

Physician Ratings of Appropriate Indications for Six Medical and Surgical Procedures

ROLLA EDWARD PARK, PHD, ARLENE FINK, PHD, ROBERT H. BROOK, MD, ScD, MARK R. CHASSIN, MD, KATHERINE L. KAHN, MD, NANCY J. MERRICK, MD, JACQUELINE KOSECOFF, PHD, AND DAVID H. SOLOMON, MD

Abstract: We convened three panels of physicians to rate the appropriateness of a large number of indications for performing a total of six medical and surgical procedures. The panels followed a modified Delphi process. Panelists separately assigned initial ratings, then met in Santa Monica, California where they received reports showing their initial ratings and the distribution of the other panelists' ratings. They discussed the indications and revised the indications

lists, then individually assigned final ratings. There was generally better agreement on the final ratings than on the initial ratings. Based on reasonable criteria for agreement and disagreement, and excluding one outlying procedure, the panelists agreed on ratings for 42 to 56 per cent of the indications, and disagreed on 11 to 29 per cent. (*Am J Public Health* 1986; 76:766-772.)

Introduction

Physicians today face mounting pressures to use procedures only when clinically valid criteria indicate that they are appropriate. The government's limits on health care financing,¹ businesses' efforts to curb health insurance costs,² increasing concern over large regional variations in procedure use,³ and the establishment of peer review organizations (PROs)⁴ all add to these pressures.

Where can physicians turn for help in deciding what procedure use is generally accepted to be appropriate? The medical literature is only a partial guide; expensive, time-consuming, randomized, clinical trials can never investigate the effectiveness of procedures in more than a small fraction of the situations in which they might be used.⁵ The National Institutes of Health (NIH) consensus statements offer some guidance for some procedures, but are generally too broadly stated to be of much help in individual cases.⁶

We report here the results of a new method for rating the appropriateness of a large number of detailed indications for the use of individual medical or surgical procedures. The method relies on a panel of expert physicians to generate and to rate the appropriateness of a list of indications. We convened three panels of physicians to rate the appropriateness of indications for performing a total of six procedures—coronary angiography, coronary artery bypass graft surgery, cholecystectomy, upper gastrointestinal endoscopy, colonoscopy, and carotid endarterectomy. We selected these six procedures because they are frequently performed, they use substantial medical resources, and they exhibit significant variation in use rates between large geographical areas of the United States.³

In this paper, we describe the composition and working of the panels; summarize the ratings, with particular emphasis on the amount of agreement or disagreement among the panelists; and discuss the implications of these results.

Address reprint requests to Rolla Edward Park, PhD, Economics Department, Rand Corporation, 1700 Main Street, Santa Monica, CA 90406-2138. Drs. Fink and Kosecoff are with Fink & Kosecoff, Inc.; Dr. Brook is with the Systems Sciences Department, the Rand Corporation, and the Departments of Medicine and Public Health, UCLA; Dr. Chassin is with the Systems Sciences Department, the Rand Corporation; Dr. Kahn is with the Systems Sciences Department, the Rand Corporation, and the Department of Medicine, UCLA; Dr. Merrick is with the Department of Medicine, UCLA, and the VA Medical Center, Sepulveda; Dr. Solomon is with the Department of Medicine, UCLA. This paper, submitted to the *Journal* October 14, 1985, was revised and accepted for publication February 26, 1986.

© 1986 American Journal of Public Health 0090-0036/86\$1.50

Methods

A cardiovascular panel rated indications for coronary angiography and coronary artery bypass graft surgery; a gastrointestinal panel rated indications for cholecystectomy, upper gastrointestinal endoscopy, and colonoscopy; a cerebrovascular panel rated indications for carotid endarterectomy. Each of the three panels was conducted as follows:

- Project staff invited nine distinguished physicians to serve on the panel.
- Staff prepared initial lists of clinical indications for performing each procedure.
- The panelists rated the appropriateness of each indication and mailed their ratings to Santa Monica, where staff tabulated them for use at the panel meeting.
- The panelists met in Santa Monica. Taking one procedure at a time, they discussed the indications, revised the indications list, and individually assigned final appropriateness ratings.
- Staff analyzed the ratings.

In selecting panelists, we attempted to keep a rough balance among different parts of the country, different relevant specialties, and academic versus private practice. Each panel had members from all four US census regions. The specialty composition of the panels is shown in Appendix A.

