Measuring Decision Quality:
Where We Stand Today

Karen R. Sepucha, PhD
Floyd J. Fowler, Jr., PhD

March 2013

Informed Medical Decisions Foundation
Partnerships for Quality Care

Introduction

Over the past few years, the Informed Medical Decisions Foundation and its associates have devoted significant resources and effort to developing and evaluating measures related to the quality of decision making. We think these can be critical patient-reported measures of how well providers are providing decision support to their patients, as well as aids to any group interested in monitoring and improving decision quality. This memorandum summarizes where we are with respect to the development of these measures and the kinds of evidence we have for their reliability and validity at the individual and practice levels.

Framing and Problem Definition

First, it is important to clarify the type of situation that is appropriate for this approach to measurement. There are certain diseases where a treatment or approach has considerable evidence of a significant benefit with minimal harm. As a result, high quality decisions in these situations are about efficiently delivering proven, effective care to all those who may benefit. A performance measure might focus on examining the percentage of eligible patients who receive the effective treatment, or the percentage of care that is "consistent" with the guidelines.

A surprising number of medical decisions do not have sufficient evidence of benefit for one option over another, or have evidence of equivalence of two or more options, or have evidence of substantial harm that accompanies the benefit. In these situations, patients and providers must make tradeoffs between competing criteria in order to arrive at a decision. Situations such as whether or not to have surgery for hip or knee osteoarthritis, herniated disc or spinal stenosis are clear examples that require tradeoffs. Examining treatment rates will not provide enough information to determine the quality of decisions. Rather, a different approach is needed—one that examines whether the right patient is being matched with the right treatment.

There is widespread support from many stakeholders—clinicians, patients, researchers and policymakers—for informing patients, engaging them in decisions about their care and ensuring that treatments reflect what's most important to patients. (National Research Council, 2008; National Breast Cancer Coalition Fund, 2002; National Priorities Partnership, 2009; Elwyn et al., 2006) Despite the remarkable consistency in the definitions of patient-centered care, quality of care, and decision quality proposed by providers, patients and researchers, there are currently no widely accepted measures in use to assess whether or not this is happening.

Decision quality is defined as the extent to which treatments reflect the considered preferences of well-informed patients and are implemented (Sepucha et al., 2004). We have developed a series of patient reported surveys, called decision quality instruments (DQIs),that include decision-specific items to assess:

1. knowledge, or the extent to which patients are informed
2. patients' goals, concerns and preferred treatment. These items can then be used to calculate concordance, or the extent to which patients' receive treatments that match their goals.

We also have developed and tested more generic items to measure the interactions between the patients and their health care providers to assess

3. decision making process, or the extent to which providers engage patients in decisions about their care.

Currently, these three aspects are reported on separately, as a total knowledge score, a concordance (or in some cases dissonance) score, and a decisions process score.

We have completed field tests for several conditions (indicated in Table 1) and have surveyed more than 1500 patients facing surgical decisions, more than 1200 who have faced decisions about medications (menopause, depression), and more than 500 facing colon cancer screening decisions. As part of these field tests we also examine the mode of administration and randomized participants to complete surveys online or by mail; over the phone with an interviewer versus automated voice recognition system. We have also surveyed underserved populations, Latina breast cancer patients (n=98) and a largely African American, low literacy population of adults for colon cancer screening (n=191).

The following sections describe the work that we have done developing and testing measures of knowledge, concordance and involvement across common preference-sensitive decisions. The instruments are available for download at http://www.massgeneral.org/decisionsciences/ and the topics are listed in Table 1. We have short versions for each that include only 5 knowledge items, 5 goals and 4 decision process items. The full version of the Decision Quality Instrument- Knee Osteoarthritis is included as Appendix A with a user guide in Appendix B.

Table 1: List of Decision Quality Instruments

	Field Tested	Instrument available for download
Breast Cancer Surgery	X	X
Breast Cancer Systematic Therapy	X	X
Breast Reconstruction	X	X
Osteoarthritis: Hip Replacement	X	X
Osteoarthritis: Knee Replacement	X	X
Colon Cancer Screening	X	X
Prostate Cancer Screening		X
Early Prostate Cancer Treatment		X
Herniated Disc	X	X
Spinal Stenosis		X
Stable Heart Disease: Revascularization		X
Benign Prostate Disease		X
Menopause	X	X
Depression	X	X

Versions of these surveys have also been used in several large retrospective studies.

1. DECISIONS study (2007): a national sample of 3,000+ adults age 40 and older, telephone survey. Early versions of the decision quality instruments for OA knee and hip, herniated disc and spinal stenosis, prostate and colon cancer screening, and depression were used in the study.
2. Medicare Surgery Study (2009): a national sample of over 2000 Medicare beneficiaries, mailed survey. Participants who had recently had prostate cancer surgery, breast cancer surgery, coronary bypass surgery or stents completed the DQIs.
3. TRENDS study (2011): a national sample of adults age 40 and older, using an online panel. Used recent versions of the decision quality instruments for OA knee and hip, herniated disc and spinal stenosis, prostate and colon cancer screening, and depression.

