The Regional Educational Impact of a Renal Stone Center

JAMES F. JEKEL, M.D., M.P.H.,a
KIMBERLY J. CHAUNCEY, B.S.N., M.P.H.,b
NICOLA L. MOORE, M.P.H.,c ARTHUR E. BROADUS, M.D., Ph.D.,d
AND DAVID P. W. GOWDY, M.D., M.P.H.e

aProfessor of Epidemiology and Public Health, Yale University School of Medicine, New Haven, Connecticut; bManagement Consultant, APM, Inc., New York, New York; cConsultant, Arthur D. Little, Inc., Boston, Massachusetts; dAssociate Professor of Medicine, Yale University School of Medicine, New Haven, Connecticut; eResident in Medicine, New Jersey Medical School, Newark, New Jersey

Received April 12, 1983

This study evaluates the impact of the Renal Stone Center (RSC) at Yale University School of Medicine on the extent of the diagnostic search for urolithiasis risk factors at the Yale-New Haven Medical Center and four affiliated community hospitals. Using a multiple time series design and a Physician Performance Index (PPI) based on multiple normative practice criteria, the hospital and outpatient charts of 1,924 patients with a primary diagnosis of urolithiasis at the five hospitals were examined to determine whether there was a change over time that could be related to the RSC. There was no statistically significant improvement in the PPI at four of the five hospitals. At the fifth, there was a statistically significant improvement that possibly was related to factors other than the RSC. Routine blood, urine, and X-ray tests usually were performed well. However, the history of urolithiasis risk factors and the examination of 24-hour urines were generally done poorly or not at all, and this generalization held true (although modified somewhat) even when physician office records were examined as well.

In September 1977, the Yale-New Haven Medical Center became one of a limited number of federally assisted specialized centers for research into the causes and treatment of urinary tract stones. This center, hereafter called the Renal Stone Center (RSC), also served as a regional referral center for patients with unusual or recurrent stones and metabolic problems related to urolithiasis. Moreover, the center was intended to have an educational impact at the medical center and in hospitals throughout the southern Connecticut region. The RSC was given a modest grant to test the hypothesis that its activities would have an educational impact on physicians and surgeons in the region, in the direction of encouraging and helping them to perform more complete hospital work-ups for underlying metabolic risk factors among patients with urolithiasis. This paper reports the results of testing that hypothesis using a multiple time series design. It does not examine the quality of the medical care for the existing stone, which seemed generally good, although the style of care varied by hospital. The first paper from this study reported details of the

This study was supported by grant AM 20570 from the National Institute of Arthritis, Metabolism, and Digestive Diseases.

Address reprint requests to: James F. Jekel, M.D., Department of Epidemiology and Public Health, Yale University School of Medicine, Box 3333, New Haven, CT 06510

Copyright © 1983 by The Yale Journal of Biology and Medicine, Inc.
All rights of reproduction in any form reserved.
measurement of physician performance and some descriptive findings regarding the patterns of diagnostic evaluation of urolithiasis risk factors [1].

**REVIEW OF THE LITERATURE**

The literature evaluating the impact of continuing medical education programs on physician care patterns is not large, and it is contradictory. Laxdal et al. [2] and Rubenstein [3], for example, found an impact on patterns of care from targeted programs based on demonstrated quality of care needs, although the duration of this effect varied. On the other hand, Sibley et al. did not find an impressive effect on physician practice patterns from their own carefully controlled educational intervention and evaluation, which, however, suffered from a small number of study subjects and a rather short duration of the intervention [4]. When they analyzed the methodologic adequacy of studies evaluating continuing medical education, they found that none satisfied key criteria for validity and generalizability.

Some believe that it is easier to show the effect of continuing medical education on prescribing behavior than on most other aspects of clinical practice, although the Rubenstein study showed impact over eight months on such behaviors as: use of packed cells rather than whole blood, use of partial thromboplastin time tests rather than the Lee-White method, use of intravenous rather than subcutaneous heparin, and increased use of lung scans and blood creatinine determinations [3]. Williamson et al. found that physicians responded to “unexpected” abnormal laboratory results with no action in 55–65 percent of the cases, and this was not improved either by a conference on this topic or by frequent warnings in a hospital newsletter [5]. It was, however, improved by placing a warning tape over the abnormal evaluations. These studies are all evaluations of specifically designed interventions in continuing education.