In developing the initial indication lists, we used as guides reviews of the medical literature on each procedure. The indications categorized patients in terms of their symptoms, past medical history, and the results of previous diagnostic tests. We tried to make the indications detailed enough so that patients presenting with a particular indication would be reasonably homogeneous, in the sense that doing the procedure would be equally appropriate (or inappropriate) for all of them. We tried to make the lists comprehensive enough so that all indications for doing the procedure that arise in practice would be included. At the same time, we tried to keep them short enough so that all of the indications could be rated by the panelists within a day. The indications were organized into "chapters" which corresponded to clinically related problems for which the procedure could be used. The total number of initial indications for each procedure ranged from 192 for cholecystectomy to 1,685 for endoscopy.

As an example, Appendix B lists the eight clinical settings (chapters) for coronary angiography. We defined indications for the use of coronary angiography in asymptomatic

matic patients, patients with chest pain of uncertain origin, and patients in six other categories. Under each chapter title are listed the factors that we used to define the indications in that clinical setting; each factor may appear at several levels and in combination with different levels of the other factors. In chapter 3 (patients with chronic stable angina), for example, contraindications to coronary artery bypass graft surgery include dementia, severely impaired one-year life expectancy, and severe, irreversible functional incapacity. Medical management is classified as none, less than maximal, or maximal. Angina is class I, II, III, or IV according to the Canadian Cardiovascular Society classification. Test results may be negative, positive, or very positive, or the test may not have been performed. We used various combinations of these factors to define the indications for coronary angiography in patients with chronic stable angina.

Comorbidity was accounted for in the indications for all of the procedures. For the cardiovascular procedures, comorbid conditions figured in the definition of contraindications to bypass graft surgery. For the gastrointestinal and cerebrovascular procedures, the panelists explicitly assigned adjustments to their ratings to account for varying levels of comorbidity.

We sent the panelists literature reviews,⁷⁻¹² rating sheets, and instructions. The literature reviews gave all panelists equal access to a central core of relevant literature. The rating sheets listed all of the indications for each procedure and provided space for an appropriateness rating on a scale of 1 to 9. Figure 1 shows a one-page example from the initial rating sheet for coronary angiography. The panelists were instructed to rate the appropriateness of each indication as of 1981 using their own best clinical judgment and considering an average group of patients presenting to an average US physician who performed the procedure during 1981. "Appropriate" was defined to mean that the expected

health benefit (i.e., increased life expectancy, relief of pain, reduction in anxiety, improved functional capacity) exceeded the expected negative consequences (i.e., mortality, morbidity, anxiety of anticipating the procedure, pain produced by the procedure, time lost from work) by a sufficiently wide margin that the procedure was worth doing. "Inappropriate" meant the opposite—the negative consequences outweighed the expected benefits. Extremely appropriate indications should be rated 9, equivocal indications should be rated 5, and extremely inappropriate indications should be rated 1. The instructions also included definitions of important terms. The panelists required about 20 to 35 minutes per 100 indications to complete the initial ratings.

The panels all met in Santa Monica and, on average, spent one day discussing and rerating each procedure. Because of a time constraint, however, rerating of colonoscopy was done by mail without group discussion two months after the gastrointestinal panel meeting.

The panelists discussed the indications for each procedure one chapter at a time. They had in front of them computer printouts that distributed their initial ratings of each item (Figure 2). Each panelist received the identical distribution of ratings, but a caret below the rating line showed the particular panelist's own initial rating.

For all procedures except colonoscopy (which was discussed only briefly at the panel meeting), the indication lists were substantially revised during discussions at the meetings. The changes were meant to tailor the indications to better describe clinically homogeneous categories. Many of the changes simply split one indication into two. Other changes adjusted the boundaries between indications. In other cases, some indications were dropped and others were added. Some initial indications were divided into multiple final indications. Other groups of initial indications were merged into a single final indication. Many of the changes were even more complex transformations. In some cases, whole chapters were divided, combined, or eliminated.

For example, in coronary angiography chapter 3 (chronic stable angina), the panelists split some of the indications by age. After the split, all of the indications for patients with Class I or II angina were rated separately for patients under age 65 and for patients 65 or older.

The printouts showing distributions of initial ratings also served as rating sheets. After discussion of each chapter, the panelists marked their final ratings directly on the printouts. Although the panelists were not required to rate all of the indications, all of them chose to rate all but a tiny fraction of the indications.

