# Assessing symptoms before hysterectomy: Is the medical record accurate?

**Michael S. Broder, MD,[a] David E. Kanouse, PhD,[b] and Steven J. Bernstein, MD[c]**

*Los Angeles and Santa Monica, Calif, and Ann Arbor, Mich*

**OBJECTIVE:** Our purpose was to evaluate the agreement between the documentation of symptoms leading to hysterectomy and the assessment of those symptoms by the patient.

**STUDY DESIGN:** A retrospective study was performed of 497 women in southern California who had hysterectomies. Sensitivity, specificity, and κ statistics were calculated for the medical records and were compared with patient interviews for the presence and severity of symptoms.

**RESULTS:** The medical record was 93% sensitive and 61% specific for identifying bleeding and 79% sensitive and 55% specific for identifying pain. Overall agreement between physician records and patient interviews was moderate for bleeding (κ, 0.55-0.58), fair for pain (κ, 0.29-0.34), and poor for impairment as a result of bleeding or pain (κ, 0.0-0.14).

**CONCLUSIONS:** Physician overestimation of symptoms could lead to overuse of hysterectomy, whereas underestimation could result in underuse. Our results suggest that both underestimation and overestimation occur for patients with abnormal bleeding, pain, or both. If physicians accurately assess symptoms but fail to document them, examinations of appropriateness will be faulty unless patients are interviewed. (Am J Obstet Gynecol 2001;185:97-102.)

**Key words:** Hysterectomy, reliability, medical record, retrospective, interview, background

Medical records, which historically served only as a place for the medical staff to record their notes, are now being used for a variety of other purposes. Physician peer review organizations use medical records to evaluate physician adherence to recommended practice.[1] Hospitals review medical records for physician credentialing and quality assurance, and some Health Plan Employer Data and Information Set measures rely on medical records to assess quality of care.[2] Medical records are also used to determine whether patients receive adequate and appropriate care.[3]

The medical record is an attractive source for evaluating care because it can be used without disrupting either the patient or the physician. For some procedures, such as coronary angiography, the medical record contains sufficient data to apply appropriateness criteria without obtaining additional information from interviews.[4] For many other procedures, including hysterectomy, the extent of agreement between medical records and other data sources is unknown.[5]

It is important to examine the quality of medical records for hysterectomy for several reasons. First, hysterectomy is the second most common major operation that women undergo in the United States. Second, there are several sets of criteria used to evaluate the use of hysterectomy that rely, either completely or in part, on data found in the medical record.[6-9] Third, because hysterectomy is usually performed to relieve symptoms and improve quality of life, it is crucial that physicians accurately assess and document patient symptoms.[10] Overestimating the impairment of patients could lead physicians to recommend hysterectomy too frequently, whereas underestimating impairment could dissuade physicians from recommending surgery sufficiently early.

In this article, we evaluate the level of agreement regarding symptoms between medical records and patient interviews with women who have undergone hysterectomy. We also examine whether physicians document the impact those symptoms had on a patient's ability to carry out her normal daily activities.

## Materials and methods

Nine capitated medical groups in southern California agreed to participate in a project designed to test implementation strategies for clinical guidelines. From the member lists of these organizations, we identified 1089 women who underwent hysterectomy before guidelines were introduced or implemented (between August 1, 1993 and July 31, 1995). We excluded 310 of these pa-

*From the Department of Obstetrics and Gynecology, University of California, Los Angeles[a]; RAND[a, b]; the Veterans Affairs Medical Center[c]; and the Departments of Internal Medicine and Health Management and Policy, University of Michigan, Ann Arbor.[c]*

*Reprint requests: Michael S. Broder, MD, Department of Obstetrics and Gynecology, University of California, Los Angeles School of Medicine, Box 951740, Los Angeles, CA 90095-1740. E-mail: mbroder@ucla.edu.*

*0002-9378/2001 $35.00 + 0* ***6/1/107098***
*doi:10.1067/mob.2001.107098*

tients because the procedure codes for their surgeries did not meet our criteria (*International Classification of Diseases, Ninth Revision*, codes 68.3-68.8), surgery was emergent or for a previously diagnosed cancer, or the patients were non-English speaking, were cognitively impaired, or were deceased. We received informed consent for participation from 539 of the 779 eligible patients. An additional 42 patients were excluded from this analysis because of incomplete data collection, leaving 497 cases in the final data set. The study was approved by the RAND Institutional Review Board. In accordance with the requirements of the Institutional Review Board, we did not collect data on patients who declined to participate in the study.

