A RAND NOTE

Obtaining Clinical Data on the Appropriateness of Medical Care in Community Practice

Jacqueline Kosecoff, Mark R. Chassin, Arlene Fink,
Mary-Frances Flynn, Lois McCloskey,
Barbara J. Genovese, Carole Oken,
David H. Solomon, Robert H. Brook

November 1987

RAND

The research described in this report was sponsored by the Commonwealth Fund, the John A. Hartford Foundation, the U.S. Health Care Financing Administration, U.S. Department of Health and Human Services, the Pew Memorial Trust, and the Robert Wood Johnson Foundation.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

A RAND NOTE

N-2750-CWF/HF/HCFA/PMT/RWJ

Obtaining Clinical Data on the Appropriateness of Medical Care in Community Practice

Jacqueline Kosecoff, Mark R. Chassin, Arlene Fink,
Mary-Frances Flynn, Lois McCloskey,
Barbara J. Genovese, Carole Oken,
David H. Solomon, Robert H. Brook

November 1987

Prepared for
The Commonwealth Fund
The John A. Hartford Foundation
The U.S. Health Care Financing Administration,
U.S. Department of Health and Human Services
The Pew Memorial Trust
The Robert Wood Johnson Foundation

RAND

Obtaining Clinical Data on the Appropriateness of Medical Care in Community Practice

Jacqueline Kosecoff, PhD; Mark R. Chassin, MD, MPP, MPH; Arlene Fink, PhD; Mary-Frances Flynn, MPH; Lois McCloskey, MPH; Barbara J. Genovese; Carole Oken, MA; David H. Solomon, MD; Robert H. Brook, MD, ScD

We sought the voluntary cooperation of a randomly selected sample of community physicians and hospitals in five states for a study of how appropriately they performed coronary angiography, carotid endarterectomy, and upper gastrointestinal tract endoscopy. Ninety percent of 913 sampled physicians (n = 819) consented to a review of up to 20 of their 1981 Medicare patients' records. These physicians represented seven different specialties and subspecialties and performed 4988 procedures, 92% of the desired sample. Only three of 230 hospitals did not participate. We attribute our method's success primarily to the formation of a network to connect the branches of the profession, respect for office and hospital practice routine, confidentiality, and the development of carefully designed medical record abstraction systems. We conclude that, with effort, cooperative research among disparate segments of the medical community can become a reality even if the topic studied is relatively sensitive.
(*JAMA* 1987;258:2538-2542)

CONTAINING the rising costs of health care has become a major national concern. Most methods of curtailment have an economic foundation and focus on controlling costs by prepaying hospitals, changing health insurance copayments and benefits, or setting limits on physicians' fees. These methods differ in form, but share a common characteristic: they are designed to reduce the use of health care services, but are not structured to reduce selectively the use of inappropriate or ineffective services while preserving the provision of appropriate and effective health care.

A different method of cost containment has a clinical foundation, and is based on the premise that the use of inappropriate services can be selectively reduced.[1] Proponents of this method maintain that quality of care can be enhanced by eliminating the adverse effect of inappropriate health care services while controlling costs. Use of this approach, however, requires the active participation of physicians and hospitals to provide the clinical information necessary to discriminate appropriate from inappropriate services. In this article, we describe how we obtained the voluntary participation of physicians and hospitals and gained access to medical records to study the relationship of appropriateness to geographic differences in the use of three common procedures.

See also pp 2533, 2543, and 2568.

To judge whether procedures are used appropriately, we first developed standards of appropriateness using a method that has been previously described.[2-4] It produces a clinically detailed set of indications for performing a procedure, with each indication rated on a nine-point scale of appropriateness, with 1 being extremely inappropriate and 9 extremely appropriate. Once the appropriateness standards were developed, we faced the challenge of obtaining clinical data to which to apply them. Acquiring these data represented a formidable methodologic task. We proposed to identify individual cases from insurance claims data and, after obtaining the consent of the relevant physicians, to abstract the corresponding medical records. The success of this strategy depended on obtaining positive responses to the following three questions.

