Retrospective Drug Utilization Review, Prescribing Errors, and Clinical Outcomes

Sean Hennessy, PharmD, PhD
Warren B. Bilker, PhD
Lan Zhou, PhD
Anita L. Weber, PhD
Colleen Brensinger, MS
Yanlin Wang, MS
Brian L. Strom, MD, MPH

Context Retrospective drug utilization review is required of all state Medicaid programs and is performed by most private-sector prescription programs. However, it has not been shown to improve clinical outcomes or reduce the rate of potential prescribing errors, known as “exceptions.”

Objective To look for an effect of retrospective drug utilization review on the rate of exceptions and of clinical outcomes in patients with an exception.

Design, Setting, and Participants Longitudinal ecologic study of the rate of exceptions, controlling for preintervention trends and calendar time; and a cohort study of all-cause and cause-specific hospitalizations in patients with an exception, controlling for potential individual-level confounders in 6 Medicaid programs using the same software in the mid-1990s.

Main Outcome Measures The rate of exceptions was examined as a function of retrospective drug utilization review implementation. In addition, before-after comparisons were made of the incidence of all-cause and cause-specific hospitalization in patients with exceptions.

Results We found no reduction in the rate of exceptions coincident with retrospective drug utilization review implementation (rate increase, 0.064 exceptions per 1000 prescriptions per month; 95% confidence interval [CI], 0.006 to 0.133). We also found no effect of retrospective drug utilization review on the incidence of all-cause hospitalization (odds ratio, 0.99; 95% CI, 0.98 to 1.00) or cause-specific hospitalization. These results persisted in multiple subgroup analyses. Study states intervened using physician alerts in between 1% and 25% of exceptions.

Conclusions We were unable to identify an effect of retrospective drug utilization review on the rate of exceptions or on clinical outcomes. Given the lack of evidence for effectiveness, and suggestions from previous research of possible harm, policy-makers should consider withdrawing the legislative mandate for retrospective drug utilization review.

JAMA. 2003;290:1494-1499
Retrospective drug utilization review programs are hypothesized to work through 2 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 to the care of future patients. Although a few studies have shown that drug utilization review programs can modestly reduce "undesirable" prescribing for identified patients (eg, reducing long-term use of histamine H2 receptor antagonists), the only study to look for clinical benefit has not appeared in the peer-reviewed literature, and it found no effect. The available studies examining spillover effects have produced conflicting results. Continuation of such programs despite a lack of evidence for their effectiveness has led some to suggest that they are a "boondoggle."1

The aims of this study were to look for an effect of retrospective drug utilization review on the rate of exceptions, as might be caused by spillover effects; to look for a direct effect of retrospective drug utilization review in reducing adverse clinical outcomes in patients with exceptions; and to report the number of alerts made by these programs, which has not been reported in the scientific literature yet is important in interpreting the results of the first 2 aims.

METHODS
Overview
This study was performed using data from 6 Medicaid programs, which used the same retrospective drug utilization review software vendor (ProVantage). We studied Medicaid programs using this product because it is commonly used, because the vendor and its Medicaid clients expressed interest in conducting a rigorous evaluation, and because the necessary archived data were available. We used all data available from the vendor prior to review implementation and had more than 1 year of data following implementation in the last state to start. The dates for which data were available for each state are shown in Figure 1.

Medicaid data are used extensively for epidemiologic and policy research. Information on dispensed prescriptions is audited and has been shown to be highly accurate. Although diagnoses are generally not recorded with the same degree of accuracy as are dispensed prescription drugs, many diagnoses have been found to be sufficiently accurate to use for research purposes. Extensive quality assurance analyses were performed on the data for this study.

This study was approved by the University of Pennsylvania’s Committee on Subjects Involving Human Beings. The committee waived patient consent.

Effect of Retrospective Drug Utilization Review on Exception Rates
Identification of Exceptions. The vendor’s software contains criteria to identify drug-disease interactions, drug-drug interactions, therapeutic duplication, excessive dose, insufficient dose, and excessive medication duration. We limited our study to drug-disease, drug-drug, and duplication criteria because these accounted for approximately 80% of alerts. We also limited the study to criteria that, according to the vendor, were used to identify exceptions in all 6 states for the entire intervention period. (These exceptions are listed at http://ceb.med.upenn.edu/cert/retrospective_dur_appendices.pdf.) Exceptions were identified retrospectively at the University of Pennsylvania using an archive of the same claims data that the programs themselves had used.

Analysis of Exceptions. For each state and for each month, we calculated the rate of exceptions per thousand prescriptions. To examine the effect of retrospective drug utilization review on the exception rate, we plotted the data visually, aligned so that implementation was considered month 0, with the preimplementation period considered as negative months and postimplementation period as positive months. We examined the plot for the presence of a negative inflection point at or following implementation, which would be indicative of a reduction in the rate of exceptions associated with retrospective drug utilization review.

