SPECIAL ARTICLE

EFFECT OF UTILIZATION REVIEW IN A FEE-FOR-SERVICE HEALTH INSURANCE PLAN

STEPHEN N. ROSENBERG, M.D., M.P.H., DAVID R. ALLEN, PH.D., JANIS S. HANDTE, M.A., THOMAS C. JACKSON, C.E.B.S., LEONARD LETO, M.S., BARRY M. RODSTEIN, M.D., M.P.H., SCOTT D. STRATTON, M.P.H., GEORGE WESTFALL, M.B.A., AND ROSLYN YASSER, M.S.W.

Abstract *Background.* Although utilization review is widely used to control health care costs, its effect on patterns of health care is uncertain.

Methods. In 1989, New York City and its unions temporarily replaced actual utilization review with sham review for half the participants in the city's fee-for-service health insurance plan. We compared the health services provided to 3702 enrollees whose requests were subjected to utilization review (the review group) with the services provided to 3743 enrollees whose requests received sham review and were automatically approved for insurance coverage (the nonreview group). The enrollees, physicians, and hospitals were all unaware of the group assignments.

Results. During the study period (mean duration, eight months), the members of the review group underwent 1255 procedures in 20 categories of procedures for which second opinions were required (such as breast, cataract, foot, hernia, and hip-replacement surgery, as well as hysterectomy and coronary bypass surgery), and the members of the nonreview group underwent 1365 procedures ($P = 0.02$). The members of the review group had 124 fewer procedures in doctors' offices and hospital outpatient departments ($P = 0.002$). In the following year, the members of the review group underwent 248 procedures from the 20 categories, and the members of the nonreview group underwent 234 ($P = 0.46$). No other differences in patterns of care were found between the groups, including rates of hospital admission to medical–surgical, substance-abuse, or psychiatric units; average lengths of hospital stay; the percentage of enrollees who received preadmission testing; or rates of use of home care. During the study period, the mean age-adjusted insurance payments per person were $7,355 in the review group and $6,858 in the nonreview group ($P = 0.06$).

Conclusions. The utilization-review program reduced the number of diagnostic and surgical procedures performed that required second opinions and did not merely delay them until the following year. Otherwise, the program had little effect. Alternatively, actual review and sham review may both have decreased the use of hospital services, with patients or their physicians choosing more efficient treatment when they believed that care would be reviewed. (N Engl J Med 1995;333:1326-30.)

MOST health maintenance organizations and over 90 percent of health insurance plans review the provision of health care to their enrollees in order to control costs.[1-4] Numbers of hospital admissions and days of inpatient care have been shown to decrease when such monitoring, known as utilization review, is introduced.[2-12] The reductions may be due to one or more of the changes that utilization-review programs strive to accomplish, including the use of ambulatory surgery and outpatient treatment for psychiatric disorders or substance abuse instead of inpatient care, preadmission testing instead of preoperative hospitalization, and home care instead of extended stays in the hospital. Utilization-review programs typically seek to lower the rate of medically unnecessary procedures by requiring patients to seek second opinions and to encourage shorter hospitalizations by fostering more efficient inpatient care. To prevent poor outcomes and complications requiring additional inpatient care, these programs provide case management in which nurses visit severely ill patients, assist in coordinating their care, and are authorized to approve insurance coverage for health services not normally included under the patients' plans, such as extended rehabilitation.[1-3,6,13-15] It is uncertain which of these objectives, if any, are actually achieved.[1,3,12]

From 1985 through 1989, New York City and its unions, represented by the Municipal Labor Committee, increasingly required municipal employees, retirees, and their dependents who were enrolled in the city's fee-for-service health insurance plan to be subject to utilization review. In 1989, the city and the unions temporarily replaced the utilization review with a sham review for half the participants in the plan, so that the hospital records of those participants could be evaluated and compared with the records of patients who had actual reviews, in order to ensure that utilization review was not affecting the quality of patient care adversely.[16] In the present study, we compared the health care provided to the two groups, to examine how the review procedures affected patterns of care.