Framing and Problem Definition

I. Approach

Reviewing the literature on medical decision making reveals there are four main elements of a good process:

a. the patient is told about the reasonable options,
b. the patient and provider discuss the pros of those options,
c. the patient and provider discuss the cons of those options,
d. the provider and patient discuss patients' goals and preferences related to the options and outcomes

Our goal was to find the shortest list of questions possible that patients could answer that would enable us to characterize each of those aspects of the decision making process. Moreover, while measures of knowledge and patient goals necessarily have to be decision specific, we wanted to develop decision process questions that could be reasonably used across all types of medical decisions.

Our first step was to identify questions that covered each of those four constructs. Candidate questions have been cognitively tested. We have used variations on core questions in a variety of kinds of surveys, both cross-section national surveys and surveys of patients sampled in practice settings. A version of these questions was also tested to be part of the ambulatory CAHPS supplement for Patient-Centered Medical Homes.

II. Status of Item Development: What We Have Learned

Targeting the items to a specific decision turns out to be an important issue. Most of our studies have been of patients who have been identified as having made a specific medical decision about taking a particular medication, having a particular screening test for cancer, or having a specific surgical intervention. Those are the contexts in which we know these questions produce valuable information about decision making. However, when they have been tried using a general referent (about "decisions you have made in the last year" or even "decisions you have made about taking medications in the last year") the variance goes down as the answers move toward the top. Thus, we think we know that the questions need to be targeted toward a specific decision.

III. Content and Wording of Items

Three questions have been used now in numerous retrospective studies where the treatment or intervention was already done and have consistently shown to provide useful information:

1. How much did a doctor (or health care provider) talk with you about the reasons you might want to (HAVE INTERVENTION)—a lot, some, a little, or not at all?
2. How much did a doctor talk with you about reasons you might not want to (HAVE INTERVENTION)—a lot, some, a little or not at all?
3. Did any of your doctors ask you if you wanted to (HAVE INTERVENTION)?

The wording can be adapted to accommodate the situation where data are being collected before the intervention has been chosen or carried out. These three questions cover three key concepts, they have proven to be readily understood and answered, and they can be used generically, without much, if any amendment, after patients have made almost any decision about having a test, taking a medication, or having a surgical intervention.

There are some variations that could be added, most of which we have tried. For example, the first two questions could be asked about each alternative, not just the chosen alternative. However, that lengthens the series and requires tailoring to a different set of alternatives for each decision. Alternatives or possible additions to question 3 include asking how much what the patient wanted was discussed and asking the patient to summarize who made the decision, from totally the doctor to totally the patient. The biggest challenge has proven to be to measure whether the reasonable alternatives are presented as options. The problem is that the appropriate options need to be tailored to the particular intervention and, in some cases, require clinical information about the patient that may be hard to get or integrate into a survey. Our preferred series asks if a doctor talked with the patient about (AN ALTERNATIVE) and if it was presented as an option to "seriously consider". The second question was found to be important because we found that options are sometimes mentioned in a way that implies they are not really options. However, those questions obviously are decision specific and require confidence that the reasonable options are known. They also will be different for every decision, making it harder to use a consistent set of questions across a variety of different decisions.

The best generic question we have found to date is:

4. Did any of your doctors explain that you could choose whether or not to (HAVE INTERVENTION)? An alternative we often have used: "Did any of your health care providers explain that there were choices in what you could do to treat your [condition]?" As worded, these only work after an intervention has been done.

This clearly is a much weaker operationalization of whether patients were presented with all the reasonable options to consider. There is a tradeoff to be made between the stronger series for decisions about a specific intervention and this weaker question that can be used more widely across a range of decisions.

IV. Scoring

We have been reporting a Decision Process Score that awards one point for each the four questions ("a lot" or "some" for the first two questions and "yes" for questions 3 and 4). (Fowler et al., 2012; Fowler et al., in press) Variations on this scoring scheme will yield highly correlated results.

IV. Evidence for Realiability and Validity

The Decision Process Score is technically a composite, with conceptual roots in what a good decision process should look like, so a calculation of Cronbach's alpha may not be an appropriate measure of reliability (Bollen and Lennox, 1991), but we have calculated them for some decisions, and they are reasonably high (often in the .6 to .8 range).

We have short term (~4 weeks) test-retest data on some variations of this measure and obtained ICC values in the .7 to .8 range.

We also have three kinds of evidence about its validity.

First, its construct validity has been examined by looking at how it relates to how likely respondents say they would be to make the same decision again, the hypothesis being that those who were more involved would be more likely to make the same decision again. For 15 different decisions, there was a statistically significant relationship for 12 of them.

Second, we compared breast cancer patient reports of their interactions with providers with coding of tape recordings of those interactions and found generally consistent significant correlations, indicating a good degree of validity of patient reports.

Third, we have used these questions in a set of clinical sites that are making a special effort to use patient decision aids in their routine care (demonstration sites). We can compare the answers of patients in those sites with the answers from patients who faced the same decisions who were part of a national cross-section sample or part of web-based samples that did not experience shared decision making as part of their routine care. While there are some issues regarding the comparability of the data (for example, the times from decision to measurement were quite different) the patients in the demonstration sites consistently reported significantly higher Decision Process Scores than cross-section samples of patients.

Thus we have quite a lot of evidence at the patient level and at the practice level that these scores are valid measures of the quality of the decision making process and that they can be used to assess the quality of decision making at a clinical site.

