



Published in final edited form as:

J Rheumatol. 2009 September ; 36(9): . doi:10.3899/jrheum090356.

Variation in Outcome Measures in Hip and Knee Arthroplasty Clinical Trials: A Proposed Approach to Achieving Consensus

Daniel L. Riddle, Paul W. Stratford, Jasvinder Singh, and Vibeke Strand

Department of Physical Therapy, Virginia Commonwealth University, Richmond, VA 23298

Daniel L. Riddle: dlriddle@vcu.edu; Paul W. Stratford: stratfor@McMaster.ca; Jasvinder Singh: jasvinder.md@gmail.com; Vibeke Strand: Vstrand@aol.com

Abstract

OMERACT began work over a decade ago on a consensus effort to identify optimal outcome measures for knee and hip osteoarthritis randomized controlled trials (RCTs). Recent evidence has indicated that extensive variation exists in the outcome measures used in RCTs of knee and hip arthroplasty published since 2000. This heterogeneity leads to confusion, not only for researchers attempting to conduct systematic reviews but also for clinicians who are trying to apply evidence to clinical practice. Given the extensive amount of psychometric research conducted in the past two decades, it would appear that the timing is ideal to design and implement a study to develop consensus on optimal outcome measures for hip and knee arthroplasty trials. A Delphi survey design is described along with an approach for synthesizing the extensive psychometric literature on the outcome measures that are used in hip and knee arthroplasty trials. Plans for dissemination of the findings are also discussed.

MeSH terms

arthroplasty; replacement; knee; outcome assessment

Lower extremity joint replacement surgery is a common and highly effective procedure for many patients with arthritis.¹ In 2002, approximately 400,000 knee arthroplasty surgeries and approximately 200,000 hip arthroplasty surgeries were conducted in the US.² Trend data reported by Kurtz and colleagues suggest that by 2010, approximately 1 million hip and knee arthroplasty surgeries will be conducted with approximately 15% of these being revision surgeries and the remaining 85% primary surgeries.³ Cost data for joint arthroplasty surgery are also impressive. Ong and colleagues determined the combined hospital and physician procedural charges for Medicare patients receiving joint arthroplasty during the years 1997 to 2003.⁴ Mean procedural charges per patient from 1997 to 2003, in 2005 dollars was \$40,000 per primary surgery and \$50,000 per revision surgery. If the reimbursement data were extrapolated to the 2006 volume data projections, hospital and physician charges for hip and knee arthroplasty would total approximately \$13 billion in the US. Given the volume and costs of hip and knee arthroplasty, it is not surprising that an extensive research emphasis has been devoted to the effort¹ with over 160 randomized trials since 2000.⁵

Recognizing the need for valid, standardized methods for comparing outcome data, the Outcome Measures in Rheumatology (OMERACT) used a consensus-based approach in

1997 to identify optimal outcome measures for knee and hip osteoarthritis clinical trials.⁶ More than 90% of the participants agreed that pain, physical function, patient global ratings of improvement, and joint imaging procedures should be included in clinical trials of patients with osteoarthritis. Participants were unable to come to consensus regarding specific measures because measures had yet to be identified in the literature as superior for the different outcomes of interest. The consensus of the participants was that in the next 3–5 years (2000 to 2003), evidence should be sufficient to identify specific measures for clinical trials.

The past 20 years have seen a tremendous growth in the development and validation of outcome measures for patients with arthritis. Researchers and clinicians now have dozens of measures to choose from when caring for patients with hip or knee arthroplasty. This diversity of measurements, however, comes with a cost. Riddle and colleagues conducted a systematic review of outcomes measures used in contemporary clinical trials and found extensive variation in the numbers and types of outcome measures used in hip and knee arthroplasty trials.⁵ For example, of the 82 hip replacement trials published since 2000, the Harris Hip Score was used in 43 (52%) but an additional 19 measures were used in the trials. There was an extensive variation across trials, not only in the specific measures used but also in the general construct that was being measured. Extensive variation also was found in knee trials.