More general association studies have shown even less effect. The classic study of medical practice by Osler Peterson found that, among other things those physicians with high scores on the quality of care in their office practices were no more likely to have attended continuing education activities than were those scoring low on quality of care [6]. Lewis and Hassanein did not find any association by area in Kansas between the frequency with which the physicians there attended continuing education programs and the reduction in infant and maternal mortality or abnormally high rates of surgery [7]. On the other hand, Roney and Roark did find on visiting individual practices that those attending many educational sessions did somewhat more thorough work-ups, although there was no evidence for a greater rate of change in practice patterns than in those physician practices where they did not attend many continuing education efforts [8]. Thus, the evidence is contradictory regarding the effectiveness of continuing medical education, whether looked at through large cross-sectional surveys or by careful evaluations of specific intervention efforts.

The general conclusion of the literature appears to be that continuing medical education is most effective when it is limited, well-defined, intensive, and continuous, especially when the definition of topics for education comes out of study of practice patterns and needs at a specific hospital or other defined practice group.

We were not able to find any studies evaluating the influence of continuing medical education on the patterns of diagnostic search for the metabolic risk factors for urolithiasis. This is not surprising, given the dearth of such evaluative studies in general. Here we are not, however, evaluating a planned, organized (imposed) continuing medical education effort. Rather, this study seeks to find if there are positive
EDUCATIONAL IMPACT OF A RENAL STONE CENTER

99

educational spillover effects from a center designed to focus on research and patient care regarding one related group of diseases. In a sense, this study is looking to see if there is a serendipitous educational benefit from an effort such as the Renal Stone Center.

METHODS

Research Design

Some form of clinical trial, in which hospital staffs were randomly exposed or not exposed to the intervention (the RSC) would be the preferred research design. However, the other hospitals in Connecticut either were affiliated with the Yale University School of Medicine or might receive educational intervention from physicians in the RSC even though randomized to the control group. Physicians from other hospitals in the region also potentially had contact with the RSC by means of referral of stone patients to the stone center’s clinic for metabolic work-up. Thus, no control hospital within the state was possible. Moreover, many referrals came from outside the state of Connecticut, so that a control hospital outside Connecticut would be methodologically, as well as logistically, difficult. Therefore, following the principles of Campbell and Stanley, we chose to use a quasi-experimental design [9]. Originally the medical center and two affiliated hospitals were selected to participate in the study, using a multiple time series design. The dependent variable was the level of physician performance in searching for urolithiasis risk factors in hospitalized patients. This was determined at two points prior to the beginning of the RSC and at two points following its inception. Improvement in the level of work-up was expected if the hypothesis of educational impact was sustained. If there was a trend toward improved work-ups before the beginning of the RSC, the expected continuation of this improvement would have to be controlled for in analyzing the degree of observed pretest–post-test improvement. This design would be strong in controlling for selection biases, as, for the most part, the same attending physicians would be present over time. It would also control for instrumentation biases, because the same laboratories and medical records would be used. The primary weakness of this quasi-experimental design would be failure to control for historical events, such as non-RSC educational efforts or scientific reports that might occur during the study period and be unrelated to the RSC efforts. Using this design, if no improvement was seen in physician performance, this would strongly suggest no effect from the RSC, but a significant improvement in performance would be more difficult to interpret. In light of these considerations, charts of all patients with a primary diagnosis of urinary tract stone at each study hospital were examined for the extent of risk factor work-up during the years 1974, 1976, 1979, and 1981. The RSC began in September 1977.

Sample

The initial stratified sample included the university medical center and two affiliated community hospitals. The community hospitals were located in urban areas; however, one urban area was served by several hospitals, whereas the other was served by only the study hospital, which had a fairly clearly defined catchment area. The two strata for affiliated hospitals were used both because of possible differences between the type of hospital and to enable some epidemiologic studies to be done in the latter type of hospital.

After the study had begun, the investigators began to wonder if the presence of
the study, including the interaction with physicians in the study hospitals regarding the details of the study, might have an impact beyond that of the RSC's other activities. Therefore, it was decided to study two other hospitals, one in each of the above strata, in a retrospective manner, so that the investigation itself would have no significant impact on the data observed. If the study itself were a major intervention, presumably this would be seen in greater before-after differences in the three hospitals studied prospectively than in the two studied retrospectively. Therefore, in the final hospital sample there were five hospitals: one medical center, two hospitals in large urban areas with competing hospitals, and two hospitals which alone served smaller urban areas.

