Evaluation Results From Prospective Drug Utilization Review: Medicaid Demonstrations

David Kidder, Ph.D., and Jay Bae, Ph.D.

In 1992 HCFA awarded two cooperative agreements for demonstrations of prospective drug utilization review (PDUR). Iowa tested an on-line prospective drug utilization review (OPDUR) system. Washington tested payments to pharmacists for providing non-dispensing "cognitive services" (CS). In this article the authors report on an evaluation of these demonstrations and on three assessments of retrospective drug utilization review (RDUR) interventions. The evaluation failed to detect effects of either State PDUR demonstration on the frequency of drug problems, utilization of prescription drugs and other health services, and clinical outcomes. However, the State RDUR interventions had immediate effects on prescribing physicians.

BACKGROUND

Following a decade-long acceleration of drug price inflation in the 1980s, Congress legislated a sweeping reform plan to contain Medicaid outpatient drug expenditures in the Omnibus Budget Reconciliation Act of 1990 (OBRA 90). OBRA 90 addressed prescription drug issues in three of its component parts.

First, the legislation's main cost-containment feature was a drug rebate program. Implemented in January of 1991, the rebate program has helped contain Medicaid drug expenditures by assuring the lowest private sector prices to the program.

The second and third components dealt with inappropriate drug therapy, considered by many to be a serious problem resulting in many avoidable illnesses and utilization of health services. As most third-party payers in the private sector had already adopted some form of drug utilization review (DUR), the Act's second drug-related component required that Medicaid adopt both PDUR and RDUR. The PDUR is initiated by the pharmacist, who reviews prescriptions at the point of sale for potential problems. If a potential problem is detected based on information available to the pharmacist (e.g., personal knowledge, research, or in-store computer screening of patient prescription records), further intervention may be made with prescribers, other pharmacists, or patients, as appropriate. Under RDUR, on the other hand, data on prescribed drugs are collected and processed into a profile to identify patterns of inappropriate drug therapy for later corrective intervention. The focus of the profile may be the physician, the pharmacist, the patient, or a drug.

The third component of Medicaid drug reform was the authorization of demonstrations of more advanced models of PDUR (§ 4401(c)(2)). OBRA 90 required that HCFA conduct demonstrations of OPDUR and payment for pharmacists' CS. OPDUR is a system that links many phar-

David Kidder is with Abt Associates Inc. Jay Bae is with the Office of Strategic Planning, Health Care Financing Administration (HCFA). The views and opinions expressed in this article are those of the authors and do not necessarily reflect the views of Abt Associates Inc., or HCFA.

macies' computers to a central screening system. Payment for CS refers to reimbursing pharmacists for patient counseling and intervention, when necessary, to ensure appropriate drug therapy. The demonstration would test if more advanced technology of on-line real-time intervention or additional payment for CS would reduce the use of unnecessary, duplicative, or inappropriate drug use and avoid costly health service utilization.

In 1992 HCFA solicited applications for demonstrations and selected two sites to test the DUR systems mandated by law. After the demonstrations had been awarded to Iowa and Washington, HCFA chose a team led by Abt Associates Inc. to conduct an external evaluation of both demonstrations. In addition, HCFA asked the evaluator to assess the effectiveness of selected RDUR programs. To enhance the precision of estimates, the evaluation team augmented the data by adding two States to the study—Maryland and Georgia.

EFFECT OF OBRA 90

Historically, State Medicaid agencies used the surveillance and utilization review program to deter fraud and abuse. With the increasing concern over the lack of drug monitoring systems to safeguard against less-than-optimal or inappropriate drug utilization among the low-income and vulnerable Medicaid beneficiaries, Congress mandated DUR to assure and improve quality in Medicaid prescription drug programs.

All State Medicaid programs implemented both PDUR and RDUR within 5 years of OBRA 90. Somewhat surprisingly, a large number of States adopted some form of OPDUR, even though OBRA 90 did not require it. In 1995 all States had DUR programs in place. Of these, 22 were running both OPDUR and RDUR programs. This represents an increase from only six States with OPDUR in 1993. Of the 25 States without OPDUR programs in 1995, all but 5 planned to implement OPDUR by fiscal year 1997.

