BJO

■ KNEE

Metabolic equivalent of task scores avoid the ceiling effect observed with conventional patient-reported outcome scores following knee arthroplasty

A. Patel,
T. C. Edwards,
G. Jones,
A. D. Liddle,
J. Cobb,
A. Garner

From Imperial College London, London, UK

Aims

The metabolic equivalent of task (MET) score examines patient performance in relation to energy expenditure before and after knee arthroplasty. This study assesses its use in a knee arthroplasty population in comparison with the widely used Oxford Knee Score (OKS) and EuroQol five-dimension index (EQ-5D), which are reported to be limited by ceiling effects.

Methods

A total of 116 patients with OKS, EQ-5D, and MET scores before, and at least six months following, unilateral primary knee arthroplasty were identified from a database. Procedures were performed by a single surgeon between 2014 and 2019 consecutively. Scores were analyzed for normality, skewness, kurtosis, and the presence of ceiling/floor effects. Concurrent validity between the MET score, OKS, and EQ-5D was assessed using Spearman's rank.

Results

Postoperatively the OKS and EQ-5D demonstrated negative skews in distribution, with high kurtosis at six months and one year. The OKS demonstrated a ceiling effect at one year (15.7%) postoperatively. The EQ-5D demonstrated a ceiling effect at six months (30.2%) and one year (39.8%) postoperatively. The MET score did not demonstrate a skewed distribution or ceiling effect either at six months or one year postoperatively. Weak-moderate correlations were noted between the MET score and conventional scores at six months and one year postoperatively.

Conclusion

In contrast to the OKS and EQ-5D, the MET score was normally distributed postoperatively with no ceiling effect. It is worth consideration as an arthroplasty outcome measure, particularly for patients with high expectations.

Introduction

Knee arthroplasty surgery is recognized to be an effective procedure which improves the quality of life for a majority of patients.[1] Two patient-reported outcome measures (PROMs) – the arthroplasty-specific Oxford Knee Score (OKS), and the quality-of-life metric Euro-Qol five-dimension index (EQ-5D) – are commonly employed and are the primary outcome measure for all knee arthroplasty procedures within the UK.[2] To maintain standards and drive improvements in care, the metric of 'improvement in PROMs score' is now used.[3]

However, previous work has suggested that both the OKS and EQ-5D are subject to ceiling effects, whereby over 15% of an assessed population score maximally.[4,5] This limits their ability to differentiate outcomes, and underestimates a patient's postoperative improvement. A ceiling effect also skews an outcome metric's distribution, invalidating the use of mean scores. Healthcare providers are currently required to demonstrate a

Correspondence should be sent to Mr Arjun Patel; email: arjun.patel09@imperial.ac.uk

```
┌─────────────────────────┐
│ Patients identified from│
│ database (n=843)        │
└─────────────────────────┘
            │          ┌──────────────────────────────┐
            ├─────────▶│ Patients with incomplete PROM│
            │          │ data excluded (n=592)        │
            │          └──────────────────────────────┘
            │
            │          ┌──────────────────────────────┐
            │          │ Patients with incomplete/    │
            ├─────────▶│ inaccurate procedural        │
            │          │ information excluded(n=105)   │
            │          └──────────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Patients with complete  │
│ data (n=146)            │
└─────────────────────────┘
            │          ┌──────────────────────────────┐
            │          │ Patients undergoing revision │
            ├─────────▶│ procedure or bilateral surgery│
            │          │ excluded (n=30)              │
            │          └──────────────────────────────┘
            ▼
┌─────────────────────────┐
│ Patients included in    │
│ final analysis (n=116)  │
└─────────────────────────┘
```
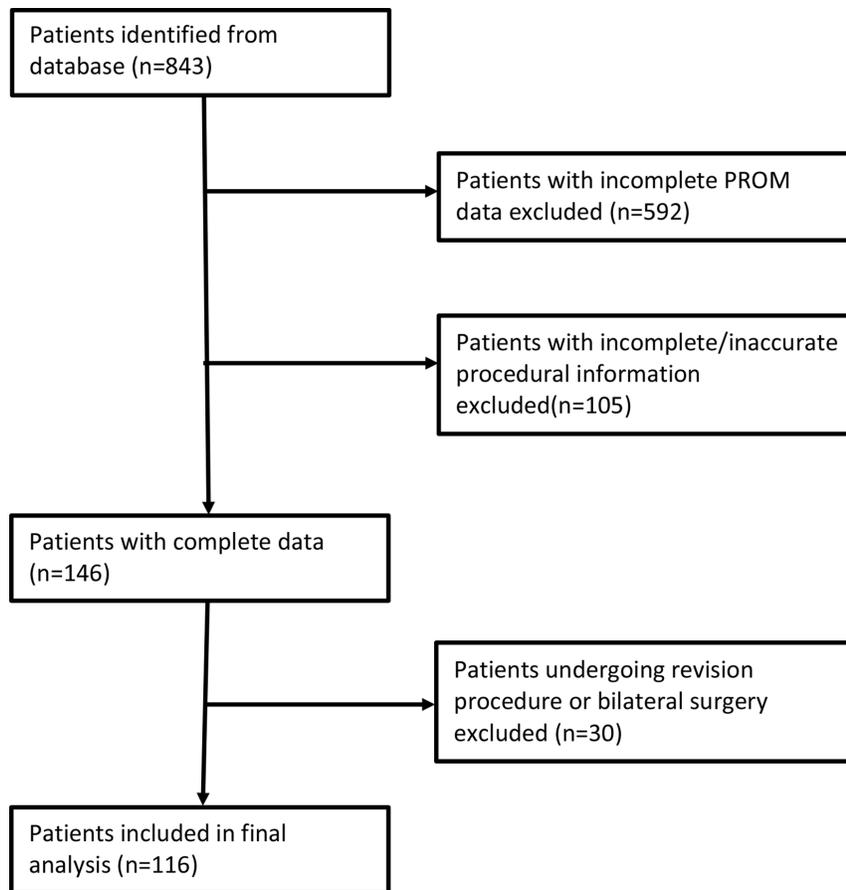
Fig. 1