Measuring Patients' Knowledge

The goal of having informed patients is well supported by consumer, provider and policy groups. It is also an ethical imperative, particularly in the case of elective surgical procedures that require patients to provide informed consent before undergoing the operation. In order to be informed, patients need to understand the situation and what would happen if they do nothing, as well as what their options are and the consequences of those options. Here we outline our approach to measuring patients' understanding of the key information needed to make informed medical decisions.

I. Approach

Different approaches for measuring knowledge exist. Many investigators ask patients to self-report their level of knowledge, for example, asking patients to rate how informed they feel or to agree/disagree that they understand their options. In the DECISIONS study, a large nationally representative study of patients who had made nine common medical decisions about testing, medication and surgery, we found no correlation between patients' reports about how informed they feel and their ability to answer specific knowledge questions correctly. (Sepucha et al., 2010) Basically, this finding reflects a pervasive problem with self assessments of knowledge, as people cannot reliably report on what they do not know. If the goal is to determine whether someone is informed, it is necessary to ask patients decision-specific knowledge questions.

II. Identifying the Content

The information that is critical for informed decisions can be grouped into the following five domains:

1. Disease (such as the prevalence, lethality, and natural history),
2. Choices (such as a description of the treatment options and what's involved with each),
3. Benefits of the choices (such as survival, symptom relief, likelihood of outcomes),
4. Harms of the choices (such as serious and permanent problems, temporary and common problems, the likelihood of problems), and
5. Decision situation (such as the urgency to treat, recognition of decision, level or strength of evidence for the options).

The process we used to develop the content begins with a detailed review of the clinical evidence (derived both from former patients and the literature) for each decision, through which a candidate set of information is distilled into a set of facts. These facts are then reviewed and rated by a convenience sample of patients who had recently made decisions about treatment and a multidisciplinary group of clinical experts. The respondents are able to comment on the accuracy and importance of items and are able to suggest additional items that they feel are missing. The resulting items are revised and expanded, as needed, until they are considered accurate, important and complete. (Sepucha et al., 2008 and Lee et al., 2010). We do not have any preset quotas for items within each domain, and the number of items varies depending on the situation and the number of options available.

III. Creating the Items

Experts in survey research methods then took the content and developed questions to cover the key facts. Then, we conducted cognitive interviews with a small sample of patients (~5) who had faced the decision. The interviews led to some revisions to improve acceptability and comprehension. The resulting set of questions was then reviewed by a physician with expertise in the particular clinical area for accuracy, importance and completeness. Once these steps were completed, we field tested the items.

IV. Scoring

A total knowledge score for the items is calculated by dividing the number of correct responses by the total number of items and multiplying by 100, resulting in scores from 0 to 100%. Open-ended items are given one point if they fall within a range that is pre-determined to be correct by medical experts based on clinical evidence and zero points otherwise. When included as an option, the response, "I am not sure," is considered incorrect. Missing responses are also considered incorrect. A knowledge score is calculated for every respondent who completes at least 50% of the items. The reason for this is that we felt that patients who refuse to complete the majority of items are more likely indicating that the survey is unacceptable as opposed to lack of knowledge.

V. Field Testing

The field tests were designed to provide data on how well the individual items work, the clinical sensibility and the psychometric properties of the scores. The psychometric properties have been published in three manuscripts (Sepucha et al., 2011; Sepucha K, Feibelmann S et al., 2012; Sepucha, Belkora et al., 2012). The key properties for the knowledge scores include:

- **Acceptability:** We examined this using response rates and length of time to complete the instrument.
- **Feasibility:** of the survey and mode of administration were examined using rates of missing data.
- **Reliability:** We examined short term (usually 4-6 weeks) retest reliability. Cronbach's alpha was not used as a measure of internal consistency for the knowledge score, as the set of knowledge items is not a measure of one underlying construct.
- **Content validity:** This was generated through the development process and the providers and patients rating of content, it was also examined explicitly in field test by asking providers how well they felt the items covered key information.
- **Discriminant validity:** This is key feature of a knowledge test and was examined using known groups comparisons (comparing scores of health care providers, patients and healthy volunteers, or comparing scores of patients who were exposed to a decision aid to others who were not exposed).

In general, the knowledge portion of the DQIs is acceptable to patients, feasible to administer, has good retest reliability and is able discriminate between groups with different knowledge. For most of the topics, there were considerable knowledge gaps identified and significant room for improvement on scores. The underserved samples tended to have lower scores, lower response rates and higher rates of missing data. The User Guide in appendix B summarizes the key results for the osteoarthritis DQI.

Measuring Value Concordance

I. Approach

A critical piece of shared decision making and patient-centered care is ensuring that treatments match patients' goals. We conducted a systematic review of methods to assess this aspect of care and did not find any consensus on how to define or measure this construct, and the approaches used varied from fairly straightforward to very complicated. (Sepucha and Ozanne 2010).

