Back to Square One: How Patient Satisfaction Survey Flaws Bias Assessments of Health Care Quality

Moses Altsech, PhD
Marketing Department Chair
Edgewood College, USA
An Introduction to
ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. The papers published in the series have not been refereed and are published as they were submitted by the author. The series serves two purposes. First, we want to disseminate the information as fast as possible. Second, by doing so, the authors can receive comments useful to revise their papers before they are considered for publication in one of ATINER's books, following our standard procedures of a blind review.

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research
This paper should be cited as follows:

Back to Square One: How Patient Satisfaction Survey Flaws Bias Assessments of Health Care Quality

Altsech, Moses Ph.D.
Marketing Department Chair
Edgewood College, Madison WI, USA

Abstract
Although opinions on how to best enact much-needed health care reform vary, calls for change have focused almost exclusively on payment reform. Yet as we try to expand health care coverage and make it more affordable for all citizens, it becomes even more important to ensure high quality of care. One of the most common ways of assessing health care quality is through the use of patient satisfaction surveys. In this study, 63 patient satisfaction measures used by a variety of health care organizations across the United States were reviewed. A considerable majority of these measures was found to have fundamental survey design flaws that would affect one’s ability to meaningfully interpret the results, and all but a handful of surveys included measures that could be modified to yield either more accurate or more detailed results. Managerial implications for health care organization decision-makers and other stakeholders are discussed, and meaningful solutions for improving patient satisfaction surveys are presented.
Customer Satisfaction in Health Care

The importance of measuring customer satisfaction has been well established, and customer satisfaction has been linked to desirable outcomes ranging from higher profits and greater loyalty to gaining new customers and benefiting from positive word-of-mouth communication (Altsech 2005; Altsech 2008; Gelb and McKeever 2006; Oliver 1997). In the context of health care, where the customer is the patient, patient satisfaction surveys have been increasingly used by providers of care. It has been argued that patient satisfaction deserves more attention from medical researchers, and that greater integration is needed between the medical and marketing approaches to quality (Bowers and Kiefe 2002). As early as 1983 a British inquiry strongly condemned the National Health Service (NHS) for not using “the well-established techniques of market research to elicit the views and experiences of its users” (Department of Health and Social Security 1983). Doing so can be an integral part of applying the marketing concept to a health care organization to make it more patient-focused (Eckrich and Schlesinger 2011).

The Meaning and Significance of Patient Satisfaction Measurement

The fact that patient satisfaction has become increasingly important should come as no surprise. In the United States, the debate over health care reform over who should pay for health care, what health care services ought to be covered, which individuals ought to be covered and by whom, includes political battles that lasted for years—decades, in fact—and is still the subject of heated political and legal confrontations (Lepore 2009; Pear 2012; Surowiecki 2009). Given the high financial cost of health care, ensuring that the services that are provided are of high quality becomes even more important. Assessment has become an integral part of health care in order to ensure that it is effective, efficient, equitable, patient-focused, timely, and safe in its delivery. The structure of healthcare organizations (and entire health care systems, in fact), the process of delivering care, and the quality of its outcomes have all become the focus of assessment efforts. Patient satisfaction measures are just one of those assessment methods, and their proliferation has increased considerably over the years across many different types of health care providers in the United States, including hospitals, clinics, hospices, private medical and dental practices, medical services firms, and others (Fitzpatrick 1991, Hopkins and Costain 1990, Maxwell 1984).

Insurance providers, government and regulatory bodies, public health advocacy groups, and the general public have taken great interest in the patient’s own voice as reflected in patient satisfaction measures. But health care organizations have incentives to measure patient satisfaction, external pressures notwithstanding: Patient satisfaction scores have been used to make improvements to the process of delivering care and the quality of care itself, to determine compensation and bonuses for medical professionals, and to showcase success stories in the context of marketing and public relations efforts. Patient satisfaction is one of the established yardsticks used to measure how successfully hospitals (and, indeed, all health care organizations) provide services, and an important measure of service efficiency (Sreenivas and Prasat 2003). Patient perceptions of quality of services and cost effectiveness have been described playing a key role in health care organizations’ competitive advantage and survival (Ford et al. 1997). Patient satisfaction measures have been used to assess consultations and patterns of communication, such as the degree to which sharing information, involving patients in decisions about their care, and offering them reassurance
are done successfully (Savage and Armstrong 1990). Patient feedback is also useful in decisions on how to structure the process of delivering care, such as the length of a consultation or the type of after-hours care options (Bollam et al. 1987).