Study flowchart. PROM, patient-reported outcome measure.

measurable health benefit for knee arthroplasty procedures to be considered 'successful', in terms of 'mean improvement'. Recent studies have shown that the commonly employed PROMs to evaluate hip arthroplasty and shoulder arthroplasty are indeed subject to a ceiling effect.[6,7]

The metabolic equivalent of task (MET) score is used as an alternative metric which considers the form and intensity of exercise or physical activity patients are able to complete. It asks patients to select three physical activities which are both important to them and affected by their joint problem. For each of these activities, respondents rate the intensity at which they are able to perform the activity on a visual scale (0 to 100). The Arizona State Universities compendium of activities provides a range of METs for each activity. Energy expenditure is quantified using the unit of MET, with one MET equating to the energy expenditure of an adult at rest, and approximately equal to 3.5 ml O_2/kg/min. A MET of 6 or more is associated with conditioning exercise: as a reference, running is calculated to expend 4.5 to 23 METs, with sprinting equivalent to a MET of 20.[8] The intensity reported by patients is used to determine where, within the reference range, the patient is currently performing their activity. The following formula is used to calculate this, where R_L is the lower value from the range of METs and R_H is the higher value. Q is the intensity. The MET score takes the maximum value scored from the three chosen activities.

$$(R_H - R_L) * Q + R_L$$

In large-scale cardiac studies and a national twin study, METs have been shown to inversely correlate with mortality.[9,10] With the ability to differentiate between a large number of activities, the MET score may be a useful tool to differentiate between patient outcomes in those with excellent postoperative function, beyond the basic activities of daily living assessed by the OKS and EQ-5D.

The primary aim of this study was to assess preoperative and postoperative OKS, EQ-5D, and MET scores in terms of their distribution, skewness, kurtosis, and whether they are subject to a significant ceiling effect. The secondary aim was to assess the usability of and validity of the MET score in a knee arthroplasty population.

Table I. Demographic data.

Variable	Total	Left	Right
Sex, n			
Male	53		
Female	63		
Mean age, yrs (SD)	65.64 (8.63)		
Mean BMI, kg/m² (range)	27.8 (17.56 to 70.40)		
Procedure n			
BCA-Lateral	3	1	2
BCA-Medial	4	1	3
Bi-UKA	2	1	1
PFA	1	0	1
TKA	7	1	6
UKA-Lateral	32	13	19
UKA-Medial	67	32	35
All procedures	116	49	67

BCA, bicompartmental arthroplasty; PFA, patellofemoral arthroplasty ; TKA, total knee arthroplasty; UKA, unicompartmental knee arthroplasty.

Fig. 2

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of Oxford Knee Scores (OKS) preoperatively. Red vertical lines represent mean values.

Methods

Study design. Data were collected prospectively from consenting patients undergoing primary knee arthroplasty procedures between 2014 and 2019. Operations were performed at two hospitals in London by a single surgeon (JC). Data were stored at Imperial College London in accordance with General Data Protection Regulation (GDPR).

Patients who had undergone primary knee arthroplasty procedures with recorded preoperative and postoperative (6 and/or 12 months) PROM data (OKS, EQ-5D, scores and MET scores) were included. Patients were excluded if they had undergone revision or bilateral procedures (Figure 1).

Patient demographic data. A total of 116 patients were included in this study. Key demographic details are listed in Table I. Overall, 99 patients underwent unicompartmental knee arthroplasty (UKA), seven total knee arthroplasty (TKA), seven bicompartmental arthroplasty (BCA), two bi-unicondylar arthroplasty (Bi-UKA), and one patellofemoral arthroplasty (PFA).[11,12]

Data collection and statistical analysis. The distribution of OKS, EQ-5D, and MET scores were analyzed and assessed for skewness, kurtosis, and ceiling and floor effects. Mean, median, and modal scores were examined. Visual assessments of normality were conducted using histograms. As per internationally accepted standards, ceiling effects were considered significant or consequential if maximum scores were recorded in > 15% of the assessed population. Similarly, floor effects were considered significant or consequential if > 15% of the cohort recorded minimum scores.[4,5,7,13] The MET was correlated with the OKS and EQ-5D Index to assess concurrent validity using Spearman's rank (STATA IC Version 13.0, Stata Corp, USA), with $p < 0.05$ considered significant. The responsiveness of scoring metrics was assessed to further examine their validity.
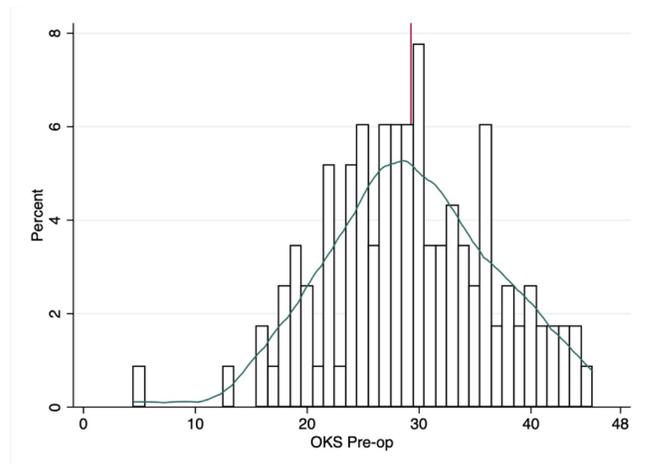
Standardized response means (SRMs) were calculated by dividing the mean change in score between two time periods by the standard deviation of that change.