We have examined three different approaches as part of the decision quality instruments:

1. Simple match: in this direct approach we assess patients' preferred treatment with a single item and then compare with treatment received to determine whether they match.
2. Dissonance score: we ask patients to rate importance of salient goals and concerns on a 0 to 10 scale. We look at those issues that are "cons" or against the treatment that the patient actually received and take the average score of the patients' ratings for those items. This provides a measure of dissonance, the higher the dissonance the less likely the patient received treatments that matched his or her goals. (Fowler et al., in press)
3. Concordance score: we ask patients to rate the importance of salient goals and concerns on a 0 to 10 scale and then use those in a multiple logistic regression model to predict treatment. The predicted probability of treatment is then used to determine whether patients received treatments that "matched" their goals. If the model predicted probability of surgery is >0.5 and the patient had surgery then we would say their treatment matched their goals (and if it was less than 0.5 and they did not have surgery we would also say it matched their goals). The summary score is the percentage of the population that received treatments that matched their goals. (Sepucha, Stacey et al., 2011, Sepucha, Feibelmann et al., 2012, Sepucha, Belkora et al., 2012)

The first approach is simple and feasible, but a criticism that has been raised is it is a fairly low bar to document that people who want to have a lumpectomy actually received it, or that people who preferred non surgical options for treating their osteoarthritis did not have a joint replacement. The second and third approaches delve deeper into the different issues that may lead a patient to choose one option over another, yet criticisms of those approaches include that the short list of goals and concerns may not include those issues that were most relevant to patients, and these methods do not allow patients to weight the relative importance of the different issues (the dissonance score assumes equal weighting and the regression model creates population weights). Further the third approach requires the decision be broken down into two options (e.g. surgery versus non surgical options or medication versus nothing) to create the logistic regression model, as well as more sophisticated computational skills and is less understandable to providers or patients.

Until the research field makes progress in the standardization of more sophisticated approaches to measure this, we believe that the first approach is best. For example, we are piloting this approach as part of assessment of surgical appropriateness, with the requirement that patients will not get through preadmission testing for a procedure if they do not indicate a clear preference for that procedure. More detailed assessments may support clinical care and quality improvement, but for performance measurement, having confidence that patients' treatment preferences are being elicited and met is an important step forward.

II. Identifying the Content

Two things are required for this measure, the patients' preferred treatment and the treatment the patient actually received. The content for the items are determined by the clinically relevant options available to the patient for the condition. In some cases, this may be fairly generic, for example, non surgical or surgical options and in others it may be more specific, lumpectomy or mastectomy. This does require having some confidence that the patients being sampled are clinically eligible for the options. The treatment received can either be self reported by the patient or taken from medical record in cases where that will be clearly documented.

III. Creating the Items

For the patients' preferred treatment, we adapted an item that has been used by O'Connor et al for choice predisposition and the items take the form of, "What did you want to do to treat your [DISEASE]? " with responses covering the available options and "not sure."

If reported by patients, the treatment received can be assessed by the following, "what did you do to treat your [DISEASE]?" In the case of chronic diseases, often there is a time frame associated, e.g. In the last three months, what have you done to treat your [DISEASE]?

IV. Scoring

A binary match variable is created that is true for patients who receive treatments that match their goals and false for all others. Patients who are unable to state a treatment preference and instead choose I am not sure are not considered to have been matched. An important part of shared decision making is helping patients come to a treatment preference and if the clinical team has failed to do this then that should be marked as a gap. This can be combined across patients in a sample to determine the percentage that receive treatments that match their goals.

V. Field Testing

We have completed field tests for several conditions and have surveyed more than 1500 patients facing surgical decisions for breast cancer, hip and knee arthritis, and herniated disc. These field tests are designed

to provide data on how well the individual items work, the clinical sensibility and the psychometric properties of the scores. The psychometric properties have been published in three manuscripts (Sepucha et al., 2011; Sepucha , Feibelmann et al., 2012; Sepucha, Belkora et al., 2012). The key properties include:

- **Acceptability:** We examined this using overall response rates and time to complete the survey
- **Feasibility:** of the goals and preferences and mode of administration was examined using rates of missing data
- **Reliability:** We examined short term (usually 4-6 weeks) retest reliability of the goals and preferred treatment
- **Construct validity:** We compared how well the model generated probability of treatment (based on the concordance score approach described above) aligned with the preferred treatment assessed directly. For example, we tested whether patients who stated that they preferred surgery had a higher model predicted probability of surgery compared to those who were unsure, and those who preferred non surgical options.
- **Predictive validity:** We tested whether patients who received treatments that matched their goals (or who had less dissonance) had less regret and more confidence in their decisions.

In general, the goals and concerns items were acceptable and feasible to administer. Some items did tend to have ceiling effects (particularly ones assessing the desire for symptom relief or reducing cancer recurrence), but were still able to discriminate among different treatment options. Even when administered retrospectively, there was evidence of mismatch between what patients said they wanted and the treatments they received. It was higher for conditions like herniated disc and osteoarthritis (where 25% of respondents did not match) than for breast cancer (where about 10% did not match). The User Guide in Appendix B summarizes the key results for the osteoarthritis DQI.

Protocols and Timing of Data Collection

In using these measures to assess the quality of decision making, the ideal would be to collect data about knowledge and preferences around the time the decision is being made, while assessing the decision making process should be done after the interactions are complete. There are major challenges to identifying when decisions are being made and collecting data then. When practices routinely use decision aids, knowledge and preferences can be collected after the decision aid is watched, but the decision making process is still in process, and possibly knowledge and preferences will change after further interactions with providers. Moreover, distribution of decision aids is not very widespread and, hence, that protocol may not be a feasible or reliable one to be used across practices.