The overall assessment of patient satisfaction is the outcome of judgments made concurrently on several aspects of the health care experience and is linked to several service quality dimensions (Choi et al. 2005; Singh 1990). Patient satisfaction includes assessments of patients’ attitudes towards doctors, medical assistants, quality of administration and quality of atmospherics (Chahal et al. 2004). Clearly the importance of these factors may carry different weights; for instance, in one study nursing care, physician care, and the outcome of care were found to be more important than the “accommodation functions” of hospital care (Carman 2000). Still, it’s important to consider all the factors due to the effect they ultimately have: Staff behavior, atmospherics, information, judgments of examination comfort and perceived worth, were all tied to patient satisfaction and willingness to recommend the provider to others (Peyrot et al. 1993). Perceived quality affects satisfaction and behavioral intentions, in turn, are affected by satisfaction (Gotlieb et al 1994). Accordingly, satisfaction and service quality have been shown to predict important health-related behavior, such as adhering to treatment recommendations, maintaining continuity of care, but also the likelihood of retaining the same provider (Fitzpatrick 1991; Hogan et al. 2007; Kincey et al. 1975; Lee and Young-Hee 2007; Mehta 2011; Roghmann et al. 1979; Weiss and Senf 1990). Just as importantly, there appears to be a link between satisfaction and improvements in patients’ health status (Fitzpatrick et al. 1983; Fitzpatrick et al. 1987).

Differences between expectations and performance result in dissatisfaction (Bearden and Teel 1983; LaBarbera and Mazursky 1983; Lee and Young-Hee 2007 Oliver 1980; Szymanski and Henard 2001). For instance, trust, staff responsiveness, and customer service were found to be significant predictors of patient satisfaction—but satisfaction was higher when patients were offered information about expected wait time, and those expectations were actually met (Mowen et al. 1993; Ross et al. 1987). This, in fact, suggests that often the challenge of increasing patient satisfaction involves some unnecessary obstacles: If patients, in some respects, have unattainable expectations regarding their health care experience, modifying patient expectations to a more realistic level may result in greater satisfaction. This could apply to expectations about prescriptions, diagnostic testing, waiting times, physicians’ assessment processes, referrals, bedside manner, location, information provided by nursing staff, privacy, billing, and more (Vukmir 2006). Since patient perceptions of quality of care, intent to recommend a hospital to others, intent to return for additional care, patient loyalty and profitability have all been strongly linked to health care employee satisfaction as well, efforts to improve employee satisfaction will also likely benefit the patient experience (Atkins et al. 1996).

**Challenges in Measuring Patient Satisfaction**

Some broad criticisms of patient satisfaction surveys have been noted: Critics say that satisfaction survey responses are spur-of-the-moment and poorly thought out, especially for people with health concerns, patients often lack the technical knowledge to make competent judgments about some aspects of their health care experience, and patient evaluations of medical staff’s competence may be largely influenced by evaluations of the staff’s “bedside manner” (Ben-Sira 1980, Fitzpatrick 1991). Those concerns, of course, apply to measuring customer satisfaction in many other contexts, and seem to put into question the very purpose
and value of measuring patient satisfaction. Given, however, the strong evidence between patient satisfaction and outcomes desirable to both patients (e.g. better health, higher likelihood of following “doctor’s orders”) and health care organizations (e.g. higher likelihood of patients returning for treatment and lower likelihood of switching providers), it is clear that the value of measuring patient satisfaction is indeed considerable for all concerned. If anything, the aforementioned concerns make it all the more important to improve both the way in which care is delivered and the way in which patient satisfaction is measured, in order to achieve even more highly diagnostic survey results.