As part of the primary study, we performed a structured chart review and conducted telephone interviews with each patient. Nurse abstractors examined inpatient and outpatient medical records for each case with a standardized abstraction form. Abstractors sought to determine whether a patient had bleeding, pain, or both noted as a reason at least in part for the procedure. Abstractors were asked to justify their answers by copying exact data from the medical record. In addition to recording detailed data from these records, abstractors reviewed and photocopied all admission and discharge notes, operative reports, laboratory results, and pathology reports. Physician reviewers then examined the chart abstractions and determined whether the treating physician considered the patient's symptoms to have a significant negative impact on her level of activity or functional ability; this significant negative impact was termed "major functional impairment."

An average of 9 months after surgery, nurses conducted patient telephone interviews with a prepared script. Telephone interviews lasted approximately 30 minutes (mean, 29.9 ± 9.8 min) and involved both open-ended and closed-ended questions. Women were asked to rate their bleeding or pain during the 3 months before the final decision to have a hysterectomy on a 5-point scale ranging from "no problem" to "big problem." Patients who reported symptoms were asked to provide the number of days per month they were unable to perform their usual activities, work, or engage in social activities because of their symptoms. On the basis of the recommendation of a 9-member multispecialty expert physician panel that convened to review the literature on hysterectomy and to help establish guidelines for its use, we considered a patient to have a major impairment if she was unable to engage in her usual activities, which were defined as activities in or out of the home, including work, social, and recreational activities, for two or more days a month because of bleeding or pain. The panel felt that, in most cases, this level of impairment constituted a significant impediment to normal activity and was a reasonable level at which to consider surgery.

Before data entry, nurses reviewed all forms for completeness and accuracy. All interviewers, nurse abstractors, and physician reviewers underwent formal training sessions and were randomly assigned to patients. All identifying patient data were removed from the records before abstraction. Five percent of the records (n = 27) were reviewed independently by two abstractors to assess interrater reliability. $\kappa$ Statistics measured the degree of agreement beyond that expected as a result of chance alone. Perfect agreement between the two data sources would yield a $\kappa$ statistic of 1, whereas agreement that was no better than chance would yield a $\kappa$ statistic of 0. $\kappa$ Scores of 0.0 to 0.20 were considered poor, scores of 0.21 to 0.4 were considered fair, scores of 0.41 to 0.60 were considered moderate, and scores of 0.61 to 0.80 were considered substantial.[11] We found perfect agreement between the two abstractors regarding the presence of bleeding ($\kappa$ = 1.0) and substantial agreement for the presence of pain ($\kappa$ = 0.74).

To assess the adequacy of the medical record to determine the presence of symptoms or conditions leading to hysterectomy, we compared documentation in the medical record of the presence or absence of bleeding or pain with the patient report of these symptoms. Similarly, we compared documentation of impairment from these symptoms with the patient report of such impairment. We then calculated the $\kappa$ statistic as a measure of agreement between the patient's report of symptoms and the medical record. With the patient interview as the criterion standard, we also calculated the sensitivity and specificity of the medical record as a test for the presence of these symptoms. In addition, to provide a standard of comparison for patient recall concerning their surgery, we compared patient interviews and medical record data on the performance of oophorectomy at the time of surgery. We performed all statistical calculations with STATA (version 5.0, Stata Corporation; College Station, Tex) statistical software.

## Results

Pain, bleeding, or both were the primary symptoms before hysterectomy in 79% of the patients (n = 394) according to the medical record. Fifty-five percent of all hysterectomies (n = 274) were performed on patients with uterine leiomyomata and some combination of pain or bleeding. Nine percent of patients (n = 43) had abnormal uterine bleeding in the absence of fibroids. Pelvic pain with adhesions, endometriosis, dysmenorrhea, and chronic pelvic pain accounted for an additional 9% of cases (Table I). Additional demographic details of the subject population have been previously reported.[12]

Surgeons performed 75% of the hysterectomies abdominally, 22% vaginally, and 3% vaginally with laparoscopic assistance. There were 366 total abdominal hysterectomies and 9 supracervical hysterectomies. Sur-