The protocol that is most feasible is to survey patients after they have an intervention. Such events are usually easy to identify in records. There are two drawbacks to that approach. First, it would better to measure knowledge and preferences before the intervention. Questions can be raised about how they might change, or be hard to recall, after experiencing the intervention. Second, that approach leaves out the people who decided not to have the intervention, the surgery or the test, or not to take the medication. The counterargument is that while this approach is imperfect, it is the people who actually get the interventions whose decision processes we are most concerned about. Most important, this protocol is easy to implement in a consistent way across practices, while we cannot say that about any of the other protocols we have tried.

Conclusion

At this point we have a lot of experience with these measures, and we have considerable evidence related to reliability and validity for the scores generated from these measures. For those parts that have to be condition specific (knowledge and goals and concerns), we have varying amounts of testing for different decisions, so some may continue to evolve. Furthermore, measuring knowledge requires continuous updating as new studies are done and new treatment options emerge. Nonetheless, we think that many of these measures are ready for use to evaluate the way decisions are being made.

References

Bollen K, Lennox R. Conventional wisdom on measurement: A structural equation perspective. *Psych Bull.* 1991;110(2):305-14.

Elwyn G, O'Connor A, Stacey D, et al. (2006). Developing a quality criteria framework for patient decision aids: Online international Delphi consensus process. *BMJ.* 2006;333-417.

Fowler FJ, Gallagher PM, Bynum JPW, et al. Decision-making process reported by Medicare patients who had coronary artery stenting or surgery for prostate cancer. *J Gen Int Med.* 2012 Aug;27(8):911-6.

Fowler FJ, Gallagher PM, Drake KM, et al. Decision dissonance: Evaluation of an approach to measuring the quality of surgical decision making. *Joint Commission Journal on Quality and Patient Safety.* (In press).

Lee C, Barry M, Cosenza C, et al. Development of instruments to measure the quality of breast cancer treatment decisions. *Health Expect.* (in press).

National Research Council. *Envisioning the National Health Care Quality Report.* Washington D.C.: The National Academy Press, 2001.

National Breast Cancer Coalition Fund. *A Guide to Quality Breast Cancer Care.* 2002. (Available online at http://www.stopbreastcancer.org/nbccf/index.html)

National Priorities Partnership (accessed on June 14, 2010 http://www.nationalprioritiespartnership.org/PriorityDetails.aspx?id=596)

Sepucha KR, Belkora J, Chang Y, et al. Measuring decision quality: Psychometric evaluation of a new instrument for breast cancer surgery. *BMC Med Inform Decis.* 2012 Jun 8;12(1):51 [Epub ahead of print].

Sepucha K, Fagerlin A, Couper M, Levin C, Singer E, Zikmund-Fisher B. How does feeling informed relate to being informed? Results from the DECISIONS study. *Med Decis Making.* 2010 Sept/Oct;30:77S-84S.

Sepucha KR, Feibelmann S, Abdu WA, et al. Psychometric evaluation of a decision quality instrument for treatment of lumbar herniated disc. *Spine.* 2012 Mar 15. [Epub ahead of print].

Sepucha KR, Fowler FJ, Mulley AG (2004). Policy support for patient-centered care: The need for measurable improvements in decision quality. *Health Aff (Millwood). 2004;*Suppl Variation: VAR54-62.

Sepucha KR, Levin C, Uzogara E, Barry M, O'Connor A, Mulley A. Developing instruments to measure the quality of decisions: Early results for a set of symptom-driven decisions. *Patient Educ Couns.* 2008;73:504-10.

Sepucha KR, Stacey D, Clay C, et al. Decision quality instrument for treatment of hip and knee osteoarthritis: A psychometric evaluation. *BMC Musculoskelet Disord* 2011 Jul 5;12(1):149.

Sepucha KR, Ozanne E.How to define and measure concordance between patients' preferences and medical treatments: A systematic review of approaches and recommendations for standardization. *Patient Educ Couns.* 2010 Jan;78(1):12-23. [Epub 2009 Jun 30].

Code:

Appendix A:

Knee Osteoarthritis
Decision Quality Instrument

INSTRUCTIONS

- The survey has questions about what it was like for you to make decisions about treating your knee osteoarthritis. It will ask about what you and your health care providers talked about, what's most important to you, and what know about knee replacement surgery and other options, such as medicines and exercise.

- Please check the box ☑ to answer each item.

- Thank you, we really appreciate your help!

Code:

SECTION 1: WHAT MATTERS MOST TO YOU

The next set of questions includes some reasons other people have given for choosing whether or not to have knee replacement surgery. We are interested in what was important to you.

Please mark on a scale from 0 to 10, where 10 is extremely important and 0 is not at all important, how important each of the following were to you for your decision about whether or not to have surgery.

