Policy Support For Patient-Centered Care: The Need For Measurable Improvements In Decision Quality

Documenting gaps in patients’ knowledge could stimulate rapid change, moving decisions and care closer to a patient-centered ideal.


ABSTRACT: The phenomenon of practice variation draws attention to the need for better management of clinical decision making as a means of ensuring quality. Different policies to address variations, including guidelines and measures of appropriateness, have had little demonstrable impact on variation itself or on the underlying quality problems. Variations in rates of interventions raise questions about the patient-centeredness of decisions that determine what care is provided to whom. Policies that support the development and routine use of measures of decision quality will provide opportunities to measurably improve the quality of decisions, thereby leading to more patient-centered and efficient health care.

Striking variations in rates of common surgical procedures among seemingly similar populations have been documented for more than sixty years. Variation is greatest when there is legitimate discretion about the best course of action. Sometimes this discretion exists because of inadequate research and the resulting collective professional uncertainty about the effectiveness of an intervention. Sometimes research is adequate but is variably interpreted or disseminated, which results in individual professional uncertainty. In some cases, two or more interventions may be equally effective. Under these circumstances, local conventional wisdom shaped by the beliefs of local medical opinion leaders can take over, masking uncertainty and driving procedure rates in one direction or the other. The result is often an idiosyncratic pattern of rates that John Wennberg has labeled the “surgical signature” of a particular geographic area. The cost and quality implications of these geographic differences are too great to be ignored.

One response of policymakers to practice variation has been to urge professional organizations and researchers to develop clinical practice guidelines to as-
sist physicians in achieving the “right” rate. A related policy initiative was the use of consensus processes to develop criteria for “appropriate” indications for procedures that exhibited high variation. These “appropriateness criteria” formed the basis for preauthorization requirements that became the most visible attempt by managed care organizations to control use and thereby reduce costs. One underlying assumption was that a major source of practice variation was the high use of inappropriate care in areas that had high rates of procedures. By increasing the proportion of appropriate care delivered, policymakers believed that guidelines and a priori judgments about appropriateness would stabilize rates and contain costs while improving quality. However, subsequent studies demonstrated that regions with high rates of procedures, such as gastrointestinal endoscopy, carotid endarterectomy, and coronary angiography, had proportions of “inappropriate” procedures similar to regions with low rates.

There were other difficulties as well with efforts to manage decisions. Doctors resisted guidelines, often citing concerns about their failure to account for differences, subjective as well as objective, among individual patients. The importance of subjective variables was also evident when different panels of physicians using the same evidence and methods produced widely diverging appropriateness criteria. The missing piece was an appreciation of warranted sources of variation—namely, patients’ subjective responses. When more than one choice is “appropriate,” the right decision cannot be determined solely by medical factors; rather, it depends critically on the specific patient’s preferences for outcomes.

The importance of patients’ preferences in understanding and responding to practice variation first became apparent in the mid-1980s. That work sought to clarify the sources for the striking variation in rates of prostate surgery for benign prostatic hyperplasia (BPH) in Maine. Men suffering from urinary problems attributable to BPH faced a dilemma: Surgery was effective in reducing urinary symptoms, but it had a negative impact on sexual functioning. For most men, the choice between surgery and watchful waiting (the only alternative then available) involved a trade-off between these competing aspects of quality of life. Men’s attitudes concerning this trade-off could only be ascertained by directly involving them in the decision—a process labeled “shared decision making.”

Failure to take account of such warranted variation among patients renders health care decision making impersonal, as interventions are provided to people who would not choose them and withheld from those who would. The shared decision-making approach to making care more patient-centered has since grown to include a number of common conditions, many involving surgery as a treatment option. This, in turn, has led to the construction of a series of patient decision aids (PtDAs) designed to promote shared decision making, many of which have been tested in clinical trials. The trend toward replacing physician autonomy by a more shared model is thus forcing changes in the way clinical decisions need to be evaluated and managed.
The Institute of Medicine's (IOM's) *Crossing the Quality Chasm* report cites this work and the resulting approach in defining patient-centeredness as one of six aims for quality improvement. The shared decision-making approach is also reflected in four of the ten “simple rules” for redesign of health care: customization based on patients’ needs and values; the patient as the source of control; shared knowledge and free flow of information; and evidence-based decision making.\(^\text{11}\) The emphasis on simple rules reflects the IOM’s characterization of health care organizations as complex, adaptive systems. In such systems the key to quality is recognizing that high levels of certainty and clinical agreement are often lacking, and therefore the most appropriate care is delivered with flexibility. Variation based on patients’ wants and needs is an essential component of quality. The converse approach, overspecification of care processes, limits the ability to customize and creates inefficiency.

