Are the Best Coronary Artery Bypass Surgeons Identified by Physician Surveys?

Arthur J. Hartz, MD, PhD, Jose S. Pulido, MD, and Evelyn M. Kuhn, PhD

ABSTRACT

Objectives. This study assessed the validity of surveys for identifying the best coronary artery bypass surgeons.

Methods. Data on physicians who performed coronary artery bypass surgery were available from New York, Pennsylvania, and Wisconsin. Data on physicians' reputation were obtained from one national and five city surveys. The measure of surgical performance was the mortality ratio (MR), that is, the ratio of the observed to the predicted patient mortality rate.

Results. Mortality ratios were very similar for the 10 722 patients treated by the 31 surgeons defined as "best" doctors in the surveys (MR = .98) and for the 74 854 patients treated by 243 other surgeons who had more than a minimal number of cases (MR = .96). The mortality ratio was 1.34 for the patients treated by surgeons with the lowest volumes and .87 for the surgeons who performed more than 400 coronary artery bypass surgeries in 3 years.

Conclusions. These results suggest that the quality of a coronary artery bypass surgeon may be more closely associated with patient volume than with the surgeon's reputation among peers. (*Am J Public Health.* 1997;87:1645–1648)

Introduction

The 1994 edition of the book *The Best Doctors in America*[1] has sold more than 17 000 copies. Many city magazines have recognized the commercial potential of articles on the "best doctors" in the community and have published issues devoted to this topic. There have been no published analyses of the impact of these articles, but they markedly increase magazine sales and no doubt affect readers' selection of physicians.

The articles on "best doctors" are based on surveys of physicians' reputation among other physicians. The validity of these surveys has not been assessed by objective measures such as patient outcomes. Unfortunately, information on outcomes relevant for a given physician specialty is usually not available, or if it is, the outcome may depend more on patient risk than on the physician's skill. Since the "best" doctors may treat higher-risk patients than other doctors, they may be penalized by outcomes not adjusted for patient risk.

Perhaps the highest-quality risk-adjusted outcome data are for physicians who perform coronary artery bypass graft surgery. These data can be used to assess the validity of the survey results with regard to the quality of the surgeons. The results of the present study may be generalizable to surveys evaluating physicians in other specialties.

Methods

Three data sets containing information on surgeons who performed coronary artery bypass graft surgery were used in this study: a New York data set for the period 1990 to 1992,[2] a Pennsylvania data set for the period 1990 to 1992,[3] and a "Wisconsin-plus" data set[4,5] that included only Medicare patients for 1989 to 1991, with most surgeries taking place in 1990. The Wisconsin-plus data set included data from Wisconsin and from one large hospital outside Wisconsin. This hospital specialized in bypass surgery and used the same data collection forms as the Wisconsin hospitals.

The available information on the individual surgeons in each data set included the number of surgeries performed, the number of deaths, and the predicted number of deaths. The predicted number of deaths was determined on the basis of patient risk factors, as has been described previously.[2–5] Although the specific risk factors differed among the three data sets, each data set included information on whether or not the patient had had a myocardial infarction immediately prior to surgery, indicators of moderate or severe impairment of ventricular function, and whether the patient had had a previous coronary artery bypass surgery.

States reported information only on individual surgeons who performed more than a specified number of surgeries. New York reported the results for individual surgeons who performed at least 200 surgeries during 3 years. Pennsylvania reported results for each of 3 years and provided information on individual surgeons who performed at least 30 surgeries during the year. We had information available for all physicians in the Wiscon-

At the time of the study, Arthur J. Hartz and Evelyn M. Kuhn were with the Department of Family and Community Medicine, Medical College of Wisconsin, Milwaukee. Jose S. Pulido is with The Eye Institute, Milwaukee, Wis.

Requests for reprints should be sent to Arthur J. Hartz, MD, PhD, Department of Family Medicine, University of Iowa College of Medicine, 2148 Steindler Bldg, Iowa City, IA 52242.

This paper was accepted December 11, 1996.

