

The EQ-5D-5L Improves on the EQ-5D-3L for Health-related Quality-of-life Assessment in Patients Undergoing Total Hip Arthroplasty

Meridith E. Greene BA, Kevin A. Rader PhD, Göran Garellick MD, PhD,
Henrik Malchau MD, PhD, Andrew A. Freiberg MD, Ola Rolfson MD, PhD

Published online: 9 December 2014
© The Association of Bone and Joint Surgeons® 2014

Abstract

Background The EQ-5D is a generic health survey that can be used to compare improvement across different interventions, measure changes in health-related quality of life over time, or to explore cost-effectiveness among treatments, hospitals, or providers. The original EQ-5D survey has three response options for each of five health dimensions; however, with so few response options, ceiling and floor effects are problematic in some populations. A new version, called the EQ-5D-5L, was developed, which gives respondents five answer options (the “5L” refers to five response levels, which is in contrast to the original survey’s three levels). However, the validity of this version

has not, to our knowledge, been evaluated in patients undergoing total hip arthroplasty (THA).

Questions/purposes The purposes of this study were (1) to characterize the redistribution of responses using the new version; (2) to describe the ceiling and floor effects in the current three-level version and identify whether the new EQ-5D-5L survey diminished these effects; and (3) to understand the convergent validity of the new version with the old and the EQ visual analog scale.

Methods Both versions of the survey were administered either preoperatively or 1 to 6 years after THA, allowing at least 2 weeks between administrations. Responses to the two versions were compared to determine response redistribution properties, ceiling and floor effects, and convergent validity. Sample sizes were determined so that the study would have 90% power to detect a Spearman correlation over 0.7 when comparing the responses of the three-level survey with the five-level survey and allowing for a rate of 25% loss to followup.

Results Most patients before surgery used the new responses in the majority of dimensions, whereas the patients taking the test after surgery used the new responses predominantly for the pain dimension. The five-level diminished ceiling effects in both groups by up to 30% and diminished floor effects in the pain dimension for patients taking the assessment before surgery by 14%, which was the only dimension in either patient group that had high

One or more of the authors (OR) has received funding from grants from the Doktor Felix Neuberghs Foundation, the Göran Bauer Grant, The Swedish Society of Medicine, and The Wangstedts Foundation. All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

The majority of this work was performed in the Harris Orthopaedic Laboratory, Massachusetts General Hospital, Boston, MA, USA.

Electronic supplementary material The online version of this article (doi:10.1007/s11999-014-4091-y) contains supplementary material, which is available to authorized users.

M. E. Greene (✉), H. Malchau, A. A. Freiberg, O. Rolfson
Harris Orthopaedic Laboratory, Massachusetts General Hospital,
55 Fruit Street, GRJ 1125, Boston, MA 02114, USA
e-mail: megreene@partners.org

M. E. Greene, G. Garellick, H. Malchau, O. Rolfson
Swedish Hip Arthroplasty Register, Gothenburg,
Sweden

M. E. Greene, G. Garellick, H. Malchau, O. Rolfson
Department of Orthopaedics, Institute of Clinical Sciences,
Sahlgrenska Academy, University of Gothenburg, Gothenburg,
Sweden

K. A. Rader
Biostatistics Department, Harvard University, Cambridge, MA,
USA

rates of floor effects. The correlation between the surveys' visual analog scale and response patterns was not different for the two versions of the survey in these populations.

Conclusions The EQ-5D five-level survey appears able to discriminate new health states indistinguishable in the original, which may allow more sensitive measurements of change in patients undergoing THA. The five-level survey should be considered for implementation in local and national registry monitoring of health-related quality of life in patients undergoing THA.

Introduction

The EuroQol Group's health-related quality-of-life (HRQoL) measure EQ-5D is a generic health survey that is part of the standard followup of pre- and postoperative patients in national and local arthroplasty registries [5, 6, 14, 17]. The HRQoL data can be used to monitor improvement after or between different interventions, measure changes in HRQoL over time, or to explore cost-effectiveness among treatments, medical centers, or healthcare providers. The survey is brief, in our experience taking approximately 3 minutes to complete, comprising five questions and a visual analog scale (EQ VAS). This makes it an attractive tool for busy clinics and patients [6].