Results

Distribution. The preoperative OKS was normally distributed, centred around a score of 29/48 (60%), (skewness -0.129, kurtosis 3.03, SD 7.45; Figure 2 and Table II). At six months postoperation, the OKS was no longer normally distributed. The median and modal scores were 43 (90%) and 44 (92%) respectively, with increased skewness and kurtosis (skewness -1.60, kurtosis 6.22; Figure 3 and Table II). By one year, the modal score was 47 points (98%), exhibiting pronounced skewness and kurtosis (skewness -3.30, kurtosis 18.01; Figure 4 and Table II).

Preoperatively, EQ-5D scores were not normally distributed, with a left skew and a leptokurtic distribution (skewness -1.41, kurtosis 4.93; Figure 5 and Table II). At six months and one year postoperatively, the distribution of EQ-5D had a more pronounced left skew and a higher kurtosis (skewness -1.91, kurtosis 9.60, Figure 6 and Table II; and skewness -2.11, kurtosis 8.57; Figure 7 and Table II, respectively) with a median of 0.84 (84%) and a modal score of 1 (100%).

Preoperative MET scores were normally distributed, centred around a mean score of 5 (SD 3.60), with a skew < 0.5 and kurtosis close to 3 (skewness 0.38, kurtosis 2.79; Figure 8, Table II). At six months' follow-up, MET scores were normally distributed around a mean score of 9 (SD 2.97) (skewness -0.25, kurtosis 4.48; Figure 9, Table II). At one-year follow-up the MET score remained normally distributed, with a mean score of 10 (SD 2.77) (skewness -0.05, kurtosis 5.76; Figure 10, Table II).

Ceiling and floor effects. No ceiling effects were seen for any metric preoperatively. The OKS did not demonstrate

Table II. Pre- and postoperative patient-reported outcome measures – mean, median, and modal scores, skewness and kurtosis values, and floor and ceiling effects.

Score	Patients, n	Mean (SD)	Median (IQR)	Modal score	Skewness	Kurtosis	Floor effect %	Ceiling effect, %
OKS								
Preoperative	116	29.3 (7.45)	29 (24.8 to 34.3)	30	-0.1	3.0	0.00	0.00
6 mths	96	41.8 (5.36)	43 (39 to 46)	44	-1.6	6.2	0.00	6.25
12 mths	83	43.9 (5.46)	45 (43.5 to 47)	47	-3.3	18.0	0.00	15.66*
EQ-5D								
Preoperative	116	0.58 (0.20)	0.64 (0.48 to 0.71)	0.74	-1.4	4.9	0.00	0.86
6 mths	96	0.82 (0.17)	0.84 (0.74 to 1)	1	-1.9	9.6	0.00	30.21*
12 mths	83	0.85 (0.19)	0.84 (0.8 to 1)	1	-2.1	8.6	0.00	39.76*
MET								
Preoperative	116	5.22 (3.60)	4.9 (3.09 to 7.28)	0	0.4	2.8	18.10†	0.00
6 mths	96	9.33 (2.97)	9.5 (7.86 to 11.02)	7	0.3	4.5	2.08	0.00
12 mths	83	9.90 (2.77)	10.2 (8.55 to 11.25)	8.55	-0.1	5.8	1.20	0.00

*Significant ceiling effects.
†Significant floor effects.
EQ-5D, EuroQol five-dimension index; MET, metabolic equivalent of task; OKS, Oxford Knee Score.
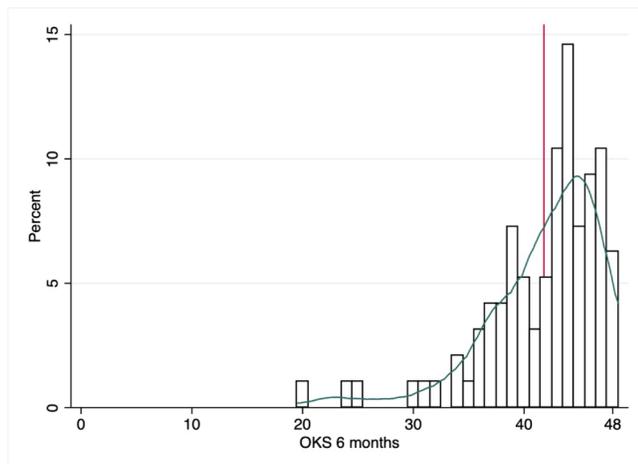
Fig. 3