TABLE 1—Mortality Ratios[a] for "Best" and Other Coronary Artery Bypass Surgeons

Physician Classification	No. Surgeons	No. Patients	Observed Mortality, %	Predicted Mortality, %	Mortality Ratio[a]	95% Confidence Interval[b]
Combined data						
"Best"	31	10 277	3.4	3.5	0.98	0.88, 1.08
US "best"	7	1 968	3.9	3.7	1.03	0.81, 1.26
Other	243	74 854	3.3	3.4	0.96	0.92, 0.99
Low-volume other	. . .	10 568	4.4	3.3	1.34*	1.24, 1.44
New York						
"Best"	13	6 096	2.8	3.0	0.95	0.81, 1.09
Other	72	31 668	2.8	3.0	0.94	0.88, 1.00
Low-volume other[c]	0	7 168	3.8	2.8	1.33*	1.19, 1.46
Pennsylvania						
"Best"	9	3 870	3.4	3.4	1.00	0.83, 1.16
Other	147	41 838	3.5	3.6	0.97	0.82, 1.02
Low-volume other[d]	0	2 892	5.2	3.5	1.47*	1.28, 1.66
Wisconsin-plus						
"Best"	9	756	8.2	7.9	1.04	0.80, 1.29
Other	24	1 348	7.5	8.0	0.94	0.75, 1.13
Low-volume other[e]	74	508	8.5	7.9	1.07	0.78, 1.37

[a]Observed mortality divided by predicted mortality. The ratio was obtained prior to rounding mortality rates.
[b]Mortality ratio $\pm 2.0 \times$ square root of the variance of the ratio, computed as described in the Methods section.
[c]Physicians with fewer than 200 patients within 3 years.
[d]Physicians with fewer than 30 patients in each of 3 years.
[e]Physicians with fewer than 30 Medicare patients reviewed.
*$P < .0001$ compared with other identified surgeons.

sin-plus data set, but we analyzed as individuals only those surgeons in that data set who performed at least 30 surgeries. Aggregate information on all physicians in each data set who did not perform enough surgeries to be evaluated individually was obtained by subtracting the total number of observed and predicted patient deaths for the individual physicians from the numbers for the state.[2,3]

We used several sources to identify the "best" coronary artery bypass surgeons by reputation. The only available national source was the book *The Best Doctors in America.*[1,6] This source identified the "best doctors" by first surveying physicians from well-known institutions regarding the best physicians in their specialty and then surveying these physicians for their recommendations of outstanding physicians. As this process continued, names were added or subtracted from the list when a consensus of the other physicians became apparent. Only surgeons or bypass surgeons were asked to identify the best bypass surgeons.

Our other sources of physicians with the best reputations were articles from city magazines. The articles were located by contacting all magazines in the larger cities in each of the three states. Articles on "best doctors" were found in *New York,*[7] *Philadelphia,*[8] *Pittsburgh,*[9] *Milwaukee Magazine,*[10] and *Madison Magazine.*[11] In contrast to the survey methods used in *The Best Doctors in America,* the community magazines surveyed physicians regardless of specialty to identify the best coronary artery bypass surgeons. Three of these magazines solicited only the opinions of physicians who had been identified as "best doctors" by other sources,[7,9,10] and two surveyed a broad sample of physicians in the area.[8,11]

We defined "best" doctors as those identified as outstanding in any of the city magazines or in the book. Seven physicians were identified in the book and 29 were identified in the magazines. The two physicians who were identified in the book but not the magazines were from the hospital in the Wisconsin-plus data set that was outside Wisconsin.

Surgeons were divided into three groups for analysis: (1) those identified individually who were listed in one of our sources for "best" physicians, (2) those listed individually who were not included in our sources for "best" physicians, and (3) those not identified individually because they treated an insufficient number of patients. For most analyses data were pooled for all physicians in a group. Our risk-adjusted measure of performance for a group of physicians was the ratio of the observed to the predicted number of deaths (i.e., the mortality ratio); the higher the mortality ratio, the worse the performance. Under the null hypothesis, the mortality ratio has an approximately normal distribution with a mean of 1 and a variance equal to $(1 - p)/np$, where p is the expected mortality rate and n is the number of patients. The statistical significance for the difference between two mortality ratios was found with the z statistic for the difference between two normally distributed variables with known variances. The association of the mortality ratio with patient volume was tested with the Pearson correlation coefficient.