Arthroplasty clinic patients at Massachusetts General Hospital complete the EQ-5D survey before their appointment as part of the standard of care providing the surgeon with the patient's self-rated HRQoL. The instrument consists of five dimensions measuring mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. In the original version of the survey, the respondent chooses from three levels which define each dimension: no, some/moderate, and extreme problems. It has been argued, however, that the EQ-5D-3L lacks descriptive richness and is limited in its ability to measure small but clinically relevant changes in the general population as well as in specific patient groups [2, 12, 13, 15, 19]. In response, the EuroQol Group developed a new five-level version of the EQ-5D (EQ-5D-5L) [7]. In the five-level version, the patient chooses from five response levels for each of the dimensions: no, minimal, some, many, and extreme problems. The assumption is that when given response options between no problems and some problems and between some problems and extreme problems, the respondent can provide a better profile of their health.

The five-level EQ-5D survey has been shown to have face validity, decrease ceiling effects, and increase discriminatory power in several patient populations over the three-level version [8, 10, 11, 16, 18]. However, the

validity of this version has not, to our knowledge, been evaluated in patients undergoing THA.

We therefore investigated the new five-level version of the EQ-5D survey in both a preoperative and a postoperative THA population. This study had three purposes: (1) to characterize the redistribution of responses using the new version; (2) to describe the ceiling and floor effects in the current three-level version and identify whether the new EQ-5D-5L survey diminished these effects; and (3) to understand the convergent validity of the new version with the old and the EQ VAS.

Patients and Methods

In this institutional review board-approved prospective study, individuals were invited to participate while in the arthroplasty clinic. New patients older than 18 years of age, with hip pain, visiting the arthroplasty clinic from October 23, 2012, through March 19, 2013, and from December 3, 2013, to March 1, 2014, who never had a hip arthroplasty were invited to participate in the study (approximately 283 patients) of whom 70 agreed (25%). Between July 16, 2013, and June 19, 2014, patients older than 18 years of age who were at least 1 year and up to 6 years after their THA were also invited to participate in the study (approximately 245 patients) of whom 96 agreed (39%). We state "approximately" because that was the total number of patients meeting the stated inclusion criteria who were seen during the periods in question; however, the actual number invited may have been lower if our study coordinator was not able to speak with every patient, as seems likely. This means that the estimates of the percentage of patients who agreed to participate given here are likely to be conservative. Early standard of care visits at our clinic occur 6 to 10 weeks and 1, 3, and 5 years after surgery. However, patients rarely come on their surgical anniversary date; therefore, our inclusion criteria considered patients up to 6 years followup. Patients returning for immediate postoperative followup were not included because their healing process would not be complete. The EQ-5D-3L responses were not different between the patients who agreed to participate in the study from those who did not indicating a representative study population (data not shown). After providing consent, patients completed each version of the EQ-5D survey at a minimum of 2 weeks apart. The first survey was completed in the office and the second survey was completed either online or by a paper form at home. A crossover design was implemented in which half of the patients took the three-level version first and the other half took the five-level

leave the setting of the clinical office. We therefore required at least 50 patients before surgery and at least 70 patients after surgery. All statistical comparisons were considered significant if $p \leq 0.05$.

Results

The new response options in the five-level survey were used by most patients before surgery in the majority of dimensions, whereas the patients who already had surgery were more likely to use the new response options in the pain dimension. The most frequent response from the patients taking the survey after THA was no problems in all five dimensions in both the three-level and five-level surveys (Supplemental Table 1 [Supplemental materials are available with the online version of CORR®.]). New health states were identified in the majority of preoperative patients in the mobility (60%), usual activities (58%), and pain/discomfort (54%) dimensions using the five-level version (Fig. 2). Most of the postoperative patients reported no problems in the majority of dimensions in both

versions of the survey excluding the pain/discomfort dimension where the new responses were chosen 43% of the time (33 of 77 patients). Some patients reported health states that were more than one response option away from their first, classified as inconsistent, in both pre- and postoperative patients. These inconsistent responses ranged from 4% to 14% in the patients before surgery and from 5% to 18% in the patients after surgery (Fig. 2).

