Comparing the Medicaid Retrospective Drug Utilization Review Program Cost-Savings Methods Used by State Agencies

Sergio I. Prada, MPA, PhD

BACKGROUND: The Medicaid Drug Utilization Review (DUR) program is a 2-phase process conducted by Medicaid state agencies. The first phase is a prospective DUR and involves electronically monitoring prescription drug claims to identify prescription-related problems, such as therapeutic duplication, contraindications, incorrect dosage, or duration of treatment. The second phase is a retrospective DUR and involves ongoing and periodic examinations of claims data to identify patterns of fraud, abuse, underutilization, drug–drug interaction, or medically unnecessary care, implementing corrective actions when needed. The Centers for Medicare & Medicaid Services requires each state to measure prescription drug cost-savings generated from its DUR programs on an annual basis, but it provides no guidance or unified methodology for doing so.

OBJECTIVES: To describe and synthesize the methodologies used by states to measure cost-savings using their Medicaid retrospective DUR program in federal fiscal years 2014 and 2015.

METHOD: For each state, the cost-savings methodologies included in the Medicaid DUR 2014 and 2015 reports were downloaded from Medicaid’s website. The reports were then reviewed and synthesized. Methods described by the states were classified according to research designs often described in evaluation textbooks.

DISCUSSION: In 2014, the most often used prescription drugs cost-savings estimation methodology for the Medicaid retrospective DUR program was a simple pre-post intervention method, without a comparison group (ie, 12 states). In 2015, the most common methodology used was a pre-post intervention method, with a comparison group (ie, 14 states). Comparisons of savings attributed to the program among states are still unreliable, because of a lack of a common methodology available for measuring cost-savings.

CONCLUSION: There is great variation among states in the methods used to measure prescription drug utilization cost-savings. This analysis suggests that there is still room for improvement in terms of methodology transparency, which is important, because lack of transparency hinders states from learning from each other. Ultimately, the federal government needs to evaluate and improve its DUR program.

KEY WORDS: cost-avoidance, cost-savings, Medicaid Drug Utilization Review program, methodology transparency, postintervention, preintervention, retrospective DUR

The Medicaid Drug Utilization Review (DUR) program promotes patient safety through state-administered drug utilization management tools and systems. Medicaid DUR is a 2-phase process that is conducted by Medicaid state agencies. The first phase is a prospective DUR, which involves electronically monitoring prescription drug claims to identify problems, such as therapeutic duplication, contraindications, incorrect dosage, or duration of treatment. The second phase is a retrospective DUR, which involves ongoing and periodic examination of claims data to identify patterns of fraud, abuse, gross overuse, or medically unnecessary care, and implements corrective action when needed. By law, each state Medicaid program is required to submit an annual report on the operation of its DUR program, including cost-savings related to the operation of such a program.

The literature on cost-savings or cost-avoidance estimations in the context of a retrospective DUR program, whether in the United States or abroad, is scarce. Of note, there is a domestic and an international industry of private companies that offer pharmacy benefit management programs, for which a retrospective DUR is part of the portfolio. These private companies do not reveal the methodology used with their clients, that is, health plans or governments. Thus, researchers or providers entering...
KEY POINTS

➤ There is great variation among US states in the methods used to measure drug utilization costs.
➤ This study reviewed the methodologies used to measure cost-savings and cost-avoidance in the Medicaid retrospective Drug Utilization Review (DUR) program for 2014 and 2015.
➤ In 2014, most states used a pre-post intervention method without a comparison group relative to the Medicaid retrospective DUR program.
➤ In 2015, the most common method was a pre-post intervention with a comparison group.
➤ Using a DUR program, states saved an average of $1,122,094 on drug utilization, but comparisons of the states’ results and descriptive statistics are unreliable.
➤ To estimate the impact of Medicaid DUR programs, a pre-post nonequivalent control group quasi-experimental design is preferred.
➤ The creation of publicly available methods for measuring cost-savings using the retrospective DUR program is also needed.