Results

Mortality ratios for the "best" coronary artery bypass surgeons and other surgeons are shown in Table 1. The data included results on more than 10 000 patients of 31 physicians identified as "best" by some source and on the subset of nearly 2000 patients of the 7 physicians identified as "best" by the national survey. The mortality ratio for both groups of "best doctors" was close to 1.00, indicating that the observed mortality rate was almost equal to the predicted rate. The group of other physicians who had enough patients to be evaluated individually had a mortality ratio that was slightly but not significantly lower than that of the "best" physicians. Physicians who did not perform enough surgeries to be consid-

ered individually had a mortality ratio of 1.34, which was significantly greater at the $P < .0001$ level than the mortality ratio for the other physicians.

Table 1 also shows results for the individual states. The number of "best" physicians in the Wisconsin-plus data set was disproportionate to the size of the state, in part because two of the physicians listed in *The Best Doctors in America* came from the hospital outside Wisconsin and in part because the "best" physicians in Wisconsin were identified in two city magazines; there was only one city magazine in New York and two in Pennsylvania. The Wisconsin-plus data set also had the highest mortality rate. One contributing factor was that the Wisconsin-plus patients were on Medicare and therefore were older than patients in the other data sets. In each of the three data sets the mortality ratio for all physicians combined was 1.00 because for a given data set the total observed and expected mortality rates were equal.

The same relationships found in the combined data set were also true for the individual data sets; that is, the "best" physicians did not have lower mortality ratios than the other physicians. For Pennsylvania and New York it was also possible to find mortality ratios for low-volume physicians, and in both states these physicians had higher mortality ratios than others. It was not possible to accurately identify low-volume physicians in Wisconsin because data were available only for a sample of Medicare patients.

It is possible that a patient's risk of mortality may be affected by his or her community in ways that are not captured by the risk-adjustment method. For this reason we compared the mortality ratios of the "best" doctors with the mortality ratios of those surgeons who practiced in the same communities or same hospitals. The surgeons who practiced in the same communities as the "best" doctors had a mortality ratio of .91 and the surgeons who practiced in the same hospitals as the "best" doctors had a mortality ratio of .92. Both of these ratios were lower than the ratio for the "best" doctors (.98), although the differences were not statistically significant at the $P < .10$ level.

The comparison of the performance of individual "best" doctors with that of the other doctors, based on mortality ratios, is shown in Table 2. The mortality ratio of one of the "best" doctors was better than the average ratio of the other doctors at a level that was of borderline statistical significance ($P < .10$). If "best doctors" had been chosen by lottery, we would expect an average of 1.5 "best" doctors to be better than the others at the $P < .10$ level of significance. One of the "best" doctors had significantly worse performance than the average of the other physicians ($P < .02$). Seven (23%) of the "best" doctors had a mortality ratio of at least 1.25 (the observed mortality was 25% higher than the predicted), compared with 70 (29%) of the 243 other doctors. The difference between 23% and 29% was not statistically significant.

TABLE 2—Performance of "Best" Coronary Artery Bypass Surgeons Compared with Others

	Overall		New York		Pennsylvania		Wisconsin-Plus	
	No.	%	No.	%	No.	%	No.	%
Better than average[a]	15	48	6	46	5	56	4	44
Worse than average[a]	16	52	7	54	4	44	5	56
Significantly better	1*		1*		0		0	
Significantly worse	1**		0		1**		0	
Mortality ratio >1.25[b]	7	23	2	15	3	33	2	22

[a]Mortality ratios of "best" surgeons were compared with the average mortality ratios of all other individually listed physicians.
[b]Of individually reported physicians not identified as "best," 29% had mortality ratios greater than 1.25.
*$P < .10$; **$P < .02$.

The data for low-volume physicians in Table 1 suggest that physicians with fewer than the minimum number of surgeries necessary to be included as individuals in the data set may have had a higher mortality ratio than other physicians. We examined whether this relationship continued for physicians who performed more than the minimum number of surgeries. Wisconsin physicians were eliminated from this analysis because the volume of surgeries for patients who were not on Medicare was unknown for these physicians. The mortality ratio was 1.07 for the 152 physicians who performed from 30 to 399 surgeries over a 3-year period and .87 for the 89 physicians who performed 400 or more surgeries ($P < .0001$). For the physicians who performed enough surgeries to be included in this data set as individuals, the correlation between surgical volume and mortality ratio was $-.27$ ($P < .0001$). The correlation would have been greater if low-volume surgeons had been identified individually and could have been included in the analysis.