**Table I.** Clinical diagnosis in 497 women who underwent hysterectomy

| *Diagnosis* | *No. of Cases (%)* |
|---|---|
| Pain | |
| Leiomyomata with pain | 66 (13) |
| Pelvic relaxation with pain or discomfort | 33 (7) |
| Endometriosis | 21 (4) |
| Chronic pelvic pain | 10 (2) |
| Dysmenorrhea | 9 (2) |
| Pelvic pain and adhesions | 4 (1) |
| Total pain | 143 (29) |
| Bleeding | |
| Abnormal uterine bleeding | 43 (9) |
| Leiomyomata with uterine bleeding | 23 (5) |
| Total bleeding | 66 (13) |
| Pain and bleeding | |
| Leiomyomata with uterine bleeding and pain | 185 (37) |
| Other diagnosis | 103 (21) |
| TOTAL | 497 (100) |

**Table II.** Agreement of medical record and patient interview for bleeding, pain, and impairment as a result of bleeding or pain

| | *Sensitivity* (%)* | *Specificity* (%)* | *κ Statistic* | *No. of patients* |
|---|---|---|---|---|
| Bleeding | | | | |
| Any problem | 89 | 67 | 0.55 | 497 |
| Medium or big problem | 93 | 61 | 0.58 | 497 |
| Major impairment | 29 | 66 | 0.00 | 355 |
| Pain | | | | |
| Any problem | 76 | 61 | 0.29 | 497 |
| Medium or big problem | 79 | 55 | 0.34 | 497 |
| Major impairment | 37 | 78 | 0.14 | 281 |
| Oophorectomy | | | | |
| Bilateral | 97† | 96† | 0.92 | 497 |
| Unilateral | 90† | 99† | 0.90 | 497 |

*Sensitivity and specificity of the medical record as a test, with the interview as the criterion standard, except as noted for oophorectomy.

†Sensitivity and specificity of patient interview with the pathology report as the criterion standard.

**Table III.** Comparison of the medical record and patient interviews for complaints of severe abnormal bleeding*

| | *Interview* | | |
|---|---|---|---|
| *Medical record* | *Problem absent* | *Problem present* | *TOTAL* |
| Problem absent | 97 | 25 | 122 |
| Problem present | 61 | 314 | 375 |
| TOTAL | 158 | 339 | 497 |

*Severe abnormal bleeding is described as a “medium” or “big” problem.

geons removed both ovaries during the hysterectomy in 277 patients and removed one ovary during the hysterectomy in 42 patients.

During the patient interview, 377 women reported that bleeding was a problem before hysterectomy. When we classified patients who said that they had had any bleeding problems, even if the problem was “very small” or “small,” as positive for bleeding, the medical record was 89% sensitive and 67% specific for identifying this problem (Table II). With a more restrictive definition, in which only patients who described their bleeding as a “medium” or “big” problem were considered positive, the sensitivity and specificity of the medical record were 93% and 61%, respectively (Task II and III). There was moderate agreement between the medical record and patient interviews both when bleeding was defined as any positive response (κ = 0.55) and when bleeding was defined with the more restrictive definition (κ = 0.58; Table II).

We also determined the agreement between the medical record and patient interviews regarding whether patients suffered from “major impairment” as a result of their bleeding. With the data-gathering algorithm, this information was available on 355 of the 377 women who had complained of bleeding before hysterectomy. The sensitivity and specificity of the chart for major impairment because of bleeding were poor (29% and 66%, respectively), with no more agreement between medical record and patient interviews than would be expected by chance alone (κ = 0.0).

Although our definition of impairment was formed on the basis of expert consensus, we performed sensitivity analyses to determine whether this poor agreement depended on our definition of impairment. We redefined major impairment as present if a patient was unable to perform her usual activities for one or more days each month because of bleeding, rather than the two days in the original definition. With this new definition, the medical record was 30% sensitive and 67% specific for impairment; however, agreement was still no better than chance alone (κ = 0.0). Even raising the threshold for impairment to 21 days of impairment per month, which excluded 90% of patients with any impairment, had no effect on the κ statistic and caused little change in sensitivity or specificity (31% and 69%, respectively).

