = 0.0243$), social functioning ($p = 0.0001$) and mental health ($p = 0.0001$). No other significant differences were observed. The authors describe this finding as 'unexpected' and suggest it may be due to 'posttraumatic growth', whereby patients feel some benefit in their disease e.g., "living more intensively and consciously"

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 2 (continued)

Authors, year	HRQoL comparison description	Type(s) of surgery included in the study	Study aim	Findings
Jing et al., 2022 [38]	HRQoL was measured six months after treatment, with SF-36- and EORTC QLQ-C30	Limb-salvage surgery only	To assess the efficacy and impact on HRQoL of neoadjuvant chemotherapy plus limb-salvage surgery, as compared to limb-salvage surgery and adjuvant chemotherapy only for a cohort of patients with osteosarcoma	Neoadjuvant chemotherapy plus limb-salvage surgery improves the post-operative limb function and long-term HRQoL of patients with osteosarcoma as compared to limb-salvage surgery plus adjuvant chemotherapy only
Malek et al., 2012 [30]	SF-36 outcome scores were compared between the limb-salvage surgery and amputation groups	Limb-salvage surgery and amputation	Compared functional status and HRQoL for patients treated with above-knee amputation versus limb-salvage surgery	No difference in the SF-36 scores between the limb-salvage surgery group and the amputation group ($p = 0.176$) or the TESS scores ($p = 0.051$). Limb-salvage surgery offers better gait efficiency and return to normal living compared with amputation, but does not improve the patient's perception of HRQoL
Mason et al., 2013 [31]	Using the QLQ questionnaire to assess whether there is a difference in HRQoL in lower extremity bone sarcoma survivors related to whether they had an amputation or a limb-salvage surgery	Limb-salvage surgery and amputation	The purpose of the study was to determine if there is a difference in overall HRQoL in lower extremity bone sarcoma survivors related to whether they had an amputation or a limb- preservation procedure while identifying psychological differences for further evaluation	The overall HRQoL of patients with limb-salvage surgery was significantly higher than patients with amputation ($p < 0.01$). The QLQ revealed significant differences ($p < 0.05$) in subgroups material well-being, occupational relations, creative-aesthetic behaviour, and sports activity
Robert et al., 2010 [32]	Compare the psychosocial and functional outcomes in osteosarcoma survivors after limb-salvage surgery and amputation using the QOL-CSS questionnaire	Limb-salvage surgery and amputation (including rotationplasty defined as amputation)	Compare the psychosocial and functional outcomes in osteosarcoma survivors after limb-salvage surgery and amputation	Lower limb function was a significant predictor of HRQoL ($p < 0.001$), whereas surgery type did not impact HRQoL. Lower limb functional outcomes as measured by TESS in the two groups were comparable ($p = 0.93$)
Silva et al., 2020 [33]	EORTC QLQ-C30 was compared between limb-salvage surgery and amputation groups	Limb-salvage surgery and amputation	To understand the perception of the HRQoL of adults with sarcomas in the lower limbs, after either amputation or conservative surgery, in addition to analysing functional and psychological aspects that may interfere with their HRQoL	The mean values of the EORTC QLQ-C30 did not differ for all dimensions except constipation, which was higher in the amputation group than in conservative surgery ($p = 0.0116$). MSTS values did not differ between procedures ($p = 0.2135$)
Solooki et al., 2018 [34]	EORTC QLQ-C30 was compared between limb-salvage surgery and amputation groups	Limb-salvage surgery and amputation	Evaluate the psychometric properties of the EORTC QLQ-C30 in Iranian osteosarcoma patients and apply this to compare the functional level, HRQoL, symptoms and financial burden of the patients who underwent amputation and limb-salvage operations	The EORTC QLQ-C30 was suitable in the Iranian setting. No significant differences in function, HRQoL and financial impact were observed between the groups studied

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 2 (continued)

Authors, year	HRQoL comparison description	Type(s) of surgery included in the study	Study aim	Findings
Suresh Nathan et al., 2021 [42]	SF-36 outcomes were compared across the different joint treatments and material used in reconstruction	Limb-salvage surgery and amputation	To investigate patient-centred outcome measures in long-term survivors of oncologic limb salvage procedures	SF-36 Mental Health and TESS degree of importance for activities of daily living identified that all categories of joint preservation were similar to a control group of people in whom there was no clinically significant extremity disease. SF-36 PCS scores were significantly better amongst joint-preserved (48.3 ± 8.5) versus joint replacement (40.1 ± 9.0) patients ($p = 0.003$). Nevertheless, there was no significant difference with respect to SF-36 PCS scores when amputees (44.7 ± 8.0) were compared to arthrodesed (46.7 ± 8.2) and joint-replaced individuals. Only for joint-preserved patients and those in whom the joint was uninvolved and required intercalary procedures (47.4 ± 10.1) did PCS approach the ideal ($p =$ ns) of the cancer patient without limb involvement (42.1 ± 7.4)
Yonemoto et al., 2007 [35]	SF-36 outcomes were compared between limb-salvage surgery and amputation groups, also taking into consideration the national Japanese average for SF-36. SF-36 was also compared to functional outcome data for any correlation	Limb-salvage surgery and amputation (including rotationplasty defined as amputation)	To evaluate the HRQoL in long-term survivors of high-grade osteosarcoma	The mean SF-36 score of physical functioning was 39.8, which was lower than the national standard of the Japanese, but the mean scores for the other seven parameters were higher than the national standards. When the HRQoL was compared between the limb-salvage and amputation groups, the score of social functioning was significantly higher in the limb-salvage group ($p = 0.023$). No significant difference was noted in the score for any of the other seven parameters. A significant positive correlation was noted between the MSTS score and the physical component, but no significant correlation was observed between the MSTS score and the mental component summary