In this paper we argue that three steps are needed to improve the quality of decisions and thereby make health care more patient-centered and efficient. First, a set of measures to assess decision quality that overcomes the current limitations needs to be designed and tested. Second, providers should develop and implement programs and processes for ensuring their routine use in the clinical setting. Third, policymakers should provide incentives and rewards for patient-centeredness by paying for performance that measurably improves decision quality. We offer observations about the challenges that need to be overcome in setting an agenda for change in this direction.

**Current Limitations Of Decision Quality Measures**

Most of what has been done measuring the quality of decisions is connected to evaluations of PtDAs. Andrew Kennedy summarized the measures used in thirty-three randomized controlled trials of decision aids and found a wide variety of measures used as either primary or secondary outcomes.\(^\text{12}\) The four most commonly used measures were patients’ self-reports of (1) satisfaction with decision making, (2) the nature of the interaction with the physician, (3) their state (for example, knowledge), and (4) decisions made. Unfortunately, each type of measure used to assess decision quality has limitations.

The first approach uses satisfaction with the decision-making process as a measure of quality. Paul Cleary provides an extensive critique of “satisfaction” as a measure of quality.\(^\text{13}\) The key weakness is that answers to satisfaction questions are driven primarily by what patients expect. In other words, in-depth testing of such questions has revealed that high satisfaction reports stem from patients’ low expectations rather than by good performance in decision support.

A second approach uses patients’ reports of the nature of the interactions. These questions need to be carefully worded and are often subject to misinterpretation. For example, many patients, when asked how much they participated in decisions about their care, say that they shared responsibility with their physician.
However, when probed to describe how, they often report that the physician provided a recommendation and the patient followed it. Because many respondents do not know how a meaningful shared decision process might progress (never having experienced it in medical care), they cannot reliably answer questions about whether or not they had such an experience.

A third approach focuses on having respondents report on their current state. Annette O’Connor’s seminal work in this area has produced the Decisional Conflict Scale (DCS), which covers four topics: perceived level of information, values clarity, support from others, and certainty about what is best. The DCS is one of the more commonly used measures in the studies of decision aids reviewed by Kennedy. Like some of the measures discussed previously, however, the value of the answers to questions about the patient’s state is limited by the patient’s perspective. Most crucially, a patient may not be able to accurately report whether or not he or she is informed and knows the relevant pros and cons. The answer to that question depends on whether or not the patient has been informed in the first place, as no one can accurately report what he or she does not know.

The last and most common primary measure of the effect of decision aids is the treatment choice that patients made, measured either by their reported intent or by their actual decision. From the payer’s point of view, it is encouraging that use of decision aids tends to reduce the rate of surgery in certain situations, but such information does not provide information sufficient for assessing decision quality. For example, Ed Wagner and colleagues reported that the use of the BPH decision aid increased knowledge and lowered the rate of surgery. However, a change in the rate of surgery does not justify a conclusion that decisions and resulting care were more responsive to the needs and values of individual patients. Such a conclusion would require evidence for concordance between individual patients’ values and the care they received.

New Measures Of Decision Quality

The quality of a clinical decision, or its patient-centeredness, is the extent to which it reflects the considered needs, values, and expressed preferences of a well-informed patient and is thus implemented. None of the available measures adequately assesses decision quality so defined. In particular, data that can be gathered from administrative or medical records alone are not sufficient to assess this concept; rather, it requires input directly from patients. To overcome the limitations of these approaches, we argue that a valid and reliable assessment of decision quality will require three different sets of information: (1) decision-specific knowledge, (2) values for the salient outcomes, and (3) treatments chosen. Items for the first two information sets can be identified through a rigorous social process, including focus groups of patients, clinicians, and researchers, to distill the key pieces of information and the key values that are likely to be most relevant. The set of knowledge questions and value-scaling tasks is not meant to be exhaus-
tive. Rather, it is meant to be a parsimonious set of items that assesses the patient’s knowledge (for example, comprising three to five questions) and values (for example, only the two to four most salient issues).