1.1. How important was it to you to relieve your knee pain?

- ☐ 10 Extremely Important to me
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5 Somewhat important to me
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 Not at all important to me

1.2. How important was it to you to not be limited in what you can do because of your knee pain?

- ☐ 10 Extremely Important to me
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5 Somewhat important to me
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 Not at all important to me

1.3. How important was it to you to avoid having knee surgery?

- ☐ 10 Extremely Important to me
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5 Somewhat important to me
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 Not at all important to me

1.4. How important was it to you to avoid taking pain medicine for a long time?

- ☐ 10 Extremely Important to me
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5 Somewhat important to me
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 Not at all important to me

1.5. How important was it to you to avoid a treatment with a long recovery time?

- ☐ 10 Extremely Important to me
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5 Somewhat important to me
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1
- ☐ 0 Not at all important to me

1.6. Which treatment did you want to do to treat your knee osteoarthritis?

- Knee replacement surgery
- Non-surgical treatment options
- I am not sure

SECTION 2: FACTS ABOUT KNEE OSTEOARTHRITIS

The next set of questions asks about your some facts doctors think are important for patients to know about knee osteoarthritis. The correct answer to each question is based on medical research. Please do your best to answer each question.

2.1. Over time, without knee replacement surgery, what usually happens to knee pain?

- Gets better
- Stays the same
- Gets worse
- I am not sure

For 2.2a to 2.2d, please mark whether or not it can help some people relieve knee pain.

2.2a. Can exercise help some people relieve knee pain?

- Yes
- No
- Not sure

2.2b. Can physical therapy help some people relieve knee pain?

- Yes
- No
- Not sure

2.2c. Can calcium pills help some people relieve knee pain?

- Yes
- No
- Not sure

2.2d. Can over-the-counter pain medicine help some people relieve knee pain?

- Yes
- No
- Not sure

2.3. Which treatment is most likely to provide relief from knee pain caused by osteoarthritis?

- Surgery
- Non-surgical treatments
- Both are about the same
- I am not sure

2.4. After knee replacement surgery, about how many months does it take most people to get back to doing their usual activities?

- Less than 2 months
- 2 to 6 months
- 7 to 12 months
- More than 12 months
- I am not sure

2.5. If 100 people have knee replacement surgery, about how many will need to have the same knee replaced again in less than 20 years?

- More than half
- About half
- Less than half
- I am not sure

2.6. If 100 people have knee replacement surgery, about how many will have less knee pain after the surgery?

- 30
- 50
- 70
- 90
- I am not sure

For 2.7a to 2.7d, mark whether or not it is a possible complication of knee replacement surgery.

2.7a. Is high blood pressure a possible complication of knee replacement surgery?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.7b. Is a blood clot in the leg a possible complication of knee replacement surgery?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.7c. Are migraine headaches a possible complication of knee replacement surgery?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.7d. Is an infection of the artificial knee a possible complication of knee replacement surgery?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.8. Serious complications can happen after knee replacement surgery including life-threatening blood clots, infections, heart attacks, and even death.

If 100 people have knee replacement surgery, about how many will have a serious complication within 3 months after surgery?

- ☐ 1
- ☐ 5
- ☐ 15
- ☐ 25
- ☐ I am not sure

For 2.9a to 2.9d, mark whether or not it is a possible side effect of using over-the-counter pain medicine for a long time. These can include medicines you can buy without a prescription like Advil, Aleve or aspirin.

2.9a. Is a stomach ulcer a possible side effect of using over-the-counter pain medicine for a long time?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.9b. Are migraine headaches a possible side effect of using over-the-counter pain medicine for a long time?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.9c. Are kidney problems a possible side effect of using over-the-counter pain medicine for a long time?

- ☐ Yes
- ☐ No
- ☐ Not sure

2.9d. Is excessive bleeding a possible side effect of using over-the-counter pain medicine for a long time?

- ☐ Yes
- ☐ No
- ☐ Not sure

SECTION 3: TALKING WITH HEALTH CARE PROVIDERS

Please answer these questions about what happened when you talked with health care providers including doctors, nurses and other health care professionals about knee replacement surgery and other non-surgical treatments, such as exercise or medicine, for knee osteoarthritis.

3.1. Did your health care providers explain that there were choices in what you could do to treat your knee osteoarthritis?

- ☐ Yes
- ☐ No

3.2. Did any of your health care providers talk about knee replacement surgery as an option for you?

- ☐ Yes
- ☐ No

3.3. How much did you and your health care providers talk about the reasons to have knee replacement surgery?

- ☐ A lot
- ☐ Some
- ☐ A little
- ☐ Not at all

3.4. How much did you and your health care providers talk about the reasons **not** to have knee replacement surgery?

- ☐ A lot
- ☐ Some
- ☐ A little
- ☐ Not at all

3.5. Did any of your health care providers talk about non-surgical treatments as something that you should seriously consider?

