

Routine Individual Feedback on Requests for Diagnostic Tests:

An Economic Evaluation

RON A. G. WINKENS, MD, PhD, ANDRÉ J. H. A. AMENT, PhD,
PETER POP, MD, PhD, PETER H. A. RENIERS,
RICHARD P. T. M. GROL, PhD, J. ANDRÉ KNOTTNERUS, MD, PhD

The authors assessed the economic consequences of routine individual feedback on test requests provided to 85 family physicians in a region with 187,000 inhabitants. In a retrospective study as part of a quasi-experiment, cost trends in a region where feedback was provided over a seven-year period were compared with cost trends elsewhere in The Netherlands without feedback. Data on variable costs were obtained for 400 individual tests that accounted for 90% of all requests. Differences in request trends thus were transformed to savings in costs of diagnostic testing, taking account of the extra costs of providing the feedback. Expenditures for diagnostic testing declined after the start of the feedback, despite the costs of providing the feedback. The savings increased as the feedback continued. Compared with the trend elsewhere without feedback, over seven years a total net sum of 1.4 million U.S. dollars was saved. Routine individual feedback is therefore economically worthwhile. *Key words:* economic evaluation; feedback; diagnostic testing; quality assurance; cost containment. (*Med Decis Making* 1996;16:309-314)

Expenditures for diagnostic testing in health care are increasing yearly. Factors such as defensive behavior of physicians, advances in technology, and the increased number of persons seeking medical care are responsible for this increase. Although increases in health care costs occur in conjunction with increases in the use of medical care and health services such as diagnostic tests, the health status of the populations of Western countries does not seem to improve correspondingly.

In the last few decades several methods have been applied to improve test-ordering behaviors and to halt the growth in expenditures for diagnostic testing. The provision of feedback about test ordering is one way to achieve this goal.¹ Various studies have demonstrated the effects of such feedback.²⁻¹⁰ In most of these reports, the results are expressed in terms of reductions in test use and corresponding

savings in the expenditures for diagnostic testing. Feedback apparently can influence test-ordering behavior, but soon after the feedback is discontinued the effects diminish.^{7,10} Hence, the long-term effect of feedback is also important.

The provision of feedback is considered to be a time-consuming, and therefore expensive, method of changing physicians' test-ordering behaviors. The benefits of feedback in financial terms are often questioned, and, based on the findings of Schroeder et al., feedback might be considered not economically worthwhile.¹¹ In Schroeder's study, the savings achieved by the reduction in test use were outweighed by the extra costs of the provision of the feedback itself. However, this conclusion need not apply to all situations. Feedback can be provided in various ways and thus can have different outcomes. Moreover, like many others, Schroeder et al. provided feedback over only a relatively short period of two years. If feedback is provided over a prolonged period as a routine procedure, there is a good chance that the savings in the costs of diagnostic testing eventually will exceed the costs of the provision of feedback. Thus, whether feedback is economically worthwhile has still not been established.

The cost-effectiveness of an intervention such as the provision of feedback about diagnostic test requests is difficult to determine, since the cost and benefit of the final outcome per patient are un-

Received October 26, 1995, from the Diagnostic Coordinating Center Maastricht (DCC) (RAGW, PP, PHAR); the Department of Health Economics, University of Limburg, Maastricht (AJHAA); the Center for Research on Quality in Family Practice, Universities of Nijmegen and Maastricht (RPTMG); and the Department of Family Practice, University of Limburg (JAK), Maastricht, The Netherlands. Revision accepted for publication November 9, 1995.

Address correspondence and reprint requests to Dr. Winkens: Diagnostic Coordinating Center Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands.

known and, moreover, depend on the underlying disorder. With regard to diagnostic testing, the underlying disorders vary widely. An alternative method to determine the economic consequences of feedback on the ordering of diagnostic actions is to evaluate expenditures for diagnostic tests before and during the period when feedback is being provided and to estimate the losses or savings in costs by comparing these expenditures with those in a concurrent control situation without feedback, taking into account factors of influence.

