

Electronic Supplementum no 362: ISAR meeting Gothenburg 2015, Sweden

Peer-reviewed article based on study presented at the 2015 meeting of the International Congress of Arthroplasty Registries

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

Ola ROLFSON¹, Kate Eresian CHENOK², Eric BOHM³, Anne LÜBBEKE⁴, Geke DENISSEN⁵, Jennifer DUNN⁶, Stephen LYMAN⁷, Patricia FRANKLIN⁸, Michael DUNBAR⁹, Søren OVERGAARD¹⁰, Göran GARELLICK¹, Jill DAWSON¹¹; the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries

¹ Swedish Hip Arthroplasty Register and Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; ² California Joint Replacement Registry, San Francisco, CA, USA; ³ Canadian Joint Replacement Registry and University of Manitoba, Winnipeg, Manitoba, Canada; ⁴ Geneva Arthroplasty Registry, Division of Orthopaedic Surgery, Geneva University Hospitals, Geneva, Switzerland; ⁵ Dutch Arthroplasty Register, Hertogenbosch, the Netherlands; ⁶ New Zealand Joint Registry and University of Otago, Christchurch, New Zealand; ⁷ Hospital for Special Surgery and Weill Cornell Medical College, New York, NY, USA; ⁸ FORCE-TJR and University of Massachusetts Medical School, Worcester, MA, USA; ⁹ Canadian Joint Replacement Registry and Dalhousie University, Halifax, Nova Scotia, Canada; ¹⁰ Danish Hip Arthroplasty Register and University of Southern Denmark, Odense, Denmark; ¹¹ Nuffield Department of Population Health and University of Oxford, Oxford, UK.

Correspondence: ola.rolfson@vgregion.se

Submitted 2015-08-15. Accepted 2015-09-17.

Abstract — The International Society of Arthroplasty Registries (ISAR) Steering Committee established the Patient-Reported Outcome Measures (PROMs) Working Group to convene, evaluate, and advise on best practices in the selection, administration, and interpretation of PROMs and to support the adoption and use of PROMs for hip and knee arthroplasty in registries worldwide. The 2 main types of PROMs include generic (general health) PROMs, which provide a measure of general health for any health state, and specific PROMs, which focus on specific symptoms, diseases, organs, body regions, or body functions. The establishment of a PROM instrument requires the fulfillment of methodological standards and rigorous testing to ensure that it is valid, reliable, responsive, and acceptable to the intended population.

A survey of the 41 ISAR member registries showed that 8 registries administered a PROMs program that covered all elective hip or knee arthroplasty patients and 6 registries collected PROMs for sample populations; 1 other registry had planned but had not started collection of PROMs. The most common generic instruments used were the EuroQol 5 dimension health outcome survey (EQ-5D) and the Short Form 12 health survey (SF-12) or the similar Veterans RAND 12-item health survey (VR-12). The most common specific PROMs were the Hip disability and Osteoarthritis Outcome Score (HOOS), the Knee injury and Osteoarthritis Outcome Score (KOOS), the Oxford Hip Score (OHS), the Oxford Knee Score (OKS), the Western Ontario and McMaster Universities Arthritis Index (WOMAC), and the University of California at Los Angeles Activity Score (UCLA). ■

Establishment of Working Group team

In early 2014, the International Society of Arthroplasty Registries (ISAR) Steering Committee established 5 working groups to further enhance the aims of the society: (1) the Bylaws and Future Funding Committee, (2) the Data Quality and Harmonization Group, (3) the Scientific Committee, (4) the Quality Improvement Committee, and (5) the Patient-Reported Outcome Measures (PROMs) Working Group. The composition of the PROMs Working Group members was balanced to cover different continents, registry affiliations, and professions. This report reviews the rationale and features of PROMs for use in arthroplasty registries, and describes the results of a survey of arthroplasty registries about their current use of PROMs for hip and knee arthroplasty.

Rationale of work

There are currently several large regional or national programs that collect and monitor PROMs before and after arthroplasty surgery (Department of Health 2008, Rolfson et al. 2011a, 2011b, Franklin et al. 2012). These PROMs complement traditional outcomes data such as complications, adverse events, reoperations, and revisions. Although primary outcomes after joint replacement include pain relief and improved function, clinician-based tests may be biased and may not be valid to describe patient self-perceptions of health status. Despite having limitations, PROMs are the best tools available to measure patient-centered outcomes objectively.