Table 1 shows the number of discharges studied by study year and hospital. A total of 1,924 discharges were examined in five hospitals at four different time periods. It should be noted that these are discharges, not separate patients. This was done for simplicity and completeness and did not greatly change the numbers, because there were few readmissions during the same calendar year. Also, if any readmissions were rapid and were clearly for the same stone event, the data were combined to make one study discharge. If the repeat discharges for one patient appeared to represent separate stone events, they were studied separately.

In each hospital, every patient discharged during the study years with a primary discharge diagnosis of ureteral or renal stone was studied, except for those patients with chronic urinary tract infection and stone formation, usually due to paraplegia. Hospital charts were the data base. Questions have been raised as to whether hospital charts give a fair representation of what is done in a hospital, but the authors hold the view that good recording is an essential part of good care, both for that admission and for subsequent admissions. Moreover, because many of the essential elements of care were X-ray or laboratory tests, these reports would be expected to be in the charts if they had been performed. It is important to remember that the primary study hypothesis concerns changes within each hospital over time, and even though there might be systematic differences in recording between hospitals, one would expect recording within hospitals to remain relatively constant over the study period. Therefore, it was assumed that the level of chart recording did not produce changes over time within each hospital, at least in those variables incorporated into the physician performance measures.

**Educational Efforts of the Renal Stone Center**

There were two major kinds of educational efforts by the RSC during the grant period. First, there were the traditional educational sessions, such as grand rounds.

| Hospital | 1974 | 1976 | 1979 | 1981 | Total |
|----------|------|------|------|------|-------|
| A        | 114  | 138  | 133  | 116  | 501   |
| B        | 85   | 95   | 94   | 101  | 375   |
| MC       | 101  | 109  | 90   | 116  | 416   |
| D        | 61   | 89   | 102  | 100  | 352   |
| E        | 70   | 62   | 71   | 77   | 280   |
| Total    | 431  | 493  | 490  | 510  | 1,924 |
and conferences. However, this type of intervention was not under the control of the RSC physicians, because the content of the formal educational sessions at each hospital was determined by the physicians at that hospital. The RSC physicians only lectured on renal stone risk factors and work-ups when requested to do so by the hospitals. Moreover, the metabolism specialists at the RSC, who were internists, were less likely to be asked to lecture to the physicians caring for the majority of stone patients, the urologists.

The other component of the educational thrust concerned referrals from Connecticut physicians to the stone clinic for metabolic work-ups of patients with recurrent or problem stones. Both telephone and written contacts with referring physicians served an educational role. Approximately 300 referrals were made to the stone clinic from physicians associated with the study hospitals between late 1977 and 1981. Table 2 shows the kinds of educational sessions given in each of the hospitals during the study period. In Hospital A the primary contact was personal discussions with the physicians and urologists most heavily involved in the care of patients with stones. In Hospital B, both individual contact and a medical grand rounds occurred, in which early study findings were reported. Here we first made an observation that may explain many of the findings: urologists and internists tend not to attend each other's conferences, even when specifically invited to do so; therefore, there is relatively little cross-fertilization in which each learns from the other. Hospital D had one grand rounds on urolithiasis, unattended by any urologists, and Hospital E did not request any educational sessions on renal lithiasis during the study period. However, during this period, Hospital E did acquire a nephrologist on the staff who had an interest in urolithiasis, and who sought to influence the care patterns for urolithiasis patients at that hospital.

| Hospital A | Three medical grand rounds given on hyperparathyroidism (1978, 1980, 1981) |
|-------------|--------------------------------------------------------------------------------|
|             | Two conferences with two urologists, the director of medical education, and the chief of medicine regarding the stone study criteria for a satisfactory work-up |
| Hospital B  | A conference in 1977 with a hospital specialist in metabolism regarding the study criteria |
|             | Medical grand rounds presenting early findings of the study at that hospital (1978) |
|             | A symposium on calcium metabolism and kidney stones (1981) and a medical grand rounds on hyperparathyroidism (1981) |
| Hospital MC | Frequent medical conferences on calcium metabolism, hyperparathyroidism, and other endocrine subjects related to renal stones; these were seldom attended by urologists |
|             | Two presentations at the weekly urology conference for the medical center and affiliated hospitals |
| Hospital D  | Three medical grand rounds on metabolic bone disease (1977, 1978, 1979) |
|             | One medical grand rounds on urolithiasis and urolithiasis risk factors, to which urologists were invited but did not attend (1979) |
| Hospital E  | A nephrologist with interest in nephrolithiasis joined the paid hospital staff during the study period (1978) |
|             | Medical grand rounds on hypercalcemia (1980, 1981), and on hyperparathyroidism (1980) |
Of considerable interest is the fact that the urologists in the affiliated hospitals regularly meet early Monday mornings at the medical center to discuss cases or hear presentations. Twice during the study period the director of the RSC spoke to this group regarding risk factors for urolithiasis and how to work up patients for these metabolic risk factors. This may have been the educational intervention with the best chance to have had an impact on the care of a large number of patients with urolithiasis, and this effect presumably would have been felt in all of the affiliated hospitals.