METHODS

Study Assignments

All non-Medicare enrollees in New York City's fee-for-service insurance plan, operated by Empire Blue Cross and Group Health, Inc., were required to telephone the city's utilization-review program, NYC Healthline, before any nonemergency admission to the hospital and any ambulatory surgery, and as soon as possible after any admission during an emergency. Failure to call in a nonemergency situation could

From the Division of Health Policy and Management, Columbia University School of Public Health (S.N.R., J.S.H., B.M.R.); the Subcommittee on Health, New York City Municipal Labor Committee (D.R.A., R.Y.); the Health Benefits Program, New York City Office of Labor Relations (S.N.R., T.C.J., L.L.); Group Health, Inc. (S.D.S.); and Empire Blue Cross and Blue Shield (G.W.) — all in New York. Address reprint requests to Dr. Rosenberg at the Division of Health Policy and Management, Columbia University School of Public Health, 600 W. 168th St., New York, NY 10032.

Supported by a grant (10964) from the New York Community Trust and the Carol and Charles Spaeth Memorial Fund of Community Funds, Inc.

result in a reduction of insurance payments for the service in question by as much as $500, with the enrollee held responsible for the balance.

Each enrollee who telephoned the utilization-review program between September 21 and December 20, 1989, was systematically assigned to one of two groups on the basis of the last digit in his or her insurance identification number (the contract holder's Social Security number). Enrollees whose identification numbers ended in one of five arbitrarily chosen digits were required to participate fully in the program of utilization review (the "review group"). Enrollees whose identification numbers ended in one of the five remaining digits were not required to participate in the program (the "nonreview group"). After information was collected on the characteristics of the enrollees and the proposed health care services, the members of the nonreview group were told that they had satisfied the program's requirements. All admissions and ambulatory care procedures they requested were approved for insurance coverage.

If an enrollee called the program again, about another admission or procedure scheduled on or before June 30, 1990, utilization review was again required for members of the review group but not for members of the nonreview group. After the end of this study period, utilization review was required for all enrollees.

Of 7445 enrollees included in this study, 3702 were assigned to the review group. When these people called the utilization-review program, they were encouraged to ask the registered nurses on the staff ("review nurses") questions about their health and their proposed treatment. The nurses and a panel of consultant physicians discussed the medical need for all nonemergency admissions (other than those for labor and delivery) with the enrollees' attending physicians. After all appeals had been considered, inpatient care was judged to be medically unnecessary in 77 of 2398 cases (3.2 percent), and the physicians and enrollees were informed that in these cases insurance coverage would be approved only for outpatient care (for example, ambulatory surgery or outpatient drug rehabilitation).

For the remaining 2321 admissions (those for which inpatient coverage was approved) and for 245 emergency and 536 obstetrical admissions, the nurses and physicians at the utilization-review program discussed the patients' proposed lengths of stay with the attending physicians. Of 22,810 hospital days requested in advance by the attending physicians, 2304 (10.1 percent) were judged medically unnecessary (after the consideration of appeals) and were not approved for coverage. (The alternatives recommended by the panel included preadmission testing and earlier discharge to home care. Nurses at the utilization-review program provided assistance with planning at discharge when necessary.) In addition to the 20,506 hospital days that were approved in advance, approval was granted for 286 hospital days at the beginning of the 245 emergency admissions, before the utilization-review program was contacted. In all, therefore, 20,792 inpatient days were approved for insurance coverage. In addition, case-management services were offered to 64 members of the review group, 60 of whom accepted the offer.

The following 20 procedures generally performed on a nonemergency inpatient basis or as invasive diagnostic or surgical outpatient procedures had been selected before the time of this study as potentially requiring second-opinion review: breast, carpal-tunnel, cataract, foot, gallbladder, knee, nose, prostate, spine, and varicose-vein surgery; coronary-artery bypass surgery; cardiac catheterization; carotid endarterectomy; dilation and curettage (excluding the termination of pregnancy); hemorrhoidectomy; hernia repair; hip replacement; hysterectomy; pacemaker implantation; and tonsillectomy, adenoidectomy, or both. Of 1388 members of the review group who were scheduled for these procedures, 265 (19.1 percent) were required to obtain a second opinion in consultation with another physician because the review nurses thought that the indications for the procedures were inconclusive and that a second opinion might lead to the choice of an alternative treatment. Voluntary consultations were also arranged for 12 patients who requested second opinions about other procedures.