Ceiling effects in the three-level survey were diminished by adding the new response options in the five-level version, but conversely, preoperative floor effects improved for only the pain dimension with the addition of the new response options. The ceiling effect diminished in just one dimension (self-care) for the 50 preoperative patients from 33 to 18 patients (66% to 36%; $p = 0.001$), whereas ceiling effects diminished in four dimensions for the 77 postoperative patients: mobility (from 42 to 30 patients), self-care (from 69 to 61 patients), usual activities (from 50 to 38 patients), and pain (from 31 to 20 patients, all $p \leq 0.03$; Table 1). It should be noted that ceiling effects were seen in more than half of the preoperative patients in only the self-care and anxiety/depression dimensions of the

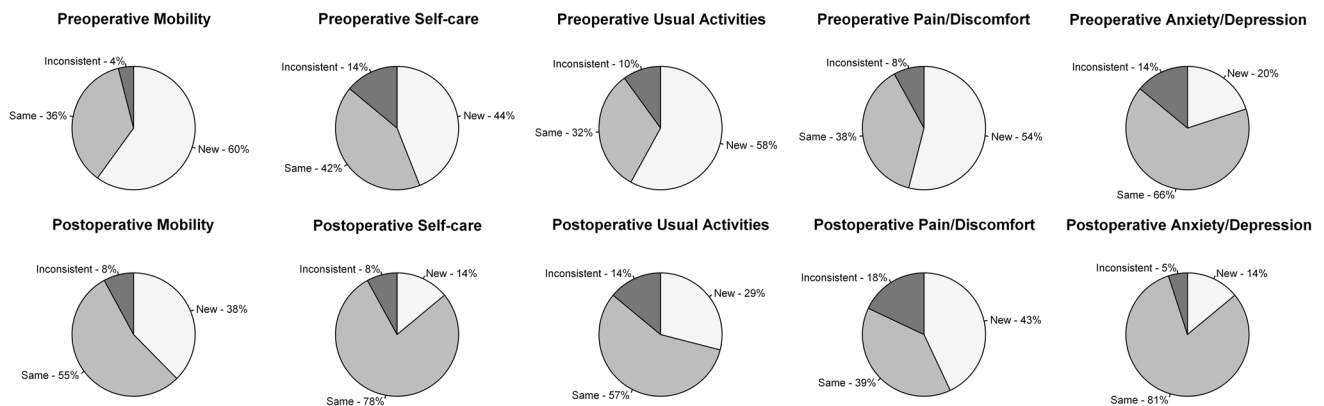


Fig. 2 Proportions of patient response trends from the three-level survey to the five-level survey for each dimension in the preoperative and postoperative patient groups are shown.

Table 1. Percentage of patients with a ceiling effect for each dimension of the EQ-5D and the complete surveys between the 3L and 5L versions for preoperative and postoperative patients using McNemar's test

Survey dimension	Preoperative patients (n = 50)			Postoperative patients (n = 77)		
	EQ-5D-3L	EQ-5D-5L	p value	EQ-5D-3L	EQ-5D-5L	p value
Mobility	16%	6%	0.06	55%	39%	< 0.01
Self-care	66%	36%	< 0.01	90%	79%	0.03
Usual activities	22%	20%	0.76	65%	49%	< 0.01
Pain	6%	4%	0.56	40%	26%	0.02
Anxiety/depression	64%	62%	0.74	88%	82%	0.06
All dimensions	2%	2%	n/a	30%	18%	0.04

Probability values presented in bold indicate significant decreases in ceiling effects.

Table 2. Percentage of patients with a floor effect for each dimension of the EQ-5D and the complete surveys between the 3L and 5L versions for preoperative and postoperative patients using McNemar's test

Survey dimensions	Preoperative patients (n = 50)			Postoperative patients (n = 77)		
	EQ-5D-3L	EQ-5D-5L	p value	EQ-5D-3L	EQ-5D-5L	p value
Mobility	0%	0%	N/A	1%	0%	N/A
Self-care	2%	0%	N/A	0%	0%	N/A
Usual activities	6%	0%	N/A	4%	0%	N/A
Pain	18%	4%	0.008	5%	3%	0.317
Anxiety/depression	2%	2%	N/A	3%	0%	N/A
All dimensions	0%	0%	N/A	0%	0%	N/A

Probability values presented in bold indicate significant decreases in ceiling effects; N/A = not applicable because test could not be performed as a result of zero values in the contingency table.