Concerns have also been expressed about the accuracy of the data that patient satisfaction measures yield; for instance, when asked, health care professionals estimate greater levels of dissatisfaction in their patients than patient surveys data reveals (Fitzpatrick 1991; Rashid et al. 1990). In Britain, where the same survey was used across the NHS and allowed for broad comparisons, just a few patients rated any particular aspect of their health care experience negatively, and as many as 80% of respondents expressed satisfaction for any given question (Fitzpatrick and Hopkins 1983). That, however, is not atypical, and a substantial acquiescent response bias was found in patient satisfaction ratings in a study of 233 adults (Ross et al. 1995). If dissatisfied customers are unlikely to complain to the service provider, this phenomenon is even more pronounced in health care settings where complaining behavior might be seen as futile or even cause retaliation against the patient (Altsech, 2008; Australian Council for Safety and Quality in Health Care 2005; Carey 2005; Goodman and Ward 2008; Payne and Fletcher 2005; Schlesinger et al. 2002). Depending on the survey delivery method, this may affect survey responses as well if patients believe their responses can be linked to their identity. It is cause for not being dismissive of negative feedback from patients since it may represent the opinions of other patients who opted to remain silent. If many are reluctant to complain, and if patient responses have been often skewed towards more positive evaluations, that suggests that not only is it important to re-examine how patient satisfaction research is conducted, but also whether the instruments used to conduct it are, in fact, free of bias themselves.

There are research firms that specialize in the creation of patient satisfaction surveys, data analysis, and reporting of data; measurement instruments have also been created, commissioned or endorsed by government agencies, regulatory bodies, professional associations, foundations, and in some instances probably created “in house” by the health care organizations intending to use them. Some provide information on the survey instrument’s validation, while others do not; generally, however, there is little evidence of reliability or validity in instruments (Sitzia 1999). A wide variation exists in the type of satisfaction measures being used, and in a review of measures spanning two decades one can readily identify some with design flaws (Haucksnecht 1990). When issues involved in designing patient satisfaction surveys are ignored, the effects are disastrous (Carr-Hill 1992).

A Study of Patient Satisfaction Surveys

Sixty three patient satisfaction forms were reviewed, including fifty five that are currently being used by a wide variety of health care organizations in at least 21 U.S. states and the District of Columbia. Since some of them are endorsed or used by professional associations, foundations, and federal government and regulatory bodies, including the US Department of Health and Human Services (USDHHS) and the Centers for Medicare & Medicaid Services (CMS), their use may well extend to all fifty states. Seven additional
surveys reviewed are sample surveys developed by well-known research firms specializing in patient satisfaction measurement, which further increases the likelihood that the surveys reviewed are used by hundreds of health care organizations across the United States. The length of these surveys varied widely, ranging from one-half to four-page surveys, and included both traditional “paper” surveys and surveys completed online.

All surveys were reviewed separately by four judges with training in survey design, who were told that these were real patient satisfaction surveys some of which, though not necessarily all, may have “room for improvement.” Judges were instructed to identify any areas for improvement, and then attempt to reach consensus on any areas of disagreement. The vague nature of the instructions was intended to give the judges greater discretion, in order to identify both flaws (i.e. errors in survey design, which were the primary purpose of this study) and items that may be correct but not optimal (e.g. a different measurement scale might yield more information rather than better information). At the end of the process there were not any points that the judges were unable to reach full consensus on.

**Important Flaws in Patient Satisfaction Questions and Measurement Scales**

The review of patient satisfaction surveys suggests that there exist serious and fundamental flaws in the questions and measurement scales embedded in a considerable number of survey instruments.

A double-barreled question asks about more than one construct in a single survey question (Lavrakas 2008). Since a single measurement scale is used to assess two constructs that should be rated separately (such as rating the promptness and courtesy of service), some respondents ultimately report an average of the two ratings, others opt to report their rating on the construct they consider more important, and others don’t respond at all—making it impossible to interpret the responses. As a result, the use of double-barreled questions has yielded some notoriously biased research reports (Babbie 2010). Of the 63 patient satisfaction surveys reviewed in this study, no fewer than 45 (71%) were found to include one or more double-barreled questions.

One of the most commonly used measurement scales in social science research is the Likert (or summated rating) scale, used to assess people’s attitudes, beliefs, emotions, feelings, perceptions, personality characteristics, and other psychological constructs along a quantitative continuum (Spector 2003). The most common types of Likert scales have 5 or 7 points, of which one is a mid-point denoting neutral, average, or uncertain judgment, surrounded by equal numbers of negative and positive points. If that mid-point on a 5-point scale is labeled using a positive label, there are now three positive points, two negative points, and no neutral point: Respondents who intend to give an “average” rating inadvertently (or inevitably) give a positive one, and reports that summarize survey results present artificially inflated positive ratings. Of the 63 patient satisfaction surveys reviewed in this study, 36 (57%) had mid-points that were decidedly not neutral. Among them, “satisfied,” “somewhat positive,” “good,” “fairly confident,” “very much,” “ok,” and “fair” (often used as a negative point label as well) were used as mid-point labels on odd-numbered Likert scales included in the survey instruments.