The 3743 members of the nonreview group were also encouraged to discuss their conditions and treatment with members of the nursing staff and could request other services from the utilization-review program, as described in a pamphlet distributed to all enrollees. Referrals for second opinions were requested by 59 members of this group (including 44 who were scheduled to undergo procedures in the 20 selected categories). No one in this group requested review of a hospital admission or length of stay, assistance with discharge planning, or case management. All 3198 admissions requested in this group were approved for insurance coverage.

The distinction between the two groups of enrollees was not made known to physicians, hospitals, city employees or retirees in general, or to those who called the utilization-review program, so the nonreview group served as a sham utilization-review group. Representatives of the city and the unions discussed the issue of informed consent when temporary suspension of the review requirement for half the enrollees was first considered. Because the effect of utilization review on the quality of care had not then been determined,[3,5-8,11,12,17] assignment to nonreview status was not thought to take away any known benefit. It was decided that there was no need to obtain consent if the members of the nonreview group were allowed to request the services of the utilization-review program voluntarily.

Analysis of Claims

We reviewed all insurance claims for services provided to members of both groups, whether or not those services had been approved for coverage. We compared the groups with respect to services provided during the study period — that is, from the time of each enrollee's call to the utilization-review program to June 30, 1990, a period lasting an average of eight months. We also compared the groups with respect to claims for services provided during the following 12 months (from July 1, 1990, to June 30, 1991), when utilization review was required for all plan participants, to determine whether the review might have delayed, rather than averted, the provision of services.[11]

This analysis was undertaken jointly by New York City, its unions and insurers, and Columbia University. To preserve confidentiality, enrollees' names were deleted and their identification numbers coded in all information provided by the insurers. The study was approved by the institutional review board of the College of Physicians and Surgeons at Columbia University.

Statistical Analysis

Unpaired t-tests were used to evaluate differences between study groups in mean age, the mean length of the study period, and the mean number of inpatient days.[18] The z statistic with correction for continuity was used to compare the groups with respect to the sex distribution, hospital admissions, procedures performed in the 20 selected categories, preadmission testing, ambulatory surgery, use of home care, and the proportion of the group whose overall insurance claims exceeded various predefined amounts.[18] To test for differences between groups in mean lengths of stay and insurance payments, analyses of covariance were performed with and without the use of age as a covariate.[19] All evaluations of probability were two-sided. P values of 0.05 or less were considered to indicate statistical significance.

Results

The mean age of the members of the review group (38.5 years) was less than that of the members of the nonreview group (39.7 years; $P = 0.003$). There was no significant difference between the review and the nonreview groups in sex distribution (63.2 and 64.3 percent, respectively, were female; $P = 0.34$) or in the average length of the study period (240 days for both; $P = 0.97$). During the study period, the utilization-review program received telephone calls with regard to 89 percent of elective admissions, 71 percent of maternity admissions, and 64 percent of emergency admissions, indicating that there was widespread awareness of the program.

During the study period, the members of the review group underwent 110 fewer procedures overall in the 20 categories subject to the second-opinion requirement

Table 1. Effect of Referral for a Second Opinion on the Incidence of Procedures Performed in 20 Selected Categories.*

Type of Procedure	Review Group (N = 3702)		Nonreview Group (N = 3743)		P Value
	No. of procedures	Rate per 1000	No. of procedures	Rate per 1000	
Study period					
Outpatient or office	789	213	913	244	0.002
Inpatient	466	126	452	121	0.52
Both	1255	339	1365	365	0.02
Following year					
Outpatient or office	180	49	162	43	0.30
Inpatient	68	18	72	19	0.85
Both	248	67	234	62	0.46

*The categories studied were as follows: breast, carpal-tunnel, cataract, foot, gallbladder, knee, nose, prostate, spine, and varicose-vein surgery; coronary-artery bypass surgery; cardiac catheterization; carotid endarterectomy; dilation and curettage (excluding the termination of pregnancy); hemorrhoidectomy; hernia repair; hip replacement; hysterectomy; pacemaker implantation; and tonsillectomy, adenoidectomy, or both.

than did the members of the nonreview group ($P = 0.02$), and the review group had 124 fewer procedures in physicians' offices and hospital outpatient departments ($P = 0.002$) (Table 1). During the following year, the members of the review group underwent 14 more procedures overall in these 20 categories than the members of the nonreview group ($P = 0.46$).

