



Outcome measures in total hip arthroplasty: have our metrics changed over 15 years?

Roy H. Lan¹ · Jack W. Bell¹ · Linsen T. Samuel² · Atul F. Kamath² 

Received: 27 September 2020 / Accepted: 25 January 2021 / Published online: 11 February 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract

Introduction Consensus has not been reached regarding ideal outcome measures for total hip arthroplasty (THA) clinical evaluation and research. The goal of this review was to analyze the trends in outcome metrics within the THA literature and to discuss the potential impact of instrument heterogeneity on clinical practice.

Materials and methods A PubMed search of all manuscripts related to THA from January 2005 to December 2019 was performed. Statistical and linear regression analyses were performed for individual outcome metrics as a proportion of total THA publications over time.

Results There was a statistically significant increase in studies utilizing outcomes metrics between 2005 and 2019 (15.1–29.5%; $P < 0.001$; $R^2 = 98.1\%$). Within the joint-specific subcategory, use of the Harris Hip Score (HHS) significantly decreased from 2005 to 2019 (82.8–57.3%; $P < 0.001$), use of the Hip Disability and Osteoarthritis Outcome Score (HOOS) significantly increased (0–6.7%; $P < 0.001$), and the modified HHS significantly increased (0–10.5%; $P < 0.001$). In the quality of life subcategory, EQ-5D demonstrated a significant increase in usage (0–34.8%; $P < 0.001$), while Short Form-36 significantly decreased (100% vs. 27.3%; $P = 0.008$).

Conclusions The utilization of outcome-reporting metrics in THA has continued to increase, resulting in added complexity within the literature. The utilization rates of individual instruments have shifted over the past 15 years. Additional study is required to determine which specific instruments are recommended.

Keywords Osteoarthritis · Total hip arthroplasty (replacement) · Patient-reported outcome measures (PROMs) · Quality · Outcomes

Introduction

Total hip arthroplasty (THA) provides significant improvement in both mobility and quality of life for patients with osteoarthritis (OA) of the hip [1]. A recent epidemiologic study projected an increase in THA by 71% to 635,000 procedures per year by 2030 [2, 3]. As the prevalence of OA and the procedure volumes of THA continue to grow, focused

use of outcome metrics will be required as clinicians and researchers seek to evaluate the effectiveness of THA.

As payers and providers seek to standardize quality and costs of care for THA, patient-reported outcome measures (PROMs) define success after THA [4, 5]. The generation of new outcome metrics has increased, as researchers seek to better understand the perceived nuances of post-surgical populations. These new metrics often refine older metrics, applying their strengths while also introducing novel methods of analysis [6]. Across the literature, both the number of individual types of outcome metrics and the frequency of their utilization continue to rise, generating increased heterogeneity that may limit the ability to generate reliable comparisons [7–9].

This increase in outcome measures may be driven in part by the Centers for Medicare & Medicaid Services (CMS) merit-based incentive payment system (MIPS) and a contemporary environment that incentivizes clinicians to collect

✉ Atul F. Kamath
kamatha@ccf.org

¹ College of Medicine, University of Tennessee Health Science Center, 910 Madison Ave, Memphis, TN 38163, USA

² Department of Orthopaedic Surgery, Center for Hip Preservation, Orthopaedic and Rheumatologic Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue Mail Code A40, Cleveland, OH 44195, USA

value-focused metrics [10, 11]. Moving forward, it is important that the THA community gain consensus regarding which outcome metrics are most preferred to generate useful measure and to facilitate clinical practice recommendations.

Although a number of studies have sought to validate individual PROMs within a population cohort, few studies have sought to understand the underlying trends in utilization of these outcome metrics [1, 5, 7]. Although the utilization rates of each of these outcome metrics continues to change in THA reporting, no review has thus far considered these trends in detail. Investigating the rise or decline of commonly used metrics may also reveal new trends in clinician preferences over time. These trends will provide a perspective on which outcome metrics have become more commonly used over time, as well as those that have been rendered obsolete or less preferred by the THA provider community.

To better understand the complexity in outcomes reporting within the THA literature, this review examined the trends in utilization of outcome metrics. We hypothesized that outcome metric utilization would demonstrate significant heterogeneity, and that utilization rates of numerous outcome metrics may have significantly changed over the past 15 years.

Materials and methods

A review of studies published between January 1, 2005 and December 31, 2019 was performed to obtain relevant quality metrics in THA. Quality metrics were separated into five subcategories: joint-specific, joint-agnostic, general health, quality of life, and patient satisfaction. These subcategories were chosen as they are commonly utilized to categorize outcomes variables when evaluating THA outcomes [12–14].

The Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Hip Disability and Osteoarthritis Outcome Score Short Form, Joint Replacement (HOOS-JR), Modified Harris Hip Score (mHHS), Oxford Hip Score (OHS), and Western Ontario McMaster Universities Arthritis Index (WOMAC) were the outcomes metrics categorized as joint-specific. These metrics are all joint-specific metrics used to evaluate THA [15–17]. The Numeric Rating Scale (NRS), University of California at Los Angeles Activity Score (UCLA), and Visual Analogue Scale for Pain (VAS) were the outcomes metrics categorized as joint-agnostic [18–20]. These metrics are all joint-agnostic metrics used to evaluate THA that are not primarily directed at general health, quality of life, or patient satisfaction. The Forgotten Joint Score-12 (FJS-12), Patient-Reported Outcome Measurement Information System Global-10 (PROMIS-10), and Veterans RAND 12-Item Survey (VR-12) were the outcomes metrics categorized under general

health [12, 21]. These metrics are all metrics used to primarily evaluate general health. The EuroQol 5-Dimension Health Outcome Survey (EQ-5D), Short Form-12 Health Survey (SF-12), and Short Form-36 Health Survey (SF-36) were the outcomes metrics categorized under quality of life [10, 22, 23]. These metrics are all metrics used to primarily evaluate quality of life. Lastly, a literature search revealed that patient satisfaction is often evaluated in THA; however, specific outcomes metrics focused solely on patient satisfaction are not yet available for THA [16, 18, 21, 24]. Thus, papers reporting “Patient Satisfaction” were also included in this analysis.