Studies are now underway to assess the validity and reliability of the ratings.

Results

Four Definitions of Agreement

We propose two "conceptions" of agreement. Both can be applied using either nine ratings or seven ratings, yielding a total of four definitions.

The stricter conception is: The raters agreed if all of the ratings were within a single three 3-point region—1 to 3, 4 to 6, or 7 to 9. We interpret this to mean that all of the raters agreed to one of the following statements: the procedure should not be done, doing it is questionable, or it should be done.

The second conception is somewhat more relaxed: The raters agreed if all of the ratings were within any 3-point range, even if that range straddles the boundary between two

INDICATIONS FOR CORONARY ANGIOGRAPHY		CHAPTER 1
<div style="border: 1px dashed black; padding: 5px;"> <p style="text-align: center;">Appropriateness Scale</p> <p style="text-align: center;">1 2 3 4 5 6 7 8 9</p> <p>1 = extremely inappropriate</p> <p>5 = equivocal (neither clearly appropriate nor clearly inappropriate)</p> <p>9 = extremely appropriate</p> </div>		Rating of Appropriateness (Circle One)
I. Asymptomatic Patients		
A. Coronary angiography (CA) is indicated in patients in high risk occupations if:		
1. No exercise ECG, no exercise thallium scan, and no exercise MUGA		1 2 3 4 5 6 7 8 9
2. Negative exercise ECG and		1 2 3 4 5 6 7 8 9
a. No or negative exercise thallium scan regardless of MUGA results, if any		1 2 3 4 5 6 7 8 9
b. Reversible defect on exercise thallium scan and		1 2 3 4 5 6 7 8 9
(i) No exercise MUGA		1 2 3 4 5 6 7 8 9
(ii) Negative exercise MUGA		1 2 3 4 5 6 7 8 9
(iii) Positive exercise MUGA		1 2 3 4 5 6 7 8 9
3. Positive exercise ECG and		1 2 3 4 5 6 7 8 9
a. No exercise thallium scan regardless of MUGA results, if any		1 2 3 4 5 6 7 8 9
b. Negative exercise thallium scan and		1 2 3 4 5 6 7 8 9
(i) No exercise MUGA		1 2 3 4 5 6 7 8 9
(ii) Negative exercise MUGA		1 2 3 4 5 6 7 8 9

FIGURE 1—Example of Form Used for Initial Ratings

CLINICAL PRESENTATION: 1. CAROTID TIA and/or AMAUROSIS FUGAX - Single Episode			
	Low Surgical Risk	Elevated Surgical Risk	High Surgical Risk
APPROPRIATENESS OF OPERATING IPSILATERALLY IF ANGIOGRAPHY SHOWS:			
Ipsi: Degree of stenosis of ipsilateral artery Contra: Degree of stenosis of contralateral artery			
1. Ipsi: Occluded Contra: None or 1-49%	9 1 2 3 4 5 6 7 8 9	9 1 2 3 4 5 6 7 8 9	9 1 2 3 4 5 6 7 8 9
2. Ipsi: Occluded Contra: 50-99%	9 1 2 3 4 5 6 7 8 9	9 1 2 3 4 5 6 7 8 9	9 1 2 3 4 5 6 7 8 9
3. Ipsi: 50-99% Contra: None, 1-49%, or 50-99%	1 1 2 1 4 1 2 3 4 5 6 7 8 9	1 1 1 3 1 2 1 2 3 4 5 6 7 8 9	4 2 1 1 1 1 2 3 4 5 6 7 8 9
4. Ipsi: 50-99% Contra: Occluded	1 1 3 4 1 2 3 4 5 6 7 8 9	1 1 1 2 1 3 1 2 3 4 5 6 7 8 9	1 4 1 1 1 1 1 2 3 4 5 6 7 8 9
5. Ipsi: 1-49% Contra: None or 1-49%	4 1 2 1 1 1 2 3 4 5 6 7 8 9	6 1 1 1 1 2 3 4 5 6 7 8 9	8 1 1 2 3 4 5 6 7 8 9
6. Ipsi: 1-49% Contra: 50-99%	3 1 1 1 1 1 1 1 2 3 4 5 6 7 8 9	5 2 1 1 1 2 3 4 5 6 7 8 9	7 1 1 1 2 3 4 5 6 7 8 9
7. Ipsi: 1-49% Contra: Occluded	3 1 1 2 1 1 1 2 3 4 5 6 7 8 9	5 3 1 1 2 3 4 5 6 7 8 9	8 1 1 2 3 4 5 6 7 8 9
8. Ipsi: 1-49% with large ulcerative lesion Contra: None, 1-49% or 50-99%	1 1 1 3 3 1 2 3 4 5 6 7 8 9	1 1 1 4 2 1 2 3 4 5 6 7 8 9	4 1 1 2 1 1 2 3 4 5 6 7 8 9
9. Ipsi: 1-49% w/multicentric ulcerative lesion Contra: None, 1-49%, or 50-99%	1 1 1 1 5 1 2 3 4 5 6 7 8 9	1 1 1 2 2 2 1 2 3 4 5 6 7 8 9	4 1 1 2 1 1 2 3 4 5 6 7 8 9