Since 1985, the Diagnostic Coordinating Center Maastricht (DCC) has been providing individual feedback about diagnostic test ordering as a routine procedure.¹² This feedback demonstrably improved the appropriateness of the test-ordering behaviors of those physicians who received it.⁶ Although the principal goal was to improve the rationality of test-ordering behaviors, the feedback also appears to affect the number of requests for diagnostic tests. Overviews showed that the start of the feedback was followed by a reduction in the total number of tests ordered. Requests for several tests that were frequently discussed in the feedback (and were designated inappropriate for the indications for which they were requested) decreased markedly, with reductions of up to 95% after two years of feedback.¹³ Requests for other tests decreased as well, probably as a result of a general learning effect. Even the last overviews (ending with 1992) showed persistent decreases in the numbers of requests. This result undoubtedly had financial implications, which had been indicated only roughly on the basis of charges for diagnostic testing.¹³ Savings in charges, however, do not reflect the economic benefits reliably. Charges include both variable and fixed costs. Fixed costs are much less relevant for assessing cost savings since they do not change when request numbers change. Therefore, we analyzed the economic consequences of feedback in more detail, looking only at variable costs of tests. The following questions were addressed:

- Is feedback economically (financially) worthwhile?
- Are the costs of the feedback itself outweighed by the savings in variable costs for diagnostic testing?

Methods

BACKGROUND

The DCC coordinates all the test requests of the 85 affiliated family physicians serving a (quite stable) catchment population of currently 187,000. Since 1985, written feedback has been provided twice each year to every individual family physician in the area,

discussing test-ordering behavior and the appropriateness of a large variety of tests. The feedback is based on analyses of request forms filled out by the individual family physician in the course of one month. Thus, the feedback concerns real cases from daily family practice. First, an overview is given of the tests requested in that month. Second, the rationality of the requests for the tests ordered is discussed. Because the request form offers clinical data about the patient (history, physical findings, suspected diagnosis, known diseases, reason for request, etc.), it is feasible to assess rationality. The requested tests are related to the clinical data of the patient provided on the request form; this combination is compared with criteria for a good diagnostic work-up as available in accepted regional guidelines and standards of the Dutch College of General Practitioners (NHG). The feedback includes comments about inappropriate requests and recommendations for rational diagnostic work-ups. Examples of requests are given, in which patients are mentioned by name and date of birth. Test results are considered only when in the feedback report the physician is asked to consider the effect of the receipt of normal or abnormal test results on his or her work-up. The comments are provided by a respected expert peer who is a specialist in internal medicine.

THE STUDY DESIGN

Retrospectively, as part of a quasi-experiment, we assessed the financial effect of feedback in trend analyses on two different levels. First, we assessed the decrease in expenses for the tests discussed in the feedback that were considered inappropriate or even unreliable. Especially for these tests, decreases in the numbers of requests, and thus cost savings, were expected. In order to assess the relative change compared with the situation should no feedback have been provided, we used data from a comparable laboratory elsewhere in The Netherlands that had not provided feedback in the same period as a control situation.

Second, we looked at our variable-cost trend for all the test requests taken together. This is relevant because a general learning effect not only decreased the numbers of requests for the tests discussed in the feedback but also appeared to decrease requests for tests not so discussed. The effects of feedback are thus not restricted to requests for tests discussed in feedback. The full extent of the economic effects was assessed by using the nationwide cost trend as a control (in The Netherlands, there is virtually no feedback about test requests outside the Maastricht region). The nationwide cost trend for all diagnostic testing by all Dutch family physicians to-

gether has been assessed in other studies.^{14,15} In this comparison, the expenditures incurred for the provision of the feedback were considered as well.