There was a relationship between classification as a "best" doctor and practice volume. The "best" physicians in Pennsylvania and New York performed more coronary artery bypass surgeries over a 3-year period than other surgeons. After adjustments were made for volume differences between the states, the "best" physicians performed 419 surgeries over a 3-year period, compared with 339 surgeries for the other physicians ($P = .06$). The mortality ratio was 1.05 for the 12 "best" physicians who performed fewer than 400 surgeries and .93 for the 10 "best" physicians who performed more than 400 surgeries. These mortality ratios did not differ significantly from the mortality ratios of other physicians who performed a similar number of surgeries.

Discussion

We used clinically risk-adjusted outcome data from three states to evaluate the validity of surveys to identify the best coronary artery bypass surgeons. In each of the three states, the surgeons selected as "best" doctors in national or local surveys had mortality ratios similar to those of other surgeons. Only 1 of 31 "best" doctors performed better than the average of the other doctors at a significance level of $P < .10$, and 7 (23%) of the "best" doctors had mortality rates more than 25% higher than predicted, compared with 29% of the other doctors. These results suggest that classification as a "best doctor" by a survey of other physicians does not ensure that a coronary artery bypass surgeon will perform better than most other doctors. Since these results were consistent for three states, it is unlikely that most current magazine surveys would provide substantially different results.

A better predictor of physician performance was surgical volume. The association of better outcomes with increased surgical volume has been found previ-

ously for surgical procedures in general[12–15] and for coronary artery bypass surgery in particular.[16–19] In the present study surgeons' mortality ratios decreased with increasing surgical volume and were lowest for physicians who performed more than 400 surgeries in a 3-year period. Surgical volume may be a proxy for experience. It may also indicate the number of referrals a surgeon receives and therefore may be a better indication of a physician's professional reputation than can be obtained from surveys. Supporting the possibility that volume is an indication of reputation is the fact that the "best" physicians performed significantly more surgeries than the other physicians.

The conclusions of this study depend on the validity of the risk-adjusted mortality ratio as an indicator of quality of care. Although the data used for this study have been shown to be generally valid and useful,[20] the risk-adjusted mortality ratio is not a perfect indicator of quality; the risk adjustment may be inadequate and mortality may be insensitive to certain aspects of quality.[21]

The limitations of risk-adjusted mortality ratios may not affect the conclusions of this study, however. Mortality is an important outcome, and the risk adjustment used in this study compares favorably with other risk-adjustment models.[22,23] In addition, our data suggest that perfect risk adjustment may not be necessary. Since the measured patient risk (i.e., the predicted mortality rate) was very similar for the "best" surgeons and other surgeons, it is unlikely that the unmeasured patient risk differed substantially.

We can only speculate on the reasons why the surveys did not identify surgeons with superior performance. In lieu of adequate direct information about a surgeon's clinical skills, other physicians may judge the surgeon on the basis of their personal interactions with the surgeon or the surgeon's prominence in research or medical organizations. It is also possible that physicians choose the "best" surgeons for political reasons, such as membership in the same health care organization. Certainly, some physicians complain about the political nature of the selection of "best" doctors.

The relationship between reputation and outcomes was tested only for physicians who perform coronary artery bypass surgery. It is possible that physician reputation is a better indicator of performance in other specialties. This is unlikely, however. For bypass surgery, mortality and severe complications that may result in mortality are very visible outcomes. In other specialties patient outcomes are often less visible to physicians without responsibility for the patient. Therefore, the weak association between reputation and performance for bypass surgeons should also hold true for physicians in other specialties.

The possibility that physician surveys are unreliable guides to physician quality will disappoint patients searching for a good doctor. The results of such surveys are in high demand by lay readers and should be more reliable than the opinion of friends, which in the past has been the most important influence on patients' choice of physician.[24] Patients have few alternative measures of physician quality. Good objective information on physicians is not generally available and is unlikely to become available in the near future. This information depends on the intellectually difficult process of defining quality of care and the costly process of collecting the data necessary to obtain risk-adjusted outcomes. Without objective information patients must rely on a physician's reputation. Unfortunately, reputation may measure a physician's skill in associating with other physicians more than it measures the physician's skill in caring for patients. Our results suggest that patient volume may be a better indicator of quality than reputation. □

Acknowledgments

Dr Pulido's effort was supported in part by an unrestricted grant from Research to Prevent Blindness, Inc, and by National Institutes of Health grant P30 EY01931.