First, are the data contained in Medicare part B insurance claims files accurate (ie, if a procedure is listed as performed by a given physician on a given date, could we find evidence in the patient's record that it actually was performed)? Although a formal reliability study of Medicare physician claims has not been done, a previous study of the

From the Health Program, The Rand Corp (Drs Chassin and Brook, and Ms Oken), and Fink and Kosecoff Inc (Drs Kosecoff and Fink, and Mss Flynn, McCloskey, and Genovese), Santa Monica, Calif; and the Departments of Medicine (Drs Kosecoff, Fink, Solomon, and Brook) and Public Health (Drs Kosecoff, Fink, and Brook), University of California at Los Angeles.

The opinions, conclusions, and proposals in the text are those of the authors alone and do not necessarily represent the views of any of these organizations, The Rand Corp, or the University of California.

Reprint requests to The Rand Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Chassin).

reliability of procedures listed on Medicare hospital claims suggested a low level of accuracy.[5]

Second, would community physicians representing many different specialties and hospitals permit us access to the medical records of their patients as part of a study of the appropriateness of the use of medical and surgical procedures? We anticipated that physicians might be reluctant to permit their patient records to be reviewed. In a national study of internal medicine and its specialties dealing with a less sensitive subject, eg, the overall response rate was 65%.[6] Also, although not much information is available about hospital participation in large studies, at least one recent investigation found that of 26 hospitals contacted for permission to review records, 18 (69%) granted it.[7]

Third, could sufficiently detailed clinical data be validly abstracted from medical records to determine specific indications for and the appropriateness of performing procedures? Clinical information about why a procedure was performed could be obtained by interviewing physicians or reviewing the contents of their patients' medical records. We compared the methods; they produced comparable data.[8] Reviewing medical records, however, has several advantages. It is economical, it is less intrusive, and it can be used retrospectively. Thus, we chose to use medical record review.

METHODS

Because of the importance to the study of a high physician participation rate and because we anticipated some difficulty in achieving it, we regarded obtaining voluntary participation of community physicians as a crucial study task. To ensure cooperation, we organized a network of physicians who, although not connected directly to our study, were asked to explain it on our behalf. The network's guiding principle and organizing feature was that physicians and hospitals should participate in studies of this kind only after they have been personally told about its purposes and methods by someone they know and whose credibility is unassailable.

The Policy Advisory Board

We organized a policy advisory board for the study. Its purpose was to evaluate the study's research design and data collection methods and to recommend physicians for participation in other phases of the study. Of 19 members, ten represented organizations such as the American Medical Association or the American Hospital Association. The board met three times: at the beginning of the study, just before the data collection began, and after it was completed. With the board's assistance, the study received 14 formal letters of support from organizations such as the American Medical Association, the American Gastroenterological Association, four local affiliates of the American Health Association, two vascular surgery societies, the American Hospital Association, and the hospital associations in each of the five states where data were to be collected. These were used to help elicit the cooperation of community physicians and hospitals.

The Steering Committees

The steering committees' responsibility was to help the research team select liaison physicians, and then to recruit their participation into the study. One steering committee was assembled for each combination of geographic area and procedure. Each consisted of five to eight physicians, for a total of 64 physicians. We identified steering committee members with the aid of the policy advisory board and through the involvement of one of the principal investigators who was formerly chairman of a department of medicine and a past president of the Association of Professors of Medicine. Steering committee members were chosen who were prominent in academic or private practice in the geographic area containing the sample of community physicians.

Almost all steering committee members were formally oriented to the study's purpose and needs in a two-hour group training session conducted in each study area; the remaining few were trained individually. At the conclusion of the training session, each committee identified potential liaison physicians from our lists of physicians sampled for inclusion in the study. Liaison physicians were selected because they were locally respected, believed to be supportive of the study's goals, and persuasive with their peers. Many steering committee members agreed to serve also as liaison physicians.

