ELSEVIER

CrossMark

Journal of Clinical Epidemiology 68 (2015) 73–79

Journal of Clinical Epidemiology

Meaningful changes for the Oxford hip and knee scores after joint replacement surgery

David J. Beard[a], Kristina Harris[a,*], Jill Dawson[b], Helen Doll[c], David W. Murray[a], Andrew J. Carr[a], Andrew J. Price[a]

[a]*Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Headington, Oxford OX3 7LD, UK*

[b]*Department of Public Health, University of Oxford, Rosemary Rue Building, Old Road Campus, Headington, Oxford OX3 7LF, UK*

[c]*Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK*

Accepted 1 August 2014; Published online 31 October 2014

Abstract

Objectives: To present estimates of clinically meaningful or minimal important changes for the Oxford Hip Score (OHS) and the Oxford Knee Score (OKS) after joint replacement surgery.

Study Design and Setting: Secondary data analysis of the NHS patient-reported outcome measures data set that included 82,415 patients listed for hip replacement surgery and 94,015 patients listed for knee replacement surgery was performed.

Results: Anchor-based methods revealed that meaningful change indices at the group level [minimal important change (MIC)], for example in cohort studies, were ~11 points for the OHS and ~9 points for the OKS. For assessment of individual patients, receiver operating characteristic analysis produced MICs of 8 and 7 points for OHS and OKS, respectively. Additionally, the between group minimal important difference (MID), which allows the estimation of a clinically relevant difference in change scores from baseline when comparing two groups, that is, for clinical trials, was estimated to be ~5 points for both the OKS and the OHS. The distribution-based minimal detectable change (MDC_{90}) estimates for the OKS and OHS were 4 and 5 points, respectively.

Conclusion: This study has produced and discussed estimates of minimal important change/difference for the OKS/OHS. These estimates should be used in the power calculations and the interpretation of studies using the OKS and OHS. The MDC_{90} (~4 points OKS and ~5 points OHS) represents the smallest possible detectable change for each of these instruments, thus indicating that any lower value would fall within measurement error.

Keywords: Minimal important change; Minimal important difference; Hip replacement; Knee replacement; Responder definition; Study designs

1. Introduction

The Oxford hip and knee scores are widely used patient-reported outcome measures (PROMs) in research, audit, and clinical practice. In addition to their appropriate development, PROMs used in the context of arthroplasty should have well-established measurement properties, assessed in that context, such as evidence of validity, reliability, and responsiveness. It is also important that scores are interpretable, that is, that qualitative meaning can be assigned to a particular quantitative score or to a difference or change in the score [1]. Determining whether statistically significant changes in the Oxford Hip Score (OHS) and the Oxford Knee Score (OKS) are also clinically meaningful is essential for judging the efficacy of joint replacement. Such interpretation is needed for assessing change in single groups of patients over time (ie, cohort studies), for differences between groups (ie, clinical trials) and for assessing changes in individual patients. Individual patient-level assessment is important as clinicians (and patients) increasingly use individual PROM scores for personal decision making. This might involve identification of appropriate

K.H. and D.J.B. contributed equally to this work and should be regarded as joint first authors.

Conflict of interest: J.D. is one of the original inventors of the Oxford Hip Score and Oxford Knee Score. She has received consultancy payments, via Isis Innovation, in relation to work involving both questionnaires. The other authors have no conflicts of interest.

Funding: This work was supported under the general program of research undertaken by the Nuffield Department of Orthopedics, Rheumatology and Musculoskeletal Sciences as a National Institute of Health Research Biomedical Research Unit- Musculoskeletal Disease.

* Corresponding author. Tel.: +44 (0)1865 223402; fax: +44 (0)1865 227966.

E-mail address: kristina.harris@ndorms.ox.ac.uk (K. Harris).

http://dx.doi.org/10.1016/j.jclinepi.2014.08.009

What is new?