Linear regression was used to calculate effect estimates and inferential statistics. We predicted a priori that any spillover effects of retrospective drug utilization review would be observed as a change in slope in the rate of exceptions vs time coincident with implementation. Therefore, we modeled the monthly rate of exceptions as a linear...
function of state (coded as indicator variables), calendar time in months, and retrospective drug utilization review intervention status. We also included state x time interaction terms to allow the preintervention slope to vary by state. Finally, we included an intervention x time interaction term as the parameter of interest. This term is interpretable as the before vs after intervention difference in slope of the rate of exceptions, assuming a common effect across states.12 The primary analysis assumed a 2-month lag between implementation and onset of effect.

We also performed secondary analyses assuming lags of 1 and 4 months. We conducted a secondary analysis that adjusted for autocorrelation by fitting a generalized linear model with an autoregressive correlation matrix.13 The results of this analysis were very similar to those of the primary analysis (data not shown). Finally, we also performed a secondary analysis using Poisson regression,14 which gave similar results (data not shown). This was expected because the Poisson distribution is well approximated by the Gaussian distribution when the number of outcomes is large.

**Individual-Level Analyses of Clinical Outcomes for Patients With an Exception**

Identification of Eligible Patients. We studied individuals for whom an exception occurred, regardless of whether an alert was sent. This was performed to permit identification of a valid comparison group for before-after comparisons because the process of flagging exceptions for alerts is done by individuals using implicit criteria and therefore is not reproducible. Exceptions occurring after retrospective drug utilization review was implemented were considered the exposed group; those occurring before implementation were considered the unexposed group. Because of the incompleteness of Medicaid claims for those aged 65 years or older,15 we also performed subanalyses that excluded this group. The primary analysis examined exceptions to the entire set of study criteria. We also examined the 5 criterion subtypes that resulted in the highest number of alerts in the states included in this analysis. This was considered a “high-dose” subgroup in which an effect might be more likely.

**Determination of Outcome.** All-cause hospitalization was the primary outcome. Therefore, we excluded from the individual-level analyses 3 states for which we could not distinguish inpatient from outpatient claims.

We also performed analyses of cause-specific hospitalization following specific exceptions. In particular, we examined the incidence of hospitalization for upper gastrointestinal bleeding following exceptions to gastrointestinal criteria (warfarin + gastrointestinal acid/peptic disorders, warfarin + aspirin, and NSAID + NSAID [nonsteroidal anti-inflammatory drug]); hospitalization for myocardial infarction following β-agonist + angina pectoris exceptions; and hospitalization for angina pectoris following β-agonist + angina pectoris exceptions. These criteria were selected because they were commonly used as the basis of alert letters in the study states, because exceptions were expected to be common, and because the expected outcomes could be reliably identified using claims data.

The specific International Classification of Diseases, Ninth Revision codes used to identify cause-specific hospitalizations are available at http://ccceb.med.upenn.edu/cert/retrospective_dur_appendices.pdf. Because of privacy concerns, medical records were not available to permit the validation of study diagnoses. However, the accuracy of administrative claims for gastrointestinal bleeding,16 myocardial infarction,17 and angina pectoris18 all exceed 85%.

The primary observation period was 120 days following each exception. Secondary analyses were performed using periods of 90 and 150 days following the exception.

We restricted the analysis to exceptions for which there was at least 1 medical claim after the 120-day observation window. This approach helps ensure that outcomes are not missed because of loss of Medicaid eligibility.19

**Outcome Analysis.** For each outcome, we first performed a contingency table analysis for the association between intervention status and all-cause or cause-specific hospitalization to determine whether the risk of adverse clinical outcomes was influenced by retrospective drug utilization review implementation. We then used logistic regression to examine the effects of potential confounding factors. State and calendar year were included in all models. We considered other potential confounding factors individually, including them in the final model if they changed the odds ratio of interest by 10% or more.20,21 These potential confounding factors included sex, age (categorized as <15, 15-34, 35-44, 45-54, 55-64, 65-74, and ≥75), number of physicians seen in the past 3 months, and the number of pharmacies dispensing medications in the past 3 months. We also examined as potential confounders drug classes included in the Chronic Disease Score22 (see http://ccceb.med.upenn.edu/cert/retrospective_dur_appendices.pdf).

**Descriptive Analyses of the Rate of Alerts**

Federal law requires each Medicaid program to provide the Centers for Medicare & Medicaid Services (CMS) with an annual report of retrospective drug utilization review activities.23 These reports typically include the number of alerts per drug class, per criterion type (eg, the number of duplication alerts for opioid analgesics) for the year of interest. We requested annual reports from the coordinator of each study state during the study period, as well as from CMS. We used these reports to calculate the average annual number of alerts to drug-disease, drug-drug, and duplication criteria in each state.