Of the 265 members of the review group who were referred for second opinions, 88 (33 percent) did not obtain them, 99 (37 percent) received opinions confirming the need for their procedures, and 78 (29 percent) received opinions that did not confirm this need (Table 2). Of 1388 procedures in the 20 selected categories that were proposed for members of the review group, 133 were not performed by the end of the study period. These included 69 (88 percent) of the 78 proposed by enrollees who received second opinions not endorsing the procedure and 29 (33 percent) of the 88 proposed by enrollees who were referred for consultations but did not obtain them.

We calculated the dollar value of all claims for institutional, professional, and ancillary services received during the study period by each enrollee. When we compared the percentage of the membership of each group whose total claims exceeded various dollar amounts, none of the differences were significant (Table 3). During the study period, the mean insurance payment per person (including institutional, professional, and ancillary services), after adjustment for age, was $7,355 in the review group and $6,858 in the nonreview group ($P = 0.06$).

Table 2. Outcome of Referral and Consultation Related to Procedures Proposed for Members of the Review Group during the Study Period.*

Category of Patients	No. of Procedures Proposed	No. Performed	No. (%) Not Performed
Referred for consultation			
Obtained no consultation	88	59	29 (33)
Received confirming opinion	99	97	2 (2)
Received nonconfirming opinion	78	9	69 (88)
Not referred	1123	1090	33 (3)
All	1388	1255	133 (10)

*The 20 categories are listed in the footnote to Table 1.

There were no significant differences between the groups during the study period in rates of admission to medical–surgical, substance-abuse, or psychiatric units; average lengths of stay; the percentage who received preadmission testing; or the rates of use of home care (Table 4). When we excluded surgical procedures in the 20 selected categories, there was no difference between the groups in the percentage of hospital surgery performed in hospital-based ambulatory care units.

Members of the review group were hospitalized for 19,248 days during the study period (Table 5), 92.6 percent of the 20,792 days that had been approved for insurance coverage. The average number of hospital days used per person did not differ significantly between groups during either the study period or the following year.

Discussion

The utilization-review program we studied was very comprehensive by 1989 standards,[16,20] but the activities it conducted are now included in most such programs.[1,2,4,10,13,14]

Referrals for a second opinion appear to have reduced the number of procedures performed in the review group among the 20 categories we studied, and these procedures were not merely delayed until the next year. The reduction probably occurred among outpatient procedures because they are more likely to be elective.[21] Because of the small number of procedures in each of the 20 categories, no procedure-specific analyses could be performed.

The receipt of a second opinion that did not recommend the requested procedure only partly explained the number of procedures in the 20 categories that were proposed but not performed. One third of the enrollees who did not obtain a second opinion also did not undergo the procedure in question, suggesting that the requirement for review acted as a deterrent, either to the physician (against performing the procedure) or to the patient (who might sometimes have been seeking cosmetic or other surgery not covered by his or her insurance).

The actual utilization review did not differ from the sham review with regard to any of the other "efficiencies" sought. The average number of hospital days per group member, a summary measure of inpatient utilization, was similar in both groups. Members of the review group actually used 7.4 percent fewer inpatient days than were approved for insurance coverage. Attending physicians may propose more days than they believe are necessary, to "bargain" with utilization reviewers.

Case-management nurses coordinate patient care (for

example, facilitating the delivery of specialized services in high-risk pregnancies) and authorize services not usually covered by insurance, in an attempt to reduce the number of patients with extremely high overall utilization and costs. There were, however, at least as many patients in the review group as in the nonreview group whose total insurance claims were higher than every dollar level used to define high cost. These findings arouse the suspicion that the services authorized by case-management nurses often supplemented rather than substituted for routine services.