Although States were required to report on costs savings from DUR programs, the diversity of methods and results makes it difficult to generalize about program cost effectiveness. Thirty-four States did cost-savings analyses of their DUR programs, using many different methodologies. In analyses of OPDUR costs savings, it was assumed that claims reversals equal program savings. Estimates of savings ranged from $500,000 in West Virginia to $22.0 million in New York. Five States used the Automated Claims Tracking System (ACTS) provided by Health Information Designs to assess decreased drug use associated with RDUR with a range of savings from $16,000 in New Hampshire to $3 million in Massachusetts. Three States computed ratios of RDUR program costs to program benefits (cost savings), with results ranging from 1:1.7 to 1:2.

DISCUSSION

Although the premise that DUR can prevent some prescription drug problems and some adverse health outcomes is plausible, as Jay, Eynon, and Javitz (1991) noted, ". . . no well-controlled evaluation of DUR systems, prospective or retrospective, has been performed." Little is known about the prevalence of drug therapy problems or about the frequency of hospital admissions attributable to preventable drug problems. To assess the benefit of OPDUR, the following questions must be answered.

- *What is the prevalence of prescribing problems?* Despite the perception that prescription drug problems are a major cause of hospital admissions, the available evidence shows a small prevalence

of preventable errors. An early analysis by Ray, Griffin, and Schorr (1990) suggested that 3-5 percent of all hospital admissions in the general population are attributable to medication toxicities. From a different perspective, Rupp et al. (1988), Rupp, DeYoung, and Schondelmeyer (1992), and Dobie and Rascati (1994), using clinical pharmacists to review prescriptions, found that 0.8-2.6 percent of all prescriptions reviewed were problematic. Further, only 0.4-0.7 percent contained errors that, if uncorrected, would lead to serious harm. This Medicaid DUR demonstration provided a valuable opportunity to investigate the prevalence of prescribing problems for a group of Medicaid recipients living in the community.

- *Can DUR change prescriber behavior?* A key assumption of DUR is that external reviewers (pharmacists and State DUR boards) can change physicians' prescribing behavior. There is a reasonably large body of research evidence on how DUR interventions affect physicians' behavior. For instance, it has been shown that more intensive interventions, led by physicians and/or clinical pharmacists, are more effective in improving prescribing practices, and most educational interventions produce at least some improvement in prescribing behavior (Davis et al., 1995). Feedback interventions were shown to be effective in modifying prescriber behavior; however, effectiveness of academic detailing, e.g., clinicians educating prescribers on prescribing practices, has mixed results. Although Avorn and Soumerai (1983) found academic detailing to reduce use of targeted drugs effective for up to 2 years, Lin et al. (1997) and Ray et al. (1987) found that detailing does not work or, when it does, it only works for a short time for antidepressants and antipsychotics. The Medicaid DUR demonstration provided an opportunity to collect data on the effectiveness of certain forms of prescriber intervention.

A DUR program can be effective when the prescribing system fails to provide safe and appropriate drug therapy. Thus, the program faces the challenge of identifying and intervening in the relatively small number of prescription problems that pose serious health threats and getting prescribers to change their behavior.

STRUCTURES

In 1992 HCFA selected two DUR experiments managed by State Medicaid agencies: Washington for CS and Iowa for OPDUR. Washington and Iowa designed the DUR demonstrations to improve drug prescribing by influencing pharmacists' behavior. However, each demonstration tested a different premise about what incentives are most effective in achieving the desired results.

Washington

Also known as Project C.A.R.E., this demonstration paid pharmacists in the treatment group for providing CS, defined as non-routine professional activities related to dispensing drugs. To test the effects of paying for CS, Project C.A.R.E. recruited 200 pharmacies and randomly assigned them in geographic clusters to two groups: 110 to a treatment group (Group A) that received payment for providing and documenting CS and 90 pharmacies to a control group (Group B) that received payment only for documenting CS. Pharmacies were recruited in three waves. Investigators also recruited a third group of pharmacies (Group C) that received no CS payments of any kind. Project C.A.R.E. operated from February 1994 through September 1995.