A review of the historic literature was carried out for the outcome metrics selected in this study [10, 12, 15, 21, 25–28]. Information regarding the year they were individually developed, their validity in measuring THA outcomes, and number of questions is listed in Table 1. It is important to note that HOOS-JR was first validated in 2008, after the first date of data collection within the study [24]. The same can be said for the FJS-12, which was validated in 2012 [29]. All other outcome metrics were validated prior to the first data collection period within this study.

The aggregate number of primary THA papers in the literature was determined using the following PubMed search: Hip Arthroplasty [Title/Abstract] AND (“2005/01/01” [Date—Publication]: “2019/12/31” [Date—Publication]) NOT Hemiarthroplasty [Title/Abstract] NOT resurfacing [Title/Abstract]. A PubMed search was performed for each individual outcome metric using the following search template: Hip Arthroplasty [Title/Abstract] AND (X) [Title/Abstract] AND (“2005/01/01” [Date—Publication]: “2019/12/31” [Date—Publication]) NOT Hemiarthroplasty [Title/Abstract] NOT resurfacing [Title/Abstract]; X = quality metric searched. Paper titles and abstracts were then manually screened to ensure that each study included within the analysis not only was related to primary THA but also contained the quality metric searched. Review papers, protocols, and articles not primarily related to THA were eliminated from the analysis.

The number of papers utilizing each outcome metric per year between January 1, 2005 and December 31, 2019 was tabulated from its specific PubMed search. The number of papers within each subcategory of outcome metrics (joint-specific, joint-agnostic, general health, and quality of life) per year was determined by adding the number of papers using each outcome metric within their respective category (Table 2). The total number of studies reporting outcomes metrics per year was found by adding the total number of articles from each outcome subcategory.

The percentage of manuscripts with each outcome metric was calculated for each category. The frequency of each outcome metric used each year was also determined. Linear regression analysis was performed to evaluate the fraction

Table 1 Summary of outcome metrics and corresponding outcome tools

Outcome metrics	Abbreviation	Year developed	Validation	Number of questions
Joint specific				
Harris Hip Score	HHS	1969	Validated	10
Hip Disability and Osteoarthritis Outcome Score	HOOS	2003	Validated	40
Hip Disability and Osteoarthritis Outcome Score Short Form (Joint Replacement)	HOOS-JR	2008	Validated	3
Modified Harris Hip Score	mHHS	1985	Validated	8
Oxford Hip Score	OHS	1996	Validated	12
Western Ontario and McMaster Universities Arthritis Index	WOMAC	1982	Validated	24
Joint-agnostic				
Numeric Rating Scale	NRS			1
University of California at Los Angeles Activity Score	UCLA	1984	Validated	1
Visual Analogue Scale for Pain	VAS	1920	Validated	1
General health				
Forgotten Joint Score-12	FJS-12	2012	Validated	12
Patient-Reported Outcome Measurement Information System Global-10	PROMIS-10	2004		10
Veterans Rand 12-Item Survey	VR-12	1997	Validated	12
Quality of life				
EuroQol 5-Dimension Health Outcome Survey	EQ-5D	1990	Validated	6
Short Form-12 Health Survey	SF-12	1996	Validated	12
Short Form-36 Health Survey	SF-36	1992	Validated	36
Patient satisfaction	N/A			Varies

of THA papers that included at least one outcomes variable. In addition, linear regression analysis was used to evaluate the fraction of all THA papers that included each outcomes subcategory, and how frequently each outcome metric within each subcategory was utilized. For all regression analyses carried out within the study, the publication year served as the independent variable. A $P < 0.01$ indicated a significant change in utilization of an outcome metric over time. This value was selected to adjust for multiple comparisons within the study.

Results

From January 1, 2005 to December 31, 2019, there were 14,744 THA studies published in the English language. Of these, 3736 (25.3%) non-duplicate articles were found to utilize at least one of the selected outcomes metrics, and therefore, met the inclusion criteria of the study (Fig. 1). There was a statistically significant increase in studies utilizing outcomes metrics between 2005 and 2019 [78/515 (15.1%) vs. 522/1772 (29.5%), respectively; $P < 0.001$; $R^2 = 98.1\%$] (Fig. 2).

Of the five outcome subcategories considered within this analysis, three demonstrated a significant change in use between 2005 and 2019 (Fig. 3). The Joint-specific