FIGURE 2—Example of Form Used for Final Ratings

of the regions specified above. If all of the ratings were within the range 3 to 5, for example, there was agreement according to this second conception, but not according to the first.

Four definitions of agreement result:

- A9S: All nine of the ratings fell within a single 3-point region—1 to 3, 4 to 6, or 7 to 9.
- A9R: All nine of the ratings fell within any 3-point range.
- A7S: After discarding one extreme high and one extreme low rating, the remaining seven ratings all fell within a single 3-point region—1 to 3, 4 to 6, or 7 to 9.
- A7R: After discarding one extreme high and one extreme low rating, the remaining seven ratings all fell within any 3-point range.

Table 1 shows the per cent of indications on which the panelists agreed, using our four definitions of agreement. On any definition, colonoscopy showed much lower agreement than any other procedure. There are several possible reasons for this. Colonoscopy was the only procedure for which final

ratings were assigned without group discussion. It had by far the largest number of final indications, possibly producing less carefully considered ratings. Or there may simply have been a lower level of consensus on appropriate indications for use of colonoscopy than there was for the other procedures, perhaps because it is a relatively new procedure. In any event, since the panel process differed substantially for colonoscopy, our discussion disregards colonoscopy and focuses on the other five procedures.

The per cent of agreement for the final ratings exceeded that for the initial ratings for most of the procedures. The increases in agreement were substantial in magnitude for coronary angiography (34 per cent on definition A7R; 95% confidence interval 26, 42), coronary artery bypass graft surgery (18 per cent; 95% CI: 12, 24), and upper gastrointestinal endoscopy (13 per cent; 95% CI: 9, 17). The changes for the other procedures were smaller or less precisely measured.

Based on the definitions that use all nine ratings (A9S and

TABLE 1—Per Cent of Indications on Which Panelists Agreed Using Four Different Definitions of Agreement (all indications)

Definition of Agreement ^b	Procedure ^a											
	ANGIO		CABS		CH		END		CO		CE	
	I ^c	F ^c	I	F	I	F	I	F	I	F	I	F
A9S: 9 ratings, strict	12.2	28.0	7.8	21.3	29.2	38.8	16.6	25.4	1.8	2.6	39.9	40.9
Standard error	(2.3)	(2.6)	(1.4)	(1.9)	(6.6)	(7.0)	(0.9)	(1.3)	(0.4)	(0.3)	(1.9)	(1.7)
A9R: 9 ratings, relaxed	12.7	28.7	7.8	22.7	29.2	38.8	16.6	25.4	1.8	2.6	39.9	40.9
Standard error	(2.3)	(2.6)	(1.4)	(1.9)	(6.6)	(7.0)	(0.9)	(1.3)	(0.4)	(0.3)	(1.9)	(1.7)
A7S: 7 ratings, strict	21.0	50.0	21.4	34.8	43.8	49.0	27.8	41.3	10.6	14.6	55.6	53.4
Standard error	(2.9)	(2.9)	(2.1)	(2.2)	(7.2)	(7.2)	(1.1)	(1.5)	(0.9)	(0.7)	(1.9)	(1.7)
A7R: 7 ratings, relaxed	22.0	56.3	23.2	41.2	43.8	53.1	28.6	41.6	11.1	14.9	55.6	53.8
Standard error	(2.9)	(2.9)	(2.2)	(2.2)	(7.2)	(7.2)	(1.1)	(1.5)	(1.0)	(0.7)	(1.9)	(1.7)

^aANGIO: coronary angiography
^aCABS: coronary artery bypass surgery
^aCH: cholecystectomy
^aEND: upper gastrointestinal endoscopy
^aCO: colonoscopy
^aCE: carotid endarterectomy
^bSee text for full definitions of agreement.
^cI indicates initial ratings, F indicates final ratings.