The greater emphasis on PROMs in routine care during the last few years has been stimulated by the introduction of the concept of value-based healthcare, in which decisions about the best ways to deliver healthcare should be based on factors that add value for the patient (Porter 2010). Traditional measures are necessary, but PROMs are fundamental in understanding and evaluating healthcare from a value-based standpoint. Although revision surgery and major adverse events are rare complications, improvements are necessary because 10% to 20% of patients who have hip or knee replacement are dissatisfied with the outcome, mainly because of persistent pain and limited function (Rolfson et al. 2011a, Dunbar et al. 2013).

Mission and aim

The ISAR PROMs Working Group is an international panel of collaborators that was organized to convene, evaluate, and advise on best practices in the selection, administration, and interpretation of PROMs and to support the adoption and use of PROMs in hip and knee arthroplasty registries worldwide. Specific aims include advising about (1) which PROMs are best and most appropriate to use, (2) when to capture PROMs in relation to treatment, (3) data collection methods, (4) approach to analysis, and (5) optimal ways to present results. The Working Group also aims to (1) promote harmonization, to allow direct comparisons within and between centers over time, (2) avoid excessive complexity, and (3) ensure that standards are guided by research agendas and not by the special interests of companies or other funders of healthcare.

Definitions

Patient-reported outcomes and patient-reported outcome measures

A patient-reported outcome is defined as any report of a patient's health status that comes directly from the patient without interpretation by others. The PROMs are standardized instruments designed to measure specific phenomena or constructs of the health status of patients in defined populations (U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research et al. 2006). The term outcome may cause confusion because it implies a measurement that occurs only after an intervention. By definition, true outcomes involve the measurement of change and typically require repeated measures before—and/or at intervals after—interventions or during the disease course. Therefore, cross-sectional patient-reported measurements after treatment do not generate true outcomes.

Generic and specific patient-reported outcome measures

There are 2 main types of PROMs that are distinguished by different levels of focus. Generic (or general health) instruments are designed to provide a measure of general health for any health state, regardless of the presence or absence

of illness, disability, or specific symptoms. Generic PROMs describe a patient's global health status that is comparable across different conditions.

Specific PROMs focus on specific symptoms, diseases, organs, body regions, or body functions. The specific PROMs may also be specifically designed to measure the effect of a specific intervention or treatment.

Methodological standards for the establishment of patient-reported outcome measures instruments

The establishment of a PROMs instrument requires the fulfillment of a set of methodological standards. The involvement of patients with methods such as interviews or focus groups to generate item content is important to ensure that the instrument will reflect the patient's standpoint and have valid content (U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research et al. 2006). The instrument must be tested rigorously to ensure that it is valid, reliable, responsive, and acceptable to the intended population. These properties are not fixed properties of an instrument but may relate to the population, the condition, or the treatment studied. The timing of questionnaire completions relative to the date of intervention may also affect the instrument's measurement properties.

After the content of a PROMs instrument has been established by the developers, none of the wording in the PROM may be changed, including word or item order. In addition to copyright issues, small changes in wording may change the perceived meaning of the items, and this may affect the measurement properties (Wild et al. 2005, Rothman et al. 2009).

Validity

Validity is the ability of an instrument to measure the intended outcome. A valid instrument adequately captures particular aspects of a person's health such as pain, mobility, or social functioning. Validity is an indication of the extent to which an assessment measures a particular construct in a particular situation, and a measure may be valid for one purpose but not another.

The different types of validity include content validity, which is the extent to which an instrument measures the intended concept. Content validity is specific to the population and treatment of interest. Content validity must be established before other measurement properties are evaluated, because other types of validity or reliability cannot overcome problems in content validity. Evidence to support content validity comes from qualitative studies that determine whether the items and domains of an instrument are appropriate and comprehensive relative to the intended measurement concept, population, and application. For PROMs, items and domains should reflect concerns of importance to most patients in the study. Content validity can be increased with (1) patient input in item generation, and (2) evaluation of patient understanding through cognitive interviews and pre-tests.

Criterion validity compares an instrument's measurement properties to a known standardized measure of the same construct, but such a measure may not exist. Construct validity is the extent to which an instrument forms pre-specified logical relations such as correlations between items or domains, and correlations to other established instruments or characteristics of patients or patient groups.