Liaison Physicians

The task of the 131 liaison physicians was to contact the sample of community physicians ("performing physicians," ie, the physicians who actually performed the study procedures) and obtain their agreement to participate in the study. As with the steering committee, we conducted training sessions to acquaint liaison physicians with their function in the study. The majority were formally trained for two hours in groups of ten to 12; some were trained individually in their offices. The liaison physicians' specific tasks were to describe the purpose of the study and its sources of support, and to attempt to convince performing physicians in the sample to allow us to review a sample of their patients' medical records in their offices or in hospitals. The liaison physicians also described our confidentiality procedures, our methods of obtaining hospital records, and our policy of remaining as unobtrusive as possible when reviewing records. More than half of the liaison physicians had to contact five or fewer physicians, 6% contacted 11 or more; almost 60% completed their work within six weeks, and 21% required more than 13 weeks.

Community Physicians

Our study was concerned with detecting differences among five different geographic areas in the appropriateness of using three procedures. Using a complete file of physicians' claims obtained from Medicare, Part B, insurance carriers in each of the five areas, we identified 1303 billing entities, defined as individual physicians or groups who billed Medicare in 1981 for performing one of the three study procedures. Our sampling methods were designed so that all patients had an equal probability of being included. In some geographic areas, we included all billing entities (ie, a 100% sample); in others, we selected a random, stratified sample so that billing entities performing a greater number of procedures had a higher probability of being included than those who claimed fewer. There were 795 billing entities in our sample, representing 913 physicians and 5411 patients. These community physicians practiced seven different specialties or subspecialties: family practice, internal medicine, cardiology, gastroenterology, general surgery, vascular surgery, and neurosurgery. They were asked to provide access to the office or hospital records of their Medicare patients receiving the study procedures in 1981. However, these physicians were not in any way directly involved in data collection. To review office records, we worked with a receptionist, nurse, or other designated person. To obtain hospital records for abstraction, we asked physicians to sign a letter (usually to the director of medical records) granting permission to our study staff to review the documents at the hospital.

Physicians and hospitals were guaranteed that we would meet the human subjects protection standards established by the US Department of Health and Human Services and guidelines for release of medical information established by the American Hospital Associ-

ation. All data collectors were required to leave a signed oath of confidentiality in each medical record they reviewed whether in an office or hospital, and to blacken out all identifying information on patients, physicians, hospitals, and laboratories from any material that they photocopied.

Hospitals

We sent each hospital descriptive information about the study and copies of letters from participating physicians requesting access to their patients' medical records. We also included the study's letters of support from the American Hospital Association and its local affiliates. Hospitals were requested to have available between five and ten records for each day of abstraction. Our data collectors were trained to find an unobtrusive place away from the routine hospital activity and to do all work without any assistance from the hospital staff.

Data Collection

After physicians agreed to participate, a member of the research team contacted each physician's office to verify that each patient listed in the claims data we sampled had undergone a procedure performed by that physician. For procedures performed in hospitals, we also verified the institution at which they were done.

We developed separate medical record abstraction forms, one each for coronary angiography, upper gastrointestinal (GI) tract endoscopy, and carotid endarterectomy. Each abstraction form had a set of guidelines specifying how to use the forms, giving synonyms for medical terms and suggesting locations in medical records where information could typically be found. These forms and guidelines have been published elsewhere.[9-11]

The majority of items in the abstraction forms were designed to measure (in both precoded and descriptive form) the indication or indications for which a procedure was performed and to assess any comorbid conditions that might affect the appropriateness of performing the procedure. The actual organization of items into an abstraction form was designed to make the abstraction process as efficient as possible and, therefore, followed the logic of how information is typically stored in a medical record. About one hour was needed to complete an abstraction. In addition, we obtained photocopies of selected test reports, such as carotid angiograms and exercise treadmills.

Fifty-four persons with a background in medical record abstraction, nursing, or utilization review were trained to collect the data. All had prior medical record review experience, passed an initial test of abstraction skills, were given 2½ days of intensive training per procedure, and successfully completed a further test at the end of training. In each of the five sites, work was supervised by a chief data collector who regularly reabstracted records to maintain quality control, as well as by research staff who operated a telephone hotline and periodically published bulletins updating all data collectors. Data collection was carried out in hospitals, physicians' offices, and clinics as necessary to obtain complete clinical information on each case.