three-level survey, but ceiling effects were seen in all dimensions of the three-level survey except the pain/discomfort dimension in the postoperative patients. Only one preoperative patient responded with no problems in all dimensions in both versions of the survey, indicating that an overall ceiling effect was uncommon in these patients. In the postoperative patients, however, the overall ceiling effect diminished with the new five-level survey from 23 patients to 14 patients of the 77 who completed both surveys (30% to 18%; $p = 0.04$). Very few patients responded with extreme problems in the majority of dimensions. It was most often selected by the preoperative patients ($n = 9$) in the three-level survey in the pain/discomfort dimension (18%), which diminished to only two patients in the five-level version (4%; $p = 0.008$). Because so few patients selected extreme problems in the five-level survey, the proportion of floor effects was unable to be compared for any dimension except for pain/discomfort (Table 2). No patients reported an overall floor effect in either version of the survey.

The convergent validity of the five-level survey with the three-level survey and the EQ VAS was favorable. The EQ VAS from the two versions were strongly correlated with one another at 0.673 (95% confidence interval [CI], 0.485–0.801) and 0.537 (95% CI, 0.355–0.679) for patients before and after surgery, respectively (both $p < 0.001$). In the preoperative group, correlations between the dimensions of the five-level survey and the EQ VAS were all either moderate or strong. Conversely, the majority of correlations between the EQ VAS and the dimensions of the three-level version were absent or weak. In the postoperative patients, the mobility, usual activities, and pain/discomfort dimensions were all moderately correlated with the EQ VAS. The correlations between the five-level survey and the EQ VAS in the postoperative patients were moderate to high in all dimensions but anxiety/depression (Table 3). In the

preoperative patients, change in the VAS score from one version to the other was directly related to response patterns in the pain/discomfort dimension. Individuals who reported greater amounts of pain from one version to the next also reported on average 13 point (95% CI, –26.0 to –0.6) worse VAS scores ($p = 0.04$). Postoperative patient-reported VAS scores tended to decrease with worsening problems in each of the five dimensions but were not found to be different in this population (all $p > 0.05$). As the time between the completion of the survey versions doubled, the preoperative EQ VAS scores decreased by an average of three units (95% CI, 0.1–6.4; $p = 0.05$). No trends were seen in the time between surveys and the response trends to the five dimensions in either group of patients.

Discussion

Patient-reported surveys are gaining popularity in the assessment of outcomes after THA [17]. Minimizing the number of questions the patient needs to answer while identifying surveys that are sensitive to changes is critical in alleviating the burden on the patient and providing usable information for the clinician. The original EQ-5D-3L is a brief survey but its sensitivity to change has been questioned [2, 12, 13, 15, 19]. The new EQ-5D-5L survey has been validated in several patient populations for improving the sensitivity of HRQoL measurements [8, 10, 11, 16, 18], but has yet to be validated in a THA population. Our work suggests that with the addition of two response options, the five-level survey may be able to better measure change in HRQoL after surgery. Patients who had yet to undergo THA made use of intermediate response options the majority of the time in the three dimensions most relevant to hip disease (mobility, usual activities, and pain/discomfort) allowing greater

Table 3. Spearman's rank correlation (r_s) between each dimension and the EQ VAS score for pre- and postoperative patients for each version of the EQ-5D survey* as well as the 95% confidence interval (CI) for each

Dimension	Preoperative (n = 50)				Postoperative (n = 77)				p value	
	r_s between 3L and VAS	95% CI	r_s between 5L and VAS	95% CI	r_s between 3L and VAS	95% CI	r_s between 5L and VAS	95% CI		
Mobility	-0.370	-0.588 to -0.102	-0.458	-0.652 to -0.205	0.616	-0.402	-0.575 to -0.196	-0.532	-0.676 to -0.350	0.318
Self-care	-0.185	-0.441 to 0.098	-0.358	-0.579 to -0.089	0.337	-0.179	-0.387 to 0.047	-0.406	-0.577 to -0.200	0.125
Usual activities	-0.349	-0.572 to -0.079	-0.501	-0.684 to -0.259	0.352	-0.429	-0.596 to -0.227	-0.420	-0.588 to -0.216	0.946
Pain/discomfort	-0.527	-0.703 to -0.292	-0.508	-0.689 to -0.268	0.902	-0.371	-0.550 to -0.161	-0.408	-0.579 to -0.202	0.782
Anxiety/depression	-0.283	-0.521 to -0.005	-0.524	-0.700 to -0.287	0.109	-0.349	-0.531 to -0.135	-0.300	-0.491 to -0.082	0.701