Reliability

Reliability has 2 elements: internal consistency and repeatability (also known as *reproducibility*). Internal consistency is the correlation between different items within a domain, and is a measure of the extent to which these items measure the concept of interest. Internal consistency is measured with Cronbach's alpha, a summary correlation; a high Cronbach alpha may indicate redundancy, and a low Cronbach alpha (< 0.70) may indicate inconsistency. However, the optimal correlation as expressed with Cronbach's alpha may be difficult to define.

Repeatability (also known as *test-retest reliability*) is the stability of a measure over time. Repeatability is a measure of the random variability in a patient's responses to the same item when repeated measures are assessed under the same conditions and no real change has occurred. Repeatability can be evaluated with different statistical tests and is considered acceptable when the intraclass correlation coefficient is > 0.70.

Responsiveness

Responsiveness is defined as an instrument's ability to detect change when change has occurred. Adequate responsiveness is important for an instrument that is used to assess clinical changes. When a patient reports that a change has occurred, but the PROMs instrument score does not change, then the instrument may have inadequate ability to detect change or questionable validity.

Assessments of responsiveness of a measure require repeated assessments over time when the patient's condition of interest has changed. Condition- or context-specific measures are more responsive to changes in the condition of interest than generic measures, which may respond to other co-existing conditions increasingly with time.

In addition to showing that an instrument may detect change, it is useful to have an estimate of the minimal change of a measure, which is the smallest change that indicates that an important or meaningful change has occurred. An estimate of minimal change is needed to determine the real clinical or subjective meaning of any observed changes, whether or not the changes are statistically significant. Statistical significance depends, in part, on sample size. A small, clinically irrelevant, change may be statistically significant when the sample size is large. A large clinically relevant change may not be statistically significant in a small study (Petersen 2001).

Minimal change can be estimated as either the smallest amount of change that is of relevance to patients (minimal clinically important difference (MCID)) or the amount of

change that is beyond simple measurement error (minimal detectable change (MDC)). The MCID is termed anchor-based, as it uses anchor questions to establish minimal change, and the MDC is distribution-based, i.e. based on statistical characteristics of the sample at issue. The MCID and MDC produce different estimates, but it is important that an estimate of minimal change that is produced using an anchor is larger than the measurement error of the instrument to confirm that the change is real (Jaeschke et al. 1989, Wyrwich et al. 1999).

Acceptability and feasibility

To optimize an instrument's properties, it is important to consider the burden to respondents and healthcare personnel in completing and administering questionnaires. Shorter measures that are written in clear, unambiguous language will encourage a higher frequency of response, which is often used as supporting evidence of the acceptability of a questionnaire.

Translations and cultural differences in response patterns

The translation of a validated PROMs instrument into different languages requires a formal methodological approach and specific expertise (Wild et al. 2005). Such methods typically include consideration and testing of cultural equivalence, especially when there is no exact word or concept available in the target language or culture to match the word or concept that is used in the primary-language version of the measure being translated. Several steps are involved in the process of translating a PROMs instrument into another language, such as forward translation, reconciliation, back translation, harmonization, review, and cognitive debriefing (Wild et al. 2005).

The use of a PROMs instrument with different platforms, such as electronic (ePRO/ePROM) and hard-copy versions, may not be straightforward and may require changes in format or minor word changes, including instructions for completion. Recommendations for best practice are available for modification of the platform, including how much change in wording is considered minor (Coons et al. 2009, Rothman et al. 2009).

The quality may vary between different translated or electronic versions of a PROMs instrument. When a translated version of a PROMs instrument is required for a specific language, it is therefore important to confirm that proper methods have been used and that the license holder supports a particular translated version. Caution is recommended in developing or adopting specific versions of ePROMs.

Other patient-reported measures related to health outcomes

Satisfaction and similar measures: In addition to measures that directly address health status, there are useful patient-reported measures related to health outcomes such as measures of satisfaction with treatment effects, fulfillment of expectations, and willingness to repeat or recommend treatment to others. These interrelated measures are not true PROMs, by definition, and can be used only after an intervention or to evaluate some-

thing that has already occurred. Nevertheless, satisfaction and similar measures are associated with changes in PROMs scores and the patient's experience of care delivery. When true PROMs have inherent limitations that fail to determine treatment success, satisfaction items may be useful adjuncts.