- ☐ Yes
- ☐ No

3.6. How much did you and your health care providers talk about how bothered you were about your knee pain?

- ☐ A lot
- ☐ Some
- ☐ A little
- ☐ Not at all

3.7. How much did you and your health care providers talk about how you felt about having surgery?

- ☐ A lot
- ☐ Some
- ☐ A little
- ☐ Not at all

3.8. Did any of your health care providers ask you whether you wanted to have knee replacement surgery or not?

- ☐ Yes
- ☐ No

SECTION 4: TREATMENT CHOICE

4.1. For each of the following, please mark whether or not **you have done** this to treat your knee osteoarthritis.

a.	Exercise and staying active……..	☐ Yes	☐ No
b.	Knee replacement surgery (**RIGHT** knee)…..	☐ Yes	☐ No
c.	Knee replacement surgery(**LEFT** knee)……..	☐ Yes	☐ No
d.	Pain medicine……..	☐ Yes	☐ No
e.	Cortisone shots into the knee…….....	☐ Yes	☐ No
f.	Weight loss……..	☐ Yes	☐ No
g.	Acupuncture…….	☐ Yes	☐ No
h.	Glucosamine and/or chondroitin……..	☐ Yes	☐ No
i.	Other treatment (please write in): _____	☐ Yes	☐ No

Appendix B: Hip and Knee Osteoarthritis Decision Quality Instrument User Guide

I. Purpose of Decision Quality Instruments:

To measure the extent to which patients are informed, involved in the decision making process and receive treatments that match their goals and preferences.

II. Versions:

- Hip Osteoarthritis Decision Quality Instrument v2.0, ©2010 [updated 2012].
- Knee Osteoarthritis Decision Quality Instrument v2.0, ©2010 [updated 2012].
- Decision Quality Worksheet: Treatments for Hip Osteoarthritis v2.0, ©2010 [updated 2012].
- Decision Quality Worksheet: Treatments for Knee Osteoarthritis v2.0, ©2010 [updated 2012].
- Hoja de Trabajo Sobre La Calidad de Decision en Tratamientos de Osteoartritis de Cadera v.2.0 ©2012 [Spanish version of Hip worksheet].
- Hoja de Trabajo Sobre La Calidad de Decision en Tratamientos de Osteoartritis de Rodilla v.2.0 ©2012 [Spanish version of Knee worksheet].

III. Timing

The decision quality instrument version is designed to be administered after a decision has been made. Modifications are required (e.g. to instructions and tenses of items) if it is to be used before a decision has been made.

The shorter worksheet version is worded to be used during the decision making process. The knowledge items and goals can be administered at any time, e.g. before or after a visit, before or after a decision aid. The decision process items need to be administered after a provider consult.

IV. Scoring:

The Hip and Knee Osteoarthritis Decision Quality Instruments (DQI) are almost identical, with "hip" being replaced with "knee," and they are scored identically. The survey contains three sets of items and results in three scores, a total knowledge score, a concordance score and a decision process score.

1. Knowledge Score: For each knowledge item, a correct response receives one point (see Table 1). Questions with multiple parts (e.g. items 2, 7 and 9 in Table 1) are scaled to total 1 point per item. Missing responses and responses of "I am not sure" receive 0 points. A total score is calculated for all patients who complete at least half of the items. Total scores are scaled from 0-100%.

Table A1: Knowledge Items (# indicates items in the worksheet version)

Question	Correct response
1. Over time, without hip/knee replacement surgery, what usually happens to hip/knee pain?	Gets worse
2. For each of the following, please mark whether or not it can help some people relieve hip/knee pain.	
a. Exercise	Yes
b. Physical therapy	Yes
c. Calcium pills	No
d. Over-the-counter pain medicine	Yes
#3. Which treatment is most likely to provide relief from hip/knee pain cause by osteoarthritis?	Surgery
# 4. After hip/knee replacement surgery, about how many months does it take most people to get back to doing their usual activities?	2 to 6 months
# 5. If 100 people have hip/knee replacement surgery, about how many will need to have the same hip/knee replaced again in less than 20 years?	Less than half
# 6. If 100 people have hip/knee replacement surgery, about how many will have less hip/knee pain after the surgery?	90
7. For each of the following, mark whether or not it is a possible complication of hip/knee replacement surgery.	
a. High blood pressure	No
b. Blood clot in the leg	Yes
c. Migraine headaches	No
d. Infection of the artificial hip/knee	Yes
# 8. Serious complications can happen after hip/knee replacement surgery including life threatening blood clots, infections, heart attacks,	5

and even death. If 100 people have hip/knee replacement surgery, about how many will have a serious complication within 3 months after surgery?	
9. For each of the following, mark whether or not it is a possible side effect of using over-the-counter pain medicine for a long time. These can include medicines you can buy without a prescription like Advil, Aleve, or aspirin.	
a. Stomach ulcer	Yes
b. Migraine headaches	No
c. Kidney problems	Yes
d. Excessive bleeding	Yes

2. Concordance score: There are multiple approaches that have been used to calculate a concordance score.

The first is a simple match, and in this direct approach, we use patients' preferred treatment (assessed with a single item) and then compare with treatment received to determine whether they match. Patient who are unsure are not considered to have treatment that matches. A summary score (0-100%) indicating the percentage of patients who received treatment that matched their stated preference can be generated.

The second approach uses patients' ratings of the importance of salient goals and concerns on a 0 to 10 scale in a multiple logistic regression model to generate a predicted probability of surgery. The dependent variable is binary: Surgery versus No Surgery and the independent variables that remained significant in multivariable analysis were: two goals (not be limited in what you can do and avoid surgery) and joint (hip/knee). Table 2 for parameter estimates for the model published in Sepucha et al 2011. Patients with a predicted probability >0.5 and who had surgery for hip/knee osteoarthritis or those with a predicted probability $\leq$ 0.5 and who did not have surgery, were classified as having treatments matching their goals. A summary score (0-100%) can be generated to reflect the percentage of patients in the sample who received treatments that matched their goals.