The Physician Performance Index

After discussion with physicians from the RSC and with internists and urologists at affiliated hospitals, normative performance indexes were developed for use in this study. The idea of a “Physician Performance Index” comes from the literature on the quality of medical care and describes a technique in which several components of what is accepted as good care are given weights depending on their perceived importance, and the performance index is the sum of the weights earned for a given patient [1]. These scores are usually expressed as a percentage of the possible score. For the purposes of this study, two basic Physician Performance Indices (PPIs) were developed, one for patients with first event, uncomplicated stone disease, and the other for patients with either complicated first stones or with recurrent stone disease. A patient’s stone disease was considered complicated if it showed as bilateral or staghorn stones, if the patient was less than 20 years of age, or if there was a diagnosis of nephrocalcinosis. Because many persons with urinary tract stones do not subsequently get others, unless there are abnormalities on a basic screen or reasons for special concern, only a relatively simple basic screening effort is all that is usually recommended. The criteria for the PPI for both first, uncomplicated stones or for recurrent or complicated stones are given in detail in the previous paper [1]. All patients with stones were expected to have: an adequate history for stone number, a family history, a history of possible risk factors for urolithiasis, an X-ray, a urinalysis, and blood tests. For recurrent or complicated stones, to these criteria were added requirements for a urine culture, and the performance of a 24-hour urine collection for calcium and uric acid (and creatinine to estimate the adequacy of the collection). For patients with first, uncomplicated stones, a perfect PPI score was eleven; for patients with complicated or recurrent stones, a perfect score was sixteen.

Methods of Analysis

The test of the hypothesis basically compared the average Physicians’ Performance Indices (PPIs) before and after the inception of the RSC. The method of analysis first determined whether there were any significant differences in the PPI between 1974 and 1976, the pre-intervention years. There were no important or statistically significant differences in the PPIs between these two years for any of the hospitals, so these data were pooled to determine the percentage adequacy on each PPI for each hospital. Likewise, the post-intervention years (1979, 1981) were compared and, due to lack of statistically significant differences, pooled at each hospital. The pooled data were, therefore, appropriate for comparison using a standard two-group (unpaired) t-test.

When comparing the before-after scores, the statistical problem of multiple comparisons arose. Because each hypothesis was tested in five hospitals, there were five
opportunities for the null hypothesis to be rejected on chance alone. Therefore, the probability that one or more of the hospitals might differ by chance alone on the before-after comparison at the 5 percent level was not 0.05 but almost five times that. Therefore, in order to reduce the danger of falsely rejecting the null hypothesis, we used the Bonferroni correction and divided the alpha level (which we set at the customary 0.05) by the number of comparisons for each hypothesis, giving us an alpha value of 0.01 for these analyses. Because of the large sample size, approximately 400 observations in each hospital, beta error was not an important problem.

**FINDINGS**

*Comparison of Overall Physician Performance Scores*

When each hospital's data were pooled into the before and after RSC periods, (i.e., 1974 and 1976 combined vs. 1979 and 1981 combined), there was a statistically significant improvement in only one hospital (Hospital E) as shown in Table 3. Of the other four hospitals, two improved slightly and two became slightly worse on the PPI scores. In these statistics, an adjustment was made to eliminate the effect of an improvement in the laboratory's performance of urine tests in Hospital A, which was not due to the RSC; the details of this are discussed in the previous report [1].