Health transition: Health transition items are another category of patient-reported measures that address a self-perceived change over a defined period for a given construct. Health transitions can be used only retrospectively. Health transition items are often used in conjunction with health status measures, and may be incorporated into questions within a PROMs instrument, such as the Veterans RAND 12-item health survey (VR-12) question: "Compared to 1 year ago, how would you rate your physical health in general now?", which has 5 response options ranging from "much better" to "much worse". However, scores of health transition items are usually not incorporated into the overall PROMs score. As indicators of clinically important improvement or deterioration, transition items are commonly used to anchor changes in a particular scale. Response bias may occur from difficulty in recalling previous states and the tendency to accommodate, or due to changes in an individual regarding internal standards, values, or conceptualization—which is also known as response shift.

Survey of collection of patient-reported outcome measures in arthroplasty registries

An e-mail survey was administered in September 2014 to full ISAR member registries ($n = 12$) and associate ISAR member registries ($n = 29$) to determine the current use of PROMs in arthroplasty registries. 8 registries administered a PROMs program that covered all elective hip or knee arthroplasty patients and 6 other registries collected PROMs for sample populations; 1 other registry had planned but had not started collection of PROMs (Table). The registries most often collected PROMs for research ($n = 10$), quality assurance ($n = 8$), or hospital assessment ($n = 8$). Some programs integrated their PROMs program in the clinic to provide input to caregivers for the care of individual patients ($n = 6$). Some registries used PROMs programs for assessment of surgeons ($n = 4$) and health regions ($n = 4$).

Different instruments were used by different registries (Table). The most common generic instruments were the EuroQol 5-dimension health outcome survey (EQ-5D) (EuroQol Group 1990, EuroQol Group 2015) and the Short Form 12 health survey (SF-12) or the VR-12, which is similar (Ware et al. 1992, Ware et al. 1996, Kosinski et al. 1999, Rolfson et al. 2011b, Boston University School of Public Health 2015, Optum Inc. 2015). 9 registries (60%) collected a satisfaction item. The most common specific PROMs were the Hip disability and Osteoarthritis Outcome Score (HOOS) (Nilsson et al. 2003), the Knee injury and Osteoarthritis Outcome Score (KOOS) (Roos et al. 1998), the Oxford Hip Score

(OHS) (Dawson et al. 1996, Murray et al. 2007), the Oxford Knee Score (OKS) (Dawson et al. 1998, Murray et al. 2007), the Western Ontario and McMaster Universities Arthritis Index (WOMAC) (Bellamy et al. 1988), and the University of California at Los Angeles Activity Score (UCLA) (Zahiri et al. 1998). Several registries reported use of a separate visual analog scale (VAS) or numeric rating scale (NRS) to measure joint pain (Breivik et al. 2000, Hawker et al. 2011, Hjerms-tad et al. 2011). Most registries collected preoperative measures and had several follow-ups, most commonly at 1 year after surgery. Preoperative PROMs were usually collected at an outpatient visit or preoperative class. Follow-up surveys were usually collected by regular mail (printed forms), but some registries also used e-mail, electronic (web) surveys, and telephone calls. Response frequencies for PROMs were in some cases approximate estimates and should be interpreted with caution. The variation in frequency of patient responses reflects the logistic challenges of collecting PROMs.

The ISAR PROMs Working Group thanks Elly Trepman for help in reorganizing the original ISAR PROMs Working Group internal report into 2 manuscripts.

All the authors participated in the conception of the study. OR and JD drafted the manuscript. KC conducted the survey. All the authors were involved in checking and revising the article critically for important intellectual content.

Bellamy N, Buchanan W W, Goldsmith C H, Campbell J, Stitt L W. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988; 15(12): 1833-40.

Boston University School of Public Health. VR-36, VR-12 and VR-6D. Boston University School of Public Health Web site. <http://www.bu.edu/sph/research/research-landing-page/vr-36-vr-12-and-vr-6d/>. 2015.

Breivik E K, Björnsson G A, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. *Clin J Pain* 2000; 16(1): 22-8.

Coons S J, Gwaltney C J, Hays R D, Lundy J J, Sloan J A, Revicki D A, Lenderking W R, Cella D, Basch E, ISPOR ePRO Task Force. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009; 12(4): 419-29.

Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br* 1996; 78(2): 185-90.

Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br* 1998; 80(1): 63-9.

Department of Health. Guidance on the routine collection of Patient-Reported Outcome Measures (PROMs) for the NHS in England 2009/10. UK: Department of Health; 2008.

Dunbar M J, Richardson G, Robertsson O. I can't get no satisfaction after my total knee replacement: rhymes and reasons. *Bone Joint J* 2013; 95-B(11 suppl A): 148-52.

EuroQol Group. EQ-5D. EuroQol Group Web site. <http://www.euroqol.org>. 2015.

Table 1. Survey responses of hip and knee arthroplasty registries that routinely collected patient-reported outcome measures ^a

Registry Type	Name	Joints included ^b		Patients included ^c	PROMs collected		Satisfaction item	Frequency of patient response		Data collection times	
		Hip	Knee		Generic	Specific		Preop.	Postop.	Preop.	Postop.
National registries											
	Swedish Hip A R	+	–	All	EQ-5D	Pain VAS	Yes VAS	86	90	+	1, 6, 10 y
	National J R (through NHS)	+	+	All	EQ-5D	OHS/OKS	Yes	80	80	+	6 mo
	Swedish Knee A R	–	+	Sample	EQ-5D	KOOS, pain VAS	Yes, expectation fulfillment	90	80	+	1 y
	New Zealand J R ^b	+	+	Random 20% sample		OHS/OKS	–	NA	70–75	–	6 mo, every 5 y
	Lithuanian A R	+	+	Sample	EQ-5D	HOOS/KOOS	–	100 (of sample)	60	+	6 mo, 1 y
	Norwegian A R	+	+	Occasional cross-sectional samples	EQ-5D	HOOS/KOOS	Yes	NA	80	–	1 to 2 y
	Dutch A R ^b	+	+	All	EQ-5D, health transition	OHS/OKS, pain NRS	Yes	50	50	+	3 and 6 mo, 1 y
National sample registries											
	FORCE-TJR	+	+	All	SF-36	HOOS/KOOS, pain VAS	–	80–85	80–85	+	6 mo, annually
	American J R	+	+	Not started ^d	EQ-5D, SF-12 or SF-36 ^{e,f}	WOMAC or HOOS/KOOS ^f	–	^d	^d	^f	^f
Local or regional registries											
	Harris J R	+	+	All	EQ-5D	Pain VAS, HHS/ KOOS, UCLA	Yes VAS	^g	^g	+	1, 3, 5, 7, and 10 y
	Register of the Orthopaedic Prosthetic Implants (Italy)–	+	–	Sample	EQ-5D	HOOS	Yes	30	80	+	1 y
	Michigan A R	+	+	All	SF-12 ^e	WOMAC, UCLA		32	12	+	Annually
	Geneva A R	+	+	All	SF-12 ^e	WOMAC, Pain VAS, UCLA	Yes, willingness to repeat	71	65	+	1, 5, 10, and 15 y
	Hospital for Special Surgery	+	+	All	SF-12 ^e	HOOS/KOOS	Yes	80	75	+	Annually
	California J R	+	+	All	SF-12 ^e	WOMAC, UCLA	–	70	70	+	6 mo, 1 y, every 2 y

^a n = 15 registries (7 national registries, 2 national samples, and 6 local or regional registries). The survey was performed in September 2014.

A R: Arthroplasty Register/Registry, J R: Joint (Replacement) Register

^b The New Zealand Joint Registry and the Dutch Arthroplasty Register were the only 2 registries that included shoulder, elbow, and ankle joint arthroplasty.

^c Patients included: all, all patients in a country, region or hospital; sample, sample of patients in a country, region, or hospital.

^d Not started; the American Joint Replacement Register were planning but had not started a PROMs program as of September 2014.

^e SF-12 or the similar Veterans RAND 12-item health survey (VR-12).

^f As preferred by healthcare provider.