To be eligible for payment under the demonstration, CS had to be related to dispensing a prescription drug to a Medicaid beneficiary but not a "requisite part of dispensing" (i.e., accepting, interpreting, and clarifying a prescription order, preparing a prescription, delivering it to a patient) and had to conform to certain "common sense" standards. Project C.A.R.E. asked pharmacists to document the problem, the intervention, and the result of each CS event. Project C.A.R.E. did not focus pharmacists' attention on a specific list of drugs.

For eligible CS, Group A pharmacists received per service payments of $4 for a "brief encounter" (6 minutes or less) or $6 for an "extended encounter" (longer than 6 minutes). Both A and B pharmacies received $40 per month for submitting CS documentation to Project C.A.R.E.

Iowa

Iowa linked all participating pharmacies in the OPDUR demonstration to an on-line DUR screener that reviewed each Medicaid prescription drug claim as it was submitted for payment electronically. Pharmacies in the treatment group received messages that identified potential prescription problems and detected when a prescription failed one or more of the OPDUR screens. Control-group pharmacies received no messages. Both treatment and control-group pharmacies were asked to document all CS provided, in order to describe demonstration effects on pharmacists' behavior. Iowa did not pay pharmacists for providing or documenting CS.

Iowa defined specific problems and drugs as the focus of its intervention on criteria developed by the University of Maryland and the Philadelphia College of Pharmacy and Science (UM/PCPS). The UM/PCPS criteria included drug-drug interactions, therapeutic duplications, and high daily dose for eight classes of drugs: angiotensin-converting enzyme (ACE) inhibitors, calcium channel blockers, cardiac glycosides, benzodiazepines, antidepressants, antipsychotics, histamine-2 receptor antagonists (H2RAs), and nonsteroidal anti-inflammatory agents (NSAIDs). Iowa also screened for early refills in all drug categories.

Iowa randomly assigned pharmacies in geographic clusters to treatment and control groups. Of 223 pharmacies recruited for the demonstration, 110 belonged to the treatment group and 113 to the control group. As in Washington, project investigators were forced to recruit in waves to achieve their initial goal of 200 pharmacies. The demonstration operated from June 1994 through June 1997.

EVALUATION

The evaluation adopted two tools to detect the frequency of prescription drug problems and to link problems to potential adverse clinical events, represented by *International Classification of Diseases, 9th Revision* (ICD-9) diagnosis codes. One was the Pennsylvania State University (PSU) screener, an expert system developed by researchers at PSU to detect suboptimal drug therapy. It incorporates updates and modifications of the UM/PCPS criteria and screens drug claims data for under- or overdosage, duration of therapy, drug-drug interactions, age-based contraindications, and duplicative therapy in eight drug categories: ACE inhibitors, antidepressants, antipsychotics, benzodiazepines, calcium channel blockers, digoxin, H2RAs, and NSAIDs. The second tool was the DUR Outcomes Bibliographic Database, a computerized bibliographic database that links PSU screener drugs and criteria to clinical outcomes. The selection of ICD-9 diagnosis

codes linked to the PSU screener criteria was done through an analysis of primary and secondary pharmaceutical and clinical literature and a subsequent review by a team of technical advisors. This review considered not only whether or not the literature suggested a link between a drug problem and a diagnosis-defined clinical event but also whether or not the data included enough cases to support statistically precise estimation.

Components and Findings

Evaluation efforts were focused on the following three areas.

- Documenting and describing the demonstrations (e.g., design and costs) and their contexts (e.g., site visits, survey of pharmacists about DUR, comparative reporting of problem drug use frequency).
- Testing the effects of demonstrations on utilization, costs, and drug problems. Statistical techniques were used to generate adjusted estimates of demonstration effects. We hypothesized that OPDUR in Iowa should reduce utilization, because early refills and duplications were explicit screening criteria, but in Washington, without explicit criteria, utilization could increase or decrease. In both demonstrations, utilization and costs of other medical services should decrease. Over the course of the demonstration, we expected the frequency of drug problems to decrease.
- Testing the effects of the demonstrations on adverse clinical outcomes. Linking medical and drug data, the project evaluated medical outcome effects of the demonstrations in five areas of potential adverse outcomes: (1) use of NSAIDs and gastrointestinal bleeding; (2) benzodiazepines and falls; (3) cardiac drugs and associated outcomes; (4) outcomes linked to beta-agonist use; and (5) antidepressants. We expected that exposure to demonstration interventions would reduce the frequency of adverse clinical outcomes.