Bt-DUX Bone tumour short form of DUCATQOL questionnaire; *EORTC QLQ-C30* European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire; *GCT* Giant cell tumour; *HRQoL* health-related quality of life; *LSS* limb-salvage surgery; *MCS* Mental Component Summary; *MSTS* Musculoskeletal Tumour Society Score; *PCS* Physical Component Summary; *PCI* Physiological Cost Index; *RNL* Reintegration to Normal Living Index; *SF-36* Short Form survey-36; *TESS* Toronto Extremity Salvage Score; *TAAQOL* TNO-AZL Adult Quality Of Life; *QLQ* Quality of Life Questionnaire; *QOL-CSS* Quality of Life Cancer Specific Scale

UNCORRECTED PROOF

significant differences in the other seven dimensions [35]. One study [38] used a randomised controlled trial to compare HRQoL as measured by SF-36 and EORTC QLQ-C30 between one group undergoing neoadjuvant chemotherapy only and limb-salvage surgery and another group treated with adjuvant chemotherapy only and limb-salvage surgery. Using SF-36, this study found significantly higher scores for role limitation due to emotional problems, mental health, physical functioning, and social functioning, and a significantly lower score for pain in the group receiving neoadjuvant chemotherapy; however, scores for role limitation due to physical problems, general health, and vitality were not reported. In addition, this study [38] found significantly lower scores for fatigue, nausea and vomiting, dyspnoea, constipation, and diarrhoea using EORTC QLQ-C30 for the group who had neoadjuvant chemotherapy; however, scores for several domains of EORTC QLQ-C30, including physical function, functional performance, emotional function, cognitive function, social function, insomnia, and financial difficulty were not reported. Another study used SF-36 to compare dimensions of HRQoL between groups of patients receiving different limb-salvage procedures and amputation [42]. This study [42] found significantly higher scores for the physical health dimension in joint preserved cases compared to those receiving joint replacement, but no significant differences in this dimension between joint replacement, arthrodesis and amputation, and no significant differences in the mental health dimension between any type of limb-salvage surgery or amputation. Two studies compared HRQoL of patients receiving limb-salvage surgery with that of healthy controls, with one finding poorer HRQoL in the limb-salvage cohort using the EORTC QLQ-C30 [44], and the other reporting no significant differences between the two cohorts using SF-36 [40]. One study reported superior scores in the SF-36 dimensions of vitality, social functioning, and mental health in a cohort of patients who underwent rotationplasty compared to the healthy controls [41]. Only one study [36] included a baseline pre-operative measure of HRQoL at the time of surgery, with this study demonstrating a statistically significant improvement in SF-12 mental and physical scores at two years post-operatively for limb-salvage surgery.

Eight studies used SF-36 to measure HRQoL in study cohorts who underwent limb-salvage surgery and amputation, with only five providing the mean values for all SF-36 dimension scores; four with their findings stratified by amputation or limb-salvage [28, 35, 37, 39]; and one which did not stratify by surgery type [29]. There was large variation in the scores for each dimension across these five studies (Fig. 2). Limb-salvage surgery appeared to be associated with slightly higher scores in most of the dimensions when compared with amputation in studies with larger sample sizes, and scores for mental dimensions were also generally higher than those for the physical dimensions across the ten groups of patients.

The SF-36 dimension scores can also be aggregated into Physical Component Summary (PCS) and Mental Component Summary (MCS) scores on a scale 0–100, with higher scores representing better health status [52]. For a reference, the US general population has mean MCS and PCS scores of 50 and a standard deviation of 10 [53]. Ten studies used SF-36, with seven producing a mean score for the PCS and MCS in their study cohorts [28, 29, 35, 37, 39, 40, 42]. The minimum mean MCS (45.7) and PCS (20.4) scores were both in a UK study of a cohort treated by either limb-salvage or amputation, without stratifying results for surgery type [29]. The maximum mean MCS score (57.9) was found in a Dutch cohort who underwent amputation [39], and the maximum mean PCS score (53.4) was found in a US cohort who underwent limb-salvage surgery [40]. Unfortunately, most studies using SF-36 did not report their results with standard deviations or standard errors, and therefore we were unable to conduct a meta-analysis to summarise the SF-36 dimension, PCS, or MCS scores. Furthermore, several studies using SF-36 provided incomplete data, including those reporting PCS and MCS scores without providing dimension scores [40, 42]; reporting dimension scores without calculation of PCS and MCS scores [38, 41]; reporting only the median of all SF-36 dimension scores [30]; and reporting only five of the eight SF-36 dimension scores [38].