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of Oxford Knee Scores (OKS) at six months postoperatively. Red vertical lines represent mean values.
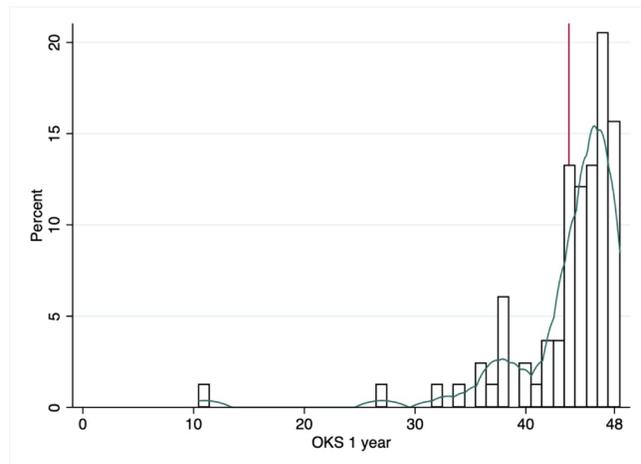
Fig. 4

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of Oxford Knee Scores (OKS) at one year postoperatively. Red vertical lines represent mean values.

a consequential ceiling effect at six months (6.2%), but did so at one-year follow-up (15.7%). The EQ-5D demonstrated a consequential ceiling effect at both six months (30.2%) and one year (39.8%) post-surgery. No ceiling effects were observed for the MET score at any timepoint.

Floor effects were not observed for the OKS or EQ-5D at any timepoint. The MET score demonstrated a consequential floor effect (18.1%) preoperatively. The MET score did not demonstrate a consequential floor effect at either six months (2.1%) or one year (1.2%) following surgery.

Validity. The MET score demonstrated weak-moderate correlations with the OKS both preoperatively ($r = 0.3159$, $p < 0.001$, Spearman's rank correlation), at six months ($r = 0.51$, $p < 0.001$, Spearman's rank correlation), and

one year postoperation ($r = 0.23$, $p = 0.030$, Spearman's rank correlation). The MET score demonstrated weak-moderate correlations with the EQ-5D preoperatively ($r = 0.36$, $p < 0.001$), at six months ($r = 0.48$, $p < 0.001$) and one year ($r = 0.29$, $p < 0.001$) postoperation.

Responsiveness. All metrics displayed high levels of responsiveness, with SRMs > 0.9 (Table III).

Discussion

This study of prospectively gathered data set out to assess the performance of the OKS, EQ-5D, and MET score in terms of distribution and ceiling effect. Both the OKS and EQ-5D demonstrated skewed and excessively kurtotic distributions at six months and, to a greater extent, at one year following surgery. By comparison, the MET score
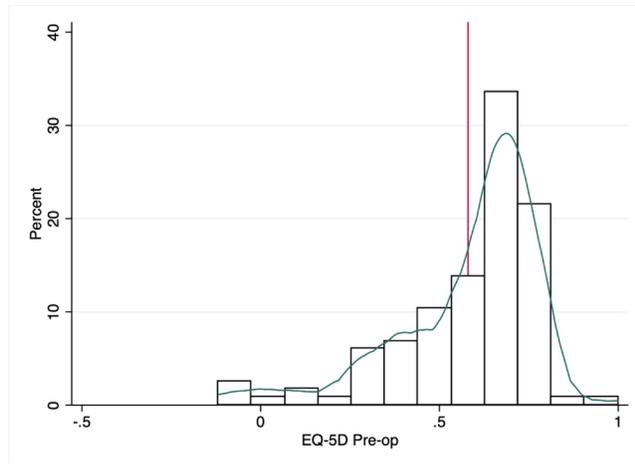
Fig. 5

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of EuroQol five-dimension (EQ-5D) scores preoperatively. Red vertical lines represent mean values.
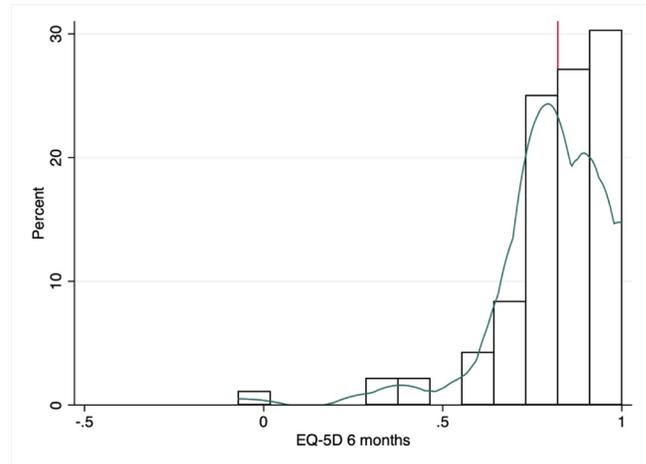

Fig. 6

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of EuroQol five-dimension (EQ-5D) scores at six months postoperatively. Red vertical lines represent mean values.

was not skewed and exhibited normal kurtosis consistent with normally distributed data both at six months and one year following surgery.