Drug Use and Prescribing Problem Rates

To provide some context for the assessment of DUR, the evaluator compared trends in Medicaid drug use and measures of prescribing problems (numbers of screen failures triggered through the PSU screener) in the four evaluation States prior to the demonstration through 1996. For Washington, Iowa, and Maryland, data were available from 1989-96. Georgia provided data from 1994-96. The evaluator estimated frequencies of drug problems using the PSU screener software.

The results revealed a picture of substantial interstate variations in many ways. Notable was the low drug-use rate and low average screen-failure rate in Georgia. Screen-failure rate was 3.1 percent of users, compared with 5.4 to 6.7 percent in the other three States. In terms of trends in screen failure rates, both Iowa and Maryland showed significant improvement—from 7.2 percent (1989) to 5.5 percent (1996) in Iowa, and from 7.0 percent (1989) to 5.7 percent (1996) in Maryland.

Failure rates for multiple-drug users and persons with diabetes were 1-2 percentage points above the average for the Medicaid population as a whole. Improvements were noted in Iowa, Maryland, and Washington for these populations as well. Some therapeutic areas have shown consistency in the extreme failure rates. Ten of 61 criteria were associated with consistently high failure rates (exceeding 8 percent of claims and persons in all States and all years—these extremes were concentrated in the

benzodiazepine, H2RA, and NSAID drug classes, and dosage and duration problems). Thirteen criteria were consistently low (less than 0.25 percent of persons and years—these criteria were clustered in the antidepressant, NSAID, and digoxin classes, and drug-drug interaction problems).

RESULTS

CS Provision

The demonstrations were designed to encourage more pharmacists' interventions when appropriate; in the process, valuable information about CS behavior was collected. Project C.A.R.E. found that paying pharmacists per CS service naturally increased the number of documented CS services, but in Iowa, where there was no explicit CS payment, we observed no effects of the intervention on CS activities. As partial evidence for a demonstration effect of the OPDUR intervention, Iowa cited a higher rate of prescription drug claim reversals for treatment than control pharmacies. Both demonstrations found evidence of substantial activities by pharmacists involving patient counseling and education.

In the Washington CS demonstration, the treatment group that received payments for CS generated more documented services, as expected, than did the control group. Although the documented CS rate for treatment pharmacists ranged from 1.3 to 2.4 per 100 dispensed prescriptions, the control group with flat payment (Group B) produced 0.7-1.0. CS rates rose over time in Project C.A.R.E. In contrast, the average documented CS rate in Iowa treatment pharmacies was 0.16 per 100 claims, compared with 0.21 for control pharmacies. It is a notable contrast that documented CS rates fell during the period of Iowa's demonstration. Both in Washington and Iowa, demonstrations showed that pharmacies with higher activity levels (workloads per prescription volume) generated lower CS rates than pharmacies with lower activity levels. In both demonstrations, about one-half of all CS was for patient-related problems. In Project C.A.R.E., control pharmacies consulted with patients more frequently than treatment pharmacies but provided patient training and assessment less frequently. In Iowa, treatment pharmacies were more likely than controls to consult patients. CS in both States resulted in higher rates of prescriber consulting; treatment pharmacies' prescriber-consulting rate was 43.7 percent, compared with 31.4 percent of control pharmacies in Washington, and 82 percent, compared with 64 percent in Iowa.

Prescription Drug Problems

The DUR sites systematically monitored and reported incidences of problematic drug use. The evaluator estimated the effects of the two demonstrations on the frequency of prescription drug problems, measured as screen failures of the 6l criteria in the PSU screener. Although both demonstrations could have reduced prescription problems, it was hypothesized that these effects would be most evident in Iowa, which used an OPDUR system targeted to the same drugs and problems captured in the PSU screener.