The medical record and patient interviews showed lower levels of agreement regarding the presence of pain than they did regarding the presence of bleeding. When any degree of pain was considered to be a problem, the medical record was 76% sensitive and 61% specific with a fair level of agreement (κ = 0.29). With a more restrictive definition that considered pain present only if it was a “medium” or “big” problem, sensitivity increased to 79%, but specificity fell to 55%; the level of agreement remained fair (κ = 0.34). Information regarding major impairment from pain was collected on 281 of the 403 patients who complained of pain before hysterectomy. Physician documentation of major impairment from pain had 37% sensitivity and 78% specificity compared with the patient report. This corresponds to a κ statistic of 0.14, which indicates a slight agreement between the patient interview and the medical record.

When we performed a sensitivity analysis, redefining major impairment as present if the patient was unable to perform her usual activities for at least one day because of pain, we found that sensitivity and specificity were almost unchanged at 35% and 78%, respectively, and that the κ statistic was slightly lower at 0.09. As we found for bleeding, even raising the threshold of impairment substantially higher had little or no effect on the κ statistic.

During the interview, we also asked the patient whether one, both, or neither of her ovaries had been removed. Of the 497 patients, 277 had bilateral oophorectomy and 42 had unilateral oophorectomy at the time of surgery. We compared patient responses with pathology reports. In this case we used the pathology report as the criterion standard and the patient response as the "test." We found that patient interviews were 97% sensitive and 96% specific for properly identifying whether both ovaries were removed at the time of surgery and that patient interviews were 90% sensitive and 99% specific for identifying removal of one ovary. The κ statistic in these cases was 0.92 and 0.90, respectively, indicating almost perfect agreement between the pathology report and the patient interview.

## Comment

We conducted this analysis to determine whether physician documentation of symptoms leading to hysterectomy agreed with a patient's estimation of those same symptoms. Perhaps not surprisingly, physician documentation of symptoms was far better than documentation of the impact of those symptoms on a patient's quality of life. There was a moderate level of agreement between physician and patient for the complaint of bleeding, whereas for pain symptoms there was a fair level of agreement. This finding is consistent with previous work that suggests that physicians underestimate pain when they assess patient symptoms.[13-16] The lowest levels of agreement concerned the effect of these symptoms on the functional status of the patient; agreement on this issue varied from poor to no better than chance alone.

Previous studies of the agreement of reports by patients and their health care providers have shown mixed results, with estimates of agreement on quality-of-life measures ranging from low to relatively high, depending on the specific measure being compared.[17] Most of these studies involved the estimation of performance status and quality of life in patients with cancer or chronic diseases; therefore they may not be directly applicable to this patient population.[18, 19] More studies will be necessary to better understand the factors that influence the accuracy with which gynecologists and other practitioners assess and document their patients' symptoms.

Many studies comparing data from medical records with data provided by patients in interviews or questionnaires have treated the medical record as the criterion standard, with the patient report regarded as an imperfect substitute that is often more readily obtainable for research purposes. In contrast, we have treated the patient report as the criterion standard for the occurrence of gynecologic symptoms of bleeding and pain that are severe enough to affect quality of life. Assessing the severity of these symptoms requires careful questioning, and the results of that questioning may not always be documented in the medical record. For both these reasons, the medical record may be inaccurate regarding the presence and severity of these symptoms. Patient reports of symptoms that affect quality of life are often difficult to validate independently, but self-reported morbidity in the general population has been found to compare favorably with physician assessment of morbidity in predicting mortality.[20] More generally, whether the medical record or patient report is the more accurate source will vary according to the types of error that affect each for a given topic.[21, 22]

It is not clear what effect the average 9-month interval between surgery and interviews has on patient responses. Our study demonstrated that on one independently confirmable measure, that is, performance of oophorectomy, patients had excellent recall at 9 months after surgery. Previous studies have shown inconsistent results regarding patient recall of health-related topics over time.[23] Emberton et al[24] showed only fair agreement for self-reported symptoms 3 months after prostatic resection, with the direction of the misclassification being random. Revicki et al,[25] however, showed very high levels of agreement for patient recall of days missed from work or usual activities after a delay of 90 days. Guadagnoli and Cleary[26] found moderate agreement between quality-of-life surveys administered in the hospital and at 3 months after discharge. For certain activities, they found that at 3 months patients tended to understate their prehospitalization level of impairment compared with similar reports made during their hospitalization. These studies suggest that, although patient reports of symptoms may change over time, this change is random and not consistently in one direction or the other. Furthermore, patients' estimates of days missed from work appear to be consistent over time, suggesting that our assessment of impairment should be little affected by the time elapsed between surgery and interview.