* The strength of the correlations was defined as: absent ($r_s < 0.20$), weak ($0.20 \leq r_s < 0.35$), moderate ($0.35 \leq r_s < 0.50$), or strong ($r_s \geq 0.50$). Randomization tests (based on 10,000 permutations) for paired data were used to determine if the correlations were significantly different between the versions of the survey with significance at $p \leq 0.05$; VAS = visual analog scale.

discrimination of their preoperative health state. The five response options allowed the postoperative patients to better define their health states as well given that the ceiling effect in this population was diminished over four of the five dimensions and for the survey as a whole. With patients reporting new health states, the presumption would be that the new five-level survey would have a greater sensitivity to change after surgical intervention.

Because patients were asked to participate rather than conducting the study on all patients undergoing arthroplasty in the clinic, there may have been some level of self-selection bias because only a minority of postoperative patients ultimately agreed to participate in the study. Patients were invited to participate as a means to increase the response rate for the second version of the survey, which could have been higher but to us seemed satisfactory. Our inclusion criteria may have allowed preoperative patients who might not ultimately get THA and postoperative patients who may have been having problems with their joint arthroplasty to be enrolled. This could be viewed as a limitation; however, these patients provide an accurate sample of the types of patients visiting our clinic. Standard patient-reported outcome measures, including the EQ-5D, are administered to all our patients, so it is important that we include all types of patients we may see in our clinic to understand how the new version of the survey may measure HRQoL in this population. We were able to show that the patients who agreed to participate in the study and those who did not, did not report different HRQoL indicating that our patient samples were representative of both groups of patients. The 2-week separation between the survey administrations may be a limitation if patient health states changed during this period. However, the time difference between the survey administrations showed no difference in the response trends to the five dimensions. By including the EQ VAS, we were able to observe that those who took longer than 2 weeks to complete the second survey had a decrease in their reported EQ VAS score. There were worsening reports of pain/discomfort in the preoperative patients who took longer than 2 weeks to complete the second survey, which suggests that patients may have been deteriorating while awaiting surgery. The EQ VAS and five dimensions of the survey were designed to measure different elements of the patient's health status. Our results suggest that although patients perceived their overall health status to have deteriorated over the longer separation period, they did not show changes in their response patterns. The association between the response trend in the pain/discomfort dimension and the EQ VAS score for the patients before surgery was found not to be significant at the 0.05 level. This may be a product of a relatively small sample size

for looking at such a trend or simply that the EQ VAS and the separate dimensions were measuring different elements of their health status.

Our patient populations had some higher rates of inconsistent responses in certain dimensions than seen in other EQ-5D-5L survey test populations [10, 11, 16, 18]. Unlike our method, the most common technique in the literature for comparing the three- and five-level surveys was to administer both surveys with demographic and additional survey questions in between the two versions during the same sitting [4, 10, 11, 16, 18]. As was seen by Janssen and collaborators [9] in 2008, this simultaneous administration of survey versions actually influenced patient responses leading to patients avoiding intermediate response options on the five-level survey if the three-level had been administered first. To avoid this bias, we administered the surveys at least 2 weeks apart, in which half of the patients completed the three-level first and the other half completed the five-level first. Our results indicated that with the separation of administrations, the order of the surveys did not influence the response patterns, which may have provided us with better measures of patient interpretation of the new five-level response options and, as explained earlier, changes in patient health states over those 2 weeks likely contributed to the inconsistent responses.

The five-level survey appears to be extremely valuable in identifying preoperative health states, whereas it appears to be slightly less so in the postoperative group. However, with the greatest proportion of postoperative patients using the new response options in the pain dimension, we gained valuable information. THA is predominantly an elective procedure undertaken to improve HRQoL through improved mobility and decreased pain/discomfort. If these areas are not improved, it is likely that patients will be dissatisfied with their treatment [1, 3]. Because approximately one-fifth of patients were able to report that they experienced more mobility and pain/discomfort problems on the five-level survey than on the three-level survey and another one-fifth were able to report experiencing less problems, clinicians got a better picture of how the patient was doing after surgery. Subsequently, clinicians may better understand how effective the THA was in improving HRQoL.