The implications of excessive skewness and kurtosis in a PROM dataset are manifold. Firstly, a mean score taken from a dataset will not be a typical representation of the dataset, necessitating the use of a median score.[14] Excessive skewness and kurtosis represent a violation of normality and contraindicate the use of parametric methods of analysis such as the *t*-test. Non-parametric tests do not require a normal distribution; however, they are significantly less powerful than parametric tests, leading to higher rates of Type 2 errors, producing results which are more difficult to interpret.[15] Funnel plots are used in the UK to identify outlying performers from PROM datasets for knee arthroplasty procedures. Deriving funnel plots from skewed datasets has been shown to increase the number of providers incorrectly identified as poor performers.[16]

A problematic ceiling effect was observed in this cohort with the OKS (16% at one year postoperatively). The modal score was close to the maximum at both six months and one year, with high levels of skewness and kurtosis. These findings are supported by National Joint Registry (NJR) data published in 2015 whereby 19,974 patients completed the OKS at one year following knee replacement, reporting a modal score of 48 and a high left skew in distribution.[17]

Others have made different observations, with Harris et al[2] concluding that the OKS did not exhibit a significant postoperative ceiling effect. Their study examined an NHS dataset from procedures between 2009 and 2011, with only 2.7% of 72,154 patients scoring maximally. However, the study did not measure scores beyond six months, which may present a significant limitation, given

that clinically important improvements in OKS scores between six and 12 months are expected.[18]

In our study, the EQ-5D demonstrated a problematic ceiling effect at both six-month (30%) and one-year (40%) follow-up. The modal postoperative score was 1 at both timepoints, with marked skew and kurtosis. The NJR report of 2015 recorded an EQ-5D modal score of 1 for 19,416 patients at one year postoperatively, supporting our findings.[17]

In contrast to the OKS and EQ-5D, we did not observe a ceiling effect with the MET score, with a normal distribution at both six months and one year postoperatively. Patients demonstrated high mean scores of 9 and 10 at these time intervals, reflecting that many were able to return to physical activity or exercise. This finding supports literature suggesting the majority of patients are able to return to sporting activity following knee arthroplasty.[19-21]

A preoperative floor effect in the MET was evidenced by the number of patients scoring 0, (i.e. unable to complete the desired activity). This limits the usefulness of the MET as a preoperative assessment tool and highlights that it will provide limited information for patients who do not engage in physical activity prior to or after knee arthroplasty surgery. In this cohort, preoperative OKS scores were normally distributed, with neither a ceiling nor floor effect. The converse preoperative floor and postoperative ceiling effects of the MET score and OKS score, respectively, indicate that it may be necessary to stratify patients (by level of physical activity or preoperative OKS, for example) in order to determine how best to evaluate patients following knee arthroplasty. For patients who are of low demand and for whom the sole aim of surgery is pain relief, a maximal score in the OKS may still represent an optimal outcome.
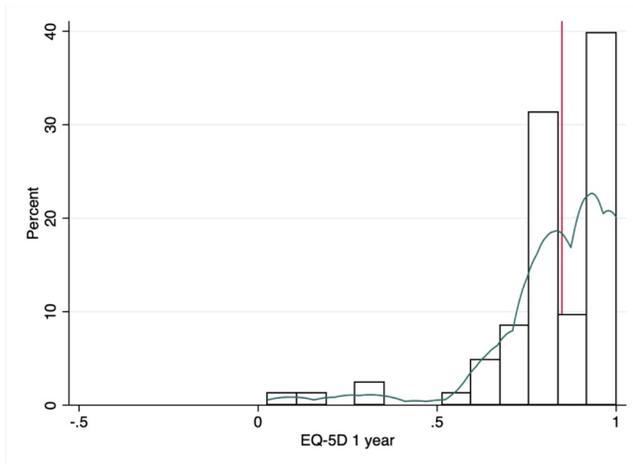
Fig. 7

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of EuroQol five-dimension (EQ-5D) scores at one year postoperatively. Red vertical lines represent mean values.
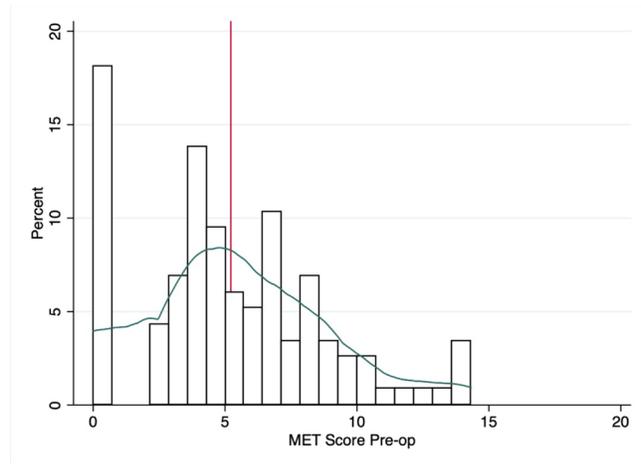
Fig. 8

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of metabolic equivalent of task (MET) scores preoperatively. Red vertical lines represent mean values.

The implications of our findings are potentially significant when considering how contemporary research uses PROMs to evaluate knee arthroplasty procedures. The Total or Partial Knee Arthroplasty Trial used the OKS as the primary outcome measure when comparing TKA versus PKA.[22] The power calculation used in the trial design was based upon detecting a difference of three points despite the modal score for TKA in the NJR being 48.[17] The skewness and kurtosis of the OKS suggest it may not have been the most appropriate outcome measure to answer the important questions posed by the study.