Decisions about treatment of BPH provide a compelling example of how these measures might work. As noted, the decision to treat urinary symptoms with surgery requires that men understand that there are treatment options and know the likelihood and degree of symptom improvement, as well as the likelihood and severity of side effects, with each. With care, this essential knowledge can be assessed in four simple multiple-choice questions. The decision also requires men to weigh the level and bother of their current symptoms against the possible risks of side effects, mainly sexual dysfunction. The values of individual men regarding the most salient trade-offs can be assessed with three category scaling tasks, each beginning with, “On a scale from 1 to 10, where 1 is the worst possible and 10 is the best possible, how would you feel if...?”

With responses to these four questions and three scaling tasks, and documentation of the treatment chosen for a particular case, one can determine whether a man was well-informed about the key issues and can assess the level of value concordance within a population of men who receive their care from a particular provider or in a particular hospital. Ideally, the provider or hospital would be able to document that the men were well informed and those who felt strongly about preserving their sexual functioning were less likely to undergo surgery for BPH and those who were very bothered by the urinary symptoms were more likely to undergo surgery. These data would provide evidence that treatments were guided by patients’ preferences, not by “surgical signatures” or other unwarranted sources of variation.

The authors and their colleagues have used this framework to assess decision quality. Michael Barry and colleagues used scaling tasks similar to those proposed and calculated the strength of association between measured values and subsequent treatment of BPH. Using a logistic regression model to control for other variables, patients who had used the BPH decision aid and were very bothered by symptoms were seven times more likely to have surgery than those not so bothered. Patients who were very bothered by the prospect of sexual dysfunction were one-fifth as likely to have surgery as those not so bothered.

Obviously, an individual patient may base a decision on some issue that is not one of the values that is selected and measured by one of the scaling tasks. Thus, an apparent mismatch between a well-informed patient’s response to the values and the treatment chosen may not indicate poor quality. (However, it would serve as a marker that further assessment may be needed.) In the aggregate and controlling for other differences, however, the better the decisions, the more that the variance in the decisions can be explained by patients’ reports of their preferences and concerns. As a result, this measure of value concordance can be used to compare decision quality and patient-centeredness of care across populations of patients, even
This same framework has been used to guide the development of a suite of measures that cover some of the key decisions for which there is considerable uncertainty and variation in patients’ preferences for the most common conditions. We have developed candidate items for fourteen other decisions including the treatment of heart disease, cancer (breast and prostate), back pain, joint replacement, and abnormal uterine bleeding. These are being evaluated.

**Getting The Measures Into Practice**

There are not very many examples of measures of quality of care that involve input from patients and are routinely integrated into medical care. The Health Plan Employer Data and Information Set (HEDIS) measures, which the National Committee for Quality Assurance (NCQA) uses to assess health plans and providers, are derived from records and are limited to a relatively small number of care processes where best practice is clear and the same for most patients. The NCQA also uses surveys to measure patients’ access to care and the character of interactions with providers and plans. Because these surveys do not measure anything that applies only to patients with specific health conditions, a cross-section of patients can be surveyed, and sample selection does not have to be integrated into the process of care.

To measure decision quality, patients need to be sampled before or shortly after a decision is made, and they need to answer a set of questions tailored to that decision. The measures must be built into the process of care itself, and this creates implementation challenges. These challenges are not dissimilar to those noted in *Crossing the Quality Chasm*, as it proposed a more sweeping agenda for quality improvement. The report noted the necessity of building organizational support for change, using information technology, and aligning payment policies with quality improvement—all guided by the recognition that health care organizations are complex adaptive systems more responsive to simple rules and system support than to overspecification of work processes. The IOM also strongly recommended a focus on common conditions.

As many health care organizations incorporate the IOM recommendations in their redesign of care, clinical data are increasingly collected electronically at the point of care. The focus on common conditions has facilitated redesign efforts including integration of decision support within institutions and across collaborating networks. For example, at Dartmouth-Hitchcock Medical Center’s (DHMC’s) Comprehensive Breast Program, newly diagnosed patients routinely view a decision aid before meeting with a surgeon to discuss treatment choices for breast cancer. These patients fill out a series of questionnaires at several points throughout the decision-making process. The same decision aid has been used extensively in many breast centers across the country, including those at Massachusetts General Hospital; Northwestern University; and the University of California, San...
Francisco. These and other institutions are now collaborating in the use of a series of decision aids addressing different decisions in breast cancer treatment as well as the development and evaluation of related decision-quality measures.

At the DHMC SPINE Center, a similar system has been implemented for patients suffering with low-back pain. Patients routinely enter quality-of-life and other data on touch-screen pads before each appointment. They use a series of decision aids, including several that have been used in a national network of collaborating sites integrating decision quality improvement with accrual to clinical trials. These examples illustrate that the technical and logistical barriers to decision-quality measurement can be readily overcome with effective clinical leadership supported by favorable incentives and rewards. More attention to better, more patient-centered decision quality is needed from payers to establish such incentives and rewards more broadly.