Completed abstraction forms were reviewed by both a physician and a nonphysician, who assessed completeness, internal consistency, and whether the coding decisions were consistent with the descriptive data offered to justify them. In addition, a physician read the photocopies of complex diagnostic studies and coded their results. Abstraction forms with missing or inconsistent data were returned for reabstraction or were supplemented by additional data from different sources (eg, the office records of the sampled or referring physicians).

Using a computerized method of analysis, data elements recorded in the abstraction forms were compared with the catalog of indications developed by the expert panels. During this process one or more indications was assigned to each case. For cases with more than one indication, the one with the highest appropriateness rating was designated the primary indication and used in subsequent analyses.

To test the validity of indications assigned by the above method, we selected 132 cases at random, representing one site each for coronary angiography and upper GI tract endoscopy, and two sites for carotid endarterectomy. Of the selected cases, 123 medical records (93%) were available for review. Without knowledge of the indication assigned by our method, a physician examined the entire medical record and recorded a conclusion as to the indication for the procedure under study. The physician then compared this implicit assignment with the results of our analysis and resolved any discrepancies.

RESULTS

Physician Response Rate

Ninety percent of a randomly selected group of physicians (819/913) participated in the research and allowed us to review up to 20 of their patients' medical records (Table). There were no significant differences in physician participation rates among procedures or sites ($P>.05$, χ^2). Of the 913 sampled physicians, 6% declined to participate, while 4% could not be located by liaison physicians or research staff. The participating physicians performed 4988 procedures, 92% of the desired sample.

Accuracy of Claims Data

In constructing our sample of patients and physicians, we noticed that a few patients in one site were listed in the claims data as having received up to ten carotid endarterectomies in 1981. On further analysis we found that four physicians had submitted claims using the procedure code for carotid endarterectomy but listing the procedure as having been performed in the office and usually charging $50 per procedure. After consulting with the local insurance carrier, we determined that these procedures were almost certainly not carotid endarterectomies but rather chelation treatments, a procedure not covered by Medicare. We excluded these physicians, who accounted for 139 procedures, from the claims data from which we sampled true carotid endarterectomies.

Of the 4988 potential cases (those listed in the claims data as having been performed by participating physicians), 239 (4.8%) were discovered during our verification process to represent errors in the claims data. The sources of these errors were varied. We ascertained that the procedure actually performed was different from the one listed on the claim in 119 cases (2.4%). The physician listed on the claim had no record of having treated the patient listed in 82 cases (1.6%). The procedure was not performed in 1981 in 23 cases (0.5%). The physician listed on the claim did not ever perform the listed procedure in 13 cases (0.3%), and the procedure was not performed within the specified geographic area in two cases (0.04%).

Hospital Response Rate

We reviewed medical records in 227 hospitals. Only three hospitals (1%) refused us access to their records. At least half of the medical records departments at participating hospitals agreed to cooperate immediately on receiving the physicians' letters granting us permission to review their patients' records; the rest had to obtain internal administrative support first.

Availability of Data From Medical Records

Of the remaining 4749 cases eligible to be abstracted and entered into our

Physician and Patient Participation Rates by Geographic Site and Procedure

Site No. and Procedure	Sample		Physician Participation No. (%)	Cases Represented by Participating Physicians No. (%)
	No. of Physicians	No. of Patients		
Site 1				
Angiography	161	747	143 (89)	671 (90)
Endoscopy	141	758	127 (90)	671 (89)
Endarterectomy	140	748	128 (91)	658 (88)
Subtotal	**442**	**2253**	**398 (90)**	**2000 (89)**
Site 2				
Angiography	52	574	50 (96)	559 (97)
Endoscopy	132	579	115 (87)	523 (90)
Endarterectomy	79	242	71 (90)	224 (93)
Subtotal	**263**	**1395**	**236 (90)**	**1306 (94)**
Site 3				
Angiography	28	581	26 (93)	580 (100)
Site 4				
Endarterectomy	67	591	57 (85)	532 (90)
Site 5				
Endoscopy	113	591	102 (90)	570 (96)
Total, All Sites	**913**	**5411**	**819 (90)**	**4988 (92)**

database, 143 medical records could not be found by participating hospitals, 24 could not be obtained because hospitals refused to participate, and for 18 cases we could not find sufficient information to determine the indication(s) for the procedure. Thus, we collected complete data for 4564 cases or 96% of those eligible.