A Norwegian study concluded that TKA offers a greater improvement in health state to patients as the radiological severity of osteoarthritis increases, using the EQ-5D as the endpoint at 12 months post-surgery.[23] The severe ceiling effect demonstrated with the EQ-5D calls this conclusion into question, as when the most common score postoperatively is 1, patients with more severe arthritis have more to gain than patients whose preoperative scores are closer to the ceiling.

Alternatives to the OKS and EQ-5D which are not subject to postoperative ceiling effects include the Forgotten Joint Score (FJS) and PROMIS Global-10.[24–26] The FJS assesses treatment success as the ability of a patient to forget about the presence of their prosthetic joint, and is designed for shoulder, hip, and knee arthroplasty. Comparative studies between the FJS and OKS have shown that the FJS is free of the ceiling effect the OKS demonstrates. However, a recent study suggested the FJS is subject to a postoperative floor effect, which limits its ability to discriminate between patients with poorer outcomes.[27] The PROMIS Global-10 is a metric performing in a similar role to the EQ-5D in capturing information about general health. A study by Khalil et al[28] demonstrates that PROMIS Global-10 scores are not subject to floor or ceiling effects in their cohort of 875 patients undergoing TKA. The PROMIS Global-10 has not yet been assessed in patients undergoing other forms of knee arthroplasty.

Our study also set out to evaluate the validity of the MET score when used in a knee arthroplasty population. MET-based scores such as the International Physical Activity Questionnaire (IPAQ) have been validated for use in the general population.[29] Here, the MET score demonstrated SRMs > 1, comparable to those for the OKS and EQ-5D, indicating a large effect size and/or excellent responsiveness.[30] The MET score displayed weak-moderate correlations between the OKS and EQ-5D at six months and one year following surgery, which may be due to the ceiling effect associated with these two scores. As the MET score is measuring an alternative aspect of patient function compared to conventional PROMs, correlation against related metrics (e.g. gait analysis or pedometer readings) may be required to further assess its validity.

Alternative rating scales have been devised to assess physical activity following arthroplasty. The Activity & Participation Questionnaire (OKS-APQ) is intended to complement the OKS by addressing issues relevant to younger, more active patients.[31] However, this score does not take any consideration of either the form or intensity of exercise performed, and no assessment has yet been made regarding the presence of ceiling effects. Similarly, the UCLA activity scale performed well when correlated to the IPAQ, but again does not make any assessment of activity intensity, nor has the presence of a ceiling effect been investigated.[32]

A strength of the current study is that data were acquired prospectively, though the study question and analysis were retrospective. Only patients for whom complete data were available were included, facilitating

Table III. Responsiveness of patient-reported outcome measures shown by standardized response means at six months and one year.

PROM	OKS (6 mths)	OKS (1 yr)	EQ-5D (6 mths)	EQ-5D (1 yr)	MET (6 mths)	MET (1 yr)
SRM	0.98	1.37	1.55	1.97	1.00	1.01

EQ-5D, EuroQol five-dimension index; MET, metabolic equivalent of task score; OKS, Oxford Knee Score; PROM, patient-reported outcome measure; SRM, standardized response mean.
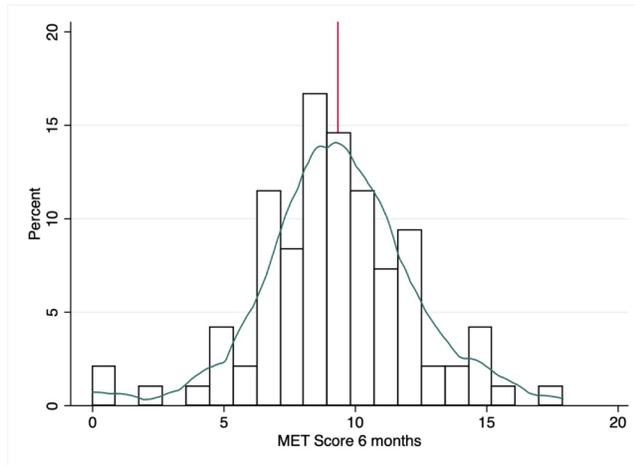
Fig. 9

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of metabolic equivalent task (MET) scores at six months postoperatively. Red vertical lines represent mean values.
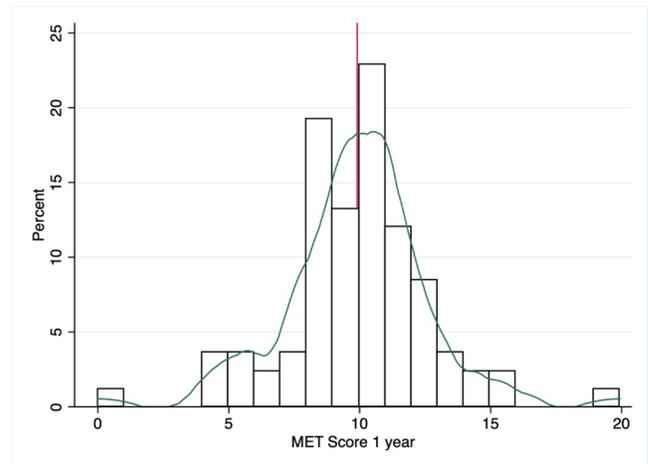
Fig. 10

Histogram with kernel density (Epanechnikov) plot demonstrating distribution of metabolic equivalent of task (MET) scores at one year postoperatively. Red vertical lines represent mean values.

a risk of selection bias, particularly given that a large number of patients (70.2%, n = 592) did not have both preoperative and postoperative PROM data available. It is certainly possible that patients with outcomes at either end of the range of satisfaction were less likely to respond to questionnaires. It should also be noted that for the MET score, patients could potentially select an activity, such as canoeing/kayaking, which does not require use of the knee joint. This did not apply to patients included in this study, and is mitigated by ensuring patients choose activities affected by their joint disease. However, some caution should be taken when interpreting the MET score without contextualizing the chosen activities.