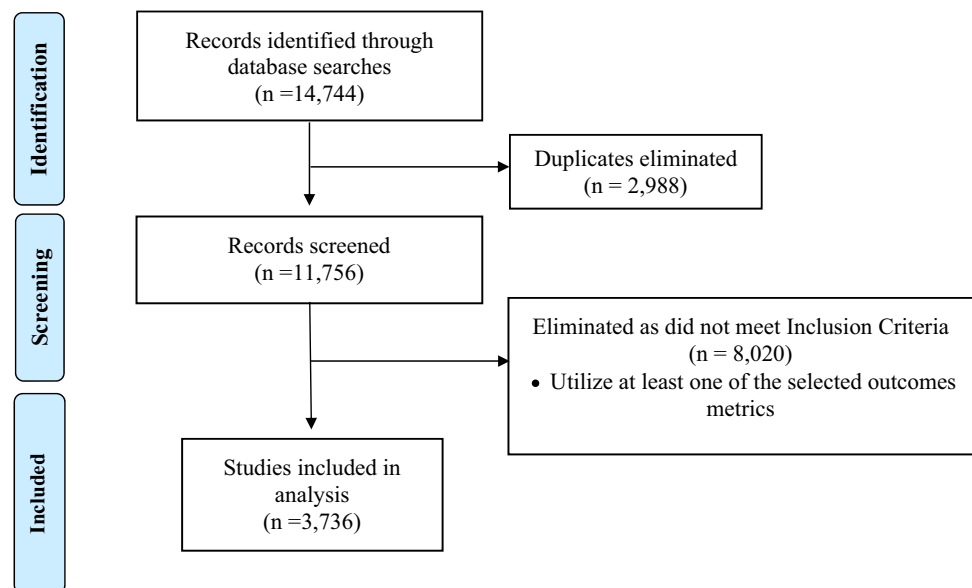
subcategory significantly decreased from 2005 to 2019 [58/78 (74.4%) vs. 314/522 (60.2%), respectively; $P < 0.001$]. In addition, joint-agnostic reporting increased from 2005 to 2019 [6/78 (7.69%) vs. 71/522 (13.6%) respectively; $P < 0.001$]. Lastly, reporting for the general health subcategory significantly increased from 2005 to 2019 [0/78 (0%) vs. 71/522 (3.3%), respectively; $P < 0.001$]. However, quality of life reporting did not significantly change from 2005 to 2019 [5/78 (6.4%) vs. 66/522 (12.6%), respectively; $P = 0.607$], nor did patient satisfaction [9/78 (11.5%) vs. 54/522 (10.3%), respectively; $P = 0.905$].

Within the joint-specific subcategory, use of HHS significantly decreased from 2005 to 2019 [48/58 (82.8%) vs. 180/314 (57.3%), respectively; $P < 0.001$], use of HOOS significantly increased from 2005 to 2019 [0/58 (0%) vs. 21/314 (6.7%), respectively; $P < 0.001$], and use of mHHS significantly increased from 2005 to 2019 [0/58 (0%) vs. 33/314 (10.5%), respectively; $P < 0.001$] (Fig. 4). Within the joint-agnostic subcategory, there were no significant changes in utilization for any of the three outcomes metrics within this study: NRS, UCLA, and VAS. Within the general health subcategory, PROMIS-10 demonstrated a significant increase in usage from 2005 to 2019 [0/0 (0%) vs. 9/17 (29.4%), respectively; $P < 0.001$]. In addition, VR-12 demonstrated a significant increase in usage from 2005 to 2019 [0/0 (0%) vs. 3/17 (17.6%), respectively; $P < 0.001$]. Lastly,

Table 2 Utilization of outcome metrics within each subcategory between 2005 and 2019

Outcome metrics	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Joint specific	58	77	82	93	117	113	147	118	156	156	226	244	244	268	314
HHS	48	57	65	67	80	84	106	77	97	101	141	145	141	137	180
HOOS	0	0	2	0	0	2	3	5	6	5	11	16	19	20	21
HOOS-JR	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5
mHHS	0	1	1	2	4	2	5	3	6	7	16	22	22	40	33
OHS	2	7	4	5	15	3	12	12	20	12	17	20	27	39	30
WOMAC	8	12	10	19	18	22	21	21	27	31	41	41	35	31	45
Joint-agnostic	6	3	9	6	13	14	21	23	34	38	36	73	74	68	71
NRS	1	0	0	1	1	0	3	0	4	0	2	4	10	5	5
UCLA	0	1	2	1	2	3	10	8	8	13	10	17	14	9	10
VAS	5	2	7	4	10	11	8	15	22	25	24	52	50	54	56
General health	0	0	0	0	0	0	0	0	0	0	1	6	9	11	17
FJS-12	0	0	0	0	0	0	0	0	0	0	1	5	6	7	9
PROMIS-10	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5
VR-12	0	0	0	0	0	0	0	0	0	0	0	1	3	2	3
Quality of life	5	11	17	21	16	21	24	23	34	38	37	59	35	61	66
EQ-5D	0	0	4	1	2	4	5	3	7	14	10	17	11	23	23
SF-12	0	3	7	2	6	7	6	7	8	11	10	15	13	16	25
SF-36	5	8	6	18	8	10	13	13	19	13	17	27	11	22	18
Patient satisfaction	9	11	14	14	12	13	14	12	13	14	19	32	43	48	54
Total	78	102	122	134	158	161	206	176	237	246	319	414	405	456	522

HHS Harris Hip Score, *HOOS* Hip Disability and Osteoarthritis Outcome Score, *OHS* Oxford Hip Score, *WOMAC* Western Ontario and McMaster Universities Arthritis Index, *NRS* Numeric Rating Scale, *UCLA* University of California at Los Angeles Activity Score, *VAS* Visual Analog Scale, *FJS-12* Forgotten Joint Score-12, *PROMIS-10* Patient-Reported Outcome Measurement Information System Global-10, *VR-12* Veterans RAND 12-Item Survey, *EQ-5D* EuroQol 5-Dimension Health Outcome Survey, *SF-12* Short Form-12 Health Survey, *SF-36* Short Form-36 Health Survey

Fig. 1 Flowchart of studies selected

in the quality of life subcategory, EQ-5D demonstrated a significant increase in usage from 2005 to 2019 [0/5 (0%) vs. 23/66 (34.8%), respectively; $P < 0.001$], while SF-36

exhibited a significant decrease in usage from 2005 to 2019 [5/5 (100%) vs. 18/66 (27.3%), respectively; $P = 0.008$].