Discussion

In this systematic review we found a large variation in HRQoL scores, the extent to which the aggregated health domains (e.g., PCS and MCS) and dimensions (e.g., physical functioning and mental health dimensions of SF-36) of HRQoL data were presented, and the choice of instrument used to measure HRQoL. Regardless of the differences in instruments used and the duration between the surgery and HRQoL measurement (3.1–13 years), most studies reported no significant difference in HRQoL between amputation and limb-salvage surgery. Possible explanations for this lack of difference may be due to the use of high quality of modern limb prosthetics which substantially improve HRQoL of amputees and the complex procedure and subsequent risk of complications such as infection, periprosthetic fracture, or aseptic loosening requiring extensive revision surgeries that result in poorer than expected HRQoL in patients undergoing limb-salvage surgery [54, 55].

We found that studies with larger sample sizes tended to yield higher SF-36 general health scores, as well as higher scores in any dimension for limb-salvage surgery when compared with amputation. Scores for mental dimensions and

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 3 Newcastle–Ottawa Scale quality assessment of selected studies

NOS criteria	Aksnes et al., 2008 [28]	Avasthi et al., 2023 [36]	Barrera et al., 2012 [37]	Bekkering et al., 2010 [39]	Bekkering et al., 2013 [29]	Fernandes et al., 2021 [44]	Fowler et al., 2021 [40]	Gradl et al., 2015 [41]	Malek et al., 2012 [30]	Mason et al., 2013 [31]	Robert et al., 2010 [32]	Silva et al., 2020 [33]	Solooki et al., 2018 [34]	Suresh Nathan et al., 2021 [42]	Yonemoto et al., 2007 [35]
Selection															
Representativeness of the exposed cohort	1	1	1	1	1	1	x	1	1	1	1	1	1	1	1
Selection of the non-exposed cohort	1	1	1	1	x	1	x	1	1	1	1	1	1	1	1
Ascertainment of exposure	1	1	1	1	1	1	x	1	1	1	1	1	1	1	1
Demonstration that outcome of interest was present at the start of the study	x	1	x	x	x	x	x	x	x	x	x	x	x	x	x
Comparability															
Comparability of cohorts on the basis of design or analysis	1	1	1	1	x	x	x	x	1	1	1	1	1	1	1
Outcome															
Assessment of outcome	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

UNCORRECTED PROOF

Table 3 (continued)

NOS criteria	Aksnes et al., 2008 [28]	Avasthi et al., 2023 [36]	Barrera et al., 2012 [37]	Bekkering et al., 2010 [39]	Bekkering et al., 2013 [29]	Fernandes et al., 2021 [44]	Fowler et al., 2021 [40]	Gradl et al., 2015 [41]	Malek et al., 2012 [30]	Mason et al., 2013 [31]	Robert et al., 2010 [32]	Silva et al., 2020 [33]	Solooki et al., 2018 [34]	Suresh Nathan et al., 2021 [42]	Yonemoto et al., 2007 [35]
Duration of follow-up	1	1	1	1	1	1	1	1	1	1	1	x	x	1	1
Adequacy of follow-up of cohorts	x	1	x	1	1	1	1	1	1	1	1	1	x	1	1
Total	6	8	6	7	5	6	3	6	7	7	7	6	5	7	7

the mean MCS score were also generally higher than those for the physical dimensions and mean PCS score, likely due to the nature of musculoskeletal morbidity resulting from bone sarcoma and its associated treatment [53]. Unfortunately, we were not able to further investigate differences in SF-36 scores because the number of studies was small, reporting of standard deviations or changes in HRQoL from the baseline was absent, and many studies did not stratify their results by surgery type or other characteristics. Several studies also used instruments other than SF-36, making it difficult to compare HRQoL across the studies. This heterogeneity in instruments used to measure HRQoL is consistent with a previous systematic review of HRQoL in adult patients with extremity sarcoma [20].

The quality of included the cohort studies was of moderate to good quality, with 12 of the 15 studies achieving a score of six or more out of maximum nine points using the NOS criteria for cohort studies (Table 3). Conversely, the study using a randomised trial design [38] adhered to only 54% of the items on the CONSORT 2010 checklist (Table 4). Only one study, which was a cohort study [36], included a baseline measure of HRQoL at the time of surgery, which makes it difficult to attribute any changes in HRQoL to surgical intervention.

Surprisingly, only one study used HRQoL instruments with preference weights (HUI2 and HUI3) [37]. As this study used a small sample ($n = 28$) of Canadian patients, the current evidence is very weak to support a direct estimation of QALYs in patients with extremity bone sarcoma after surgical treatment, especially for those in countries other than Canada.