In both trend analyses, several factors were considered. In The Netherlands there are fixed charges per test, but it is generally assumed that these charges do not reflect the actual situation, especially when the numbers of tests change. In previous analyses, however, only these charges were used.^{12,13} When the volume of tests changes, some cost factors, such as manpower, remain relatively constant. Other costs vary with the numbers of tests. In order to determine the actual change in costs due to the feedback, we considered only these variable costs, among which were the costs of materials (disposables, chemicals, etc.) and equipment (by interest and depreciation). Cost of manpower was not considered. This may have led to an underestimation of the savings achieved.

In the assessment of costs and savings incurred through feedback, the expense of the provision of the feedback itself was considered. For the hours spent in developing the feedback, the salary costs of an expert reviewer and administrative support were calculated. Only small proportions of the costs of materials and postal charges were taken into account. The costs of providing the feedback amounted to approximately \$55,000 per year (for peer reviewer, secretarial assistance, equipment, and materials) and were relatively stable during the study period.

DATA COLLECTION

Data were available from our center for requests received over the period 1983–1991. Control data from another Dutch laboratory were available over the same period. This laboratory was comparable to our center with regard to degree of urbanization and average family practice size. Data could be compared for 44 tests that could be requested in both regions (predominantly hematologic, serologic, and clinical chemistry tests), representing approximately 70% of all requests for diagnostic tests made by family physicians in our region. These 44 tests include virtually all of the tests frequently discussed in the feedback since 1985. However, far more tests than these 44 are discussed in the feedback.

Comparable data reflecting the overall request trends for primary care in The Netherlands were obtained from the literature.^{14,15} Variable costs per test were obtained from the administrative department of the University Hospital Maastricht. These variable costs were determined at one single point of time. Changes may occur over time in the variable costs per test, but consideration of such changes was beyond the scope of this investigation.

THE ANALYSIS

The numbers of requests per test were multiplied by the variable cost per test. To correct for differences in the numbers of patients, control laboratory data were indexed to the Maastricht level in 1984. Thus, an indexed total sum per test per year was obtained. From the total savings, the yearly costs of providing the feedback were subtracted. The savings due to the feedback were determined on every level by comparing actual expenditure (including the costs of provision of the feedback) per year with the expenditure if no feedback had been provided.

Results

Variable costs per test unfortunately could not be obtained for all tests available at the Maastricht diagnostic center. In total, 630 tests are available at our center. Variable costs were available for more than 400 different tests (65%). However, these 400 tests represent more than 90% of the total number of all requests by family physicians. Only about 190 of these tests are requested regularly. Virtually all of these 190 tests are included in the 400 different tests for which variable costs were available. The tests for which variable costs could not be obtained were several (seldom requested) radiologic tests and all endoscopic test procedures. For all 44 tests for which comparable data were available from the control laboratory, variable costs were obtained.

Test-ordering behaviors changed considerably beginning with the start of the feedback in 1985.¹³ Reductions in the numbers of requests were especially evident for hematology and clinical chemistry (table 1).

Compared with the control laboratory, the costs resulting from requests for the 44 tests at our center decreased beginning in 1985, whereas the corresponding costs increased almost continuously at the control laboratory (table 2). In 1985, the variable costs for the 44 tests decreased by 12.5% with feedback, compared with the trend without feedback. In 1991, this decrease was 57%. However, the marked reduction in the volume of requests was accompanied by only a small economic benefit: the savings were outweighed by the costs of providing the feedback.

At the DCC, far more tests can be requested (and are discussed in the feedback) than the 44 that are comparable with the data available from the control laboratory. To gauge the full economic consequences of routine individual feedback, the complete package of all requests must be considered. The overview of the total costs incurred when feedback was used to change test-ordering behaviors

Table 1 • Number of Requests in the Maastricht Region per Category of Test per Year, 1983 through 1991