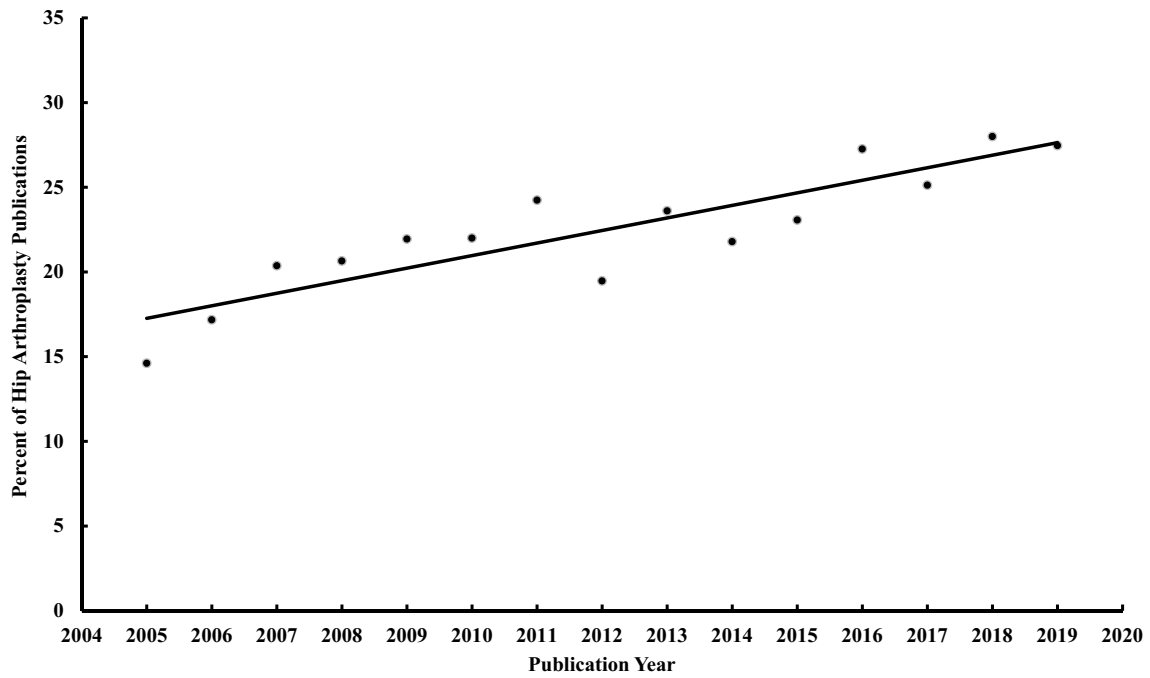


Fig. 2 Percentage of hip arthroplasty publications per year utilizing at least one outcome metric