While the generic SF-36 or disease-specific EORTC QLQ-C30 provided informative numerical scores of HRQoL regarding the effectiveness of surgical management, the decisions made by patients when responding to these instruments is not based on any value or desirability (i.e., preference) for comparing one health state to another. This is because the scoring algorithms for health states in these simple instruments infer that the response choices (e.g., 'Have you had pain?' vs. 'Have you vomited?') and the intervals between the responses (e.g., 'Not at All' vs. 'A Little') are of equal importance. Hence, measures of HRQoL whose scoring systems are not weighted by preference do not allow for comparisons to be made across different areas of healthcare [18]. Because healthcare resources are finite, there is a requirement to assess the value of competing interventions by means of economic evaluation. Our results have shown that there is an important gap in the use of HRQoL instruments that facilitate estimation of QALYs. This deficiency of preference-weighted HRQoL evidence hinders the economic evaluation needed to inform decision on application of novel treatments for surgically excising extremity bone sarcomas, and also prevents an understanding of the effectiveness and

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 4 CONSORT 2010 [27] checklist of information to include when reporting a randomised trial for Jing et al. 2022 [38]

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	N/A
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	1 & 2
	2b	Specific objectives or hypotheses	2
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	2
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	N/A
Participants	4a	Eligibility criteria for participants	2
	4b	Settings and locations where the data were collected	2
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	2
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	2 & 3
	6b	Any changes to trial outcomes after the trial commenced, with reasons	N/A
Sample size	7a	How sample size was determined	N/A
	7b	When applicable, explanation of any interim analyses and stopping guidelines	N/A
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	2
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	2
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	N/A
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	N/A
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	2
	11b	If relevant, description of the similarity of interventions	2
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	2
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	N/A
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	3 & 4
	13b	For each group, losses and exclusions after randomisation, together with reasons	N/A
Recruitment	14a	Dates defining the periods of recruitment and follow-up	2
	14b	Why the trial ended or was stopped	N/A
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	4
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	3 & 4
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	3 & 4
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	N/A
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	N/A
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	N/A

UNCORRECTED PROOF

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

Table 4 (continued)

Section/Topic	Item No	Checklist item	Reported on page No
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	4
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	N/A
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	3–5
Other information			
Registration	23	Registration number and name of trial registry	N/A
Protocol	24	Where the full trial protocol can be accessed, if available	N/A
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	N/A

UNCORRECTED PROOF

cost-effectiveness of existing treatments. This is particularly relevant given the growing prevalence of costly, but potentially effective approaches in orthopaedic surgery [56–59].

The lack of preference-weighted HRQoL could be improved by mapping the common existing non preference-weighted measures, including SF-36 and EORTC QLC-C30, to preference-weighted instruments such as SF-6D, EQ-5D, or EORTC-8D [18, 60, 61]. Given SF-36 was shown to be a very common tool for measuring HRQoL, there was a noticeable absence of conversion to SF-6D using well-established approaches such as the method proposed by Brazier et al. [18]. However, mapping can be erroneous when generic preference-weighted measures such as SF-6D or EQ-5D are used in a disease-specific population because the instruments may not be sensitive to certain disease characteristics [62]. While there is a disease-specific