The statistical analysis was performed using SAS version 8.1 (SAS Institute Inc, Cary, NC). P<.05 was set as the level of significance.

**RESULTS**

TABLE 1 lists the characteristics of the study states. The study sample in-
cluded states from all 4 geographic regions of the United States and both large and small Medicaid programs. The timing of the available data with respect to state, calendar time, and implementation is shown in Figure 1.

**Effect of Retrospective Drug Utilization Review on Exception Rates**

Figure 2 shows the combined monthly exception rate per 1000 prescriptions, stratified by state, with implementation positioned at month 0. Overall, there was no visually apparent decline in the exception rates coincident with retrospective drug utilization review implementation. This observation was confirmed by the results of the regression models. In particular, in the primary analysis that assumed a 2-month lag, the slope of exceptions over time showed an increase, rather than a decline, in exceptions associated with retrospective drug utilization review of 0.064 exceptions per 1000 prescriptions per month, which was consistent with random variation (95% confidence interval [CI], -0.006 to 0.133). There was also no apparent decline in the slope using a 1-month lag (slope increase, 0.053 exceptions per 1000 prescriptions per month; 95% CI, -0.015 to 0.121) or a 4-month lag (slope increase, 0.001 exceptions per 1000 prescriptions per month; 95% CI, -0.067 to 0.070). A statistical test failed to reject the null hypothesis of a common effect across states (2-month lag, \[P = .31\]).

None of the subanalyses based on exception category showed a decline in slope coincident with retrospective drug utilization review. The slope for drug-disease exceptions increased by 0.015 exceptions per 1000 prescriptions per month (95% CI, -0.032 to 0.063); the slope for drug-drug exceptions increased by 0.006 exceptions per 1000 prescriptions per month (95% CI, -0.001 to 0.011); and the slope for duplication exceptions increased by 0.043 exceptions per 1000 prescriptions per month (95% CI, 0.011 to 0.075).

In the subanalysis of the 5 subgroups of criteria that resulted in the highest number of alerts (ie, the “high-dose” subgroup), the slope of the rate of exceptions associated with retrospective drug utilization review increased by 0.022 exceptions per 1000 prescriptions per month (95% CI, 0.009 to 0.034).

**Individual-Level Analyses of Clinical Outcomes for Patients With an Exception**

Table 2 summarizes the association between retrospective drug utilization review implementation and the incidence of all-cause and cause-specific hospitalization in those with an exception, using a 120-day observation window. All of the odds ratios were very close to 1.0. The results using 90-day and 150-day time windows were very similar (data not shown). Age, sex, and any of potential drug confounders did not change the odds ratios of interest by 10% or more. However, adjusting for these factors individually tended to increase rather than reduce the odds ratios associated with retrospective drug utilization review. Analyses restricted to those younger than 65 years produced results similar to the overall analysis (data not shown).

**Descriptive Analyses of the Rate of Alerts**

Table 3 shows the average annual number of prescriptions, exceptions, and alerts to drug-drug, drug-disease, and duplication criteria, with rate of exceptions per 1000 prescriptions, and number of alerts expressed as a percentage.
of exceptions. The average exception rate was similar among states, ranging from 8 to 13 exceptions per 1000 prescriptions.

The programs studied issued from 761 to 3236 alerts per year. This translates into an alert rate that ranged nearly 25-fold from 1% to 25%, with smaller states having the highest rates.

This study has a number of strengths. It is the first study to examine the effects of retrospective drug utilization review programs on improving clinical outcomes in individuals with exceptions. It based inclusion on computerized criteria, avoiding lack of comparability inherent in manual methods. The ecologic analyses accounted for potential differences among states in the rate and scope of exceptions and for secular trends. Results were consistent across subgroups, including high-dose subgroups where one would expect to find the largest effect.

The only previous study to examine clinical effects of retrospective drug utilization review programs not included in the current study. However, because the overall structure of programs appears similar, dramatic differences in program effectiveness seem unlikely. Furthermore, although data for this study ended in 1996, a recent article suggests that the function of Medicaid retrospective drug utilization review programs has changed little since then.