Table A2: Concordance model: analysis of maximum likelihood estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.2500	1.1940	12.6705	0.0004
Not be limited in what you can do (0-10)		1	0.5844	0.1219	22.9774	<.0001
Avoid surgery (0-10)		1	-0.2290	0.0369	38.5472	<.0001
Joint (Hip=1, Knee=0)	Hip	1	0.9681	0.2514	14.8343	0.0001

3. Decision Process Score: Participants receive 1 point for a response of "yes" or "a lot/some." The total points are summed and then divided by the total number of items to result in scores from 0-100%, with higher scores indicated a more shared decision making process.

V. Development Process:

This has been described in detail in Sepucha et al (2008), briefly to generate the survey we:

- Conducted a review of the clinical evidence & of focus groups and interviews with patients to generate a candidate set of facts and goals salient to the decision
- Surveyed a convenience sample of patients (n=88) and a multidisciplinary group of clinical experts (n=51) to rate the facts and goals for importance, completeness, and accuracy.
- Drafted the instrument and then conducted cognitive interviews with patients who had knee or hip osteoarthritis (n=10) to evaluate items for acceptability and comprehension
- Conducted field test to evaluate the instruments

Three studies were used to evaluate psychometric properties:

- A cross-sectional study with 382 adults with knee or hip osteoarthritis in the U.S.
- A survey of 45 primary care providers and specialists in the U.S.
- A randomized controlled trial comparing use of knee and hip osteoarthritis decision aids to control with 127 patients in Canada

VI. Psychometric Properties:

These data are taken from Sepucha et al (2011).

Reliability:

- Knowledge score short term (~4 week) retest reliability ICC=0.80 (95% CI 0.69 to 0.87), n=91
- The short term (~4 week) retest reliability for the goals were ICC > 0.72 for all except "avoid treatment that has a long recovery time" (ICC=0.55).
- Decision Process score: internal consistency Cronbach alpha=0.78 and short term (~4 week) retest reliability ICC=0.78 (95% CI 0.67, 0.86)

Note: we did not calculate the internal consistency of the knowledge score because the items do not draw from a single underlying construct.

Validity

- Discriminant validity:
  - The total knowledge score discriminated between patients and providers mean differences of 21%, 95% CI (12%, 30.5%), p<0.001 for knee and 18%, 95% CI (8.5%, 27%), p<0.001 for hip
  - The total knowledge score also discriminated between patients who had seen a decision aid and those who had not, mean difference of 14%, 95%CI (8% to 21%), p<0.001.
  - The concordance model was able to discriminate among patients who stated a preference for surgery, those who were unsure and those who stated a preference for non-surgical options (model predicted probability of surgery 0.74 vs. 0.59 vs. 0.40, respectively, p<0.001 for all comparisons).
- Content validity was confirmed through the extensive feedback from patients and providers in the development process as well as in the field test.
- Predictive validity: Patients who received care that was concordant with their goals had higher decision confidence and less regret

VII. Appropriate Use

The DQIs are protected by copyright. They are available to use at no cost, provided that you:

- Cite the reference in any questionnaires or publications
- Do not charge for or profit from them
- Do not alter them except for customization for a specific condition and reformatting

Suggested Citations for the DQIs:

Sepucha KR. Knee [or Hip] Osteoarthritis Decision Quality Instrument v.2.0. ©Massachusetts General Hospital, 2010 [updated 2012]. Downloaded from: http://www.massgeneral.org/decisionsciences/research/DQ_Instrument_List.aspx.

Sepucha KR. Decision Quality Worksheet: Treatments for Knee [or Hip] Osteoarthritis. v.2.0. ©Massachusetts General Hospital, 2010 [updated 2012]. Downloaded from: http://www.massgeneral.org/decisionsciences/research/DQ_Instrument_List.aspx.

Suggested Citation of the User Guide:

Sepucha KR and Feibelmann S. Hip and Knee Osteoarthritis Decision Quality Instrument User Guide. © 2013. Available from: http://www.massgeneral.org/decisionsciences/research/

VIII. Selected References

Sepucha K, Fowler F, Mulley A. Policy Support For Patient-Centered Care: The Need For Measurable Improvements In Decision Quality. *Health Affairs*. 2004 Oct 7 [web publication].

Sepucha K, Levin C, Uzogara E, Barry M, O'Connor A, Mulley A. Developing instruments to measure the quality of decisions: Early results for a set of symptom-driven decisions. *Patient Education and Counseling* 2008 73:504-510.

Sepucha K, Stacey D, Clay C, Chang Y, Cosenza C, Dervin G, Dorrwachter J, Feibelmann S, Katz JN, Kearing S, Malchau H, Taljaard M, Tomek I, Tugwell P, Levin C. Decision quality instrument for treatment of hip and knee osteoarthritis: a psychometric evaluation. *BMC Musculoskelet Disord*. 2011; 12(1):149.

IX. Questions or comments? Please contact us at decisions@partners.org or visit our website at http://www.massgeneral.org/decisionsciences/