Validity

Of the 123 patient records used in the validity analysis, there were 13 (11%) errors in abstraction or analysis leading to disagreement with indications implicitly assigned by the physician member of the research team. Of the 13 errors, nine resulted in changes in appropriateness scores of one point or less on the nine-point scale; seven resulted in higher scores, two in lower scores, and four in no change. Of the 123 cases, 121 were correctly classified as appropriate (ie, median ratings of 7 to 9 by both methods), equivocal (ratings of 4 to 6 by both methods), or inappropriate (ratings of 1 to 3 by both methods). Two cases were incorrectly classified, both as appropriate. Thus, virtually all errors resulted in higher appropriateness ratings than resulted from the physician directly reading the medical record.

COMMENT

We obtained the cooperation of 90% of physicians and 99% of hospitals in a study to examine the relationship between the rates of use of selected medical and surgical procedures and the appropriateness of their use. Over 800 randomly selected physicians of diverse specialties and subspecialties granted us permission to review their patients' records in their offices and 227 hospitals. The physicians and hospitals who participated in the research were located in five widely dispersed states. We found no differences in physician participation rates among sites or procedures. Furthermore, because physicians participated at an overall rate of 90%, and these physicians represented 92% of cases, we know that high-volume physicians did not decline to participate in disproportionate numbers. The most often stated reason for nonparticipation by the 6% of physicians who declined was an unwillingness to allow their records to be reviewed without individual patient consent. The most frequent reason for our inability to contact the remaining 4% of physicians was that they had moved from the area and could not be traced. Physician claims data proved to be an accurate source from which to identify samples of patients and physicians.

Office and hospital medical records proved to contain complete enough clinical data for us to determine the indications for the three procedures we studied; only 18 cases (0.4%) could not be analyzed because data were incomplete. Our study was not designed to assess the relative completeness of hospital compared with office medical records. We can, however, make a few nonquantitative observations. Most often, the hospital medical records were complete enough to determine indications. For a significant minority of cases, perhaps 20%, at least one important piece of information was missing from the hospital record. Usually the missing data were results of tests performed prior to the procedure under study (eg, an exercise treadmill, an upper GI tract roentgenographic series, or a carotid arteriogram). Usually, the office records of the physician who performed the procedure contained the missing data. Rarely, we obtained the necessary information from the primary care physician who referred the patient for the procedure.

We believe several factors contributed to the high rates of physician and hospital participation we observed. We did not directly approach the practitioners whose patients' records were necessary for our sample; instead, we relied on other physicians who knew the practitioners and whom we had previously trained to understand and explain the study's goals and methods. We made every attempt to fit our study's needs into the practice routine; for example, we required little or no effort from community physicians or hospital staff. We trained our data collectors to work unobtrusively and to respect the confidentiality of information in a medical record. We were aware also that no matter how quiet and understanding we were, the presence of outsiders would cause disruption, and so we provided modest financial compensation to participating physicians. Physicians with specific tasks were oriented formally to the study's purposes and needs; at the orientation, we tried to anticipate questions and provide written explanatory materials for physicians and hospitals.

We also demonstrated that Medicare Part B physician claims data are valid sources of information about health care services. They identify with a high degree of accuracy specific procedures performed by specific physicians for specific patients. They therefore can be used to select a sample frame for studies of specific health care services. Of the errors we found in the claims data, the most frequent was a discrepancy between the procedure listed on the claim and the procedure we found had actually been performed. In most instances the discrepancy was minor. For example, a cardiac catheterization with an aortogram appeared as a catheterization with coronary angiography. We did not find a systematic pattern in which more complex or expensive procedures were claimed than were actually performed. The most likely explanation for this high degree of accuracy is that procedures listed on physician claims are usually coded directly by the physician who performed them or by office personnel intimately familiar with the relevant procedure codes. In contrast, procedure coding on Medicare hospital claims, which was found in the 1970s not to be highly accurate,[5] was often performed by personnel not directly associated with the hospitals where the care was rendered. Nor, until recently, did the data serve as the basis for payment.