A9R), the panelists agreed on the final ratings for 25 to 40 per cent of the indications, depending on the procedure. Using A9R rather than A9S makes surprisingly little difference; almost all of the ratings that clustered in any 3-point range fell within a single one of our regions (1 to 3, 4 to 6, or 7 to 9). If we do not require unanimity (definitions A7S and A7R), agreement jumps by 16 percentage points on average, reaching 50 per cent or better for three of the procedures.

Four Definitions of Disagreement

As with agreement, we propose two conceptions of disagreement that can be used with either nine or seven ratings to yield four definitions: The first conception is: The raters disagreed if at least one assigned a rating of 1 and at least one assigned a rating of 9. We interpret this extreme polarization as disagreement. The second conception is more relaxed: The raters disagreed if at least one rating fell in the lowest 3-point region (1 to 3) and at least one in the highest (7 to 9). We interpret this to mean that one rater thought the

procedure should not be done for this indication and one thought that it should be done.

Four definitions of disagreement result:

- D9S: Considering all nine ratings, at least one was a 1 and at least one was a 9.
- D9R: Considering all nine ratings, at least one fell in the lowest 3-point region (1 to 3) and at least one fell in the highest (7 to 9).
- D7S: After discarding one extreme high and one extreme low rating, at least one of the remaining seven ratings was a 1 and at least one was a 9.
- D7R: After discarding one extreme high and one extreme low rating, at least one of the remaining seven ratings fell in the lowest 3-point region (1 to 3) and at least one fell in the highest (7 to 9).

Table 2 shows the per cent of indications that satisfied each of the four definitions of disagreement. Colonoscopy showed a high level of disagreement by any of the definitions, corresponding to the low level of agreement noted above.

TABLE 2—Per Cent of Indications on Which Panelists Disagreed Using Four Different Definitions of Agreement (all indications)

Definition of Agreement ^b	Procedure ^a											
	ANGIO		CABS		CH		END		CO		CE	
	I ^c	F ^c	I	F	I	F	I	F	I	F	I	F
D9S: 9 ratings, strict	9.3	2.0	20.5	8.8	16.7	20.4	37.9	30.2	49.2	60.9	28.4	14.9
Standard error	(2.0)	(0.8)	(2.1)	(1.3)	(5.4)	(5.8)	(1.2)	(1.4)	(1.5)	(0.9)	(1.7)	(1.2)
D9R: 9 ratings, relaxed	60.0	30.0	59.5	39.5	47.9	42.9	63.8	48.5	82.0	80.8	40.9	34.0
Standard error	(3.4)	(2.7)	(2.6)	(2.2)	(7.3)	(7.1)	(1.2)	(1.5)	(1.2)	(0.7)	(1.9)	(1.6)
D7S: 7 ratings, strict	0.0	0.3	5.4	2.3	0.0	2.0	2.8	7.0	9.9	21.7	4.7	2.3
Standard error	(0.0)	(0.3)	(1.2)	(0.7)	(0.0)	(2.0)	(0.4)	(0.8)	(0.9)	(0.8)	(0.8)	(0.5)
D7R: 7 ratings, relaxed	31.7	11.0	29.7	16.6	33.3	20.4	36.0	28.9	55.2	48.1	21.6	18.1
Standard error	(3.3)	(1.8)	(2.4)	(1.7)	(6.9)	(5.8)	(1.2)	(1.4)	(1.5)	(0.9)	(1.6)	(1.3)

^a ANGIO: coronary angiography
^a CABS: coronary artery bypass surgery
^a CH: cholecystectomy
^a END: upper gastrointestinal endoscopy
^a CO: colonoscopy
^a CE: carotid endarterectomy
^b See text for full definitions of agreement.
^c I indicates initial ratings, F indicates final ratings.

Disagreement decreased substantially between the initial and final ratings for angiography (21 per cent; 95% confidence interval 13, 28), coronary artery bypass graft surgery (13 per cent; 95% CI: 7, 19), and endoscopy (7 per cent; 95% CI: 3, 11), according to definition D7R. There was no consistent change in disagreement for the other procedures.