It is both a strength and a limitation that all types of primary knee arthroplasty were eligible for inclusion. The range of procedures suggests that the findings are not confined to one particular procedure, though a much larger cohort would be required for detailed subgroup analysis of this heterogenous group. That said, the focus of this study was to test the performance of PROMs rather than a particular type of knee arthroplasty.[33,34] The numbers of patients used in this study to assess the MET are relatively small in comparison to those used to validate widely used PROMS such as the OKS or EQ-5D. The findings of this study would carry more weight with a larger cohort. The authors plan to further assess and validate the MET in future analyses.

Postoperative MET scores were normally distributed, without a ceiling effect in our cohort of patients undergoing primary knee arthroplasty. This contrasts with the OKS and EQ-5D, which were observed to have negatively skewed and highly kurtotic distributions, with significant ceiling effects. The MET score may be a valuable adjunct to these established scores, capturing and quantifying health gains in high-functioning patients whose procedures would otherwise be classified as delivering an insufficient health gain. It may also allow healthcare providers to make a more accurate assessment of a knee arthroplasty procedure's value in relation to general health and, indeed, mortality.

Take home message

- Postoperatively, the Oxford Knee Score and EuroQol five-dimension index demonstrate ceiling effects, and skewed and highly kurtotic distributions, failing to fully capture health gains.
- The metabolic equivalent of task (MET) is an alternative metric, assessing the form and intensity of exercise and not subject to a postoperative ceiling effect.
- To make a more accurate assessment of a knee arthroplasty procedure's health value, clinicians and researchers should consider using the MET as an adjunct to conventional patient-reported outcome measures.

Twitter

Follow A. Patel @ArjunP2291
Follow T. C. Edwards @edwards_tomc
Follow J. Cobb @orthorobodoc

Follow A. Garner @DrAmyGarner

References

1. **Price AJ**, **Alvand A**, **Troelsen A**, **et al**. Knee replacement. *Lancet*. 2018;392(10158):1672–1682.

2. **Harris K**, **Lim CR**, **Dawson J**, **Fitzpatrick R**, **Beard DJ**, **Price AJ**. The Oxford knee score and its subscales do not exhibit a ceiling or a floor effect in knee arthroplasty patients: an analysis of the National Health Service PROMs data set. *Knee Surg Sports Traumatol Arthrosc*. 2017;25(9):2736–2742.

3. **Briggs T**. A national review of adult elective orthopaedic services in England: Getting it right first time. British Orthopaedic Association. 2015. https://gettingitrighttfirsttime.co.uk/wp-content/uploads/2018/07/GIRFT-National-Report-Mar15-Web.pdf (date last accessed 8 February 2023).

4. **McHorney CA**, **Tarlov AR**. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293–307.

5. **Terwee CB**, **Bot SDM**, **de Boer MR**, **et al**. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.

6. **Singh HP**, **Haque A**, **Taub N**, **et al**. Floor and ceiling effects in the Oxford Shoulder Score: an analysis from the National Joint Registry. *Bone Joint J*. 2021;103-B(11):1717–1724.

7. **Edwards TC**, **Guest B**, **Garner A**, **Logishetty K**, **Liddle AD**, **Cobb JP**. The metabolic equivalent of task score: a useful metric for comparing high-functioning hip arthroplasty patients. *Bone Joint Res*. 2022;11(5):317–326.

8. **Kujala UM**, **Kaprio J**, **Sarna S**, **Koskenvuo M**. Relationship of leisure-time physical activity and mortality: the Finnish twin cohort. *JAMA*. 1998;279(6):440–444.

9. **Booth FW**, **Laye MJ**, **Roberts MD**. Lifetime sedentary living accelerates some aspects of secondary aging. *J Appl Physiol (1985)*. 2011;111(5):1497–1504.

10. **Kokkinos P**, **Faselis C**, **Sidossis L**, **et al**. Exercise blood pressure, cardiorespiratory fitness and mortality risk. *Prog Cardiovasc Dis*. 2021;67:11–17.

11. **Garner A**, **van Arkel RJ**, **Cobb J**. Classification of combined partial knee arthroplasty. *Bone Joint J*. 2019;101-B(8):922–928.

12. **Garner A**, **Cobb J**. Combined partial knee arthroplasty. In: *Personalized Hip and Knee Joint Replacement*. Springer International Publishing, 2020: 243–254.

13. **Lim CR**, **Harris K**, **Dawson J**, **Beard DJ**, **Fitzpatrick R**, **Price AJ**. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open*. 2015;5(7):e007765.

14. An introduction to analysing quality improvement & assurance data. Healthcare Quality Improvement Partnership. 2018. https://www.hqip.org.uk/wp-content/uploads/2018/10/final-an-introduction-to-data-analysis-october-2018.pdf (date last accessed 8 February 2023).

15. **Cappelleri JC**, **Zou KH**, **Bushmakin AG**, **Alvir Jm**, **Alemayehu D**, **Symonds T**. *Patient-Reported Outcomes: Measurement, Implementation and Interpretation*. New York, USA: Chapman and Hall, 2014.