Heterogeneity in outcome measures across trials potentially leads to several problems. Clinicians who are attempting to integrate findings from multiple trials of an intervention cannot readily interpret findings when different outcome measures are used. Researchers conducting systematic reviews cannot calculate summary measures of effect if measures across trials are different. Finally, some outcome measures have superior measurement properties compared to others and measures with weaker psychometric properties continue to be used in hip and knee arthroplasty trials.⁵

Since the work of OMERACT in 1997 the Osteoarthritis Research Society International (OARSI), has also worked to improve clinical trial reporting for patients with osteoarthritis. OARSI has collaborated with OMERACT in the establishment of criteria for interpreting patient response in OA drug trials.^{7,8} The authors established criteria for judging the magnitude of treatment effect but did not identify the specific measures that should be used in OA drug trials. Neither OMERACT nor OARSI have developed a consensus for identifying specific measures for use in randomized controlled trials (RCTs) of patients with hip or knee arthroplasty.

Separately, the World Health Organization used a worldwide consensus-based approach in 2001 to develop the International Classification of Functioning Disability and Health (ICF).⁹ We believe that the ICF provides an ideal framework for conceptualizing outcome after knee and hip arthroplasty surgery from a biological, individual and societal perspective.

We propose to extend and bridge these initiatives through a multi-staged approach to establish consensus-based recommendations of specific post-hospitalization outcome measures after hospitalization for knee and hip arthroplasty trials. The purpose of our paper is to describe a research design and method for achieving consensus on optimal outcome measures for knee and hip arthroplasty randomized trials.

Methods

ICF as a conceptual framework for categorizing outcome measures

The World Health Organization formally adopted the *International Classification of Functioning, Disability and Health*, (ICF), as the standard language for describing health related states and conditions. The ICF is now the internationally agreed upon standard language to describe health.

The framework for the ICF model is illustrated in Figure 1. Each major component within the ICF model will be briefly defined. For a more thorough examination of ICF, several comprehensive descriptions are available. As applied to patients with joint arthroplasty, health conditions, the component at the top of the figure, equate to the arthritis, and any complications arising from treatment such as infection or deep vein thrombosis.

Body function and structure refers to the functioning and structural integrity of specific body organs and systems. Patients with hip or knee arthroplasty may, for example, have reduced muscle strength, joint swelling, pain and psychological distress.^{11,12} Activity is defined as the completion of a task or action by an individual. Limitations in commonly performed tasks for patients with hip or knee arthroplasty may be walking, bending, sitting and stairclimbing.^{13,14}

Participation is the term used to describe a patient's involvement in everyday life. When a person's everyday life is disrupted, the person's participation is restricted. For example, if a patient's ability to attend religious services was compromised, that person would have a participation restriction. In addition to the four components described above, there are two additional components that impact directly on Body Structures, Functions and Activity and Participation. These are termed contextual factors. Two contextual factors describe the complete background of an individual and their daily life—this is a poor introduction to contextual factors and should be reworded e.g. Two broad categories of contextual factors address the interacting influence of personal and environmental factors. Environmental factors are external to the person and influence that person's daily life. These factors include all features of the environment including policies, laws and values in that person's environment. Personal factors include gender, race, lifestyle and daily routines. The ICF is receiving worldwide support from a variety of areas in medicine and seems the ideal conceptual model to frame a study of the variation in outcome measures used in hip and knee arthroplasty trials.^{10,11,12}

Overview of proposed study design

The proposed study will focus on four of the six main components of the ICF model. The four components are body structure and function, activity, participation and personal factors. For the personal factors component, this study will focus on the measurement of patient satisfaction, the most commonly measured personal factor outcome in the hip and knee arthroplasty literature.⁵ Health conditions will not be examined because these outcomes are most commonly assessed during inpatient care and the proposed study design is focused on outcomes after hospitalization—the context of this is unclear. What health conditions? Does this fall under “body structures”? why is it brought up at all?. The proposed study also will not address outcome measures related to biomechanical issues such as prosthetic loosening. Environmental factors are not included primarily because they are rarely measured in hip and knee arthroplasty trials.⁵

Hip and knee replacement outcome measures will be considered separately. In addition, primary replacement will be considered as separate and distinct from revision surgery. Trials will be considered in three separate categories much like that described by Riddle and

colleagues.⁵ Optimal outcome measures for trials of surgical interventions, trials of non-surgical physical interventions (ie. physical therapy) and trials of non-surgical medical interventions (i.e. medication) will each be identified.