Disregarding colonoscopy, the percentage of indications with at least one 1 and at least one 9 (D9S) in the final ratings ranged from 2 to 30; depending on the procedure, requiring at least two ratings at each extreme (D7S) reduced these percentages substantially, to a maximum of 7 for endoscopy. Between 30 and 50 per cent of all indications had at least one 1 to 3 rating and at least one 7 to 9 (D9R). Requiring two ratings at each extreme (D7R) reduced the range to between 10 and 30 per cent.

Differences in Ratings by Specialty

Table 3 reports on median ratings by three specialty groupings: medical generalists, medical specialists, and surgeons (see Appendix A for more information on the composition of the specialty groups). The radiologist on each panel constituted a fourth group of one; we do not report his ratings so as not to disclose information on individuals. The radiologists' ratings are included in the overall median.

The ratings are shown as average deviations of the median rating for each specialty from the overall median for all the panelists. Thus, for example, the median initial rating by medical generalists was, on average, 1.49 points lower than the overall median for the initial coronary angiography indications.

The ratings by specialty group were frequently different from the overall ratings, but consistent patterns are hard to discern. Two observations stand out, however: Final ratings by the medical specialists were never more than half a rating point from the overall median, and the surgeons rated all three surgical procedures substantially higher than the overall median.

Discussion

Our panels demonstrated that physicians can rate the appropriateness of large numbers of indications for performing medical and surgical procedures. The lists of indications that they rated were much more detailed and comprehensive than any that have been attempted in the past.

The detailed lists of rated indications⁷⁻¹² may interest some researchers, but they are not a valid guide to current medical practice. We plan to use the ratings to investigate the relation between appropriateness and geographical variations in procedures done during 1981.¹³ Thus we instructed the panelists to rate the indications in light of 1981 knowledge. Rapid development of some treatments (for example, intracoronary streptokinase in heart attacks and balloon angioplasty for obstruction of coronary arteries) has surely altered the appropriateness of some of the indications since 1981.

The panel process clearly brought the individual panelists' ratings closer together. Whether we compare per cent agreement or per cent disagreement, the final ratings tended to be closer together than the initial ratings.

Even in the final ratings, however, the panelists at best agreed on a little over half of the indications, and they clearly disagreed on up to 30 per cent of them. These statements disregard colonoscopy, which had a much higher level of disagreement. Perhaps the appearance reflects reality; there may really be substantial disagreement about the value of these procedures for many indications. Other possibilities include the following:

- Perhaps panelists had in mind different "average groups of patients presenting to an average US physician." This is plausible, since their perceptions of what is an average group presumably depend to some extent on what the panelists see in their own practices.
- Perhaps the indications were not properly framed—for example, they did not match clinically relevant categories, or were not tightly enough defined to avoid the "different group" problem. However, the fact that the panelists

TABLE 3—Average Deviation from Overall Median by Specialty

	Procedures					
	ANGIO	CABS	CH	END	CO	CE
Medical Generalists						
Initial rating	-1.49	-1.07	-0.66	-0.42	0.90	0.26
Standard error	(0.09)	(0.08)	(0.15)	(0.04)	(0.06)	(0.06)
Final rating	-0.60	-0.40	-0.27	0.26	0.94	-0.43
Standard error	(0.05)	(0.04)	(0.14)	(0.04)	(0.03)	(0.04)
Medical Specialists						
Initial rating	0.65	0.01	0.16	0.47	-0.86	-0.39
Standard error	(0.83)	(0.07)	(0.17)	(0.04)	(0.05)	(0.06)
Final rating	0.11	-0.02	0.18	0.32	-0.10	-0.48
Standard error	(0.06)	(0.04)	(0.16)	(0.04)	(0.03)	(0.05)
Surgeons						
Initial rating	-0.50	0.93	0.49	0.38	0.44	0.45
Standard error	(0.08)	(0.09)	(0.16)	(0.04)	(0.05)	(0.04)
Final rating	0.09	0.65	1.02	-0.21	-1.94	0.46
Standard error	(0.07)	(0.07)	(0.25)	(0.04)	(0.04)	(0.03)

ANGIO = coronary angiography
 CABS = coronary artery bypass surgery
 CH = cholecystectomy
 END = upper gastrointestinal endoscopy
 CO = colonoscopy
 CE = carotid endarterectomy

substantially revised the initial lists and unanimously accepted the final lists suggests that the final lists should define indications that are well suited for rating their appropriateness.