16. **Neuburger J**, **Cromwell DA**, **Hutchings A**, **Black N**, **van der Meulen JH**. Funnel plots for comparing provider performance based on patient-reported outcome measures. *BMJ Qual Saf*. 2011;20(12):1020–1026.

17. **NJR Editorial Board**. 12th Annual Report 2015 National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. 2015. https://www.hqip.org.uk/wp-content/uploads/2018/02/national-joint-registry-12th-annual-report-2015.pdf (date last accessed 8 February 2023).

18. **Browne JP**, **Bastaki H**, **Dawson J**. What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. *Health Qual Life Outcomes*. 2013;11(1):1–7.

19. **Hopper GP**, **Leach WJ**. Participation in sporting activities following knee replacement: total versus unicompartmental. *Knee Surg Sports Traumatol Arthrosc*. 2008;16(10):973–979.

20. **Bradbury N**, **Borton D**, **Spoo G**, **Cross MJ**. Participation in sports after total knee replacement. *Am J Sports Med*. 1998;26(4):530–535.

21. **Fisher N**, **Agarwal M**, **Reuben SF**, **Johnson DS**, **Turner PG**. Sporting and physical activity following Oxford medial unicompartmental knee arthroplasty. *Knee*. 2006;13(4):296–300.

22. **Beard DJ**, **Davies LJ**, **Cook JA**, **et al**. The clinical and cost-effectiveness of total versus partial knee replacement in patients with medial compartment osteoarthritis (TOPKAT): 5-year outcomes of a randomised controlled trial. *Lancet*. 2019;394(10200):746–756.

23. **Rehman Y**, **Lindberg MF**, **Arnljot K**, **Gay CL**, **Lerdal A**, **Aamodt A**. More severe radiographic osteoarthritis is associated with increased improvement in patients' health state following a total knee arthroplasty. *J Arthroplasty*. 2020;35(11):3131–3137.

24. **Robinson PG**, **Rankin CS**, **Lavery J**, **Anthony I**, **Blyth M**, **Jones B**. The validity and reliability of the modified forgotten joint score. *J Orthop*. 2018;15(2):480–485.

25. **Thomsen MG**, **Latifi R**, **Kallemose T**, **Barfod KW**, **Husted H**, **Troelsen A**. Good validity and reliability of the forgotten joint score in evaluating the outcome of total knee arthroplasty: A retrospective cross-sectional survey-based study. *Acta Orthop*. 2016;87(3):280.

26. **Hays RD**, **Bjorner JB**, **Revicki DA**, **Spritzer KL**, **Cella D**. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Qual Life Res*. 2009;18(7):873–880.

27. **Lee QJ**, **Chang WYE**, **Wong YC**. Forgotten Joint Score for early outcome assessment after total knee arthroplasty: Is it really useful? *Knee Surg Relat Res*. 2020;32(1):37.

28. **Khalil LS**, **Darrith B**, **Franovic S**, **Davis JJ**, **Weir RM**, **Banka TR**. Patient-Reported Outcomes Measurement Information System (PROMIS) global health short forms demonstrate responsiveness in patients undergoing knee arthroplasty. *J Arthroplasty*. 2020;35(6):1540–1544.

29. **Hagströmer M**, **Oja P**, **Sjöström M**. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr*. 2006;9(6):755–762.

30. **Liang MH**, **Fossel AH**, **Larson MG**. Comparisons of five health status instruments for orthopedic evaluation. *Med Care*. 1990;28(7):632–642.

31. **Dawson J**, **Beard DJ**, **McKibbin H**, **Harris K**, **Jenkinson C**, **Price AJ**. Development of a patient-reported outcome measure of activity and participation (the OKS-APQ) to supplement the Oxford Knee Score. *Bone Joint J*. 2014;96-B(3):332–338.

32. **Naal FD**, **Impellizzeri FM**, **Leunig M**. Which is the best activity rating scale for patients undergoing total joint arthroplasty? *Clin Orthop Relat Res*. 2009;467(4):958–965.

33. **Sephton BM**, **Edwards TC**, **Bakhshayesh P**, **Nathwani D**. Should we routinely perform a post-operative hemoglobin check following unicompartmental knee arthroplasty? *Knee*. 2020;27(1):249–256.

34. **Sephton BM**, **Bakhshayesh P**, **Edwards TC**, **Ali A**, **Kumar Singh V**, **Nathwani D**. Predictors of extended length of stay after unicompartmental knee arthroplasty. *J Clin Orthop Trauma*. 2020;11(Suppl 2):S239–S245.

Author information:
- A. Patel, MBBS, BSc, MRCS, Clinical Research Fellow
- T. C. Edwards, MBBS, BSc, MRCS, Clinical Research Fellow
- G. Jones, BSc(Hons), MBBS, FRCS (Tr&Orth), PhD, Clinical Senior Lecturer
- A. D. Liddle, MBBS, BSc, DPhil, FRCS (Orth), Clinical Senior Lecturer
- J. Cobb, BM, BCh, MCh, FRCS, Chair in Orthopaedic Surgery
- A. Garner, BMBCh, MA, FRCS, PhD, Clinical Research Fellow
  MSk Lab, Sir Michael Uren Biomedical Engineering Research Hub, Imperial College London, London, UK.