The converse of our finding that the claims data are accurate is also impor-

tant, namely, that fraudulent billing, at least in regard to the three procedures studied here, is infrequent. We did find that four physicians misused the procedure code for carotid endarterectomy in one site. We suspect that they were performing chelation therapy although we did not directly verify this suspicion by medical record review.

These physicians represented 0.3% of all billing entities in the universe from which we sampled. They accounted for 0.7% of procedures in this universe. Although we clearly did not perform a financial audit, this is the first large-scale study in which Medicare physician claims have been compared with medical record data. Over 95% of the time we found documentation in medical records that the procedure listed on the claim was performed by the physician listed.

Finally, this research has shown that a clinically detailed medical record abstraction system that is largely dependent on nonphysician abstractors can produce valid information about appropriateness. Contributing to this result was redundancy, in that we required written justification for certain statements on the forms, and extensive training and monitoring of data collectors.

By participating in this study, physicians have helped provide new and essential information on the appropriateness of care given in large areas of this country. This is a necessary component of a clinical, medical record–based approach to cost containment, which focuses on preserving beneficial services and advocates eliminating inappropriate ones. Based on our experience, we are convinced that, with effort, cooperative research on relatively controversial topics using clinically detailed medical record review systems can be done.

This work was supported by the Commonwealth Fund, New York, the John A. Hartford Foundation, New York, the Health Care Financing Administration of the US Department of Health and Human Services, Washington, DC, the Pew Memorial Trust, Philadelphia, and The Robert Wood Johnson Foundation, Princeton, NJ.

The authors are indebted to the many organizations and individuals whose cooperation has been essential to the success of this study. In particular, we are most grateful to our policy advisory board, steering committees, liaison physicians, and the hundreds of physicians and hospitals who participated.

References

1. Brook RH, Lohr KN, Chassin MR, et al: Geographic variations in the use of services: Do they have any clinical significance? *Health Aff* 1984;3:63-73.

2. Park RE, Fink A, Brook RH, et al: Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health* 1986;76:766-772.

3. Park RE, Fink A, Brook RH, et al: *Physician Ratings of Appropriate Indications for Six Medical and Surgical Procedures*, publication R-3280-CWF-HF-PMT-RWJ. Santa Monica, Calif, The Rand Corp, 1986.

4. Chassin MR, Kosecoff J, Park RE, et al: Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA* 1987;258:2533-2537.

5. Institute of Medicine: *Reliability of Medicare Hospital Discharge Records*. Washington, DC, National Academy of Sciences, 1977.

6. Mendenhall RC, Tarlov AR, Girard RA, et al: A national study of internal medicine and its specialties: II. Primary care in internal medicine. *Ann Intern Med* 1979;91:275-287.

7. Savitz DA, Grace C: Determinants of medical record access for an epidemiologic study. *Am J Public Health* 1985;75:1425-1426.

8. Kosecoff J, Fink A, Brook RH, et al: The appropriateness of using a medical procedure: Is information in the medical record valid? *Med Care* 1987;25:196-201.

9. Chassin MR, Kosecoff J, Park RE, et al: *Indications for Selected Medical and Surgical Procedures: A Literature Review and Ratings of Appropriateness: Coronary Angiography*, publication R-3204/1-CWF-HF-HCFA-PMT-RWJ. Santa Monica, Calif, The Rand Corp, 1986.

10. Merrick NJ, Fink A, Brook RH, et al: *Indications for Selected Medical and Surgical Procedures: A Literature Review and Ratings of Appropriateness: Carotid Endarterectomy*, publication R-3204/6-CWF-HF-HCFA-PMT-RWJ. Santa Monica, Calif, The Rand Corp, 1986.

11. Kahn KL, Roth CP, Kosecoff J, et al: *Indications for Selected Medical and Surgical Procedures: A Literature Review and Ratings of Appropriateness: Diagnostic Upper Gastrointestinal Endoscopy*, publication R-3204/4-CWF-HF-HCFA-PMT-RWJ. Santa Monica, Calif, The Rand Corp, 1986.