Overall, the study hypothesis was supported by data from only one hospital, and there the gain in percentage score was relatively modest, from 59.3 percent of "perfect" average score before to 64.2 percent after the RSC, a gain of about 5 percent. Despite the fact this gain was statistically significant, even at the level set by the Bonferroni adjustment, the practical importance of this 5 percent improvement is

| HOSPITAL | All Patients | Pre-RSC | Post-RSC | t  | p     |
|----------|--------------|--------|----------|----|-------|
| A        | 54.8         | 56.7   | 1.44     | 0.150 |
| B        | 71.2         | 69.6   | 0.99     | 0.321 |
| MC       | 63.6         | 62.7   | 0.73     | 0.464 |
| D        | 57.8         | 57.7   | 0.10     | 0.922 |
| E        | 59.3         | 64.2   | 2.83     | 0.005*|

*Statistically significant at p < .01

| HOSPITAL | First, Uncomplicated Stones | Pre-RSC | Post-RSC | t  | p     |
|----------|-----------------------------|--------|----------|----|-------|
| A        | 54.4                        | 58.4   | 1.93     | 0.055 |
| B        | 71.2                        | 70.1   | 0.51     | 0.612 |
| MC       | 67.4                        | 61.6   | 2.94     | 0.004*|
| D        | 57.1                        | 59.2   | 0.93     | 0.353 |
| E        | 60.9                        | 66.4   | 2.32     | 0.022*|

| HOSPITAL | Recurrent or Complicated Stones | Pre-RSC | Post-RSC | t  | p     |
|----------|--------------------------------|--------|----------|----|-------|
| A        | 55.2                          | 54.5   | 0.47     | 0.639 |
| B        | 71.2                          | 68.8   | 1.01     | 0.315 |
| MC       | 60.2                          | 63.3   | 1.80     | 0.737 |
| D        | 59.5                          | 56.1   | 1.75     | 0.827 |
| E        | 56.0                          | 60.2   | 2.27     | 0.026*|

*Statistically significant at p < .01
questionable. For example, Sibley et al. suggest that an improvement of less than 10 percent is clinically unimportant [4].

Similar findings are seen in the remaining parts of Table 3, where the total results are subdivided into scores for patients with first, uncomplicated stones and those with recurrent or complicated stones. For patients with first stones, there was a statistically significant drop in score at the medical center in patients with first stones, but this was balanced by an increase in the score for patients with complicated or recurrent stones, so that the overall change in score was not statistically significant. Hospital E's overall gain was reflected in the scores from both the first and recurrent stone work-ups, as both were statistically significant at the .03 level.

In an attempt to ascertain the source of the improvement for Hospital E, both those with first and recurrent stones were evaluated on only the “basic” or eleven-point index used for the patients with first stones. This was to see if the severity of the stone disease influenced the extent of the basic work-up. There was no significant difference between the scores for patients with first and recurrent stones on this eleven-point scale, but there was a gain over time when all patients were compared on this scale, from 62.7 percent before the RSC to 68.2 percent after its inception ($p = 0.003$). Likewise, in Hospital E, when all patients were evaluated on the 16-point scale designed for patients with complicated or recurrent stones, no difference was seen between the average scores for first stone work-ups vs. recurrent stone work-ups, but a statistically significant gain was seen from before to after the inception of the RSC, from 53.9 percent to 59.0 percent ($p = 0.001$). Examining only the five extra points added to the basic work-up for patients with recurrent or complicated stones, there was no difference at any point in time between first stone and recurrent stone work-ups, but there was an improvement over time when all patients were compared only on these five points, where the percentage of possible score increased from 34.7 percent to 39.1 percent. Regarding individual tests, Hospital E improved most on the performance of routine blood work upon admission and in obtaining a serum BUN or creatinine in order to estimate kidney function.

Why should Hospital E be the only hospital to improve significantly, and why should this improvement persist across different performance criteria and with different types of patients? This was especially troublesome to the original study hypothesis, because Hospital E appeared to have the fewest relevant clinical conferences and appeared to have fewer referrals to the stone clinic than did most of the other study hospitals. One factor may explain most of the observed difference: Hospital E obtained a paid hospital nephrologist with an interest in kidney stone diagnosis and treatment. He worked with the urologists and internists to improve care for stone patients. One might wonder if the improvement was due to the fact that the patients of this specialist were worked up more thoroughly, thereby enriching the average score. However, this physician did not admit many patients with stones to the hospital, and for those he did treat in the hospital, his scores were not significantly different from those of the other physicians. His own tendency was to do the advanced stone work-ups on an ambulatory basis. Therefore, the improvement in scores is due to a general increase among the scores of practitioners and not due to this specialist's patient work-ups. It is likely that his contribution was seen especially in the improvement in the proportion of patients having renal function tested (from 84 percent to 100 percent) and the proportion who had renal tubular acidosis ruled out (from 77 percent to 90 percent).
Another way to analyze the data is to consider the median score and range for each item in the five hospitals and see whether the range and median scores changed over time. Table 4 shows that neither the median score nor the range changed very much, with two exceptions, which are related to each other. As previously mentioned, Hospital A scored very poorly on urinalysis in 1974 due to the laboratory, which did not report the pH values in urinalysis, and this also made its score for ruling out RTA rather low. This was corrected by 1979, after the RSC, but the change was not due to the RSC, and, therefore, the effects of this change were removed from the analysis. The only other definite before-after change was in the kidney function tests in one hospital (E), which is the reason for the improvement in this range and median. Looking at the data in this way confirms the general conclusion that improvement in scores was, at best, modest. When the non-RSC related improvement in Hospital A was eliminated from consideration, no important improvements were found in the ranges of PPI scores from before to after the RSC.