We have designed the study using a multi-staged process. For stage 1, literature that has examined the psychometric properties of the various outcome measures identified in the study by Riddle and colleagues⁵ will be identified and summarized. All instruments will be categorized into one of the four key domains of the ICF. For stage 2, a multi-step Delphi survey will be conducted. Experts participating in the Delphi survey will be provided succinct summaries of the psychometric properties of all instruments to guide them during the Delphi survey process. See Figure 2 for a summary of the flow of the study---there is too little information on this figure to make it really helpful.

The final outcome of the Delphi survey will be a consensus summary that identifies, 1) which of the four ICF components should be measured for each of the three types of RCTs conducted on patients with hip arthroplasty, knee arthroplasty and revision hip or knee arthroplasty, 2) the optimal outcome measures for each ICF component for primary and revision hip and knee arthroplasty. The consensus summary will be presented at an upcoming OMERACT conference to determine if international consensus can be achieved.

Stage 1 – Synthesis of psychometric literature—The goal of the literature synthesis is to locate relevant articles reporting the psychometric properties of the outcome measures identified in the systematic review of Riddle et al.⁵ We will search the Medline database and limit our search to articles presented in English. The general search strategy will be to present the name of the instrument, the terms *arthroplasty* or *replacement*, the location of *hip* or *knee*, and a comprehensive set of measurement terms. We expect to conduct searches for approximately 60 outcome measures. All searches will be conducted by the investigative team.

We provide an example to illustrate the approach that will be used. For example, the following are search examples for the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). The WOMAC search (“Western Ontario and McMaster Universities Osteoarthritis Index” OR WOMAC) AND (arthroplasty OR replacement) AND (hip OR knee) AND (change OR valid* OR reliab* OR sensitiv* OR responsive* OR psychometric OR clinimetric) yielded 97 articles.

Data from the searches will be abstracted and critiqued by trained abstractors please clarify how this will be standardized. Abstractors will be six graduate students or clinicians with a background in rehabilitation science and familiarity with the concepts of reliability, validity, and responsiveness. Candidate abstractors fulfilling these requirements will undergo a training program which will consist of instruction in the ICF model, sample exercises and sample critiques with guided feedback.

Following the training program, candidate abstractors will be evaluated on 6 research papers, three of which will be head-to-head comparison studies of competing measures. The gold standard for these critiques will be set by the investigator. It is not clear how the investigator will evaluate the candidate abstractors. Please clarify what will be the specific benchmarks by which each abstractor will be evaluated in a uniform manner. The gold standard set of 6 papers will be reviewed by the investigator prior to the start of the study. Only candidate abstractors whose responses are consistent with the gold standard on all 6 papers will be accepted as abstractors for this component of the study.

The results will be compiled at two levels. The first level will group the information by ICF categories; the second level will summarize the material for competing measures within each ICF component—this is too brief. Please elaborate because it is not clear to me how the two levels are distinct. The second level summary also will present the psychometric properties of a measure, and when available, the results from head-to-head comparison studies of competing measures. The results of these summaries will be presented in Rounds 2 and 3 of the consensus exercise. The summaries will present “typical” values or information and not an exhaustive review of the literature concerning a measure. A glossary of terms will accompany the psychometric summaries.

Stage 2 - Consensus Delphi approach—The Delphi procedure is a method for achieving consensus of opinion among a panel of experts on a topic when there is lack of agreement or an incomplete state of knowledge.^{13,14} The classic Delphi method usually consists of three rounds of questions with the format of the first round questions being open-ended.^{13,14} However, modified versions of this method abound and they are often characterized by semi-structured or structured first round questions.^{13,15,16} Our proposed consensus design will apply a modified Delphi method which combines structured and open-ended questions in the first round with a total of three rounds to achieve consensus. All communication with Delphi participants will be via email correspondence.

The principal determinant of sample composition is credibility with the target audience.¹⁴ We believe the target audience to be those who conduct clinical trials in the field of hip or knee arthroplasty and consumers of their work. It is with this group in mind that we define our expert panel. Jairah and Weinstein suggest that experts should not only be knowledgeable with the topic area, but also that they must be impartial to the findings.¹⁷ Delbecq et al have noted that heterogeneous groups, characterized by panel members with different perspectives produce a higher proportion of high quality solutions than homogeneous groups.¹⁸