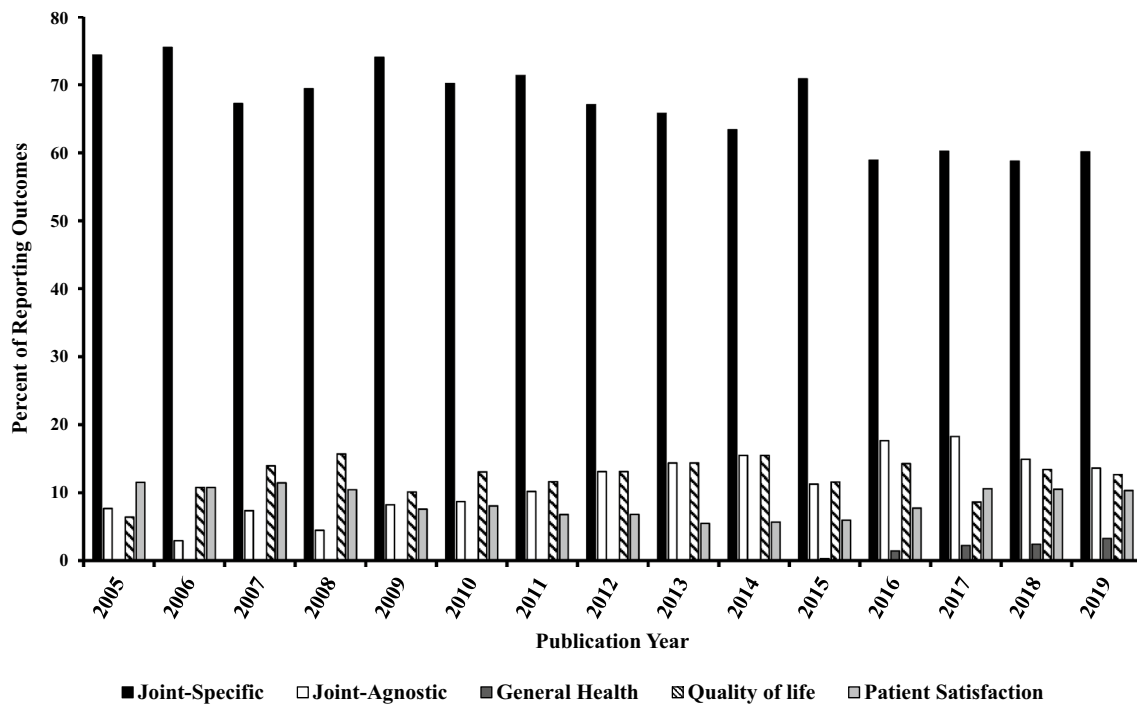


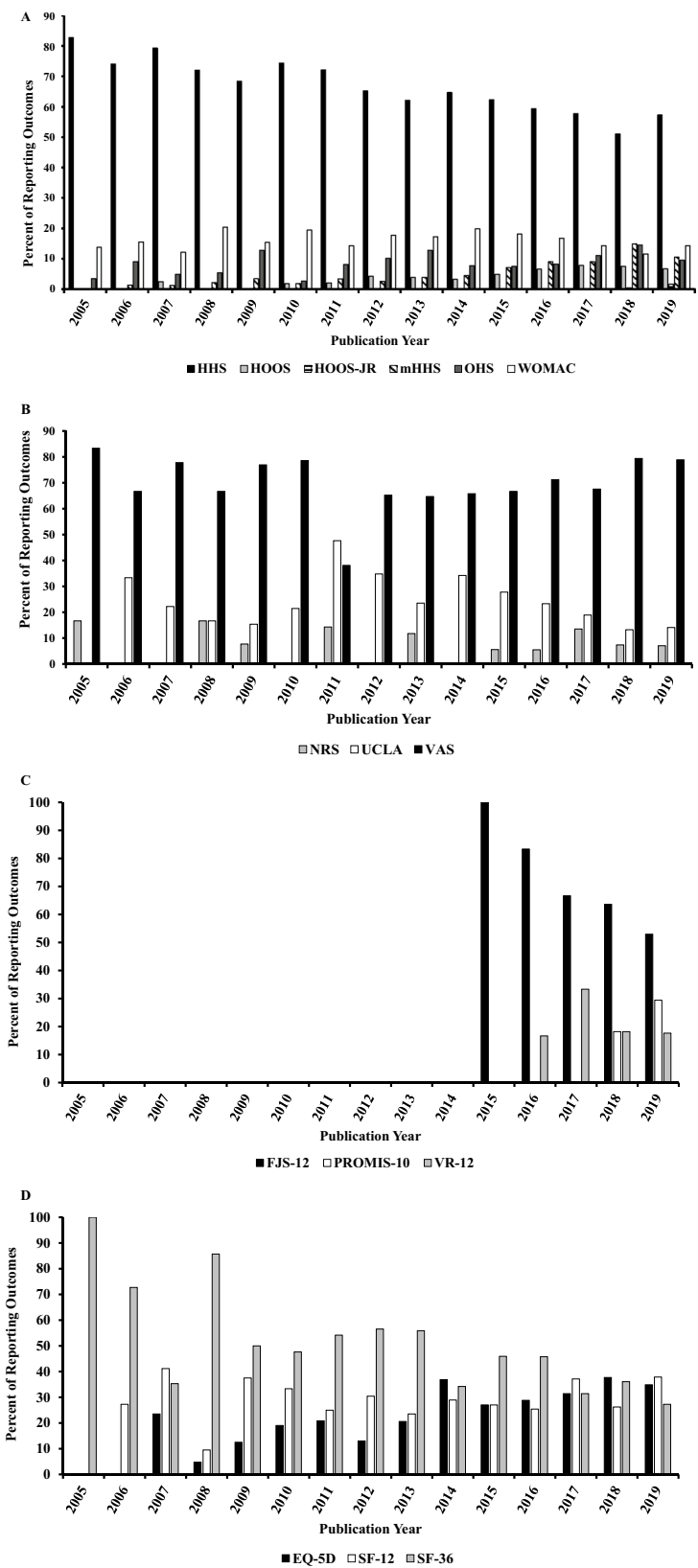
Fig. 3 Utilization of outcome metrics within each subcategory between 2005 and 2019

Discussion

The review considered five categories of outcome metrics

as well as individual instruments within each category from 2005 to 2019. This is the first study to consider how the trends in reporting of individual outcomes metrics in THA have changed over time. We hypothesized that outcome

Fig. 4 Changes in utilization of outcome metrics by subcategory between 2005 and 2019. **a** Individual metric use in the joint-specific subcategory. **b** Individual metric use in the joint-agnostic subcategory. **c** Individual metric use in the general health subcategory. **d** Individual metric use in the quality of life subcategory



metric utilization rates would demonstrate evolving clinician preferences over time, resulting in heterogeneity throughout the literature. There was a statistically significant increase in studies utilizing outcome metrics between 2005 and 2019.

We found that the number of publications per year utilizing at least one outcome metric increased significantly between 2005 and 2019. Recent studies reported that, while the reporting of outcome metrics continues to increase, clinician choice of individual instruments is often highly variable [8, 30]. Despite this heterogeneity, the increase in instrument utilization may demonstrate that both payers and providers have developed a strategy of cost reduction and patient care improvement. While increased reporting of outcome metrics generates additional data on patient outcomes, it should be done wisely, especially because additional metrics may introduce increased heterogeneity within the already complex literature [31]. In fact, this analysis shows that there are 16 widely used outcome metrics used to evaluate THA, across five subcategories. This heterogeneity in outcomes reporting may lead to unnecessary confusion within the industry and may ultimately hinder subsequent research and analysis. A consensus among experts may streamline data reporting and ultimately improve quality of care.

Among the subcategories of outcome metrics, the study found that from 2005 to 2019, joint-agnostic reporting significantly increased, the general health subcategory significantly increased, and the joint-specific subcategory significantly decreased. The decline in utilization in the joint-specific category, accompanied by the increase in more patient-centered metrics (e.g., general health and joint-agnostic categories), may indicate a transition in clinician preference towards monitoring the health of the patient as a whole over time, rather than just the pain and function of a single joint [1, 6].

With respect to individual outcomes measures within the joint-specific subcategory, three outcomes changed significantly from 2005 to 2019. The use of HOOS and mHHS significantly increased, while the use of HHS significantly decreased. Numerous studies illustrate the validity of HHS, as well as its ability to coordinate comparable data across many different countries and languages, which may contribute to its increased utilization [32–35]. This review found that as HHS decreased significantly, mHHS increased significantly. The results of our study indicate that clinicians may have shifted their preferences away from HHS to mHHS, falling in line with several studies which have recommended the use of mHHS due to its increased simplicity, reliability, and reproducibility in a variety of clinical settings [36, 37].

HOOS-JR, OHS, and WOMAC saw no significant change in utilization within the joint-specific subcategory. The lack of change may be due to its recent introduction, first appearing in the results of our study in 2018. Although OHS has been found to be both reliable and valid, it did not show a

significant change in utilization from 2005 to 2019 [38]. While WOMAC utilization rates showed no significant change, the instrument is already the second most utilized metric within the joint-specific subcategory at 16.1% of all joint-specific metrics between 2005 and 2019. The lack of significant change in utilization could be due to an unchanging physician preference for the metric.