The pharmacy benefit management and healthcare fields have very few sources to which they could resort. As health systems around the world mature and incorporate or increase demand for DUR programs, the lack of literature becomes critical for an international audience. A seminal report proposed guidelines for estimating the impact of Medicaid DUR programs, but this 1994 report is not retrievable electronically.1 Several authors have advocated for rigorous impact evaluation studies of retrospective DUR programs using quasi-experimental designs.2-4 A case study of histamine antagonists using a quasi-experimental preintervention-postintervention (henceforth “pre-post,” with a comparison group) design, in which drug utilization data are collected before and after the intervention, with a comparison group, concluded that the use of a comparison group is critical in the evaluation of the impact of DUR programs.2

The results of a study of the effectiveness of an intervention with a DUR letter using a pre-post, nonequivalent control group and a quasi-experimental design showed that interventions by physicians and pharmacists reduced drug spending.3 In addition, 3 separate evaluations of retrospective DUR interventions were reported—2 used a pre-post estimator and 1 used a control group selected based on computer algorithms that mimicked the target patients for the intervention, but none of these 3 reports estimate cost-savings.4

A 2005 report prepared for the Kaiser Family Foundation provides detailed descriptions of more than 30 specific cost-control strategies related to drug utilization.5 This report examines each strategy independently, providing a description of how each approach functions and is used by private or public payers, and discusses the effectiveness and cost-savings potential of the approaches when available; however, the report does not discuss cost-savings measurement methods.5

This article is the result of a problem I encountered when I was looking for articles that describe methods to measure cost-savings and cost-avoidance in a retrospective DUR program, but I could not find any. I then looked at the Medicaid DUR program and found a variety of approaches by different states, from no disclosure of methods to very detailed methods. There is a remarkable lack of consistency in assessing cost-savings methodology, which makes comparisons of monetary savings impossible.

The goal of this article is to review and synthesize the current state of cost-savings methodologies used by states in the Medicaid retrospective DUR program. In addition, the article provides a brief summary of the current methods used for cost-savings estimates in these programs to provide support for healthcare researchers and providers entering the DUR programs field.

Methods

For each state, I downloaded the cost-savings methodology from the DUR 2014 and DUR 2015 reports to the Centers for Medicare & Medicaid Services (CMS) from the Medicaid website.6,7 A total of 100 reports were downloaded, 2 for each state; however, reports from Arizona were unavailable, because almost all of its Medicaid program beneficiaries are enrolled in managed care organizations. I reviewed documents to extract the methodology used by each state to estimate the retrospective cost-savings for the DUR program. Next, I grouped the methodologies into categories using criteria such as the inclusion of comparison or control groups, the period of intervention, therapy groups, and the type of drug.

For all states, except Wyoming, only expenditures in pharmacy services were reported. In its federal fiscal year (FFY) 2014 report, Wyoming was the only state to measure medical and pharmacy cost-avoidances separately.6

Discussion

In FFY 2014, 35 of the 50 states and the District of Columbia reported having a retrospective DUR program, 13 reported using other DUR programs (ie, a prospective DUR), 2 reported no DUR program, and 1 state’s report (ie, Arizona) was unavailable.6,7

In FFY 2015, the situation improved slightly, with 36 states reporting having a retrospective DUR program, 13
Medicaid Retrospective DUR Program Cost-Savings

reported using other DUR programs (ie, a prospective DUR), 1 reporting having no DUR program, and 1 state’s (ie, Arizona) report was unavailable.\(^5,^6\)

In 2014, of the 35 states that reported cost-savings as a result of having a retrospective DUR program, 28 shared details regarding the methodology used, and 7 did not. In 2015, 30 states shared details on their methodologies, and 6 did not.

After reviewing the 58 retrospective DUR reports (Table 1) that included methodologic details, it was evident that 3 different methods were used to estimate cost-savings or cost-avoidance, including (1) a pre-post method with a comparison group, (2) a pre-post method without a comparison group, and (3) direct and indirect effects. The order in which the methods are shown in Table 1 represents their frequency of use in 2015 and does not imply rigorousness. Of note, in 2014 most states used a pre-post intervention without a comparison group method, whereas the most frequently used method was a pre-post intervention with a comparison group in 2015.\(^5,^6\)

Table 1 provides details on which method each state used by fiscal year. Most states continued to use the same method in 2014 and in 2015.\(^5,^6\) The state of Ohio, which did not report results in FFY 2014, subsequently used a combination of retrospective and prospective DUR methodologies in FFY 2015. Similarly, in 2015, Pennsylvania, Rhode Island, and West Virginia began using the pre-post intervention with a comparison group method, which suggests an improvement in methodology, because it is a more rigorous technique. Different methodologies are used by different states.\(^5,^6\)

Table 2 shows the cost-savings estimation methodologies that were used.