Our expert panel will consist of a North American representation of orthopaedic surgeons, rheumatologists, and physical therapists who have published a minimum of five peer reviewed papers relevant to the assessment of patients post hip or knee arthroplasty---does this include co-authors? Is a ranking of authorship appropriate e.g. preference for those who are first/senior authors?. At least half the experts will meet the criterion of five papers for the hip, the knee or both. Our rationale for only selecting experts who have published peer reviewed work is that these individuals have been evaluated by their peers and found to produce credible work. We chose this approach over, for example, select members of professional societies because we did not want potential selection to be politically driven. In addition, we will exclude (please use consistent tense) excluded those who published work describing the development of an outcome measure. Developers of outcome instruments may be particularly biased toward their own measures. Panel members will be identified from the body of literature dealing with hip and knee arthroplasty. The proportion of experts sampled from a specific discipline will be representative of the number of studies authored by members from that discipline with the maximum representation from any one discipline being 66% (what about studies with numerous co-authors from different disciplines?). Authors will be stratified by discipline and geographical location and purposely sampled within strata. Preference will be given to authors who have contributed the most to the literature, as determined by publication counts of researchers identified by the research team and by medline searches using key words (replacement OR arthroplasty) AND (hip OR knee).??.regardless of authorship rank. I think many will have concerns with this approach.

Delphi exercise panel sizes have varied widely (e.g., 10 to over 1000) and appear to be driven by available resources including the pool of likely participants, and the time and cost

associated with managing and summarizing data.¹⁹ Given the lack of consensus on a method for estimating the requisite sample size, our goal is to have complete data on 30 panel members. A sample of this size is consistent with other reported Delphi exercises in related areas.^{20,21} In addition, sample sizes greater than 30 have seldom been found to improve results.^{22,23}

The goal of Round 1 is to initiate the consensus process regarding “what should be measured?” This includes both ICF components and outcomes (not specific outcome measures) within components. Specifically Round 1 questions will address: 1) the ICF components relevant to hip or knee arthroplasty with respect to surgical technique studies, non-surgical physical interventions and non-surgical medical interventions; and 2) the relevant outcomes within ICF components. The expert panel will be provided with an introduction to the task and a summary of the information obtained from the Phase 1 literature review. Specifically, the information will list outcomes (e.g., pain, range of motion, functional status) under the appropriate ICF components and not the measures (e.g., WOMAC pain subscale, goniometer, 6-minute walk test) used to assess these outcomes. Using a structured question format, experts will be asked to rate on a 7-point Likert scale the extent to which each ICF component is essential to RCTs targeting surgical techniques, non-surgical physical interventions and non surgical medical interventions of patients undergoing hip or knee arthroplasty (see Figure 3). Hip arthroplasty will be assessed separately followed by knee arthroplasty. After primary arthroplasty is completed, the same approach will be applied to revision surgeries for the hip and then for the knee.?? This could be a lot clearer. Is the plan here to develop a core set based on the ICF? If so, please state this. Also, will there be two separate core sets-one for clinical research/trials and a more limited one for clinical practice-an approach that has been advocated in other areas.

Round 2 has two goals: (1) to achieve consensus on “what outcomes should be measured?” ???isn’t this being addressed in round 1. Please see introductory sentence of preceding paragraph. and (2) to begin the consensus process on “how should the outcomes be assessed?” The latter question will be specific to those outcomes for which consensus was achieved following Round 1. The Round 2 information package will contain a summary of the Round 1 results and a review of the psychometric properties of measures that assess outcomes for which consensus was achieved from Round 1. Item summary information will contain a ranking of outcomes within each ICF component. In addition the median score, inter-quartile range, and a histogram of responses will be provided for each item. Each expert will be shown his/her Round 1 item responses relative to group summary data. With respect to the question “how should the outcomes be assessed?” the psychometric properties of interest will include reliability and validity (cross-sectional and longitudinal). Also, a summary of the results from head-to-head comparison studies of competing measures will be provided. Once again, expert panel members will be asked to respond on the 7-point Likert scale described previously. Also, following each question a space will be provided for experts to provide clarifying comments should they wish. In addition, expert pane members will also be offered the opportunity to add measures not identified in the Phase 2 literature review.

The goal of Round 3 is to continue building consensus for the question “how should the outcomes be assessed?” The Round 3 information package will contain a summary of the Round 2 results and a review of the psychometric properties of measures for which consensus was not achieved in Round 2. Item summary information will contain a ranking of measures within each ICF component. In addition the median score, inter-quartile range, and a histogram of responses will be provided for each item. Each expert will be shown his/her Round 2 item responses relative to the group summary data. The Round 3 administration

will replicate that described in Round 2 pertaining to question “how should the outcomes be assessed?”