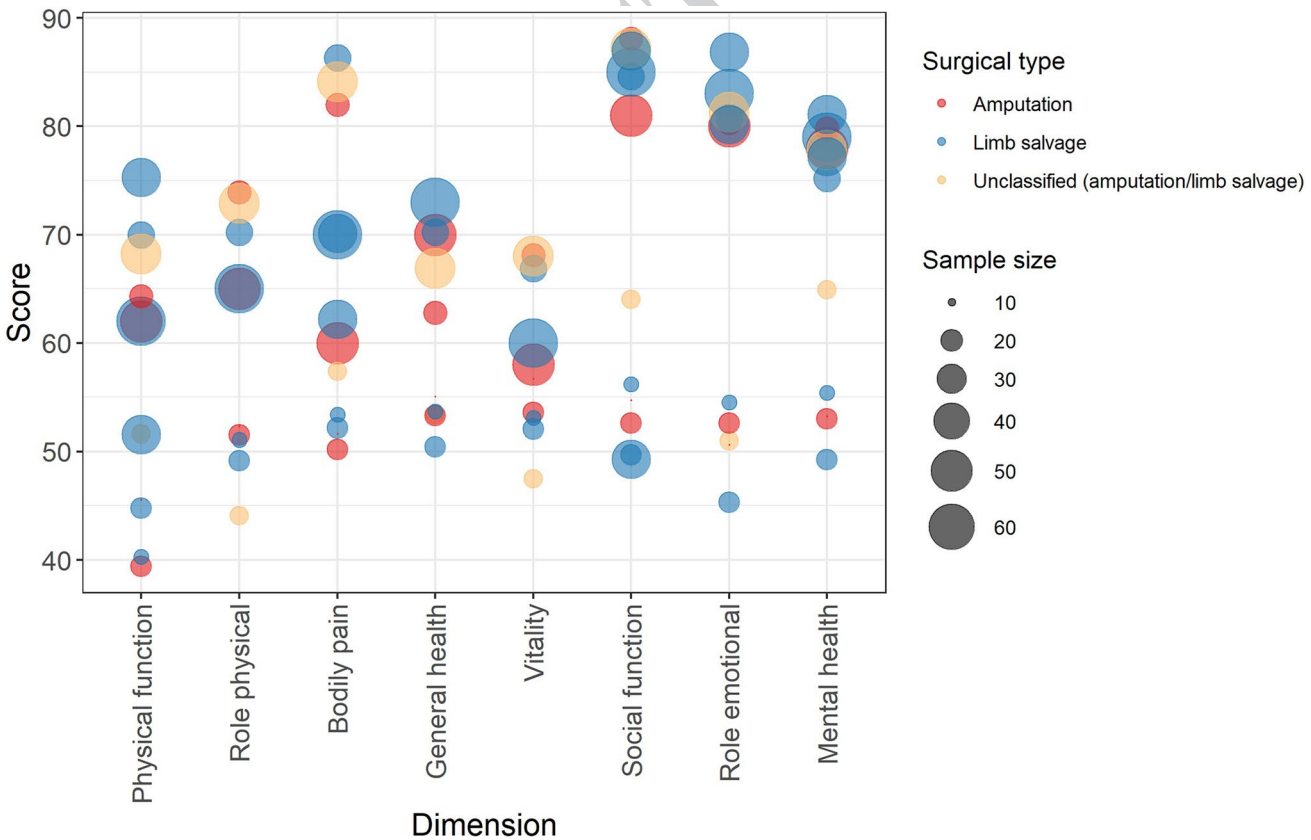


Fig. 2 SF-36 dimension scores in cohorts undergoing amputation, limb-salvage or unclassified surgery for treatment of extremity bone sarcoma

Journal : Large 11136 | Article No : 3554 | Pages : 24 | MS Code : 3554 | Dispatch : 11-11-2023

preference-weighted instrument for cancer in the EORTC-8D, which is mapped from EORTC QLC-C30, it was derived from patients with multiple myeloma. EORTC-8D may therefore be insensitive to certain characteristics associated with sarcomas [63]. A new cancer-specific instrument mapped from EORTC QLC-C30, the Quality of Life Utility Measure-Core 10 dimensions (QLU-C10D), has however recently been developed from data of patients with multiple types of cancer, including sarcoma [64, 65]. Nonetheless, policy makers in many countries require the use of generic preference-weighted instruments in economic evaluations. Given that our results have highlighted the lack of studies using these instruments, it is essential to include preference-weighted measures to improve evidence for future research in extremity bone sarcoma.

Our study has several limitations. No meta-analysis could be conducted because of substantial heterogeneities in patient characteristics, follow-up periods, and the absence of reported variation in the HRQoL values. For the studies that used the SF-36 to measure HRQoL, we were not able to obtain the raw scores to convert them to the SF-6D health utility values. We included studies involving patients who underwent rotationplasty, but classification of rotationplasty was unclear as some authors defined it as limb-salvage surgery [28, 41], while others considered it a form of amputation [32, 35, 37].

Conclusions

We completed a systematic review and found 16 studies measuring HRQoL in adult patients with extremity bone sarcomas who had been treated with limb-salvage surgery or amputation. We found only one study using preference-weighted instruments for HRQoL measurement in a small sample of patients. The approaches used to measure HRQoL were inconsistent and the outcome scores varied substantially. The most commonly used instrument among the existing studies was the generic SF-36 questionnaire, but no studies reported the preference-weighted HRQoL (SF-6D) derived from SF-36. Our findings showed that the current evidence is very weak to support a direct estimation of QALYs in patients with extremity bone sarcoma after surgical treatment. As the armamentarium of surgical treatments available to orthopaedic oncologists expands, juxtaposed with the expectations of regulatory bodies that thorough HTA should be conducted for these new interventions, future research into the surgical treatment of extremity bone sarcomas should consider the use of well-established preference-weighted instruments to measure HRQoL.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11136-023-03554-3.

Author contributions Conceptualisation: ATD; Methodology: ATD and ND; Literature search and data extraction: LFC and PA; Formal analysis and Investigation: All authors; Visualisation: LFC and ATD; Writing—original draft: LFC and ATD; Writing—review & editing: All authors; Funding acquisition: ND and ATD; Supervision: ATD ~~and ND~~.

Funding This study was partly funded by the Innovative Manufacturing Cooperative Research Centre (IMCRC). IMCRC had no role in the design or conduct of the study, the collection, management, analyses, and interpretation of the data, or the preparation and review of the manuscript. An Tran-Duy is partly supported by the Methods and Implementation Support for Clinical and Health research (MISCH) Hub at the University of Melbourne.