The previous study to examine clinical effects of retrospective drug utilization review has not appeared in a peer-reviewed publication and it also failed to detect an effect. One study has

### Table 2. Adjusted Odds Ratios for All-Cause or Cause-Specific Hospitalization Associated With Having an Exception Identified in the Postintervention Period vs the Preintervention Period

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Outcome</th>
<th>Before Retrospective Drug Utilization Review</th>
<th>After Retrospective Drug Utilization Review</th>
<th>Odds Ratio (95% CI) for Outcome After Implementation*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Exceptions</td>
<td>Exceptions With Outcome, No. (%)</td>
<td>No. of Exceptions</td>
<td>Exceptions With Outcome, No. (%)</td>
</tr>
<tr>
<td>All study</td>
<td>488 198</td>
<td>106 167 (22)</td>
<td>1 618 323</td>
<td>373 997 (23)</td>
</tr>
<tr>
<td>Top 5 by number of alerts</td>
<td>68 274</td>
<td>15 196 (22)</td>
<td>264 511</td>
<td>57 481 (22)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>61 408</td>
<td>219 (0.4)</td>
<td>149 512</td>
<td>636 (0.4)</td>
</tr>
<tr>
<td>β-Agonist + angina</td>
<td>11 581</td>
<td>79 (0.7)</td>
<td>46 208</td>
<td>96 (0.2)</td>
</tr>
<tr>
<td>Hospital admission for angina</td>
<td>11 581</td>
<td>543 (5)</td>
<td>46 208</td>
<td>1686 (4)</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.

*Adjusted for state and calendar year using an observation period of 120 days following the exception (defined as potential prescribing errors).

### Table 3. Average Number of Prescriptions, Exceptions, and Alerts per Year

<table>
<thead>
<tr>
<th>State</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average annual No. of prescriptions</td>
<td>32 481 1567</td>
<td>16 118 898</td>
<td>10 045 008</td>
<td>2 885 989</td>
<td>992 124</td>
<td>380 859</td>
</tr>
<tr>
<td>Average annual No. of exceptions to drug-drug, drug-disease, and duplication criteria per year*</td>
<td>265 174</td>
<td>180 879</td>
<td>127 245</td>
<td>22 579</td>
<td>13 097</td>
<td>3918</td>
</tr>
<tr>
<td>Average annual rate of exceptions per 1000 prescriptions</td>
<td>8</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Average annual No. of alerts for drug-drug, drug-disease, and duplication criteria</td>
<td>3236</td>
<td>3200</td>
<td>1347</td>
<td>761</td>
<td>1034</td>
<td>969</td>
</tr>
<tr>
<td>Alerts, % of exceptions</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>25</td>
</tr>
</tbody>
</table>

*Includes only 1 exception per criterion per person per month.
even suggested that retrospective drug utilization review programs could cause harm, such as discontinuation of drugs in patients who still need them, abrupt discontinuation of drugs that should be tapered, and use of undesirable alternatives.27

Potential reasons for the apparent lack of effect of retrospective drug utilization review 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 may have had for prescribing the medication.28

Given the lack of evidence for the effectiveness of retrospective drug utilization review, and the suggestion of possible harm, policymakers should consider withdrawing the legislative mandate for retrospective drug utilization review. This would agree with the recommendations of the 1969 Task Force on Prescription Drugs that drug review programs should not be implemented in a widespread manner until they are shown to be effective. It is possible that future drug utilization review programs can be improved to confer clinical benefits. However, programmatic efforts to improve prescribing should be shown to be effective and free of major detrimental effects before being widely adopted.

Author Contributions: Study concept and design: Hennessy, Blikер, Strom. Acquisition of data: Hennessy, Strom. Analysis and interpretation of data: Hennessy, Bliker, Zhou, Weber, Brensinger, Wang, Strom. Drafting of the manuscript: Hennessy.

Critical revision of the manuscript for important intellectual content: Hennessy, Bliker, Zhou, Weber, Brensinger, Wang, Strom.


Obtained funding: Hennessy, Strom.

Study supervision: Hennessy, Strom.

Funding/Support: This research was funded by the National Institute on Aging (1R01 AG14601), the Agency for Healthcare Research and Quality (K12 HS00066 and U18-HS10399), and the Pharmark Corporation. Dr Hennessy is supported by a career development award from the National Institute on Aging (K123AG003987).

Acknowledgment: Helene L. Lipton, PhD, Richard Platt, MD, MSc, Wayne A. Ray, PhD, Stephen B. Soumerai, ScD, and Andy S. Stergachis, PhD, served on the study’s academic advisory panel. Jesse A. Ber- lin, ScD, Jeffrey L. Carson, MD, Sandra A. Norman, MD, and Wendy P. Stephenson, MD, MS, MPH, pro- vided methodologic input. Stephen Durbor pro- vided database support. M. Lee Morse, Aida LeRoy, PharmD, Peter F. Hoffman, MD, and Kevin Hall, MD provided data and technical support from the software and data vendor. Drug utilization review coordi- nators from the study states provided valuable in- put. Cynthia Lyon provided reports on behalf of CMS. Maureen Moffett provided research assistance. Cathy Mulvaney and Sherri Ghee provided se- cretarial assistance.

REFERENCES