The Iowa data showed that 7.4 percent of all prescription drug claims in Iowa demonstration pharmacies generated at least one screen failure, with the highest failure rate in the early-refill criteria, at 4.1 percent of all claims. Among elderly Medicaid recipients in the demonstration, 36.9 percent had at least one screen failure each quarter, with early refills being the most common problem. The rate of prescription drug claim reversals as a drug

therapy change showed an increase during the period of the demonstration and was higher for treatment pharmacies (4.2 percent) than controls (3.7 percent), a difference that was statistically significant. Areas where the demonstration produced statistically significant increases include the reversals related to drug-drug interaction screens and early-refill screens.

The Iowa DUR data showed reductions of drug problems in 26 out of 45 estimates, however, only 3 estimates were statistically significant at the 10-percent level. Further, analysis controlling for other factors found similar results for demonstration effects on the number of prescription drug problems. No patterns within specific criteria were detected.

The Washington data also showed mixed results, where 14 out of 45 estimates suggested a reduction in frequency of problems, but only 1 of these was statistically significant at the 5-percent level. On the other hand, there were three statistically significant estimates of increased frequency of drug problems. Analysis of the number of problems produced no significant negative estimates. Within specific criteria, only 3 of 146 estimates were statistically significant, and only 1 of these suggested fewer problems. Given the evidence, neither demonstration seems to have reduced prescription problems, and Iowa OPDUR performed no better than Washington CS in this regard.

Effect on Drug Expenditures and Utilization

Two important questions asked by the evaluator were (1) whether or not the DUR would generate savings either directly (by reducing expenditures for prescription drugs) or (2) indirectly (by preventing "downstream" utilization of health care services and reducing the frequency of clinical outcomes associated with prescribing problems). First, we report the direct drug-cost effect of the demonstrations.

The data show that drug costs varied widely across the Medicaid population in the demonstrations. In Washington, compared with an average $633 for dually eligible elderly individuals, costs ranged from $1,236 for persons with diabetes to $2,100 for persons with 10 or more prescriptions. In Iowa, the comparable range was from $916 (for persons with diabetes) to $1,510 (for persons with multiple prescriptions), compared with $498 for all elderly.

The evaluator computed adjusted estimates of expenditures for expensive drugs and the eight PSU screener drug categories by using three measures of prescription drug utilization: percent using drugs, a count of drugs used, and total costs of drugs used. Although DUR interventions might increase or decrease drug costs, the evaluator hypothesized that in general DUR would produce reductions in drug utilization. The analysis results showed reduced utilization in many areas. However, few estimates reached levels of statistical significance sufficient to conclude that there was any pattern of effects in any of the populations or in either of the drug category sets.

Effect on Use of Health Services

To assess the comprehensive financial impact of the demonstrations, we needed to look beyond the drug expenditures and utilization and study the entire health service expenditures and utilizations of Medicaid recipients in control and experimental groups. The evaluation selected high-risk Medicaid populations for analysis: the dually eligible elderly, subpopulations of this group (persons with diabetes, multiple prescribers, or multiple prescriptions), and dually eligible non-elderly per-

sons with disabilities. Controlling for demographic and other influences, the evaluation examined utilization of hospital services, professional services, emergency room, and any medical services reimbursed by Medicaid or Medicare. The evaluation defined two measurement levels: all utilization and expenditure, and utilization and expenditure associated with targeted diagnoses.

Regression analyses of the Washington data showed 16 of the 40 estimates with the predicted lower utilization effects, but only two of these approached the 5-percent level of statistical significance. On the other hand, a significant estimate—of effects on payments for emergency room treatments—was positive. In Iowa, 28 of 40 models produced negative estimates, but only 6 of these were statistically significant. When we singled out only the targeted diagnoses, the Washington demonstration showed no consistent pattern in the estimated direction of effects. Ten of 54 estimates approached standard levels of significance, but 3 of these were positive. Similar results were found in Iowa. The empirical analysis seems to indicate that there was no general pattern of reduction in either the utilization or expenditures of drugs or other health services.

Retrospective DUR Intervention

In addition to the core evaluation of PDUR demonstrations, HCFA asked the evaluator to conduct targeted studies of State RDUR interventions. RDUR screens for problems and intervenes after prescriptions have been filled. RDUR programs generally have specific targets (toward specific drugs and problems) and are limited in frequency (often to a single intervention) and scope (to a short list of prescribers, pharmacies, and occasionally patients). Three separate studies were completed, two in Iowa and one in Maryland.