- Perhaps panelists differed in their understandings of what the indications meant even though the definitions were supplied both in writing and orally at the meetings, and there was plenty of opportunity to discuss questions and clarify the definitions.

- Perhaps the indications with disagreement do not arise in practice. They may be hypothetical situations with which the panelists have had no experience. There may be a much

higher degree of agreement on the indications that do occur frequently in practice. However, the panelists did drop some indications from the initial lists on the grounds that they never occur in practice, suggesting that most of the final indications do occur at least occasionally.

Our tentative conclusion is that there really is disagreement about the value of these procedures for many indications, and that the disagreement reflects the lack of detailed evidence about the circumstances in which these procedures are efficacious. This conclusion is consistent with the view that uncertainty underlies much of the geographic variation in procedure use.^{14, 15}

**APPENDIX A
PANEL MEMBERSHIP**

Specialty	Panels		
	Cardiovascular	Gastroenterological	Cerebrovascular
Medical generalists	1 family physician 2 internists	1 family physician 2 internists	1 family physician 1 internist
Medical specialists	3 cardiologists	3 gastroenterologists	2 neurologists
Surgeons	2 cardiothoracic surgeons	2 general surgeons	3 vascular surgeons 1 neurosurgeon
Radiologist	1 radiologist	1 radiologist	1 neuroradiologist

**APPENDIX B
CLINICAL SETTINGS AND FACTORS:
A FRAMEWORK FOR THE INDICATIONS FOR CORONARY ANGIOGRAPHY**

- | | |
|---|--|
| <ol style="list-style-type: none"> Asymptomatic patients <ul style="list-style-type: none"> Occupation Exercise ECG test Exercise thallium scan Exercise MUGA test Chest pain of uncertain origin <ul style="list-style-type: none"> Chest pain with exertion Exercise ECG test Exercise thallium scan Exercise MUGA test Chronic stable angina <ul style="list-style-type: none"> Contraindications to coronary artery bypass graft surgery Medical management of angina Degree of angina Exercise ECG test Exercise thallium scan Exercise MUGA test Unstable angina <ul style="list-style-type: none"> Contraindications to coronary artery bypass graft surgery Medical management of angina Degree of angina Exercise ECG test Exercise thallium scan Exercise MUGA test During an acute MI <ul style="list-style-type: none"> Contraindications to coronary artery bypass graft surgery Complications of MI | <ol style="list-style-type: none"> Within six months of an acute MI <ul style="list-style-type: none"> Contraindications to coronary artery bypass graft surgery Type of MI Presence of angina post MI Medical management of angina Degree of angina Exercise ECG test Exercise thallium scan Exercise MUGA test Sudden death survivors <ul style="list-style-type: none"> Contraindications to coronary artery bypass graft surgery Relation of sudden death episode to an MI Presence of angina post MI Medical management of angina Degree of angina Exercise ECG test Exercise thallium scan Exercise MUGA test Following coronary artery bypass graft surgery <ul style="list-style-type: none"> Timing of coronary angiography Presence of angina post CABG Medical management of angina Degree of angina Exercise ECG test Exercise thallium scan Exercise MUGA test |
|---|--|

ACKNOWLEDGMENTS

We gratefully acknowledge the essential contribution of the technical panel members.

Cardiovascular Panel

- Mortimer Buckley, MD
- Ernie Chaney, MD
- Gottlieb Friesinger, MD
- Ward Kennedy, MD
- John Kirklin, MD
- Ceylon Lewis, MD
- Joe Noble, MD
- Harold Sox, MD
- William Wilson, MD

Gastrointestinal Panel

- John R. Amberg, MD
- Arthur Feinberg, MD
- Martin Greene, MD
- Thomas Hendrix, MD
- Edwin Maynard, III, MD
- Robert McCurdy, MD
- John Morrissey, MD
- Robert Tyson, MD
- James Weber, MD

Carotid Endarterectomy Panel

- William H. Baker, MD
- Richard R. Brummett, MD
- Randall M. Cebul, MD
- John Paul Conomy, MD
- Robert W. Hobson, II, MD
- James T. Robertson, MD
- John O. F. Roehm, Jr., MD
- Hugh Henry Trout, III, MD
- Jack Page Whisnant, MD

This work was supported in part by grants from the Commonwealth Fund (Grant No. 6137), the John A. Hartford Foundation (Grant No. 83172-2H), the Pew Memorial Trust (Glenmede Trust), and the Robert Wood Johnson Foundation (Grant No. 7309). Any opinions, conclusions, and proposals in the text are those of the authors and do not necessarily represent the views of any of these foundations or of the Rand Corporation.

