The Ineffectiveness of Retrospective Drug Utilization Review

Editor's note: As policymakers debate adding a prescription drug benefit to Medicare, they must also seek ways to promote cost-effective use of drugs and minimize inappropriate prescribing. For more than a decade, all state Medicaid agencies and most private insurers have used computerized drug utilization review (DUR) programs to prevent or rectify potential prescribing errors. DUR can be retrospective, in which claims data are reviewed to identify patterns of drug use, or prospective, in which prescriptions are reviewed before a drug is dispensed. This Issue Brief summarizes a landmark study that suggests that retrospective DUR has had no measurable effects on outpatient drug use or clinical outcomes in the Medicaid program.

Retrospective drug utilization review (DUR) is widely used, but its effectiveness is unknown

In 1990, Congress required all state Medicaid agencies to implement DUR programs by 1993. The legislative goal was to assure that outpatient prescriptions were appropriate, medically necessary, and unlikely to have adverse medical consequences. One component of the mandated program is retrospective DUR, in which claims are reviewed to identify prescribing problems or errors. Software is used to identify problems such as drug-drug interactions, overuse (for example, early refills), drug-disease interactions, duplicate therapy, excessive or insufficient dose, and drug pregnancy contraindications.

• A typical retrospective DUR process involves computerized screening of drug claims each month to detect “exceptions” that appear to violate predetermined criteria for appropriate prescribing. If the criteria are valid, then exceptions represent prescribing errors. For example, one criterion is that a given patient should not receive more than one narcotic analgesic at one time.

• Exceptions are reviewed manually to determine whether a physician alert should be issued. Alerts are typically made through the mail, although a few programs do so by telephone. Typical alert letters include the patient’s name, the criterion that has been violated (sometimes with a reference supporting the validity of the criterion), and a statement that the clinical care of the patient remains at the discretion of the physician.
In theory, retrospective DUR programs might work through two mechanisms: direct effects and spillover effects. Direct effects apply to patients who are identified in alerts and benefit from a change in therapy. Spillover effects refer to the possibility that prescribers might apply lessons learned from receiving an alert about one patient to the care of other patients in the future.

In practice, however, it is not clear that these retrospective DUR programs work. Despite a decade of implementation of retrospective DUR in all Medicaid programs and in most private sector drug benefit programs, the effectiveness of this mechanism has not been evaluated.

Hennessy and colleagues used data from six Medicaid programs over four years to assess the direct effects of retrospective DUR on clinical outcomes in patients with exceptions, and to assess changes in the rate of exceptions over time, as might be expected if there were spillover effects. Exceptions prior to DUR implementation were identified using archived claims data and the same algorithms currently used.

The study sample included states from the four geographic regions of the U.S., and both large and small Medicaid programs. These states used the same DUR software vendor. The analysis was limited to criteria that identified drug-disease, drug-drug, and duplication problems because these criteria account for about 80% of alerts issued.

For each state, the investigators calculated the rate of exceptions per thousand prescriptions. The primary analysis assumed a two-month lag between implementation and effects, although lags of one month and four months were also considered.

Clinical outcomes were measured by all-cause hospitalization rates, and cause-specific hospitalizations following specific exceptions (for example, hospitalizations for upper gastrointestinal bleeding after exceptions to gastrointestinal criteria). The primary observation period after each exception was 120 days, although analyses were also performed using periods of 90 and 150 days.

The study included patients for whom an exception was issued, regardless of whether an alert was sent. This was done to permit identification of a valid comparison group for before-after comparisons, because the process of flagging exceptions for alerts is done using individual judgment and implicit criteria, and therefore is not reproducible.

Changes in the rate of exceptions, which represent potential prescribing problems, were calculated before and after implementation of retrospective DUR.
The average rate of exceptions was similar among states, ranging from 8 to 13 exceptions per 1000 prescriptions.

Between 1% and 25% of exceptions identified with the software resulted in an alert. Each year, the programs issued 761 to 3236 alerts.

Smaller states showed the highest alert rates. The alert rate per state was not associated with changes in the exception rate coincident with retrospective DUR implementation.

After adjusting for prior trends in drug use, the investigators found no significant changes in the rate of exceptions associated with implementation of retrospective DUR. No effect on exceptions was evident when considering a 1-month, 2-month or 4-month lag time between DUR implementation and any effects.

The rate of hospitalizations in patients with exceptions was measured before and after DUR implementation to ascertain any direct clinical benefit attributable to the program. The investigators also examined whether retrospective DUR might have an effect primarily on exceptions based on criteria that resulted in the highest number of physician alerts.

The incidence of hospitalization of patients within 120 days of an exception was virtually identical before and after implementation of retrospective DUR (22% vs. 23%). After adjusting for potential individual-level factors, retrospective DUR had no effect on hospitalizations in the 90, 120 or 150 days after the exception.

Retrospective DUR had no effect on hospitalizations in the subgroup of patients identified by the top five criteria that produced the highest number of physician alerts. Although no direct before and after comparison of patients with alerts was possible, this suggests that retrospective DUR did not affect hospitalizations even in the group where one might expect to find the largest effect.

Even when the analysis was limited to specific exceptions thought to be related to specific causes for hospitalizations (such as upper gastrointestinal bleeding, myocardial infarction or angina), no effect of retrospective DUR was detected.

This study included millions of Medicaid patients in six states that paid for about 63 million prescriptions each study year, and identified more than 600,000 exceptions by retrospective DUR each year. The study methods adjusted for secular and time trends in drug use, and produced results consistent across states, exception categories, and types of patients. Retrospective DUR appears to have no measurable effects on prescribing errors or clinical outcomes in the Medicaid program.
POLICY IMPLICATIONS

Continued

- Given the lack of evidence for the effectiveness of retrospective DUR, policymakers should consider withdrawing the legislative mandate for such review in the Medicaid program.

- As legislators discuss a new drug benefit in the Medicare program, they should not include requirements for retrospective DUR as it is currently implemented in Medicaid. It is possible that future DUR programs can be improved to confer clinical benefits, but efforts to improve prescribing should be shown to be effective before being widely adopted.

- The reasons for the apparent lack of effect of retrospective DUR need exploration. These reasons might include the unknown validity of many criteria, the low alert rate, the time lag from the exception to the alert, and the modest effectiveness of alert letters in changing prescribing, particularly when letters are not based on any underlying reasons physicians might have had for prescribing the medication.


This Issue Brief was supported by the Agency for Healthcare Research and Quality, Center for Education and Research on Therapeutics, University of Pennsylvania School of Medicine (U18 HS10399). The opinions expressed are those of the authors and not the funding agency.

Published by the Leonard Davis Institute of Health Economics, University of Pennsylvania, 3641 Locust Walk, Philadelphia, PA 19104-6218.

Janet Weiner, MPH, Associate Director for Health Policy, Editor
David A. Asch, MD, MBA, Executive Director

Visit us on the web at www.upenn.edu/ldi

Issue Briefs synthesize the results of research by LDI’s Senior Fellows, a consortium of Penn scholars studying medical, economic, and social and ethical issues that influence how health care is organized, financed, managed, and delivered in the United States and internationally. The LDI is a cooperative venture among Penn schools including Dental Medicine, Medicine, Nursing and Wharton, and the Children’s Hospital of Philadelphia. For additional information on this or other Issue Briefs, contact Janet Weiner (e-mail: weinerja@mail.med.upenn.edu; 215-573-9374).

© 2003 Leonard Davis Institute