Because this study focused on medical records rather than physician interviews, we cannot determine whether physicians' assessment of their patients' symptoms was more accurate than is reflected in the medical record. Gynecologists may assess their patients' symptoms more accurately than they document them. Physician training has emphasized the recording of observable findings in the medical record and placed less emphasis on the patient's perspective on her illness and her symptoms. However, the gynecologic symptoms considered here are examples of an important class of conditions that require taking a medical history to elicit the critical information from the patient for proper assessment. In the case of hysterectomy and other treatments and procedures that relate more directly to improving quality rather than length of life, physicians may need to be better trained in eliciting and recording relevant quality-of-life data, rather than simply recording physical examination findings. The determination of whether the problem resides in charting deficiencies or in the assessment of symptoms is beyond the scope of this study. However, if the problem were entirely in the failure to document known symptoms, we would expect the primary deficiency in the medical record to be low sensitivity. Instead, we found that physicians also often recorded symptoms that patients reported they did not have, which suggests that some of the problem lies in inaccurate assessment.

We have no reason to believe that physicians falsified records to establish a reason for performing surgery. It seems more likely that physicians simply failed to ask patients directly about the impact of their symptoms and inferred that patients had problems on the basis of standard definitions. For example, a gynecologist may record a bleeding problem on the basis of whether a patient reports bleeding for more than 8 days per month, regardless of whether she feels this is a problem. Although 8 days of bleeding may make it impossible for one patient to function normally, it may have no effect on another patient. It is also possible that the inaccuracies that we reported may have been the result of patients who reported problems before surgery that they then forgot by the time of the interview, although this seems less likely in light of prior work that does not demonstrate systematic forgetting of symptoms.[22-25]

Regardless of whether physicians knew about but did not record symptoms or did not know about symptoms at all, these data highlight several potential problems with medical care for women with gynecologic problems. First, the inaccuracy of the medical record with respect to the presence of certain types of symptoms that primarily affect quality of life makes appropriateness assessments suspect if they fail to include data obtained directly from patients. American College of Obstetricians and Gynecologists criteria sets rely in part on the presence and impact of symptoms to determine whether a case should undergo further review for appropriateness.[6-8] The attempt to assess compliance with these criteria without speaking to patients may lead to the misestimation of the number of inappropriate procedures performed. Furthermore, if impairment because of symptoms is a crucial variable in the decision-making process leading to hysterectomy, as it should be when surgery is performed to improve quality of life, then failure to document that impairment represents a quality problem. Symptom impact cannot be accurately inferred from measures such as frequency or duration of bleeding; it must be directly ascertained from the patient.

Although these data represent care at only 9 medical groups in one geographic area, we have no reason to believe that these groups differed systematically from other groups in ways that would limit the ability to generalize our findings. There were no significant demographic differences between those women whose physicians accurately recorded their symptoms and those whose physicians did not, and accuracy did not differ significantly among the 9 medical groups.

If documentation accurately reflects gynecologists' assessments of their patients' symptoms, then our findings suggest that these physicians have a poor understanding of the symptoms that lead their patients to have hysterectomies performed. As a result, physicians may recommend hysterectomy to some women whose symptoms are not severe enough to warrant such surgery. Although data from this study were collected only from women who actually had hysterectomies performed, the low sensitivity of the medical record for impairment as a result of bleeding or pain suggests that gynecologists may fail to identify substantial numbers of women who are suffering from disabling pain and bleeding. If further studies demonstrate that this is also true of women who present for evaluation of gynecologic complaints, including women who never have surgery, this would suggest that hysterectomy may be underused as well as overused, that is, that women with disabling symptoms, particularly those with pain stemming from gynecologic problems, are not being properly identified and thus do not receive treatments that could relieve their suffering.