	1983	1984	1985	1986	1987	1988	1989	1990	1991	Decrease* in 1991 (%)
Hematology	53,298	57,344	50,384	43,538	41,556	41,842	44,418	40,685	34,899	39
Serology	6,980	7,194	6,262	4,700	4,318	3,556	3,328	3,137	2,903	60
Clinical chemistry	46,456	50,980	44,312	37,238	36,498	38,696	39,499	39,328	35,591	30
Urine/feces tests	2,336	2,700	2,080	2,068	1,910	1,866	1,698	1,621	1,389	48
Bacteriology/virology	6,676	6,880	4,622	4,916	5,216	5,658	6,296	6,297	5,738	17
ECG	1,712	2,028	2,066	1,822	1,900	1,960	1,702	1,924	1,939	4
Endoscopy	264	408	394	496	916	902	1,133	1,452	1,466	-259
Radiology	13,210	14,096	13,864	12,974	12,820	13,840	13,658	14,700	14,154	0
Histology	4,898	4,838	5,124	5,066	5,066	5,746	5,787	5,312	4,849	0

*Decrease in 1991 compared with the level in 1984.

Table 2 • Costs and Savings in the Diagnostic Coordinating Center (DCC) in the Maastricht Region for 44 Tests Compared with Those in a Control Laboratory*

	Cost Trend, 44 Tests, Control Laboratory	Real Costs, 44 Tests, DCC	Reduction Year to Year
1983	\$76,573 (96)	\$72,198 (91)	—
1984	\$79,826 (100)	\$79,826 (100)	—
1985	\$80,799 (101)	\$70,753 (87)	\$10,046
1986	\$80,858 (102)	\$59,808 (75)	\$21,050
1987	\$83,487 (105)	\$52,157 (65)	\$31,330
1988	\$86,784 (109)	\$51,664 (64)	\$35,120
1989	\$89,586 (112)	\$55,766 (69)	\$33,820
1990	\$97,256 (122)	\$49,544 (62)	\$47,712
1991	\$98,011 (123)	\$42,234 (53)	\$55,777

*Costs of feedback are not included. Costs at the control laboratory are indexed, as percentages (in parentheses), to the level of costs at the DCC in 1984 (100%).

shows that the usual increase in costs per year in the Maastricht region became a decrease beginning in 1985. Based on the nationwide trend of yearly increases of 7–8% in the costs of diagnostic testing in primary care, we estimated that the costs of diagnostic testing in the Maastricht region would have increased 51% from 1984 to 1991 if no feedback had been provided (table 3).¹⁵ Despite the costs of providing the feedback, money was saved in every year since 1985, resulting in total savings of more than 1.4 million U.S. dollars after seven years of provision of feedback.

Discussion

The results show that the provision of feedback was accompanied by a lower expenditure for diagnostic testing, even when the extra costs incurred in providing the feedback are considered. The saving in variable costs increases when the feedback is continued. With regard to these findings, a number of aspects are worth elucidating.

The economic consequences of feedback are not easy to assess. A cost-effectiveness study is difficult to perform. An economic evaluation can show only part of the economic effects. Moreover, we had no data about patient outcomes. The possible side effects of feedback on test-ordering behaviors are important. For example, a physician might refer more patients to a hospital when the use of outpatient diagnostic tests is discouraged. A separate study showed that the number of hospital referrals in Maastricht did not increase after the start of the feedback in 1985. The trend in hospital referrals was comparable to the nationwide referral trend.¹⁶ Also, we faced the problem of missing data. Variable costs could not be obtained for all the tests that can be requested by family physicians. Unfortunately, so far in the literature costs or savings have been based on charges and not on variable costs.

Ideally, the study design would have been a prospective, randomized controlled trial. In our assessment of the economic effects of feedback on test-ordering behaviors, such a design was no longer feasible. The feedback procedure started in 1985, and, due to lack of manpower, a prospective ran-

Table 3 • Variable Costs and Savings for All Tests except Endoscopies That Can Be Requested at the Diagnostic Coordinating Center (DCC)*

	Costs according to National Trend	Costs at the DCC	Reduction Year to Year	Savings† through Feedback Year to Year
1983	\$494,350	\$512,075	—	—
1984	\$529,938	\$529,938	—	—
1985	\$567,842	\$488,081	\$79,761	\$24,761
1986	\$608,455	\$435,243	\$173,212	\$118,212
1987	\$651,779	\$436,000	\$215,779	\$160,779
1988	\$698,269	\$450,844	\$247,425	\$192,425
1989	\$747,971	\$463,240	\$284,731	\$229,731
1990	\$801,357	\$471,558	\$329,799	\$274,799
1991	\$858,401	\$434,487	\$423,914	\$368,914

*Costs for the DCC region according to the trends before and after 1984 are estimated on the basis of the nationwide trend.