**Further Improving Patient Satisfaction Surveys and Data Collection**
It is a common practice to choose a type of Likert scale (e.g. a 5 or 7-point scale) and use it consistently throughout a survey. Although that consistency is not necessary per se, it does make the reporting of results more user-friendly, since the reader doesn’t have to wonder whether a 4.5 rating is sometimes very high (on a 5-point scale) and sometimes just above average (on a 7-point scale). Among the surveys reviewed, most used a 5-point scale consistently throughout the survey; some used both 5-point and 7-point scales, and others used both 4-point and 5-point scales. In one instance 4, 5 and 11-point scales were used in the same survey, and in another 3, 4, 7 and 10-point scales were used, likely making any report on the results a test of both the reader’s attention and patience.

Dichotomous questions (e.g. “yes/no”) are also commonly used, and in many cases appropriately so. Asking respondents whether they’ve been hospitalized within the last 12 months does warrant a “yes” or “no” response. Asking them whether they would recommend a particular doctor to a friend, however, could be answered in a similar manner—but respondents who are entirely undecided, or are leaning one way or the other without having made a firm decision, would be forced to either not respond at all or give a response that does not truly reflect their intentions. Measuring attitudes and intentions using dichotomous questions often yields less information than Likert-scale questions would elicit. Several of the surveys reviewed would benefit from replacing yes/no questions with Likert scale-rated questions.

On one instance, patients were asked to indicate whether during their hospital stay they had been given any medication that they had not taken before (a “yes/no” question). This is an example of a question that respondents may not be in a position to answer (not having been told, being unable to recall, or not having the technical knowledge to respond). Considering the patient’s ability to respond to a question is important if the results are to be meaningful.

There were a few isolated cases where the patient satisfaction form could be submitted through the health care organization’s web site without requiring access a secure patients-only area. Therefore virtually anyone (a patient, an employee, a competitor, a person studying patient satisfaction forms) could go online, complete, and submit a Survey Monkey-based survey, rendering the aggregate patient satisfaction survey results entirely meaningless.

Ultimately, three of the 63 surveys were found to be flawless; one used by a surgical practice in Florida, and two provided by the reputable non-profit Rand Corporation. Oddly, two other surveys listed by Rand suffered from some of the flaws discussed extensively above, as did the CMS-created HCAHPS survey, a USDHHS survey, the FEHC survey used by several members of the National Hospice and Palliative Care Organization, and many others, including those created and sold by research firms whose very business is survey design and data analysis.

**Managerial Implications for Health Care Organizations**

The survey design concerns identified in this study don’t involve complex or obscure methodological tidbits that could easily be overlooked or misinterpreted; instead they involve very basic, fundamental concepts and long-established best practices in survey design. The fact that some of these surveys are used (and were likely developed) by small private practices makes it reasonable to attribute the inadequacies to a lack of resources or, most likely, to the notion that the skills required for survey design are limited to the ability to use online search engines, use Word for Windows, and form complete sentences—in short, the
belief that neither expertise nor experience are required to complete the task. The fact that some of these surveys, however, are used by large health systems, hospitals, clinics, and other organizations with more resources and access to expertise, is indicative of the fact that they see no perceived need (as well as lack the ability) to evaluate the instruments they use: If, as is almost certainly the case, the patient satisfaction surveys in use were developed or endorsed by government agencies, regulatory bodies and professional associations, and created by well-known research firms that claim both experience and expertise in survey design, then these health care organizations have little reason to doubt the quality of the surveys. It is a quintessential example of the blind leading the blind.

In one instance, a network of clinics and a large hospital that were part of the same health system used one of the worst and one of the better patient satisfaction surveys respectively, suggesting that pooling—and, even more importantly, coordinating—resources is particularly important. One of them likely used the wrong vendor, and “caveat emptor” never sounded more relevant.

The implications, of course, are far-reaching: Health care organizations may not be measuring what they intend to measure in the first place—or they may be measuring it inaccurately or inadequately or both. Worse yet, those organizations whose patient satisfaction data collection is not just driven by reporting requirements, regulations, or keeping up appearances, but instead by a genuine desire to continuously improve the quality of care they provide to the people they serve, may ultimately make decisions about processes, services, personnel, resource allocation, and marketing efforts based at least partly on data that is misleading and biased. Not only will they not achieve any desirable outcomes, but they will waste scarce time and resources in the process. Less scrupulous health care organizations may even use the opportunity to gain an unfair competitive advantage by public reporting (and promotion) of patient satisfaction data with artificially inflated positive ratings.