There are examples of payers’ efforts to alter incentives that could actually work against decision quality. The private sector, in large part through the Leapfrog initiative, has set standards with regard to the minimum volume of surgical procedures. This is a well-intended response to worrisome outcome variation including high operative mortality rates that have been associated with low surgical volume. However, these quality standards create incentives for provider organizations that fall below the minimum to increase the number of procedures they perform. The danger is that this may result in more patients having procedures than would otherwise choose them.

The complementary and corrective next step in strategies to pay for performance should focus on decision quality, to ensure that individual patients are getting the care they want and need, no more or less. Medicare, as the largest payer, may have a special responsibility in this regard. It is encouraging that recent legislation has called for the implementation of shared decision making as part of a demonstration project to improve the quality of care. Section 646 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003, which also contains provisions for reform of the reimbursement system to pay for quality, could serve as the means for galvanizing leadership responsibility for decision quality and provide a laboratory for developing cost-effective methods for its measurement.

It is axiomatic that improvements seldom occur unless the desired performance is routinely measured. Conversely, when measurements of performance are routinely conducted, they create pressure for improvement in and of themselves. Documenting gaps in patients’ knowledge and lack of concordance between patients’ values and preferences and the care received could stimulate rapid change, moving decisions and care closer to the patient-centered ideal advocated by policymakers.

Such measures could provide an important lever for change to leaders of organi-
zations committed to quality improvement. They could invoke simple rules for the complex adaptive systems in which decisions are made. Decisions made when patients are uninformed about the most highly relevant facts could be discouraged or even prohibited. Personalization of care by attending to differences in patients’ values could be encouraged and rewarded. Information technology could be brought to bear to better support doctors and patients in their decision-making roles beginning with common conditions. Because such practical and widespread application of decision theory would be new to the clinical setting, ongoing evaluation and research would be essential.

The persistent widespread variation in rates of procedures will continue until there is a concerted effort to attend to the quality of individual decisions. We recommend that improvement in the quality of patient decision making be given highest priority on the pay-for-performance agenda of private and public payers.

The authors thank John Wennberg and Vickki Entwistle for their helpful comments and suggestions. Floyd Fowler is president of the Foundation for Informed Medical Decision Making, a not-for-profit organization dedicated to improving the quality of medical decisions through developing and disseminating decision support tools and methods. Karen Sepucha and Al Mulley receive financial support from the Foundation for Informed Medical Decision Making for their work in the design and editing of decision aids. Mulley also receives royalties from the foundation.

NOTES


6. For example, the efforts of back surgeons who disagreed with guidelines based on the work of the Patient Outcome Research Team (PORT) funded by the Agency for Health Care Policy and Research (AHCPR) contributed to AHCPR’s near demise in 1995. See B.H. Gray et al., “AHCPR and the Changing Politics of Health Services Research,” Health Affairs, 25 June 2003, content.healthaffairs.org/cgi/content/abstract/hlthaff.w3.283 (29 July 2004).


10. For example, the 1990 National Institutes of Health (NIH) Consensus Conference on the treatment of early-stage breast cancer concluded that lumpectomy plus radiation “is preferable because it provides survival equivalent to total mastectomy...while preserving the breast.” This guideline, however, ignored outcomes of importance to patients as indicated in the following letter written to the editor of the *New York Times*, 20 October 2002: "The decision about treatment for breast cancer remains an intensely personal one. The mastectomy I choose...felt a lot less invasive than the prospect of six weeks of daily radiation, not to mention the 14% risk of local recurrence." In more recent versions of guidelines for breast cancer as well as other conditions, this shortfall has been addressed with specific language highlighting the need to inform patients of benefits and risks and to incorporate patients’ preferences in the choice of treatments.


18. The authors conducted interviews with more than 500 men with BPH, participated in extended collaboration with urologists, and used the application of formal decision analysis. See Note 8.


21. Whether the benefits of implementing decision-quality measures would justify the costs remains a question in need of further research. Widespread implementation of these measures will likely highlight important gaps in patients’ knowledge and a lack of connection between patients’ preferences and the care they receive. However, even more revealing would be studies that randomize decision-support interventions and decision-quality measures in high-use and low-use areas. This type of trial may help determine whether attention to decision quality can help support warranted sources of variation in care, while minimizing unwarranted sources.