†After subtraction of the costs of providing feedback.

domized study could not be implemented at that time. The provision of the feedback has since led to considerable changes. Further changes are likely to be much more difficult to achieve, invoking the law of "diminishing results." In our region a "new" prospective study of the economic effects of the provision of feedback on the totality of requests would have been overtaken by events, and therefore it might have traced only minor additional effects.

The difference between our results and those of other studies of feedback is remarkable. Our feedback is actually the first that has resulted in long-lasting and wide effects and major savings. This can be explained by two features of our feedback: first, the feedback covers a wide variety of tests; and second, the feedback has been provided as a routine procedure over many years. In an earlier study, feedback reduced the costs (charges) of diagnostic testing over two years, but the money spent on providing the feedback outweighed the savings.¹¹ The savings might have increased if the feedback had been continued in that study, leading, eventually, to a different conclusion. Moreover, the feedback appeared to have a general learning effect on the ordering of tests that it did not cover. This may have been due to the decision-analytic approach of the feedback and to general remarks about the value of diagnostic testing. For example, test requests to confirm a diagnosis are not recommended when complaints, signs, and symptoms make the pretest probability of a positive result high. Such general remarks apply to all testing.

The savings found for the 44 tests that were comparable with the control laboratory's tests were small, despite the considerable changes in the numbers of requests published earlier.¹³ This is because these 44 tests are predominantly low-cost tests. The determination of the economic feasibility of routine individual feedback can, however, be obtained only by looking at the overall picture of all requests. The results suggest that even when the extra costs of providing the feedback are considered, the feedback has been economically worthwhile from the start.

Our calculations take no account of changes in laboratory staffing. Theoretically, a reduction in the number of requests will in time lead to a cutback in staff.

The variable costs were held stable in all calculations over the period 1983-1991. Undoubtedly, due to factors such as inflation, the variable costs would have increased slightly over this nine-year period. On the other hand, due to automation, some variable costs might have decreased. Altogether, we did not expect any substantial effect on the results of the study when the costs per test are assumed to be stable. Also, the costs of providing the feedback turned out to be stable between 1985 and 1991. Any

potential reduction in cost that might have been expected to accrue from more experience with the procedure (greater efficiency) was outweighed by the involvement of more tests, making the cost of feedback relatively independent of the number of tests discussed.

As already mentioned, variable costs were available for tests accounting for 90% of all requests. The remaining 10% included all endoscopies. They are the most expensive tests that can be requested at our center. The total costs in the Maastricht region therefore are likely to be higher than those shown in table 3. The absence of data about endoscopies is inconsequential in the comparison with the nationwide trend, since in the nationwide data, costs of endoscopies are included only to a small extent. There are only a few regions in The Netherlands (one of them being Maastricht) where family physicians can request endoscopy; in the other regions, referral to a specialist is needed. Consequently, for the comparison of the Maastricht region's cost trend with the nationwide cost trend, endoscopy is relatively unimportant.

During 1990 and 1991, in an investigation of the rationality of requests for diagnostic tests, feedback was provided for tests not discussed before, including several radiologic tests. The numbers of requests of these radiologic tests decreased, especially in the second year of the trial. This outcome could explain the inordinate increase in the savings in 1991.⁶ The feedback provided during this period was merely an extension of the usual feedback and therefore did not increase the cost of providing the feedback.

In all calculations, variable costs were considered without possible fees for any specialist involved. These fees are in fact variable costs as well. However, the true costs of specialist involvement (e.g., for x-ray interpretation) are not known. The estimated extra savings would have been considerable.