In Iowa, RDUR was conducted for two groups of patients. The first group consisted of 268 patients who received both NSAIDs and maintenance doses of a gastrointestinal drug (evidence of a possible NSAID-induced gastric ulcer) who could benefit from misoprostol. The second group of 379 patients were identified as users of a long-acting bronchodilator, salmeterol, with no concurrent use of a short-acting bronchodilator. The Medicaid DUR board sent letters to physicians and pharmacists to inform them about potential problems and recommendations. The evaluator selected control groups of patients, using a computer algorithm that mimicked the target patients for the intervention.

In the misoprostol RDUR intervention, 8 percent of patients named in the letters started using misoprostol, compared with 1 percent of controls, a small but statistically significant effect. Of those who maintained contact with the same physician both before and after the intervention, 12 percent were associated with physicians who received letters started using misoprostol, compared with 2 percent of controls. In the salmeterol RDUR, there was an increase from 3 to 12 percent of targeted patients taking the suggested combination therapy after prescribers received the letter. In the control group, the intervention produced a 6 to 8-percentage-point increase in the proportion of patients using the combination therapy with varying degrees of significance.

In Maryland, the RDUR intervention was conducted by the State Pharmacists' Association targeting 81 patients on acute dosage levels of Zantac for more than 56

days during a review period. The evaluation selected a control group of 96 at-risk patients who were unaffected by the RDUR. The program reduced Zantac refills substantially (by 18 percentage points) in the treatment group, compared with a 2-percentage-point decline among comparison group patients. When adjusted for potential confounding influences, the effect increased to 28 percentage points. Somewhat counter to intuition, Zantac dosage levels in the treatment group actually increased, perhaps reflecting the fact that physicians removed their healthiest patients from Zantac. There was no evidence that prescribers substituted other drugs for Zantac. However, there was suggestive evidence of a large and significant spillover effect (on at-risk patients not targeted in the letters), though information about Maryland's targeting criteria was limited, making it difficult to be confident of correctly identifying other at-risk patients. These findings are generally consistent with other studies of RDUR interventions targeted at anti-ulcer drugs.

These studies found that all three RDUR interventions achieved the expected effects—either a reduction or an increase in use of a specified drug, depending on the interventions.

PROGRAM COSTS

Because none of the studies conducted for this evaluation show evidence that prospective DUR is effective, we did not attempt to conduct cost-benefit analysis. However, for future comparisons, we developed benchmark estimates of program costs. The evaluator worked with investigators in Washington and Iowa, using a common framework for separating research costs (associated specifically with the demonstration and unlikely to be part of an ongoing, statewide program) from operational costs. In Washington it proved difficult to trace actual expenditures and to clearly separate the components of costs. Instead, we simply report that the aggregate budget for Project C.A.R.E. included roughly $52,000 (81 percent) for CS payments and $12,000 (19 percent) for data processing.

Iowa provided information on operations expenditures for two models of OPDUR: a basic version that Iowa investigators believed to be characteristic of the average State program and an enhanced version, based on Iowa's model and including substantial expert input into testing and validating software updates. Average yearly costs were estimated to be about $160,000 for the basic model and $200,000 for the enhanced model. Based on volume of prescriptions screened, this implies costs of $.043 per prescription (basic) and $.055 per prescription (enhanced). Varying certain assumptions used in the estimation (for example, the size of a markup for indirect costs) made little difference in these estimates.

CONCLUSION

The demonstrations tested two models of DUR and gathered valuable data, which were subsequently analyzed for evaluation of financial and outcome effects. The evaluation results show no evidence that the DUR programs, as tested in the two State Medicaid programs, had any measurable effects in reducing the frequency of drug problems or on utilization of and expenditures for prescription drugs and other medical services. In a series of carefully designed epidemiological studies, the evaluation looked for effects of the demonstrations on reducing adverse clinical outcomes that were linked to the misuse of specific drugs. However, no evidence of a reduction in problems was found in any of the five problem areas. On the other hand, our analyses of three RDUR interventions found statistically

significant effects on prescription drug utilization in the hypothesized directions.