There are several possible explanations for our inability to find differences between the effects of actual review and those of sham review. Utilization review may have little effect on the use of health care, or this particular review program may not have been effective. It is also possible that this program of utilization review and others, actual or sham, influence use of services, but achieve their results primarily through the deterrent effect that their mere existence exerts, rather than through changes in the care of individual patients.[1,3,15] Patients, their physicians, or both may choose more efficient options for care when they expect that care to be reviewed. Several studies have documented substantial reductions in hospital use when review programs are first implemented but have not detected any effect on the rate at which such use increases in subsequent years.[6,7,10] Unfortunately, we could not measure the deterrent effect of the utilization-review program we studied by comparing hospital use before and after the introduction of the program, because its components (second opinions, case management, and the like) had been implemented progressively over a four-year period.[22]

We analyzed utilization review in New York City at a time when hospital occupancy exceeded 87 percent[23] and most hospitals were reimbursed on a per-admission basis, so that they had little incentive to increase the number of admissions or prolong hospital stays. In an area with low occupancy rates and per diem payments, actual and sham review might have measurably different effects. It also may be that the effects of utilization review on fee-for-service health care differ from its effects in health maintenance organizations.

Because referrals for a second opinion for outpatient procedures appeared to produce measurable results, New York City and its unions expanded their second-opinion requirements to include additional outpatient procedures (for example, magnetic resonance imaging). Since deterrence may also have had an effect, the city continued all its other utilization-review requirements but limited them to a periodically revised subgroup of diagnostic and procedural categories. This reduced the inconveniences imposed on patients and their physicians and decreased the program's operational cost per insurance contract per month from $1.34 in 1991 to $1.19 in 1992. We believe second opinions and deterrence are the least coercive mechanisms by which utilization review may affect care. Second opinions provide information that most patients consider an aid to decision making.[22] Deterrence provides physicians with an incentive to choose more cost-effective alternatives.

Table 3. Proportion of Enrollees Whose Total Insurance Claims during the Study Period Exceeded Various Levels, According to Group.

Total Amount of Claims ($)	Review Group (N = 3702)	Nonreview Group (N = 3743)	P Value
	no. (%)		
≥10,000	699 (18.9)	689 (18.4)	0.62
≥20,000	291 (7.9)	282 (7.5)	0.63
≥30,000	156 (4.2)	144 (3.8)	0.46
≥40,000	85 (2.3)	70 (1.9)	0.23
≥50,000	54 (1.5)	36 (1.0)	0.06
≥100,000	10 (0.3)	4 (0.1)	0.17
≥150,000	1 (<0.05)	1 (<0.05)	1.00
≥200,000	0	1 (<0.05)	1.00

Table 4. Additional Indicators of the Effect of Utilization Review during the Study Period.

Indicator	Review Group (N = 3702)	Nonreview Group (N = 3743)	P Value
Admissions per 1000 group members			
Medical–surgical unit	641.8	625.4	0.15
Substance-abuse unit	7.3	9.4	0.40
Psychiatric unit	17.6	16.8	0.88
Average stay (days)			
Medical–surgical or obstetrical unit			
Before age adjustment	6.1	5.9	0.37
After adjustment	6.1	5.8	0.18
Substance-abuse unit			
Before age adjustment	9.7	10.2	0.86
After adjustment	9.3	10.6	0.62
Psychiatric unit			
Before age adjustment	18.1	19.6	0.62
After adjustment	18.1	19.6	0.62
Preadmission testing (% of surgical admissions)	41.1	42.3	0.48
Ambulatory hospital-based surgery (% of all hospital surgery)*	13.9	14.9	0.46
Home care claims per 1000 group members	32.1	32.1	0.94

*Excludes surgical procedures in the 20 selected categories.

Table 5. Overall Use of Inpatient Care during and after the Study.