During the study, we were the only center in The Netherlands that was using request forms on which detailed clinical data could be mentioned. Such a form is a prerequisite for assessing the rationality of requests for diagnostic tests. The form could influence test-ordering behavior as well, but this possible effect should not be overestimated. In our region, the trend in the numbers of requests for tests between 1978 and 1984 (increase) was similar to the nationwide trend, despite the introduction of our request form in 1980.

Aside from all reductions in test use and (after taking into account the extra costs of providing the feedback) the corresponding savings in costs, it should be emphasized that the main purpose of our feedback was to improve the rationality of requests for diagnostic tests and make test use appropriate. In our opinion, feedback should always focus on

quality of care, apart from the possible economic benefits.

On the basis of the size of our catchment population, we can roughly estimate what such feedback might have achieved on a national level, treating variable costs as equal throughout the country. The population of our catchment area is 187,000, which is 1.32% of the Dutch population. Had feedback been provided nationwide, a total amount of more than 106 million U.S. dollars would have been saved in The Netherlands after seven years.

We conclude that routine individual feedback about the test requests of family physicians is economically worthwhile. The longer the feedback is continued, the greater the savings are. This applies only when the feedback is focused on a wide range of diagnostic tests.

References

1. Eisenberg JM. Physician utilization. The state of research about physicians' practice patterns. *Med Care*. 1985;23:461-83.
2. Schectman JM, Elinsky EG, Pawlson LG. Effect of education and feedback on thyroid function testing strategies of primary care clinicians. *Arch Intern Med*. 1991;151:2163-6.
3. Martin AR, Wolf MA, Thibodeau LA, Dzau V, Braunwald E. A trial of two strategies to modify the test-ordering behavior of medical residents. *N Engl J Med*. 1980;303:1330-6.
4. Fowkes FGR, Evans KT, Hartley G, et al. Multicentre trial of four strategies to reduce use of a radiological test. *Lancet*. 1986;i:367-70.
5. Marton KI, Tul V, Sox HC. Modifying test-ordering behavior in the outpatient medical clinic. A controlled trial of two educational interventions. *Arch Intern Med*. 1985;145:816-21.
6. Winkens RAG, Pop P, Bugter-Maessen AMA, et al. Randomised controlled trial of routine individual feedback to improve rationality and reduce numbers of test requests. *Lancet*. 1995;345:498-502.
7. Eisenberg JM. An educational program to modify laboratory use by house staff. *J Med Educ*. 1977;52:578-81.
8. Berwick DM, Coltin KL. Feedback reduces test use in a health maintenance organization. *JAMA*. 1986;255:1450-4.
9. Everett GD, de Bois CS, Chang PF, Holets T. Effect of cost education, cost audits, and faculty chart review on the use of laboratory services. *Arch Intern Med*. 1983;143:942-4.
10. Sommers LS, Sholtz R, Shepherd RM, Starkweather DB. Physician involvement in quality assurance. *Med Care*. 1984;22:1115-38.
11. Schroeder SA, Myers LP, McPhee SJ, et al. The failure of physician education as a cost containment strategy. *JAMA*. 1984;252:225-30.
12. Pop P, Winkens RAG. A diagnostic centre for general practitioners: results of individual feedback on diagnostic actions. *J R Coll Gen Pract*. 1989;39:507-8.
13. Winkens RAG, Pop P, Grol RPTM, Kester ADM, Knottnerus JA. Effect of feedback on test ordering behaviour of general practitioners. *Br Med J*. 1992;304:1093-6.
14. Dunning AJ. Growth in number and costs. On diagnostic and therapeutic actions in health care. The Netherlands: The Hague, 1986.
15. Financial Overview of Care. Report no. 21310. The Netherlands: The Hague, 1990.
16. Winkens RAG, Grol RPTM, Beusmans GHMI, Kester ADM, Knottnerus JA, Pop P. Does a reduction in general practitioners' use of diagnostic tests lead to more hospital referrals? *Br J Gen Pract*. 1995;45:289-92.