Information on patient mortality for New York surgeons was provided by the New York Department of Health, and information for Pennsylvania surgeons was provided by the Pennsylvania Health Care Cost Containment Council. We would like to thank Teri Wermager for locating the articles on best doctors and preparing the manuscript.

References

1. Naifeh S, Smith GW, Stec LP, Arnold SF, Greame CJ, eds. *The Best Doctors in America 1994–1995*. Aiken, SC: Woodward/White; 1994.
2. *Coronary Artery Bypass Surgery in New York State 1990–1992*. Albany, NY: New York State Department of Health, December 1993:7–11.
3. *Coronary Artery Bypass Graft Surgery*. Vol. 3, 1992. Harrisburg, Pa: The Pennsylvania Health Care Cost Containment Council; December 1994. Technical Report.
4. Hartz AJ, Kuhn EM, Kayser KL, Pryor DP, Green R, Rimm AA. Assessing providers of coronary revascularization: a method for peer review organizations. *Am J Public Health*. 1992;82:1631–1640.
5. Hartz AJ, Kuhn EM. Comparing hospitals that perform coronary artery bypass surgery: the effect of outcome measures and data sources. *Am J Public Health*. 1994;84:1609–1614.
6. Naifeh S, Smith GW, Stec LP, eds. *The Best Doctors in America*. 1st ed. Aiken, SC: Woodward/White; 1992.
7. Tanne JH. The best doctors in New York. *New York*. November 1991 (special issue).
8. Saline C. Top docs. *Philadelphia*. March 1994:65–104.
9. Webb S. Top doctors: Pittsburgh physicians select the best of their peers. *Pittsburgh*. October 1992:40–45.
10. Woods MM. The doctors' advice. *Milwaukee Magazine*. March 1991:38–44.
11. Nowlen C, Muckian M. Top doctors. *Madison Magazine*. May 1993:62–76.
12. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med*. 1979;301:1364–1369.
13. Luft HS. The relation between surgical volume and mortality: an exploration of causal factors and alternative models. *Med Care*. 1980;18:940–959.
14. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health Serv Res*. 1987;22:158–182.
15. Flood AB, Scott WB, Ewy W. Does practice make perfect? Part I: the relation between volume and outcomes for selected diagnostic categories. *Med Care*. 1984;22:98–114.
16. Showstack JA, Rosenfeld KE, Garnick DW, et al. Association of volume with outcome of coronary artery bypass graft surgery. Scheduled vs nonscheduled operations. *JAMA*. 1987;257:785–789.
17. Hannan EL, Kilburn H, Bernard H, O'Donnell JF, Lukacik G, Shields EP. Coronary artery bypass surgery. The relationship between inhospital mortality rate and surgical volume after controlling for clinical risk factors. *Med Care*. 1991;29:1094–1107.
18. Hannan EL, Kumar D, Racz M, Siu AL, Chassin MR. New York State's cardiac surgery reporting system: four years later. *Ann Thorac Surg*. 1994;58:1852–1857.
19. Hannan EL, Siu AL, Kumar D, Kilburn H, Chassin MR. The decline in coronary artery bypass graft surgery mortality in New York State. The role of surgeon volume. *JAMA*. 1995;273:209–213.
20. Hannan EL. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA*. 1994;271:761–766.
21. Green J, Wintfeld N. Report cards on cardiac surgeons: assessing New York State's approach. *N Engl J Med*. 1995;332:1229–1232.
22. Landon B, Iezzoni LI, Ash AS, et al. Judging hospitals by severity-adjusted mortality rates: the case of CABG surgery. *Inquiry*. 1996;33:155–166.
23. Omoigui NA, Miller DP, Brown KJ, et al. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. *Circulation*. 1996;93:27–33.
24. Americans rate the U.S. health care system: cite reasons for choosing doctors. Fenton, Miss: Maritz Marketing Research Inc; September 7, 1989. News release.