**Pre-Post Intervention with a Comparison Group**

In this design, the intervention starts with an alert (ie, a letter) sent to the prescribing physician or to the pharmacy provider, because a potential drug therapy problem was identified for a recipient. The potential drug therapy problems include drug–disease interactions, drug–drug interactions, or the overutilization, underutilization, and appropriateness of therapies.

In this design, the intervention group includes patients who were reviewed and had their clinical significance risk for potential drug utilization problems confirmed by a clinical pharmacist, and their provider receives an intervention letter. Then, a control or comparison group is formed from a random group of recipients who have had at least 1 claim for any drug in the preintervention and postintervention periods (to ensure that Medicaid beneficiaries remain members, and that they have not died during the period of the analysis) and have not been chosen for intervention letters.

<table>
<thead>
<tr>
<th>Table 1 Reporting Status of States, by Medicaid DUR Annual Survey Report</th>
<th>Federal fiscal year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methodology</strong></td>
<td><strong>2014</strong></td>
</tr>
<tr>
<td><strong>Retrospective DUR program in use</strong></td>
<td></td>
</tr>
<tr>
<td>Cost-savings reported, methodology available</td>
<td></td>
</tr>
<tr>
<td>Pre-post intervention with a comparison group</td>
<td>AL, AR, CT, DE, IN, KS, MT, NY, ND, SD, WI</td>
</tr>
<tr>
<td>Pre-post intervention without a comparison group</td>
<td>DC, FL, IA, LA, MD, MO, MN, NE, RI, TX, WV, WY</td>
</tr>
<tr>
<td>1. By specific medical diagnostic or intervention</td>
<td>DC, MO, MN, TX, WV</td>
</tr>
<tr>
<td>2. Gross difference</td>
<td>FL, MD, NE, RI, WY</td>
</tr>
<tr>
<td>3. By specific medications</td>
<td>LA</td>
</tr>
<tr>
<td>4. Patient- and problem-focused DUR</td>
<td>IA</td>
</tr>
<tr>
<td>Direct and indirect effects</td>
<td>KY, MI, NH, TN, VA</td>
</tr>
<tr>
<td>Cost-savings reported, methodology not available</td>
<td>ID, NM, NC, UT, MS, PA, OK</td>
</tr>
<tr>
<td>Other DUR programs used</td>
<td>CO, MA, AK, CA, GA, IL, ME, NV, NJ, OR, SC, VT, WA</td>
</tr>
<tr>
<td>No DUR program</td>
<td>HI, OH</td>
</tr>
</tbody>
</table>

NOTE: A report is available on the Medicaid website for all states, except for Arizona.

\(^a\)Change in methodology reported or reporting status compared with 2014. Ohio used a combination of retrospective and prospective DUR methodologies in 2015.

\(^b\)MS reported educational retrospective DUR without cost-savings estimate in both years.

\(^c\)Prospective DUR, maximum allowable cost, preferred drug list, dose optimization, maximum quantity limit analysis, among others.

\(^d\)DUR indicates drug utilization review.

Sources: References 6, 7.

<table>
<thead>
<tr>
<th>Table 2 Cost-Savings Estimation Methodologies</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-post intervention with a comparison group</strong></td>
<td></td>
</tr>
<tr>
<td>Total estimated cost-savings(^1)</td>
<td>((C(^{pre}) – C(^{post})) + (C(^{pre}) – C(^{post}))) – (C(^{pre}) – C(^{post}))</td>
</tr>
<tr>
<td>Annualized cost-savings for postintervention-eligible patients (estimated)(^2)</td>
<td>[(S(^{pre}) – S(^{post})) * N(^{pre})] / 4</td>
</tr>
<tr>
<td>Annual cost-savings for postintervention-eligible patients (estimated)(^3)</td>
<td>[(S(^{pre}) – S(^{post})) * N(^{pre})] + [(S(^{pre}) – S(^{post})) * N(^{post})] + [(S(^{pre}) – S(^{post})) * N(^{pre})] + [(S(^{pre}) – S(^{post})) * N(^{post})]</td>
</tr>
<tr>
<td><strong>Pre-post intervention without a comparison group, by specific medications</strong></td>
<td></td>
</tr>
<tr>
<td>Quarterly drug cost-savings per recipient (estimated)(^4)</td>
<td>Q(^{post}) – Q(^{pre})</td>
</tr>
</tbody>
</table>

\(^1\)Total cost-savings are calculated separately by single and multiple interventions.