^g Not Known

EQ-5D, EuroQol 5-dimension health outcome survey; FORCE-TJR, Function and Outcomes Research for Comparative Effectiveness in Total Joint Replacement; NA, not applicable, NHS, National Health Service, England; HOOS, Hip disability and Osteoarthritis Outcome Score; KOOS, Knee injury and Osteoarthritis Outcome Score; NRS, numeric rating scale; OHS, Oxford Hip Score; OKS, Oxford Knee Score; SF-12 (or 36), Short Form 12 (or 36) health survey; PROMs, patient-reported outcome measures; UCLA, University of California at Los Angeles Activity Score; VAS, Visual Analog Scale; WOMAC, Western Ontario and McMaster Universities Arthritis Index.

EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16(3): 199-208.

Franklin P D, Allison J J, Ayers D C. Beyond joint implant registries: a patient-centered research consortium for comparative effectiveness in total joint replacement. *JAMA* 2012; 308(12): 1217-8.

Hawker G A, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken)* 2011; 63 (suppl 11): S240-52.

- Hjermstad M J, Fayers P M, Haugen D F, Caraceni A, Hanks G W, Loge J H, Fainsinger R, Aass N, Kaasa S, European Palliative Care Research Collaborative (EPCRC). Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage* 2011; 41(6): 1073-93.
- Jaeschke R, Singer J, Guyatt G H. Measurement of health status. Ascertain- ing the minimal clinically important difference. *Control Clin Trials* 1989; 10(4): 407-15.
- Kosinski M, Keller SD, Hatoum H T, Kong S X, Ware J E Jr. The SF-36 Health Survey as a generic outcome measure in clinical trials of patients with osteoarthritis and rheumatoid arthritis: tests of data quality, scaling assumptions and score reliability. *Med Care* 1999; 37(5 suppl): MS10-22.
- Murray D W, Fitzpatrick R, Rogers K, Pandit H, Beard D J, Carr A J, Dawson J. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br* 2007; 89(8): 1010-4.
- Nilsdotter A K, Lohmander L S, Klässbo M, Roos E M. Hip disability and osteoarthritis outcome score (HOOS) – validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003; 4: 10.
- Optum Inc. SF-36.org: a community for measuring health outcomes using SF tools. Optum Inc. Web site. <http://www.sf-36.org>. 2015.
- Petersen J H. Book review: *Quality of Life. Assessment, Analysis and Interpretation*. P. M. Fayers and D. Machin, Wiley, Chichester, 2000. No. of pages: xii +404. Price: £60.00. ISBN 0-471-96861-7. *Stat Med* 2001; 20(14): 2214-6.
- Porter M E. What is value in health care? *N Engl J Med* 2010; 363(26): 2477-81.
- Rolfson O, Kärrholm J, Dahlberg L E, Garellick G. Patient-reported outcomes in the Swedish Hip Arthroplasty Register: results of a nationwide prospec- tive observational study. *J Bone Joint Surg Br* 2011a; 93(7): 867-75.
- Rolfson O, Rothwell A, Sedrakyan A, Chenok K E, Bohm E, Bozic K J, Garellick G. Use of patient-reported outcomes in the context of different levels of data. *J Bone Joint Surg Am* 2011b; 93 (suppl 3): 66-71.
- Roos E M, Roos H P, Lohmander L S, Ekdahl C, Beynnon B D. Knee Injury and Osteoarthritis Outcome Score (KOOS) – development of a self-admin- istered outcome measure. *J Orthop Sports Phys Ther* 1998; 28(2): 88-96.
- Rothman M, Burke L, Erickson P, Leidy NK, Patrick D L, Petrie C D. Use of existing patient-reported outcome (PRO) instruments and their modifica- tion: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modifica- tion PRO Task Force Report. *Value Health* 2009; 12(8): 1075-83.
- U.S. Department of Health and Human Services FDA Center for Drug Evalu- ation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medi- cal product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006; 4: 79.
- Ware J E Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF- 36). I. Conceptual framework and item selection. *Med Care* 1992; 30(6): 473-83.
- Ware J Jr, Kosinski M, Keller S D. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; 34(3): 220-33.
- Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Eri- kson P, ISPOR Task Force for Translation and Cultural Adaptation. Prin- ciples of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005; 8(2): 94-104.
- Wyrwich K W, Tierney W M, Wolinsky F D. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999; 52(9): 861-73.
- Zahiri C A, Schmalzried T P, Szuszczewicz E S, Amstutz H C. Assessing activity in joint replacement patients. *J Arthroplasty* 1998; 13(8): 890-5.