There remains the question of whether this evaluation could have failed to detect any effects. Had there been stronger evidence that the interventions caused major changes in pharmacists' behavior, it would be more difficult to dismiss the argument that there were effects but that they were simply too small to be detected. However, such evidence was lacking, so that the conclusion that there were no effects to be found seems a reasonable one.

A primary lesson we learned is that OPDUR interventions and CS payment did not measurably add to the underlying process of quality review. However, these findings, in our opinion, do not prove that PDUR is unnecessary. Most pharmacies already have computer-based drug screening systems. With the spread of managed care and provider-sponsored organizations, centralized, on-line screening will undoubtedly spread as well. Even if pharmacists weigh and evaluate the messages they receive about potential drug problems against their own experience and information from other sources, a well-designed screening system is still valuable, as this evaluation found in surveys and focus groups with pharmacists.

The evaluation showed that, with or without payments, pharmacists routinely provide similar levels of CS. This finding casts doubt about using financial incentives to generate more CS. The participating pharmacists in Washington complained that CS payments were too low. Indeed, the payment may have been too low to adequately compensate for the time and effort of interventions.

The RDUR was relatively effective, but here too, caution would be advisable. This evaluation only reported on three interventions. The tests for effectiveness used in each of the three studies were based on quasi-experimental designs, raising questions about whether or not differences between treatment and control patients might be attributable to factors other than the interventions. Because of small sample sizes, no attempts were made to assess downstream effects of the RDUR interventions on other health services utilization and on clinical outcomes. Nonetheless, positive findings on RDUR from this evaluation support the conclusions reached by other investigators (Guo et al., 1995; Zimmerman et al., 1994; Collins et al., 1997; and Brufsky et al., 1998).

What do these findings mean for Medicaid, which is operating both prospective and retrospective DUR programs under mandate from OBRA 90? In one sense, the question of DUR policy has been rendered moot for a substantial proportion of Medicaid's client base through the spread of managed care. State Medicaid directors can debate whether or not to require specific DUR models in contracts with managed care organizations, but the operational details will generally be left to the plans. However, in most States, Medicaid remains the responsible fee-for-service payer for costly and vulnerable client populations, the adult disabled and dually eligible elderly. Therefore, policymakers need to decide whether or not to fund further demonstrations and studies, to advocate changes in DUR programs based on current knowledge, or to leave the present system in place.

The case for more research can be made on the following arguments. Based on feedback gained during the demonstrations, designers of the interventions might have done things differently. For example, although the $4 payment for a brief CS was viewed as fair by most pharmacists, the $6 payment for extended CS was not. Secondly, even on a priori grounds, the interventions tested may not have been the most promising

models of prospective DUR. For example, it would have been instructive to assess the relative performance of Project C.A.R.E. and Iowa OPDUR against a third model that combined OPDUR capabilities with CS payment.

Given the prevalent use of DUR technology, a reasonable strategy might be to propose marginal changes to current programs. Information gained in how pharmacists deliver CS and which interventions produce more results could help improve DUR policy. It was evident in both Iowa and Washington that, whatever incentives were introduced, pharmacists tended to spend a substantial amount of time counseling and educating patients. It was also clear in the three RDUR studies that prescribers responded to letters from State DUR boards targeted to specific drugs and patients. Previous research demonstrates the effectiveness, if not the cost-effectiveness, of more intensive interventions such as academic detailing. Because surveys and focus groups conducted in this evaluation depict an evolving but still distant relationship between pharmacists and prescribers, an effective use of Medicaid funds might include encouraging more RDUR and/or academic detailing programs targeted at drugs and patients where the potential harm from misprescribing is well documented. In fact RDUR programs may provide critical information that can be converted into valid screens in OPDUR systems.

Finally, a decision to stay the present course also seems reasonable given that we already have functioning DUR programs in Medicaid. Our evaluation did not identify any changes that would enhance the existing Medicaid DUR programs. Neither do the demonstration findings support adopting the CS payment as a part of the Medicaid DUR program, as it was not proven to reduce cost and use of health services.