Declarations

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

References

1. Hameed, M., & Dorfman, H. (2011). Primary malignant bone tumors—recent developments. *Seminars in Diagnostic Pathology, 28*(1), 86–101.
2. Luetke, A., Meyers, P. A., Lewis, I., & Juergens, H. (2014). Osteosarcoma treatment–where do we stand? A state of the art review. *Cancer Treatment Reviews, 40*(4), 523–532.
3. Bielack, S. S., Kempf-Bielack, B., Gn, D., Exner, G. U., Flege, S., Helmke, K., Kotz, R., Salzer-Kuntschik, M., Werner, M., & Winkelmann, W. (2002). Prognostic factors in high-grade osteosarcoma of the extremities or trunk: An analysis of 1702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *Journal of Clinical Oncology, 20*(3), 776–790.
4. Errani, C., Longhi, A., Rossi, G., Rimondi, E., Biazzo, A., Toscano, A., Alì, N., Ruggieri, P., Alberghini, M., & Picci, P. (2011). Palliative therapy for osteosarcoma. *Expert Review of Anticancer Therapy, 11*(2), 217–227.
5. Misaghi, A., Goldin, A., Awad, M., & Kulidjian, A. A. (2018). Osteosarcoma: A comprehensive review. *Sicot-J, 4*, 12.
6. Bielack, S., Jürgens, H., Jundt, G., Kevric, M., Kühne, T., Reichardt, P., Zoubek, A., Werner, M., Winkelmann, W., & Kotz, R. (2009). Osteosarcoma: the COSS experience. *Pediatric and adolescent osteosarcoma* (pp. 289–308). Springer.
7. Jaffe, N. (2009). Osteosarcoma: review of the past, impact on the future. The American experience. *Pediatric and adolescent osteosarcoma* (pp. 239–262). Springer.
8. Longhi, A., Errani, C., De Paolis, M., Mercuri, M., & Bacci, G. (2006). Primary bone osteosarcoma in the pediatric age: State of the art. *Cancer Treatment Reviews, 32*(6), 423–436.
9. IMCRC. (2017). Just in time implants. Available from https://www.imcrc.org/just-in-time-implants/ Accessed 10 Feb 2023.
10. Young, P., Bell, S., & Mahendra, A. (2015). The evolving role of computer-assisted navigation in musculoskeletal oncology. *The Bone & Joint Journal, 97*(2), 258–264.
11. Dong, C., Beglinger, I., & Krieg, A. H. (2022). Personalized 3D-printed guide in malignant bone tumor resection and following reconstruction–17 cases in pelvic and extremities. *Surgical Oncology, 42*, 101733.
12. Angelini, A., Trovarelli, G., Berizzi, A., Pala, E., Breda, A., & Ruggieri, P. (2019). Three-dimension-printed custom-made