## REFERENCES

1. Dept of Health and Human Services (US). Medicare peer review organization manual. Washington: Dept of Health and Human Services (US); 1988. HCFA-Pub 19.
2. Lohr KN. Committee to design a strategy for quality review and assurance in medicare division of health care services. In: Lohr KN, editor. Medicare: a strategy for quality assurance. Washington: National Academy Press; 1990.
3. Donabedian A. The definition of quality and approaches to its assessment: explorations in quality assessment and monitoring. Ann Arbor (MI): Health Administration Press; 1980. p. 163.
4. Kosecoff J, Fink A, Brook R, Chassin M. The appropriateness of using a medical procedure: is information in the medical record valid? Med Care 1987;25:196-201.
5. Dresser MV, Feingold L, Rosenkranz SL, Coltin KL. Clinical quality measurement: comparing chart review and automated methodologies. Med Care 1997;35:539-52.
6. ACOG criteria set. Quality evaluation and improvement in practice: abdominal hysterectomy with or without adnexectomy for endometriosis. Number 27, October 1997. Int J Gynaecol Obstet 1998;60:92-3.
7. ACOG criteria set. Hysterectomy, abdominal or vaginal for chronic pelvic pain. Number 29, November 1997. Int J Gynaecol Obstet 1998;60:316-7.
8. ACOG criteria set. Hysterectomy, abdominal or vaginal for ab-

normal uterine bleeding. Number 28, November 1997. Int J Gynaecol Obstet 1998;60:314-5.

9. Leape LL, Bernstein S J, Bohon CJ, Dickerson VM, Ling FW, Shiffman RN, et al. Hysterectomy: clinical recommendations and indications for use. Santa Monica (CA): RAND; 1997. Publication MR-592/1-AHCPR.
10. Bernstein S, Fiske M, McGlynn E, Gifford D. Hysterectomy: a review of the literature on indications, effectiveness, and risks. Santa Monica (CA): RAND; 1996. Publication MR-592/2-AHCPR.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
12. Rowe MK, Kanouse DE, Mittman BS, Bernstein SJ. Quality of life among women undergoing hysterectomies. Obstet Gynecol 1999;93:915-21.
13. Symptom management for acute pain. Bethesda (MD): Public Health Service (US), National Institutes of Health; 1994. NIH Pub No. 94-2421.
14. Donovan M, Dillon P, McGuire L. Incidence and characteristics of pain in a sample of medical-surgical patients. Pain 1987;30: 69-78.
15. Marks RM, Sachar EJ. Undertreatment of medical inpatients with narcotic analgesics. Ann Intern Med 1973;78:173-81.
16. McCormack JP, Li R, Zarowny D, Singer J. Inadequate treatment of pain in ambulatory HIV patients. Clin J Pain 1993;9:279-83.
17. Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol 1992;45:743-60.
18. Aaronson NK, Bakker W, Stewart AL. Multidimensional approach to the measurement of quality of life in lung cancer clinical trials. In: Aaronson NK, Beckmann J, editors. The quality of life of cancer patients. New York: Raven Press; 1987. pp. 63-82.
19. Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM. Who should measure quality of life, the doctor or the patient? Br J Cancer 1988;57:109-12.
20. Ferraro KF, Farmer MM. Utility of health data from social surveys: is there a gold standard for measuring morbidity? Am Sociol Rev 1999;64:303-15.
21. Hewson D, Bennett A. Childbirth research data: medical records or women's reports? Am J Epidemiol 1987;125:484-91.
22. Westbrook JI, McIntosh JH, Rushworth RL, Berry G, Duggan JM. Agreement between medical record data and patients' accounts of their medical history and treatment for dyspepsia. J Clin Epidemiol 1998;51:237-44.
23. Harlow SD, Linet MS. Agreement between questionnaire data and medical records. Am J Epidemiol 1989;129:233-48.
24. Emberton M, Challands A, Styles RA, Wightman JAK, Black N. Recollected versus contemporary patient reports of pre-operative symptoms in men undergoing transurethral prostatic resection for benign disease. J Clin Epidemiol 1995;48:749-56.
25. Revicki DA, Irwin D, Reblando J, Simon G. The accuracy of self-reported disability days. Med Care 1994;32:401-4.
26. Guadagnoli E, Cleary PD. How consistent is patient-reported pre-admission health status when collected during and after hospital stay. Med Care 1995;33:106-12.