Concerning individual outcome measures within the joint-agnostic subcategory, no individual metric showed a significant change in utilization. While the review found no significant change in utilization rate, use was consistent at 23.1% of all joint-agnostic metrics from 2005 to 2019.

Within the general health category, PROMIS-10 demonstrated a significant increase in usage from 2005 to 2019 while VR-12 demonstrated a significant increase in usage from 2005 to 2019. PROMIS-10 is a relatively new metric, first appearing within the search parameters in 2018 and was rapidly adopted (29.4% of general health metrics in 2019). PROMIS-10 has been more heavily studied in TKA, but has not been as well reviewed in THA [27]. VR-12 is able to monitor the effects of comorbidities such as depression and smoking on THA outcomes over time, which may explain its increase in utilization [39]. Despite these results, several studies have shown that scale tends to provide insight into the complex relationship between the psychologic and physical factors impacting patient [39–42].

Within the quality of life subcategory, EQ-5D demonstrated a significant increase in usage from 2005 to 2019, while SF-36 exhibited a significant decrease in usage. SF-12 demonstrated no significant change in utilization over time. Several studies have found that both SF-36 and SF-12 are acceptable quality of life metrics for the majority of patients [10, 30, 43]. Our results may demonstrate that clinicians tend to prefer shorter quality of life instruments, leading to the utilization rate of SF-36 to decrease while the use of SF-12 has increased [44].

Regarding the utilization of patient satisfaction as an outcome metric, this review saw no significant change from 2005 to 2019. Patient satisfaction is difficult to measure because it encompasses the entire patient experience in THA, from initial communication of preoperative expectations to postoperative follow-up and recovery. Several studies have shown that long-term patient satisfaction in THA is best correlated with pain during activity and lifestyle factors [45, 46]. Despite these findings, the review saw no significant increase in reporting of patient satisfaction, which could be due to the measurement of patient satisfaction as a subsection of additional outcome metrics in the quality of life and joint-agnostic subcategories.

Similar complexity in outcome reporting exists within several fields of orthopedics. A review of reporting outcomes in total knee arthroplasty procedures found significant complexity in the reporting of outcome measures, presenting

considerable difficulties in the attempt to compare outcomes in total knee arthroplasty across fields [47]. The current review similarly analyzed the trends in reporting outcomes, which demonstrated similar complexity in THA. As the use of outcome metrics, both in frequency of utilization and in number of individual instruments, continues to increase, we recommend that THA community seek to form a consensus on the preferred outcome metrics. Such an agreement will allow clinicians and researchers to work together to deliver quality patient care through efficient utilization of the vast amount of THA outcome data generated each year. As the transition from volume-focused to value-focused health care continues, a more comprehensive understanding of these metrics will be necessary. Thus, the THA community should strive to reach an agreement on the most useful subcategories, as well as the most reliable and valid outcome metrics within each subcategory. Such a consensus will improve uniformity in the THA literature, and may translate into unified patient management strategies. These changes may ultimately lead to improved quality of care and outcomes for patients receiving THAs. In addition, such changes may be further implemented in similar fields, such as total knee arthroplasty, or spinal surgery.

A number of limitations must be kept in mind when interpreting the results of this study. For instance, not every outcome metric was considered within this study. The vast number of outcomes metrics that are currently being used to evaluate THA indicates that significant heterogeneity exists within the literature. To address this limitation, we performed a comprehensive literature search prior to conducting our analyses to identify the most prevalent outcomes metrics reported within the literature and used by clinicians. This search included analysis of past reviewer articles for relevant outcome metrics, as well as evaluation of articles based on inclusion criteria, outlined in Fig. 1. In addition, this study did not incorporate geographical data; thus, country-level analysis is outside of the scope this paper, but may be an important area of future investigation.

Conclusions

Outcome metric reporting in THA has increased over the last 15 years, accompanying by a rise in both aggregate outcomes reporting and in the number of individual instruments used to evaluate THA. As researchers and clinicians seek to improve patient care, the complexity of reporting outcomes may limit the ability to analyze, compare, and extrapolate study results. We recommend a rigorous evaluation of outcome metrics, which will permit a community-wide consolidation towards the most preferred, valid, and reliable metrics, permitting cross-study comparison and easier generation of clinical guidelines for practice.

Acknowledgements There are no acknowledgements within this study.

Funding There is no funding source.

Compliance with ethical standards

Conflict of interest RHL, JWB, and LTS have nothing to declare. AFK declares the following: AAOS: Board or committee member; American Association of Hip and Knee Surgeons: Board or committee member; Anterior Hip Foundation: Board or committee member; BodyCad: paid consultant; DePuy, A Johnson & Johnson Company: paid consultant; paid presenter or speaker; Innomed: IP royalties; Johnson & Johnson: stock or stock options; Procter & Gamble: stock or stock options; Signature Orthopaedics: research support; Zimmer: paid consultant; paid presenter or speaker; stock or stock options.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was not required for this study.