UNCORRECTED PROOF

prosthetic reconstructions: From revision surgery to oncologic reconstructions. *International Orthopaedics, 43*, 123–132.
13. Wong, R. M. Y., Wong, P. Y., Liu, C., Chung, Y. L., Wong, K. C., Tso, C. Y., Chow, S.K.-H., Cheung, W.-H., Yung, P.S.-H., & Chui, C. S. (2021). 3D printing in orthopaedic surgery: A scoping review of randomized controlled trials. *Bone & Joint Research, 10*(12), 807–819.
14. Whitehead, S. J., & Ali, S. (2010). Health outcomes in economic evaluation: The QALY and utilities. *British Medical Bulletin, 96*(1), 5–21.
15. Torrance, G. W., Furlong, W., & Feeny, D. (2002). Health utility estimation. *Expert Review of Pharmacoeconomics & Outcomes Research, 2*(2), 99–108.
16. Devlin, N. J., & Brooks, R. (2017). EQ-5D and the EuroQol group: Past, present and future. *Applied Health economics and Health policy, 15*(2), 127–137.
17. Furlong, W. J., Feeny, D. H., Torrance, G. W., & Barr, R. D. (2001). The Health Utilities Index (HUI®) system for assessing health-related quality of life in clinical studies. *Annals of Medicine, 33*(5), 375–384.
18. Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics, 21*(2), 271–292.
19. Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care, 30*(6), 473–483.
20. Tang, M. H., Pan, D. J., Castle, D. J., & Choong, P. F. (2012). A systematic review of the recent quality of life studies in adult extremity sarcoma survivors. *Sarcoma, 2012*, 171342.
21. Higgins, J. P. T. T. J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). John Wiley & Sons.
22. Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *British Medical Journal, 339*, b2535.
23. (2023) What is sarcoma? Available from https://sarcoma.org.au/pages/about-sarcoma/what-is-sarcoma
24. Adolescent, Group YAOPR. (2006). Closing the gap: Research and care imperatives for adolescents and young adults with cancer. Bethesda, MD: National Institutes of Health.
25. Covidence Covidence systematic review software. Available from www.covidence.org Accessed 10 February 2023.
26. Wells, GA., Shea, B., O'connell, D., Peterson, J., Welch, V., Losos, M., Tugwell P. (2012). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp Accessed 10 Feb 2023.
27. Schulz, K. F., Altman, D. G., Moher, D., & the CG,. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*(1), 18.
28. Aksnes, L. H., Bauer, H. C. F., Jebsen, N. L., Folleras, G., Allert, C., Haugen, G. S., & Hall, K. S. (2008). Limb-sparing surgery preserves more function than amputation: A Scandinavian sarcoma group study of 118 patients. *The Journal of Bone & Joint Surgery British Volume, 90*(6), 786–794.
29. Bekkering, W. P., Billing, L., Grimer, R. J., Vlieland, T. P. M. V., Koopman, H. M., Nelissen, R. G. H. H., & Taminiau, A. H. M. (2013). Translation and preliminary validation of the English version of the DUX questionnaire for lower extremity bone tumor patients (Bt-DUX): A disease-specific measure for quality of life. *Journal of Surgical Oncology, 107*(4), 353–359.
30. Malek, F., Somerson, J. S., Mitchel, S., & Williams, R. P. (2012). Does limb-salvage surgery offer patients better quality of life and functional capacity than amputation? *Clinical Orthopaedics and Related Research, 470*(7), 2000–2006.
31. Mason, G. E., Aung, L., Gall, S., Meyers, P. A., Butler, R., Krug, S., Kim, M., Healey, J. H., & Gorlick, R. (2013). Quality of life following amputation or limb preservation in patients with lower extremity bone sarcoma. *Frontiers in Oncology, 3*, 210.
32. Robert, R. S., Ottaviani, G., Huh, W. W., Palla, S., & Jaffe, N. (2010). Psychosocial and functional outcomes in long-term survivors of osteosarcoma: A comparison of limb-salvage surgery and amputation. *Pediatric Blood & Cancer, 54*(7), 990–999.
33. Silva, R. S., Tabet, L. P., Batista, K. T., Naves, J. F., Viana, E. D. E. M. B. O., & Guilhem, D. B. (2020). Quality of life in adults with sarcomas under conservative surgery or amputation. *Acta Ortopedica Brasileira, 28*(5), 236–242.
34. Solooki, S., Mostafavizadeh Ardestani, S. M., Mahdaviazad, H., & Kardeh, B. (2018). Function and quality of life among primary osteosarcoma survivors in Iran: Amputation versus limb salvage. *Musculoskeletal Surgery, 102*(2), 147–151.
35. Yonemoto, T., Ishii, T., Takeuchi, Y., Kimura, K., Hagiwara, Y., Iwata, S., & Tatezaki, S.-i. (2007). Evaluation of quality of life (QOL) in long-term survivors of high-grade osteosarcoma: A Japanese single center experience. *Anticancer Research, 27*(5B), 3621–3624.
36. Sachin, A., Swagat, M., Pankaj, A., Vineet, K., Ammar, A., Prabhat, K., Madhusudan, M. (2023). Outcome analysis of En-bloc excision and endoprosthetic replacement among the cases operated for distal femoral and proximal tibial giant cell tumour around the knee: A retrospective study. *Journal of Clinical and Diagnostic Research*, *Vol 17*. JCDR Research and Publications Private Limited
37. Barrera, M., Teall, T., Barr, R., Silva, M., & Greenberg, M. (2012). Health related quality of life in adolescent and young adult survivors of lower extremity bone tumors. *Pediatric Blood and Cancer, 58*(2), 265–273.
38. Jing, S., Ding, F., Yuan, Y., An, J., & He, Q. (2022). Efficacy of neoadjuvant chemotherapy plus limb-sparing surgery for osteosarcoma and its impact on long-term quality of life. *Evidence-Based Complementary and Alternative Medicine, 2022*, 1–6.
39. Bekkering, W. P., Vliet Vlieland, T. P. M., Koopman, H. M., Schaap, G. R., Schreuder, H. W. B., Beishuizen, A., Tissing, W. J. E., Hoogerbrugge, P. M., Anninga, J. K., & Taminiau, A. H. M. (2010). Quality of life in young patients after bone tumor surgery around the knee joint and comparison with healthy controls. *Pediatric Blood & Cancer, 54*(5), 738–745.
40. Fowler, E. G., Eckardt, J. J., Vuong, A., Greenberg, M. B., Yamaguchi, K., Bukata, S. V., & Bernthal, N. (2021). Gait outcomes following proximal tibial tumor resection and endoprosthetic reconstruction. *Gait & Posture, 88*, 167–173.
41. Gradl, G., Postl, L. K., Lenze, U., Stolberg-Stolberg, J., Pohlig, F., Rechl, H., Schmitt-Sody, M., von Eisenhart-Rothe, R., & Kirchhoff, C. (2015). Long-term functional outcome and quality of life following rotationplasty for treatment of malignant tumors. *BMC Musculoskeletal Disorders, 16*, 262.
42. Suresh Nathan, S., Tan Lay Hua, G., Mei Yoke, C., Mann Hong, T., & Peter Pereira, B. (2021). Outcome satisfaction in long-term survivors of oncologic limb salvage procedures. *European Journal of Cancer Care, 30*(2), e13377.
43. Ware, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220–233.
44. Fernandes, L., Holm, C. E., Villadsen, A., Sorensen, M. S., Zebis, M. K., & Petersen, M. M. (2021). 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*. https://doi.org/10.1097/CORR.0000000000001797
45. Stokke, J., Sung, L., Gupta, A., Lindberg, A., & Rosenberg, A. R. (2015). Systematic review and meta-analysis of objective and