- An array of values representing clinically meaningful changes has been provided for the Oxford Knee Score (OKS) and the Oxford hip Score (OHS).
- This study builds on a previous study that provided only minimal important difference (MID) values for the OKS/OHS (which were based on a much smaller sample) and provides and discusses an array of minimal important change/MID values.
- The results of this study will enable researchers and clinicians to better interpret changes in the OKS and OHS, after joint replacement surgery both in the research setting and in the clinical practice.

timing of an intervention or assessment of progress and/or deterioration after intervention or delay in treatment (such as hip replacement).

Two different approaches can be used to estimate and interpret the smallest amount of change in a score that could be considered to have clinical importance [2]. These are commonly termed as ''anchor-based'' or ''distribution-based'' methods [3].

Anchor-based methods explore how an observed change or difference in the score on the instrument relates to an external criterion or relevant anchor (eg, responses on a global transition item). The anchor can be rated or set by the patient, clinician, or other stakeholder. Anchor-based methods are, by definition, more likely to be clinically relevant as they relate the score or changes in the score to a clinically meaningful reference measure (a little better, about the same, and so forth). Anchor-based methods can provide information at both group and individual levels. For example, a hypothesis may explore (1) the *change* in health status in a single group or a single individual over time [often referred to as the minimal important change (MIC)] or (2) the *difference* in health gain or loss between two independent groups of patients [the minimal important difference (MID)].

Distribution-based methods are based on the statistical characteristics of the sample in a particular study [2–4]. The observed change is expressed as a standardized metric with examples including the effect size (ES), the standard error of measurement (SEM), and the minimal detectable change (MDC). Apart from the ES, which is normally applied to comparisons between groups, distribution-based measures can provide information at an individual patient level. For example, the MDC is the smallest change, *for an individual*, that is likely to be beyond the measurement error of the instrument and therefore to represent a true change.

It is acknowledged that these definitions and their use can cause confusion [3,4], and in the literature, several terms are often used interchangeably.

In this article, we calculate and describe estimates of meaningful change and difference for the Oxford hip and knee scores and discuss how they should be used.

2. Materials and methods

An analysis of the NHS PROMS data set of all hip and knee replacements undertaken from January 1, 2009, to December 31, 2011, in England and Wales was performed. Full PROMs data reports and methodology guides can be found online [5].

2.1. Subjects and/or assessments

As a part of the NHS PROMS program, patients who were listed for primary joint replacement surgery completed a set of preoperative questionnaires including the OHS or the OKS. The OHS and the OKS are both 12-item questionnaires that address pain and functional disability in relationship to the patient's hip or knee problems, respectively. Items were originally devised using interviews with patients undergoing joint replacement surgery, so that they would reflect the patient's perspective. In each case, item responses have five categories and are Likert scaled. The original scoring system was from 1 to 5, with a summary score ranging from 12 (best) to 60 (worst). The recommended scoring system has since changed with items now scored from 0 to 4, with a summary score range of 0 (worst) to 48 (best) [6]. Although both questionnaires contain the same number of items and are scored similarly, their scales are not equivalent to each other (ie, a score of 10 on the OHS cannot be assumed to represent the same level of severity as a score of 10 on the OKS).

At 6 months after surgery, the measures were repeated and patients also completed a global transition item (''overall, how are your <hip/knee> problems now, compared to before your operation?'') with five response categories: ''much better'' (scored 1), ''a little better'' (2), ''about the same'' (3), ''a little worse'' (4), and ''much worse'' (5).

2.2. Statistical methods

Data were analyzed using IBM SPSS Statistics version 20.0 (SPSS Inc., Chicago, IL, USA). Change scores followed a normal distribution, which allowed the use of parametric statistics. Change-related parameters for group-level estimates were examined using both anchor- and distribution-based methods.

2.2.1. Anchor-based method

Initially, the appropriateness of the anchor item to record change in the OKS and OHS was first assessed by