	Review Group (N = 3702)	Nonreview Group (N = 3743)	P Value
Study period			
Inpatient days approved	20,792	Not applicable	—
Inpatient days used	19,248	18,445	—
Days used per group member	5.20	4.93	0.49
Following year			
Inpatient days used			
All	6873	6223	—
Per group member	1.86	1.66	0.33

We are indebted to Allen Brawer, Jeffrey Goldstein, David Hickey, Jeffrey Kahn, Donna Lynne, Antonio Morelli, Hal Perlis, Abraham Steinberg, Caroline Sullivan, Joseph Wohl, and Dorothy Wolfe for assistance in the design and implementation of this study; to Joseph Angarola, George Bopp, Kevin Connors, Christopher Davis, Helene Hellman, Agnes Kammerer, Lorrayne Lindsey, Robert O'Brien, Natalie Richstone, David Sellar, Bette Senne, Carol Shanik, Thomas Voss, Jr., and Charlene Weigel for assistance in creating and operating data-collection systems; and to Joseph Fleiss, Sherry Glied, Sheila Gorman, and Alex Tytun for advice on the analysis and presentation of data.

REFERENCES

1. Managed health care: effect on employers' costs difficult to measure. Washington, D.C.: General Accounting Office, 1993. (Publication no. GAO/HRD-94-3).
2. Khandker RK, Manning WG, Ahmed T. Utilization review savings at the micro level. Med Care 1992;30:1043-52.
3. Ermann D. Hospital utilization review: past experience, future directions. J Health Polit Policy Law 1988;13:683-704.
4. Bailit HL, Sennett C. Utilization management as a cost-containment strategy. Health Care Financ Rev 1991;Annu Suppl:87-93.
5. Henderson MG, Souder BA, Bergman A, Collard AF. Private sector initiatives in case management. Health Care Financ Rev 1988;Annu Suppl:89-95.
6. Wickizer TM, Wheeler JRC, Feldstein PJ. Does utilization review reduce unnecessary hospital care and contain costs? Med Care 1989;27:632-47.
7. Feldstein PJ, Wickizer TM, Wheeler JRC. Private cost containment: the effects of utilization review programs on health care use and expenditures. N Engl J Med 1988;318:1310-4.
8. Scheffler RM, Sullivan SD, Ko TH. The impact of Blue Cross and Blue Shield plan utilization management programs, 1980–1988. Inquiry 1991; 28:263-75.
9. Wickizer TM. Effect of hospital utilization review on medical expenditures in selected diagnostic areas: an exploratory study. Am J Public Health 1991; 81:482-4.
10. Payne SMC, Campbell D, Penzias BG, Socholitzky E. New methods for evaluating utilization management programs. Qual Rev Bull 1992;18:340-7.
11. Wickizer TM. The effects of utilization review on hospital use and expenditures: a covariance analysis. Health Serv Res 1991;27:103-21.
12. *Idem.* The effect of utilization review on hospital use and expenditures: a review of the literature and an update on recent findings. Med Care Rev 1990;47:327-63.
13. Korda H. Utilization review for Medicaid diagnosis-related group systems: practice, innovation, and lessons of experience. Am J Med Qual 1994;9:54-67.
14. Mechanic D. Managed care: rhetoric and realities. Inquiry 1994;31:124-8.
15. Zusman J. Utilization review: theory, practice, and issues. Hosp Community Psychiatry 1990;41:531-6.
16. Rosenberg SN, Stamm JB, Dowling S, Lynne D. Combining prospective managed care and retrospective review to contain cost. Managing Employee Health Benefits 1992;1:27-33.
17. Epstein AM. The outcomes movement — will it get us where we want to go? N Engl J Med 1990;323:266-70.
18. Daniel WW. Biostatistics: a foundation for analysis in the health sciences. 5th ed. New York: John Wiley, 1991:121-43.
19. Analysis of covariance. In: Winer BJ. Statistical principles in experimental design. New York: McGraw-Hill, 1971:752-812.
20. Health Insurance Association of America. Providing employee health benefits: how firms differ. Washington, D.C.: Health Insurance Association of America, 1990.
21. Lagoe RJ, Milliren JH. Changes in ambulatory surgery utilization 1983–88: a community-based analysis. Am J Public Health 1990;80:869-71.
22. Rosenberg SN, Gorman SA, Snitzer S, Herbst EV, Lynne D. Patients' reactions and physician-patient communication in a mandatory surgical second-opinion program. Med Care 1989;27:466-77.
23. AHA hospital statistics: a comprehensive summary of U.S. hospitals. 1991–92 ed. Chicago: American Hospital Association, 1991:156.