UNCORRECTED PROOF

subjective quality of life among pediatric, adolescent, and young adult bone tumor survivors. *Pediatric Blood & Cancer, 62*(9), 1616–1629.
46. Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., Filiberti, A., Flechtner, H., Fleishman, S. B., Haes, J. C. JMd., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P. B., Schraub, S., Sneeuw, K., Sullivan, M., & Takeda, F. (1993). The European organization for research and treatment of cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute, 85*(5), 365–376.
47. Bekkering, W. P., Vlieland, T. P. M. V., Koopman, H. M., Schaap, G. R., Schreuder, H. W. B., Beishuizen, A., Tissing, W. J. E., Hoogerbrugge, P. M., Anninga, J. K., & Taminiau, A. H. M. (2009). The Bt-DUX: Development of a subjective measure of health-related quality of life in patients who underwent surgery for lower extremity malignant bone tumor. *Pediatric Blood & Cancer, 53*(3), 348–355.
48. Fekkes, M., Kamphuis, R. P., Ottenkamp, J., Verrips, E., Vogels, T., Kamphuis, M., & Verloove-vanhorick, S. P. (2001). Health-related quality of life in young adults with minor congenital heart disease. *Psychology & Health, 16*(2), 239–250.
49. Evans. DR., Cope, WE. (1989). Manual for the quality of life questionnaire. Multy-Health Systems.
50. Ferrell, B. R., Hassey Dow, K., & Grant, M. (1995). Measurement of the quality of life in cancer survivors. *Quality of Life Research, 4*(6), 523–531.
51. Choong, P. F. M. (2021). *Sarcoma : A practical guide to multidisciplinary management*. Springer.
52. Ware, J., Kosinski, M., Keller, S. (2001). SF-36 physical and mental health summary scales. A user's manual 1994.
53. Laucis, N. C., Hays, R. D., & Bhattacharyya, T. (2015). Scoring the SF-36 in orthopaedics: A brief guide. *The Journal of Bone and Joint Surgery American, 97*(19), 1628.
54. Thornley, P., Vicente, M., MacDonald, A., Evaniew, N., Ghert, M., & Velez, R. (2019). Causes and frequencies of reoperations after endoprosthetic reconstructions for extremity tumor surgery: A systematic review. *Clinical Orthopaedics and Related Research, 477*(4), 894–902.
55. Lathouwers, E., Díaz, M. A., Maricot, A., Tassignon, B., Cherelle, C., Cherelle, P., Meeusen, R., & De Pauw, K. (2023). Therapeutic benefits of lower limb prostheses: A systematic review. *Journal of Neuroengineering and Rehabilitation, 20*(1), 4.
56. Chen, A. F., Kazarian, G. S., Jessop, G. W., & Makhdom, A. (2018). Robotic technology in orthopaedic surgery. *JBJS, 100*(22), 1984–1992.
57. Moschetti, W. E., Konopka, J. F., Rubash, H. E., & Genuario, J. W. (2016). Can robot-assisted unicompartmental knee arthroplasty be cost-effective? A Markov decision analysis. *The Journal of Arthroplasty, 31*(4), 759–765.
58. Khan, F., Pearle, A., Lightcap, C., Boland, P. J., & Healey, J. H. (2013). Haptic robot-assisted surgery improves accuracy of wide resection of bone tumors: A pilot study. *Clinical Orthopaedics and Related Research, 471*(3), 851–859.
59. Sherman, W. F., & Wu, V. J. (2020). Robotic surgery in total joint arthroplasty: A survey of the AAHKS membership to understand the utilization, motivations, and perceptions of total joint surgeons. *The Journal of Arthroplasty, 35*(12), 3474-3481.e3472.
60. Beck, A. C. C., Kieffer, J. M., Retèl, V. P., van Overveld, L. F. J., Takes, R. P., van den Brekel, M. W. M., van Harten, W. H., & Stuiver, M. M. (2019). Mapping the EORTC QLQ-C30 and QLQ-H&N35 to the EQ-5D for head and neck cancer: Can disease-specific utilities be obtained? *PLoS ONE, 14*(12), e0226077.
61. Rowen, D., Brazier, J., Young, T., Gaugris, S., Craig, B. M., King, M. T., & Velikova, G. (2011). Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value Health, 14*(5), 721–731.
62. Brazier, J. E., Yang, Y., Tsuchiya, A., & Rowen, D. L. (2010). A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *The European Journal of Health Economics, 11*(2), 215–225.
63. Bray, N., Spencer, L. H., & Edwards, R. T. (2020). Preference-based measures of health-related quality of life in congenital mobility impairment: A systematic review of validity and responsiveness. *Health Econ Rev, 10*(1), 9.
64. King, M. T., Costa, D. S. J., Aaronson, N. K., Brazier, J. E., Cella, D. F., Fayers, P. M., Grimison, P., Janda, M., Kemmler, G., Norman, R., Pickard, A. S., Rowen, D., Velikova, G., Young, T. A., & Viney, R. (2016). QLU-C10D: A health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. *Quality of Life Research, 25*(3), 625–636.
65. Gamper, E.-M., King, M. T., Norman, R., Loth, F. L. C., Holzner, B., Kemmler, G., & the EQoLG,. (2022). The EORTC QLU-C10D discrete choice experiment for cancer patients: A first step towards patient utility weights. *Journal of Patient-Reported Outcomes, 6*(1), 42.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

UNCORRECTED PROOF

Journal:	**11136**
Article:	**3554**

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details Required	Author's Response
AQ2	Please confirm the section headings are correctly identified.	Yes they are correct
AQ3	Please provide complete details for the references [23, 36].	See comments
AQ1	Inclusion of a data availability statement is preferred for this journal. If applicable, please provide one.	Not applicable. All data are presented in the paper.