References

1. Arbab D, van Ochten JHM, Schnurr C et al (2017) Assessment of reliability, validity, responsiveness and minimally important change of the German hip dysfunction and osteoarthritis outcome score (HOOS) in patients with osteoarthritis of the hip. *Rheumatol Int* 37:2005–2011. <https://doi.org/10.1007/s00296-017-3834-y>
2. Pabinger C, Lothaller H, Portner N, Geissler A (2018) Projections of hip arthroplasty in OECD countries up to 2050. *Hip Int* 28:498–506. <https://doi.org/10.1177/1120700018757940>
3. Sloan M, Premkumar A, Sheth NP (2018) Projected volume of primary total joint arthroplasty in the U.S., 2014 to 2030. *J Bone Jt Surg Am Vol.* <https://doi.org/10.2106/JBJS.17.01617>
4. Michael E, Porter P (2010) Perspective - what is value in health care? *N Engl J Med* 363:1–3. <https://doi.org/10.1056/NEJMp1002530>
5. Collins NJ, Roos EM (2012) Patient-reported outcomes for total hip and knee arthroplasty. Commonly used instruments and attributes of a “good” measure. *Clin Geriatr Med* 28:367–394. <https://doi.org/10.1016/j.cger.2012.05.007>
6. Hung M, Saltzman CL, Greene T et al (2018) Evaluating instrument responsiveness in joint function: the HOOS JR, the KOOS JR, and the PROMIS PF CAT. *J Orthop Res* 36:1178–1184. <https://doi.org/10.1002/jor.23739>
7. Alviar MJ, Olver J, Brand C et al (2011) Do patient-reported outcome measures in hip and knee arthroplasty rehabilitation have robust measurement attributes? A systematic review. *J Rehabil Med* 43:572–583. <https://doi.org/10.2340/16501977-0828>
8. Wilson I, Bohm E, Lübbecke A et al (2019) Orthopaedic registries with patient-reported outcome measures. *EFORT Open Rev* 4:357–367. <https://doi.org/10.1302/2058-5241.4.180080>
9. Vajapey SP, Morris J, Spitzer AI et al (2020) Outcome reporting patterns in total knee arthroplasty: a systematic review. *J Clin Orthop Trauma.* <https://doi.org/10.1016/j.jcot.2020.05.014>
10. Konopka JF, Lee Y-Y, Su EP, McLawhorn AS (2018) Quality-adjusted life years after hip and knee arthroplasty: health-related quality of life after 12,782 joint replacements. *JBJS Open Access* 3:e0007. <https://doi.org/10.2106/JBJS.OA.18.00007>
11. Murphy L, Helmick CG (2012) The impact of osteoarthritis in the United States: a population-health perspective. *Am J Nurs.* <https://doi.org/10.1097/01.NAJ.0000412646.80054.21>