REFERENCES

1. Aaron JH, Schwartz WB: The painful prescription: rationing health care. Washington, DC: Brookings Institution, 1984.
2. Thurow LC: Medicine versus economics. *N Engl J Med* 1985; 313:611-614.
3. Chassin MR, Brook RH, Park RE, *et al*: Variations in the use of medical and surgical services by the Medicare population. *N Engl J Med* 1986; 314:285-290.
4. Lohr K: Peer review organizations: quality assurance in Medicare. (Rand P-7125.) Santa Monica, CA: Rand Corp, 1985.
5. Fink A, Brook RH, Kosecoff J, *et al*: The sufficiency of the clinical literature for learning about the appropriate uses of six medical and surgical procedures. Santa Monica, CA: Rand Corp, 1985.
6. Rennie, D: Consensus statements. *N Engl J Med* 1981; 304:665-666.
7. Chassin MR, Kosecoff J, Park RE, *et al*: Indications for selected medical and surgical procedures—A literature review and ratings of appropriateness: coronary angiography. (Rand R-3204/1-CWF/HF/PMT/RWJ.) Santa Monica, CA: Rand Corp, 1986.
8. Chassin MR, Park RE, Fink A, *et al*: Indications for selected medical and surgical procedures—A literature review and ratings of appropriateness: coronary artery bypass graft surgery. (Rand R-3204/2CWF/HF/PMT/RWJ.) Santa Monica, CA: Rand Corp, 1986.
9. Solomon DH, Brook RH, Fink A, *et al*: Indications for selected medical and surgical procedures—A literature review and ratings of appropriateness: cholecystectomy. (Rand R-3204/3-CWF/HF/PMT/RWJ.) Santa Monica, CA: Rand Corp, 1986.
10. Kahn KL, Roth CP, Kosecoff J, *et al*: Indications for selected medical and surgical procedures—A literature review and ratings of appropriateness: gastrointestinal endoscopy. (Rand R-3204/4-CWF/HF/PMT/RWJ.) Santa Monica, CA: Rand Corp, 1986.
11. Roth CP, Kahn KL, Fink A, *et al*: Indications for selected medical and surgical procedures—A literature review and ratings of appropriateness: fiberoptic colonoscopy. (Rand R-3204/5-CWF/HF/PMT/RWJ.) Santa Monica, CA: Rand Corp, 1986.
12. Merrick NJ, Park RE, Kosecoff J, *et al*: Indications for selected medical and surgical procedures—A literature review and ratings of appropriateness: carotid endarterectomy. (Rand R-3204/6-CWF/HF/PMT/RWJ.) Santa Monica, CA: Rand Corp, 1986.
13. Brook RH, Lohr KN, Chassin MR, *et al*: Geographic variations in the use of services: do they have any clinical significance? *Health Affairs* 1984; 3(4):63.
14. Eddy DM: Variations in physician practice: the role of uncertainty. *Health Affairs* 1984; 3(4):74-89.
15. Wennberg JE: Dealing with medical practice variations: a proposal for action. *Health Affairs* 1984; 3(4):6-32.

**Availability of Postdoctoral and Senior Fellowships
in the Behavioral Sciences**

The National Institutes of Health (NIH) Working Group on Health and Behavior wishes to bring to the attention of the research community the fact that the NIH National Research Service Award (NRSA) program provides postdoctoral and senior fellowships in the behavioral sciences as well as in the biomedical sciences (Section 472 of the Public Health Service Act). Fellowships are available to postdoctoral and senior scientists for up to three years depending upon their prior research experience.

The NIH supports research fellowships in a wide range of behavioral sciences topics, including research on basic processes as well as research applied to behavioral factors in health and illness. The NRSA program offers an opportunity to research scientists and clinicians to broaden their scientific background or to extend their potential for research in areas supported by the various Institutes at the NIH.

Receipt dates for applications are January 10, May 10, and September 10 of each year. For information about NRSA postdoctoral fellowships, including application procedures, applicable regulations, suitability of research areas, stipends, and allowable costs, please contact:

Office of Grant Inquiries
Division of Research Grants
National Institutes of Health
Westwood Building 449
Bethesda, MD 20892
Telephone: 301 496-7441