Coming to consensus and dissemination of the findings—Although there is no uniformly agreed upon method or standard for defining consensus or convergence of opinion, often a percentage level is applied when considering an item for inclusion.^{13,14} Clearly, the choice of the percentage cut-off value for inclusion of an item is arbitrary and values have varied widely (e.g., 55% to 100%).^{14,24} For the proposed study we define consensus as having been met if 70% of the responding expert panel members endorse an item at the “agree” or “strongly agree” level (or in the negative “disagree” or “strongly disagree”). We chose the 70% criterion because this is the criterion generally supported by OMERACT.^{25,26}

Dissemination to clinical researchers will be accomplished via publication in the peer-reviewed literature. Results will be distributed to and discussed with the Centers for Medicare and Medicaid Services (CMS), the key US policy maker for patients with hip and knee arthroplasty, and to The American Academy of Hip and Knee Surgeons, The Hip Society, and The Knee Society. Availability of these standardized outcome measures will be critical for future NIH consensus and state-of-the-science conferences related to joint arthroplasty surgery and will enhance the application and generalizability of data collected through future federally and privately funded research.

I think that what is lacking in this group is individuals from WHO with expertise in the development of core sets based on the ICF. What these investigators aim to accomplish is the development of a core set of outcomes for arthroplasty based on the ICF. The manuscript should be revised to emphasize this and should also elaborate on how this has been accomplished for OA in some detail and for other rheumatic conditions more briefly. How does the methodological approach differ and what is the rationale for choosing a different approach by this OMERACT group. I am surprised that there is no mention in this manuscript of the exercise that was undertaken to develop a core set for OA based on the ICF.

Acknowledgments

Source of support: None

References

1. NIH Consensus Statement on total knee replacement December 8–10, 2003. *J Bone Joint Surg Am.* 2004; 86-A:1328–1335. [PubMed: 15173310]
2. Kurtz S, Mowat F, Ong K, Chan N, Lau E, Halpern M. Prevalence of primary and revision total hip and knee arthroplasty in the United States from 1990 through 2002. *J Bone Joint Surg Am.* 2005; 87:1487–1497. [PubMed: 15995115]
3. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am.* 2007; 89:780–785. [PubMed: 17403800]
4. Ong KL, Mowat FS, Chan N, Lau E, Halpern MT, Kurtz SM. Economic burden of revision hip and knee arthroplasty in Medicare enrollees. *Clin Orthop Relat Res.* 2006; 446:22–28. [PubMed: 16672867]
5. Riddle DL, Stratford PW, Bowman DH. Findings of extensive variation in the types of outcome measures used in hip and knee replacement clinical trials: A systematic review. *Arthritis Rheum.* 2008; 59(6):876–83. [PubMed: 18512725]

6. Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol.* 1997; 24:799–802. [PubMed: 9101522]
7. Pham T, van der HD, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarthritis Cartilage.* 2004; 12:389–399. [PubMed: 15094138]
8. Pham T, van der HD, Lassere M, et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. *J Rheumatol.* 2003; 30:1648–1654. [PubMed: 12858473]
9. World Health Organization. International classification of functioning, disability and health (ICF). Geneva, Switzerland: 2001.
10. Cieza A, Ewert T, Ustun TB, Chatterji S, Kostanjsek N, Stucki G. Development of ICF Core Sets for patients with chronic conditions. *J Rehabil Med.* 2004:9–11. [PubMed: 15370742]
11. Grill E, Huber EO, Stucki G, Herceg M, Fialka-Moser V, Quittan M. Identification of relevant ICF categories by patients in the acute hospital. *Disabil Rehabil.* 2005; 27:447–458. [PubMed: 16040548]
12. Grill E, Stucki G, Boldt C, Joisten S, Swoboda W. Identification of relevant ICF categories by geriatric patients in an early post-acute rehabilitation facility. *Disabil Rehabil.* 2005; 27:467–473. [PubMed: 16040550]
13. Linstone, HA.; Turoff, M. *The Delphi Method: Techniques and Applications.* Don Mills, ON: Addison-Wesley Publishing Company; 1975.
14. Powell C. The Delphi technique: myths and realities. *J Adv Nurs.* 2003; 41:376–382. [PubMed: 12581103]
15. Duffield C. The Delphi technique: a comparison of results obtained using two expert panels. *Int J Nurs Stud.* 1993; 30:227–237. [PubMed: 8335432]
16. Bond S, Bond J. A Delphi survey of clinical nursing research priorities. *J Adv Nurs.* 1982; 7:565–575. [PubMed: 6759554]
17. Jairath N, Weinstein J. The Delphi methodology (Part one): A useful administrative approach. *Can J Nurs Adm.* 1994; 7:29–42. [PubMed: 7880844]
18. Delbecq, AL.; Van de Ven, AH.; Gustafson, DH. *Group Techniques for Program Planning: A Guide to Nominal and Delphi Processes.* Glenview, IL: Scott, Foresman and Company; 1975.
19. Reid, N. The Delphi technique: its contributions to the evaluation of professional practice. In: Ellis, R., editor. *Professional Competence and Quality Assurance in the Caring Professions.* London: Chapman and Hall; 1988.
20. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol.* 2006; 6:2. [PubMed: 16433905]
21. Weigl M, Cieza A, Andersen C, Kollerits B, Amann E, Stucki G. Identification of relevant ICF categories in patients with chronic health conditions: a Delphi exercise. *J Rehabil Med.* 2004:12–21. [PubMed: 15370743]
22. de Villiers MR, de Villiers PJ, Kent AP. The Delphi technique in health sciences education research. *Med Teach.* 2005; 27:639–643. [PubMed: 16332558]
23. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health.* 1984; 74:979–983. [PubMed: 6380323]
24. Williams PL, Webb C. The Delphi technique: a methodological discussion. *J Adv Nurs.* 1994; 19:180–186. [PubMed: 8138622]
25. Gladman DD, Strand V, Mease PJ, Antoni C, Nash P, Kavanaugh A. OMERACT 7 psoriatic arthritis workshop: synopsis. *Ann Rheum Dis.* 2005; 64(Suppl 2):ii115–ii116. [PubMed: 15708925]
26. Kirwan J, Heiberg T, Hewlett S, et al. Outcomes from the Patient Perspective Workshop at OMERACT 6. *J Rheumatol.* 2003; 30:868–872. [PubMed: 12672218]

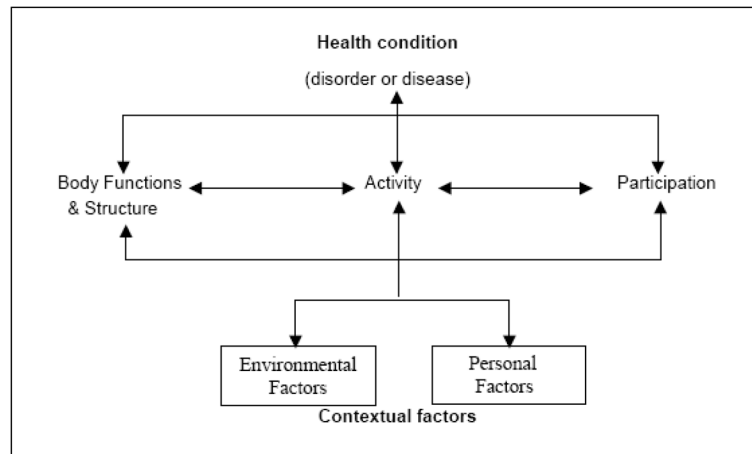


Figure 1.
The World Health Organization's ICF model of health and health conditions