Increased understanding of the association of inappropriate drug utilization and adverse clinical outcomes would inform any improvements to DUR, but this requires basic research. Alternatively, systems for monitoring DUR program performance would be of immediate practical use to policymakers. Because OPDUR is an operating part of many State Medicaid programs, it would be useful to develop data capture systems that, at a minimum, monitor prescription drug utilization in detail that matches the drug categories and screening criteria in operation. Similarly, standardized drug utilization indicators may be designed for tracking RDUR interventions, providing State DUR Boards and Medicaid officials with the tools needed to assess and improve program performance.

ACKNOWLEDGMENTS

The authors wish to acknowledge our many colleagues who contributed, as researchers and as advisers, to the evaluationof the Medicaid DUR Demonstrations. Authors of the evaluation final report included Marian Wrobel, Susan Foster, Pinka Chatterji, Susan Massey, Karen Wilcock, Lois Olinger (Abt Associates), Bruce Stuart (University of Maryland), Frank Ahern (Pennsylvania State University), and Dan Gilden (JEN Associates). Many others provided valuable assistance in designing analyses and reviewing project reports. These included our project consultants, HCFA staff, staff from the Georgia, Iowa, Maryland, and Washingtion Medicaid agencies and from the demonstrations, and members fo the evaluation's Technical Advisory Group.

REFERENCES

Avorn, J., and Soumerai, S.B.: A Randomized Trial of a Program to Reduce the Use of Psychoactive Drugs in Nursing Homes. *The New England Journal of Medicine* 308(24):1457-1463, 1983.

Brufsky, J.W., Ross-Degnan, D., Calabrese, D., et al.: Shifting Physician Prescribing to a Preferred Histamine-2-Receptor Antagonist. *Medical Care* 36(3):321-331, 1998.

Collins, T.M., Mott, D.A., Bigelow, W.E., and Zimmerman, D.R.: A Controlled Letter Intervention to Change Prescribing Behavior: Results of a Dual-Targeted Approach. *Health Services Research* 32(4):471-489, 1997.

Davis, D.A., Thomson, M.A., Oxman, A.D., and Haynes, R.B.: Changing Physician Performance: A Systematic Review of the Effect of Continuing Medical Education Strategies. *Journal of the American Medical Association* 274(9):700-705, 1995.

Dobie, M., and Rascati, K.L.: Documenting the Value of Pharmacist Interventions. *American Pharmacy* NS34(5):50-54, 1994.

Guo, J.J., Gibson, J.T., Hancock, G.R., and Barker, K.N..: Retrospective Drug Utilization Review and the Behavior of Medicaid Prescribers: An Empirical Marginal Analysis. *Clinical Therapeutics* 17:1174-1187, 1995.

Jay, E.D., Eynon, B.P., and Javitz, H.S.: The *Medi-Cal Therapeutic Drug Utilization Review Project: Quarterly Report and Final Analysis Report*. Menlo Park, CA: SRI International, July 1991.

Lin, E.H., Katon, W.J., Simon, G.E., et al.: Achieving Guidelines for the Treatment of Depression in Primary Care: Is Physician Education Enough? *Medical Care* 35(8):831-842, 1997.

Ray, W.A., Griffin, M.R., and Schorr, R.I.: Adverse Drug Reactions and the Elderly, *Health Affairs* 9 (Fall):114-122, 1990.

Ray, W.A., Griffin, M.R., Shaffner, W., et al.: Psychotropic Drug Use and the Risk of Hip Fracture. *The New England Journal of Medicine* 316:363-369, 1987.

Rupp, M.T., Schondelmeyer, S.W., Wilson, G.T., and Krause, J.E.: Documenting Prescribing Errors and Pharmacist Interventions in Community Practice. *American Pharmacy* NS28(9):574-580, 1988.

Rupp, M.T., DeYoung, M., and Schondelmeyer, S.W.: Prescribing Problems and Pharmacist Interventions in Community Practice. *Medical Care* 30(10):926-994, 1992.

Zimmerman, D.R., Collins, T.M., Lipowski, E.E., and Sainfort, F.: Evaluation of a DUR Intervention: A Case Study of Histamine Antagonists. *Inquiry* 31:89-101, 1994.

Reprint Requests: David Kidder, Ph.D., Abt Associates Inc., 55 Wheeler Street, Cambridge, MA 02138-1168. E-mail: david_kidder@abtassoc.com