12. Molloy IB, Yong TM, Keswani A et al (2020) Do medicare's patient-reported outcome measures collection windows accurately reflect academic clinical practice? *J Arthroplasty* 35:911–917. <https://doi.org/10.1016/j.arth.2019.11.006>
13. Bozic KJ, Pui CM, Ludeman MJ et al (2010) Do the potential benefits of metal-on-metal hip resurfacing justify the increased cost and risk of complications? *Clin Orthop Relat Res* 468:2301–2312. <https://doi.org/10.1007/s11999-010-1301-0>
14. Mújica Mota RE (2013) Cost-effectiveness analysis of early versus late total hip replacement in Italy. *Value Health* 16:267–279. <https://doi.org/10.1016/j.jval.2012.10.020>
15. Lodhia P, Gui C, Chandrasekaran S et al (2016) The economic impact of acetabular labral tears. *Am J Sports Med* 44:1771–1780. <https://doi.org/10.1177/0363546516645532>
16. Siljander MP, McQuivey KS, Fahs AM et al (2018) Current trends in patient-reported outcome measures in total joint arthroplasty: a study of 4 major orthopaedic journals. *J Arthroplasty* 33:3416–3421. <https://doi.org/10.1016/j.arth.2018.06.034>
17. Pulikottil-Jacob R, Connock M, Kandala NB et al (2015) Cost effectiveness of total hip arthroplasty in osteoarthritis: comparison of devices with differing bearing surfaces and modes of fixation. *Bone Jt J* 97-B:449–457. <https://doi.org/10.1302/0301-620X.97B4.34242>
18. Varnum C (2017) Outcomes of different bearings in total hip arthroplasty - implant survival, revision causes, and patient-reported outcome. *Dan Med J* 64(3):B5350
19. Ponnusamy KE, Vasarhelyi EM, Somerville L et al (2018) Cost-effectiveness of total knee arthroplasty vs nonoperative management in normal, overweight, obese, severely obese, morbidly obese, and super-obese patients: a markov model. *J Arthroplasty* 33:S32–S38. <https://doi.org/10.1016/j.arth.2018.02.031>
20. Ponnusamy KE, Vasarhelyi EM, McCalden RW et al (2018) Cost-effectiveness of total hip arthroplasty versus nonoperative management in normal, overweight, obese, severely obese, morbidly obese, and super obese patients: a Markov model. *J Arthroplasty* 33:3629–3636. <https://doi.org/10.1016/j.arth.2018.08.023>
21. Gaffney CJ, Pelt CE, Gililland JM, Peters CL (2017) Perioperative pain management in hip and knee arthroplasty. *Orthop Clin N Am* 48:407–419
22. Pennington M, Grieve R, Sekhon JS et al (2013) Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis. *BMJ*. <https://doi.org/10.1136/bmj.f1026>
23. Jansen JA, Kruidenier J, Spek B, Snoeker BAM (2020) A cost-effectiveness analysis after implementation of a fast-track protocol for total knee arthroplasty. *Knee*. <https://doi.org/10.1016/j.knee.2019.09.014>
24. Lovelock TM, Broughton NS, Williams CM (2018) The popularity of outcome measures for hip and knee arthroplasties. *J Arthroplasty* 33:273–276
25. Kamaruzaman H, Kinghorn P, Oppong R (2017) Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature. *BMC Musculoskelet Disord* 18:1–17. <https://doi.org/10.1186/s12891-017-1540-2>
26. Rosenlund S, Broeng L, Holsgaard-Larsen A et al (2017) Patient-reported outcome after total hip arthroplasty: comparison between lateral and posterior approach: a randomized controlled trial in 80 patients with 12-month follow-up. *Acta Orthop* 88:239–247. <https://doi.org/10.1080/17453674.2017.1291100>
27. Shim J, Hamilton DF (2019) Comparative responsiveness of the PROMIS-10 global health and EQ-5D questionnaires in patients undergoing total knee arthroplasty. *Bone Jt J* 101 B:832–837. <https://doi.org/10.1302/0301-620X.101B7.BJJ-2018-1543.R1>
28. Finch DJ, Pellegrini VD, Franklin PD et al (2020) The effects of bundled payment programs for hip and knee arthroplasty on patient-reported outcomes. *J Arthroplasty* 35:918–925.e7. <https://doi.org/10.1016/j.arth.2019.11.028>
29. Faschingbauer M, Kasperek M, Schadler P et al (2017) Predictive values of WOMAC, KOOS, and SF-12 score for knee arthroplasty: data from the OAL. *Knee Surg Sport Traumatol Arthrosc* 25:3333–3339. <https://doi.org/10.1007/s00167-016-4369-6>
30. Rolfson O, Eresian Chenok K, Bohm E et al (2016) Patient-reported outcome measures in arthroplasty registries: report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries: part I. Overview and rationale for patient-reported outcome measures. *Acta Orthop* 87:3–8. <https://doi.org/10.1080/17453674.2016.1181815>
31. Cooper ME, Torre-Healy LA, Alentado VJ et al (2018) Heterogeneity of reporting outcomes in the spine surgery literature. *Clin Spine Surg* 31:E221–E229. <https://doi.org/10.1097/BSD.0000000000000578>
32. Torre M, Luzzi I, Mirabella F et al (2018) Cross-cultural adaptation and validation of the Italian version of the Hip disability and Osteoarthritis Outcome Score (HOOS). *Health Qual Life Outcomes* 16:1–9. <https://doi.org/10.1186/s12955-018-0935-6>
33. Lyman S, Lee YY, Franklin PD et al (2016) Validation of the HOOS, JR: a short-form hip replacement survey. *Clin Orthop Relat Res* 474:1472–1482. <https://doi.org/10.1007/s11999-016-4718-2>
34. Nilsson A, Bremander A (2011) Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) hip and knee questionnaire. *Arthritis Care Res* 63:200–207. <https://doi.org/10.1002/acr.20549>
35. Trathitiphan W, Paholpak P, Sirichativapee W et al (2016) Cross-cultural adaptation and validation of the reliability of the Thai version of the Hip disability and Osteoarthritis Outcome Score (HOOS). *Rheumatol Int* 36:1455–1458. <https://doi.org/10.1007/s00296-016-3505-4>
36. Kumar P, Sen R, Aggarwal S et al (2019) Reliability of modified Harris Hip Score as a tool for outcome evaluation of total hip replacements in Indian population. *J Clin Orthop Trauma* 10:128–130. <https://doi.org/10.1016/j.jcot.2017.11.019>
37. Ramisetty N, Kwon Y, Mohtadi N (2015) Patient-reported outcome measures for hip preservation surgery—a systematic review of the literature. *J Hip Preserv Surg* 2:15–27. <https://doi.org/10.1093/jhps/hnv002>
38. Quah C, Holmes D, Khan T et al (2018) The variability in Oxford hip and knee scores in the preoperative period: Is there an ideal time to score? *Ann R Coll Surg Engl* 100:16–20. <https://doi.org/10.1308/rcsann.2017.0090>
39. Bienstock DM, Snyder DJ, Kroshus TR et al (2019) Relationship between baseline patient-reported outcomes and demographic, psychosocial, and clinical characteristics. *JAAOS Glob Res Rev* 3:e039. <https://doi.org/10.5435/jaasglobal-d-19-00039>
40. Matsumoto M, Baba T, Ochi H et al (2017) Influence of the contralateral hip state after total hip arthroplasty on patient-reported outcomes measured with the Forgotten Joint Score-12. *Eur J Orthop Surg Traumatol* 27:929–936. <https://doi.org/10.1007/s00590-017-1963-3>
41. Hamilton DF, Giesinger JM, MacDonald DJ et al (2016) Responsiveness and ceiling effects of the Forgotten Joint Score-12 following total hip arthroplasty. *Bone Jt Res* 5:87–91. <https://doi.org/10.1302/2046-3758.53.2000480>
42. Behrend H, Zdravkovic V, Giesinger J, Giesinger K (2016) Factors predicting the forgotten joint score after total knee arthroplasty. *J Arthroplasty* 31:1927–1932. <https://doi.org/10.1016/j.arth.2016.02.035>
43. Webster KE, Feller JA (2016) Comparison of the short form-12 (SF-12) health status questionnaire with the SF-36 in patients with knee osteoarthritis who have replacement surgery. *Knee Surg*

- Sport Traumatol Arthrosc 24:2620–2626. <https://doi.org/10.1007/s00167-015-3904-1>
44. Jenkinson C, Layte R, Jenkinson D et al (1997) A shorter form health survey: Can the sf-12 replicate results from the sf-36 in longitudinal studies? *J Public Health* 19:179–186. <https://doi.org/10.1093/oxfordjournals.pubmed.a024606>
 45. Schmitz PP, van Susante JLC, Hol A et al (2019) No decline in high patient satisfaction after total hip arthroplasty at long-term follow-up. *Eur J Orthop Surg Traumatol* 29:91–95. <https://doi.org/10.1007/s00590-018-2243-6>
 46. Naal FD, Impellizzeri FM, Leunig M (2009) Which is the best activity rating scale for patients undergoing total joint arthroplasty? *Clin Orthop Relat Res* 467:958–965. <https://doi.org/10.1007/s11999-008-0358-5>
 47. Lan RH, Bell JW, Samuel LT, Kamath AF (2020) Evolving Outcome Measures in Total Knee Arthroplasty: Trends and Utilization Rates Over the Past 15 Years. *J Arthroplasty*. 35(11):3375–3382. <https://doi.org/10.1016/j.arth.2020.06.036>

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