Resistance to change is often based on the argument that adopting a widely used patient satisfaction survey provides the added benefit of benchmarking data, and that making changes will make it difficult to compare current and future performance to that of previous years. Ultimately, the decision will depend on what the organization values the most: When accurate assessments of patient satisfaction and a continuous improvement in the quality of care are the ultimate goals, continuing to use survey instruments tainted by measurement error and knowingly using the biased data they yield to make both strategic and tactical decisions is not a viable option.

Relying entirely on a marketing research firm that sells ready-made surveys, data collection and data analysis services (often for a period of many years) as the sole arbiter of survey design quality is not much different from relying entirely on a seafood merchant as the sole arbiter of whether that seafood is fresh. In the latter case, fortunately, customers can additionally rely on their own senses of sight, smell and taste (which, even after the fact, will detect poor quality). In the former case, the situation is decidedly more complicated: Even large health care organizations with entire marketing departments have in-house expertise in public relations, social and traditional media, advertising and promotion, but typically not marketing research. The proverbial “second pair of eyes” invariably has to come from hiring a full-time staff member with both special training in and experience with marketing research methods, or from a contracted employee or consultant with extensive training, experience and established credentials specific to marketing research methods who can either create surveys or at least be the organization’s advocate when surveys are other research services are purchased from third parties. Given the importance of that advocate’s role, this is one instance where being very picky in the selection process would serve the organization well.
Advocacy groups, non-profit organizations, and coordinated initiatives by groups of health care organizations working together on payment reform can also make a difference in improving patient satisfaction measures by emphasizing the pivotal role that improved quality assessments can play in the context of providing high quality, affordable care.

Limitations and Suggestions for Future Research

As many and as varied the surveys reviewed in this study may have been, they include a high number of instruments—as many as half—used by private medical practices (some small, and some very large). The surveys cannot be described as a representative sample, since only those that were available online or could be readily obtained in hard copy were reviewed, and none of the surveys reviewed were telephone surveys; all were self-administered online or paper surveys. Whether different types or harder to obtain surveys are of higher quality is difficult to ascertain. A deliberate attempt was made to include patient satisfaction surveys from many different types of health care organizations; whether assessing a large number of surveys specific only to hospitals or clinics or dental practices would yield different insights remains uncertain. Having information on the process of how these surveys were developed, access to reports presenting patient satisfaction survey results and information on how those results were used could put many of the concerns raised by the present study in a more complete context.

Further research could illustrate why these survey design flaws are so widespread, and what consequences they have for health care organizations. One way to accomplish that would be a study of the process used by health care organizations (in aggregate and by type) to develop, commission, adopt, or purchase the patient satisfaction forms they use. Investigating the range of motives for measuring patient satisfaction in the first place, as well as the managerial uses of patient satisfaction survey results, would be of particular interest as well. Perhaps most intriguing of all would be to assess health care organization decision makers’ attitudes towards patient satisfaction research, and their reactions to the evidence of widespread survey design flaws in surveys used by their own and other organizations.

In what would be a departure for some (but not all) types of health care organizations, expanding the unit of care to include not just patients, but their families as well, would also necessitate the creation of additional high-quality measures of satisfaction. This may be relevant to some types of organizations, but not others: Most people go to the dentist alone, but many of those who are hospitalized may be less likely to have a direct experience with the billing department or the gift shop at the hospital than their spouse. Since the experiences and opinions of significant others have been shown to influence patients’ behavioral intentions as they relate to future care, capturing those significant others’ feedback would be meaningful for hospitals. Hospices, on the other hand, have a long history of capturing family members’ feedback—although, astonishingly, they fail to capture feedback from patients themselves, who are often perfectly capable of providing it. The opportunities for improvement are considerable.

Ultimately, the goal would be to create buy-in among decision makers that the use of scientifically sound measurement is of paramount importance in measuring patient satisfaction, and reassessing the surveys currently in use is as meaningful as it is urgent. Having a few major health care organizations lead the way in improving the quality of patient satisfaction surveys could eventually result in the kind of standards consistency that would benefit both the organizations and the patients they serve.
References


of Satisfaction Decisions,” *Journal of Marketing Research*, 17 (November), 460-69.