Flow of the proposed study

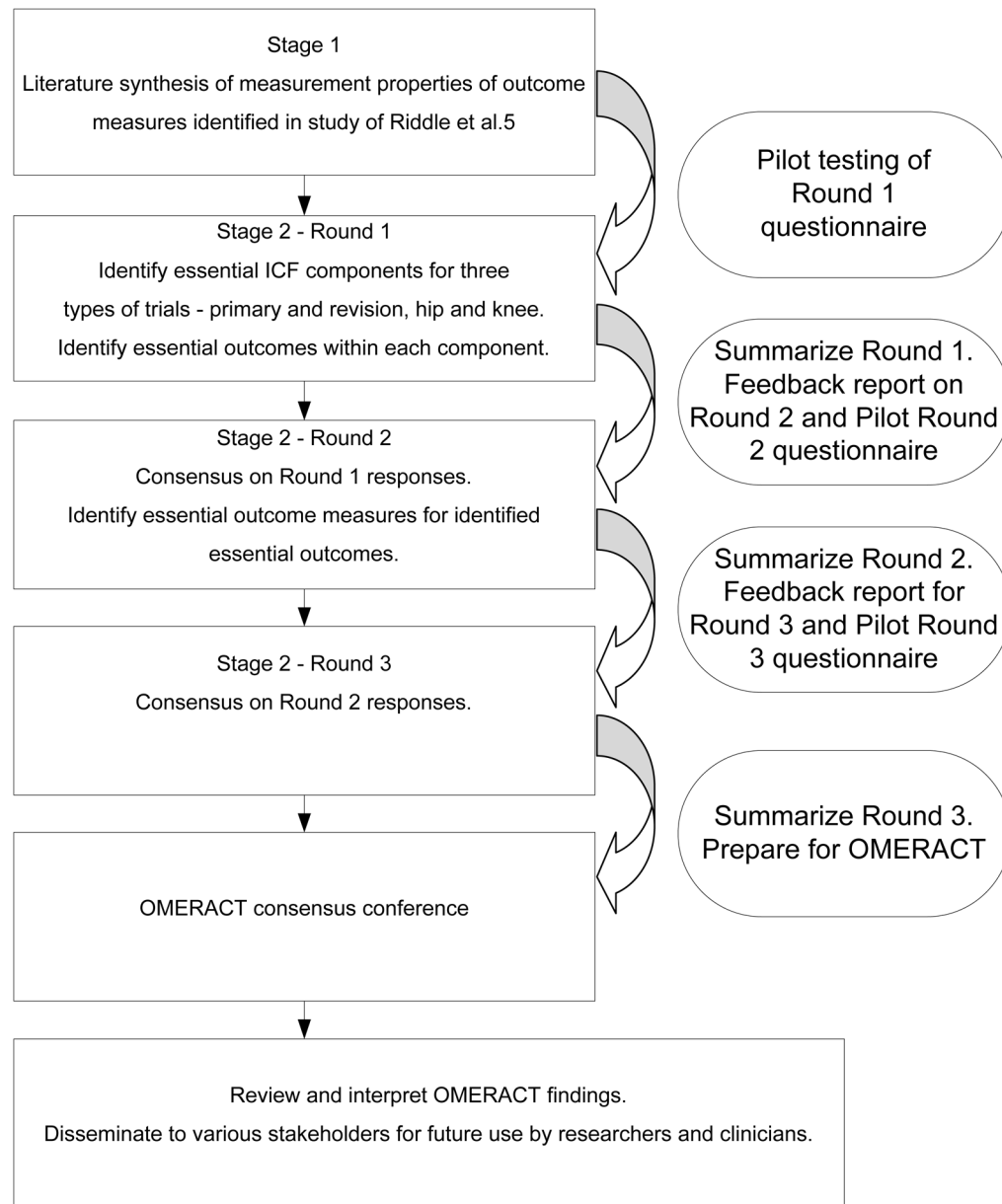


Figure 2.
Flow of the proposed study

(Circle the number in the box which best expresses your opinion of the following statement.)

1. In a study to compare 2 medications to reduce pain following knee arthroplasty, the assessment of outcomes from the ICF Activity/Participation Component is essential.

1 strongly disagree	2 disagree	3 mildly disagree	4 no opinion	5 mildly agree	6 agree	7 strongly agree
---------------------------	---------------	-------------------------	--------------------	----------------------	------------	------------------------

2. In a study to compare uncemented stems to cemented stems in patients with osteoarthritis of the hip, the assessment of outcomes from the ICF Structure and Function Component is essential.

1 strongly disagree	2 disagree	3 mildly disagree	4 no opinion	5 mildly agree	6 agree	7 strongly agree
---------------------------	---------------	-------------------------	--------------------	----------------------	------------	------------------------

3. In a study to compare different approaches to outpatient rehabilitation for patients with osteoarthritis of the hip, the assessment of outcomes from the ICF Personal Factor (patient satisfaction) Component is essential.

1 strongly disagree	2 disagree	3 mildly disagree	4 no opinion	5 mildly agree	6 agree	7 strongly agree
---------------------------	---------------	-------------------------	--------------------	----------------------	------------	------------------------

Figure 3.
Examples of Potential Round 1 Delphi Questions