The five-level version of the EQ-5D survey performed well in this population of pre- and postoperative patients undergoing THA by decreasing ceiling and floor effects and increasing the discriminatory power of the instrument over the original version. Particularly in the preoperative patients, we were able to discriminate new health states by implementing the five-level version of the survey. Despite most of the postoperative patients reporting no problems in the majority of dimensions in both versions of the survey, they more commonly made use of the new response options in the pain dimension.

Because patients undergoing hip arthroplasty can be vastly different from one another, it is important to have assessment tools capable of capturing the nuanced differences among them whether in a single clinic or at the national level. Incorporation of the five-level version of the EQ-5D survey into national followup programs in arthroplasty registries may mitigate previous concerns about ceiling effects in the earlier version of the survey [12, 15, 17, 19] and will enable clinicians to better serve patient needs by allowing researchers to identify risk factors for more or less improvement after treatment.

By adding intermediate response options to the EQ-5D survey, patients were allowed more flexibility in reporting their health states. They were not forced to upgrade or downgrade their current health to fit the parameters of the survey. With this new but familiar tool, clinicians and researchers may more accurately assess changes in HRQoL. The EQ-5D-5L survey should be considered for implementation in place of the original three-level survey to assess patient HRQoL before and after THA. Future studies of the EQ-5D-5L survey in THA populations should explore if the survey's improved sensitivity will allow for improved measurements of change in patient HRQoL either from the preoperative to postoperative period or over time postoperatively.

Acknowledgments We thank Sandra Pierre for dedicated assistance in recruiting patients into this study.

References

1. Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. *J Arthroplasty*. 2011;26:209–213.
2. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ*. 2004;13:873–884.
3. Clement ND, Muzammil A, Macdonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. *J Bone Joint Surg Br*. 2011;93:464–469.
4. Craig BM, Pickard AS, Lubetkin EI. Health problems are more common, but less severe when measured using newer EQ-5D versions. *J Clin Epidemiol*. 2014;67:93–99.
5. Devlin NJ, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Econ*. 2010;19:886–905.
6. EuroQolGroup. EuroQol—a new facility for the measurement of health related quality of life. *Health Policy*. 1990;16:3.
7. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20:1727–1736.
8. Hinz A, Kohlmann T, Stobel-Richter Y, Zenger M, Brahler E. The quality of life questionnaire EQ-5D-5L: psychometric properties and normative values for the general German population. *Qual Life Res*. 2014;23:443–447.
9. Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. *Value Health*. 2008;11:275–284.

10. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, Swinburn P, Busschbach J. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res.* 2013;22:1717–1727.
11. Jia YX, Cui FQ, Li L, Zhang DL, Zhang GM, Wang FZ, Gong XH, Zheng H, Wu ZH, Miao N, Sun XJ, Zhang L, Lv JJ, Yang F. Comparison between the EQ-5D-5L and the EQ-5D-3L in patients with hepatitis B. *Qual Life Res.* 2014;8:2355–2363.
12. Johnson JA, Pickard AS. Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta, Canada. *Med Care.* 2000;38:115–121.
13. Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol.* 2003;56:317–325.
14. Lawless BM, Greene M, Slover J, Kwon YM, Malchau H. Does age or bilateral disease influence the value of hip arthroplasty? *Clin Orthop Relat Res.* 2012;470:1073–1078.
15. Luo N, Johnson JA, Shaw JW, Feeny D, Coons SJ. Self-reported health status of the general adult US population as assessed by the EQ-5D and Health Utilities Index. *Med Care.* 2005;43:1078–1086.
16. Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. *Med Care.* 2007;45:259–263.
17. Rolfson O, Karrholm J, Dahlberg LE, Garellick G. Patient-reported outcomes in the Swedish Hip Arthroplasty Register: results of a nationwide prospective observational study. *J Bone Joint Surg Br.* 2011;93:867–875.
18. Scalone L, Ciampichini R, Fagioli S, Gardini I, Fusco F, Gaeta L, Del Prete A, Cesana G, Mantovani LG. Comparing the performance of the standard EQ-5D 3L with the new version EQ-5D 5L in patients with chronic hepatic diseases. *Qual Life Res.* 2013;22:1707–1716.
19. Sullivan PW, Lawrence WF, Ghushchyan V. A national catalog of preference-based scores for chronic conditions in the United States. *Med Care.* 2005;43:736–749.