Another observation from Table 4 is that the scores were fairly good except in the areas of medical stone history, the performance of a 24-hour urine, and the ade-

| Item                      | Time  | Range of Scores (%) | Median Score (%) |
|---------------------------|-------|---------------------|------------------|
| Stone history             | Before| 62-86               | 73               |
|                           | After | 57-82               | 66               |
| Other history items       | Before| 7-27                | 17               |
|                           | After | 11-34               | 18               |
| Blood tests for calcium and UA | Before| 74-92               | 84               |
|                           | After | 62-93               | 86               |
| Urinalysis                | Before| (7)*                | 96               |
|                           | After | 94-99               | 96               |
| Rule out RTA              | Before| (62)*               | 92               |
|                           | After | 82-100              | 95               |
| Stone analysis            | Before| 68-89               | 88               |
|                           | After | 76-92               | 85               |
| X-ray exams               | Before| 90-98               | 97               |
|                           | After | 88-99               | 98               |
| Performance of 24-hour urines | Before| 0-27                | 8                |
|                           | After | 2-16                | 15               |
| Adequacy of 24-hour urines | Before| 0-16                | 1                |
|                           | After | 0-13                | 3                |
| Adequacy of urine culture | Before| 80-92               | 89               |
|                           | After | 79-90               | 87               |
| Kidney function tests     | Before| 80-98               | 94               |
|                           | After | 91-99               | 98               |

*Unusually low due to laboratory problem in Hospital A, because the laboratory did not report the urine pHs during the pretest years, and the improvement for the post-test years was not due to the RSC. For this reason, this value was indicated separately. More details are given in [1].
quacy of that performance. Thus, of necessity, most of the improvement in overall scores would have had to come from these weak areas, and, in fact, little improvement actually did come from them.

**Outpatient Work-ups**

Some physicians felt that the optimal time for doing many of the chemical investigations for stone risk factors was on an outpatient basis, when the patients had returned to normal diet and activity. This argument was reasonable, but we saw little evidence from the hospital charts that much outpatient work-up was done, even though outpatient chemistries done at the hospital clinics would have been entered into the hospital charts at most of the hospitals, and seldom were such studies found. Moreover, data obtained on an outpatient basis prior to an admission should have been mentioned (and even the data recorded) in the hospital chart for subsequent admissions, and this was almost never found in patients with several admissions, even when readmitted by the same physician. Therefore, with permission, we examined the office charts of the three urologists at one of the study's community hospitals, examining only those patients' records for whom an indicated 24-hour urine was not performed in the hospital during the index admission nor was it subsequently recorded in the hospital record.

A total of 87 office charts were examined for evidence of outpatient 24-hour urines for calcium, uric acid, and other indicated chemistries. Although the extent of work-up and the degree of change over time varied among the three physicians, overall, 21 of 87 patients (24 percent) did have the indicated work-up done on an ambulatory basis (Table 5). This statistic improved from before the RSC (17 percent) to after the RSC (32 percent). This change, though suggestive, was not statistically significant (chi-square on one d.f. = 2.42, N.S.). Although the office study implies that incorporating office data into the PPI scores would have improved the PPI scores somewhat, the lack of statistically significant improvement in office diagnosis over time implies that omitting office data from the PPIs does not bias the before-after comparison. Therefore, apart from a suggestion that the impact of the RSC might have been greater on outpatient practice than on inpatient work-ups, the conclusion remains that patients with recurrent or complicated stones often do not get fully worked up for metabolic risk factors.

**TABLE 5**

Analysis of Office Performance of 24-Hour Urine Collections for Calcium and Uric Acid among Patients with Such Collection Indicated But Not Done in the Hospital, Three Urologists, Affiliated Community Hospital

| 24-Hour Urine Collection Done? | Before Renal Stone Center | After Renal Stone Center | Total |
|-------------------------------|--------------------------|--------------------------|-------|
|                               | No. (%)                  | No. (%)                  | No. (%) |
| Yes                           | 8 (17)                   | 13 (32)                  | 21 (24) |
| No                            | 38 (83)                  | 28 (68)                  | 66 (76) |
| Total                         | 46 (100)                 | 41 (100)                 | 87 (100) |

Chi-square = 2.42 on 1 d.f. \( p > 0.05 \) N.S.
DISCUSSION

The lack of greater evidence of RSC impact on physician diagnostic practices in urolithiasis was a disappointment, but it probably should not have been a surprise. It is perhaps naive to assume that a program of any kind with first priorities in one direction (here, research and patient care) should have a profound impact in a somewhat different area (education). There was no lack of effort, and there was considerable teaching contact. Nevertheless, a consistent lesson in many areas of social intervention is that it is difficult to achieve progress in planned interventions even when the resources and program direction are under the control of those doing the intervention. Success is even less likely if, as here, those conducting the intervention must be in the role of "offering services" and do not actually have the authority to develop and execute a clearly planned educational intervention from the beginning. The RSC staff had the responsibility for deciding neither the basic topic, nor the audience or frequency of their talks, seminars, rounds, or grand rounds at the affiliated hospitals. Nor did they control who referred patients to them from Connecticut and the adjoining states. Therefore, although isolated teaching sessions were provided with reasonable frequency, these could not be considered a carefully integrated educational program.

Another important issue of medical education concerns who attends the sessions that are offered. Here there is even less control, even by the hospitals. It became obvious that the majority of patients with urolithiasis were primarily cared for by urologists (70–90 percent, depending on the hospital), and yet, with rare exceptions, the staff of the RSC were asked to speak to educational sessions primarily planned for internists, where urologists were almost never in attendance. Therefore, one of the weaknesses of the educational intervention was simply that an important target group was largely missed by the intervention, at least by the formal educational sessions. On the other hand, a number of urologists referred one or more patients to the RSC for metabolic work-up, and undoubtedly the communication with the RSC staff was an intervention of sorts. The most effective behavior change appeared to occur in the situation where a committed individual had the best opportunity for individual discussion and intervention.

We believe this experience should be a warning regarding educational expectations. Unless there is control over what is taught, to whom, and how the education is done, it is probably unreasonable to expect a favorable outcome from the effort. The one instance here where a positive impact was seen was characterized by education from within, on a one-to-one basis. Although this cannot be considered proof, it would seem to lend some support to the educational value of service chiefs within hospitals. There was also the suggestion that the impact of the RSC on office practice may have been greater than on inpatient care patterns.

REFERENCES

1. Jekel JF, Chauncey KJ, Moore NL, et al: The diagnostic evaluation of risk factors for urinary tract stones: an analysis of care patterns in five hospitals. Yale J Biol Med 55:89–100, 1982
2. Laxdal OE, Jennett PA, Wilson TW, et al: Improving physician performance by continuing medical education. CMA J 118:1051–1058, 1978
3. Rubenstein E: Continuing medical education at Stanford: the back-to-medical-school program. J Med Educ 48:911–918, 1973
4. Sibley JC, Sackett DL, Neufeld V, et al: A randomized trial of continuing medical education. New Eng J Med 306(9):511–515, 1982
5. Williamson JW, Alexander M, Miller GE: Continuing education and patient care research. JAMA 201(12):938–942, 1967
6. Peterson OL, Andrews LP, Spain RS, et al: An analytic study of North Carolina general practice: 1953–1954. J Med Educ 31(12–part 2):1–165, 1956
7. Lewis CE, Hassanein RS: Continuing medical education—an epidemiologic evaluation. New Eng J Med 282(5):254–259, 1970
8. Roney JG, Roark GM: Continuing Education of Physicians in Kansas: an Exploratory End Result Study. Menlo Park, CA, Stanford Research Institute, 1975
9. Campbell DT, Stanley JC: Experimental and Quasi-Experimental Designs for Research. Chicago, Rand McNally College Publishing Company, 1963