\(^2\)Total cost-savings are calculated separately by intervention category (medical, diagnostic).

\(^3\)Total cost-savings are calculated as the sum of quarterly savings (t = 1, 2, 3, and 4).

\(^4\)Cost-savings are estimated quarterly by clinical groups (all patients, improved patients, and unimproved patients).

\(C\) indicates total cost per period; CM, added cost of patients with only multiple interventions; CU, added cost of patients with only 1 intervention; I, intervention group; N, number of eligible patients; POST, postintervention period; PRE, preintervention period; R, comparison group; S, total cost per eligible patient by period.

Source: Reference 7.
Patients who received only 1 alert (ie, a single intervention) and who received more than 1 alert (ie, multiple interventions) are separated by intervention and control groups, because a smaller effect on cost-savings is expected on patients who have had an intervention than on those with multiple interventions.

In addition, recipients are analyzed using 180 days (ie, 6 months) of claims data before and after the intervention date. In addition, a null period of 14 days was included between the preanalysis and postanalysis periods to allow for the delivery and circulation of intervention letters. The final step in the process is that a difference-in-difference estimator is used.

For a detailed calculation used by any of the states using the methodology, please refer to the FFY 2014 Alabama Medicaid DUR report.\(^6\)

1. Pre-Post Intervention without a Comparison Group. This methodology uses 2 periods: the number of months before the implementation of edit or criteria, and the number of months after implementation of the DUR program. In the DUR program jargon, edit or criteria refers to the events that trigger a drug use recommendation, such as avoiding, replacing, or lowering its doses. The estimator for the intervention’s effect on drug expenditures is the difference in the amount paid between the preimplementation and postimplementation periods.

The most common time intervals pre- and postimplementation are 6 months preintervention and 6 months postintervention. Other states used a shorter time span of 3 months. Although the method must include a null period, this was not explicit in any of the states’ reports, except for Florida, in which a nonstrict period of 3 months after the deployment of the edit was reported.

Although the estimator for cost-savings is the same in all states, each state differs in its analysis of the cost-savings.

2. Pre-Post Intervention for Specific Medical Diagnoses or Interventions. Some states report their results according to specific medical diagnoses. The savings are calculated per targeted member (ie, patient) per time frame, by the type of intervention. The results are estimated by the type of population-based intervention. Some examples include patients receiving treatment for diabetes, hyperlipidemia, drug abuse, or gastrointestinal disorders, or patients receiving antibiotics, polypharmacy, or anticonvulsant therapies.

In these reports, the overall estimate is projected monthly or quarterly, and is then annualized. Later, this result is divided by the number of targeted recipients and by the total months of intervention to get the per-member per-year numbers.

Wyoming uses a variation on this method.\(^6\) The per-patient difference between the preintervention and postintervention drug expenditures was calculated, and this difference was multiplied by the number of patients who were still in the program during the postintervention period to reach the quarterly cost-savings estimates. These numbers were then annualized. Finally, the sum of the annual cost-savings by each type of intervention was obtained to find the annual estimate (Table 1). For the most detailed numeric example, see Wyoming’s FFY 2015 Medicaid DUR report.\(^7\)

3. Pre-Post Intervention Gross Difference. In this simple methodology, the cost-savings are calculated as the difference between the total drug utilization expenditures before and after the intervention in the targeted population. Depending on the time of the pre-post measurement, the annual cost-savings are projected. For instance, cost-savings that are measured in 6 months are then multiplied by 2 to estimate the full-year savings.

Some states estimate the gross difference for patients who received only 1 alert (ie, a single intervention) and those who received more than 1 alert (ie, multiple interventions). In 2014, Wyoming reported cost-savings by medical and pharmacy costs (Table 1). For a detailed numeric example, see the FFY 2014 Medicaid DUR report for Wyoming on the Medicaid website.\(^6\)

4. Pre-Post Intervention for Specific Medications. Some states reported savings by specific medications, with 3 months as the most common intervention period. Ohio identified different intervention groups by pharmaceutical class, and patients were classified into 3 groups, including (1) all patients associated with letters of intervention, (2) patients showing clinical improvement, and (3) unimproved patients; however, no explanation was given regarding what criteria were used to classify patients in the latter 2 groups (Table 1). For a detailed numeric example, see the FFY 2014 Medicaid DUR report for Ohio on Medicaid’s website.\(^6\)

4. Patient- and Problem-Focused DUR. Iowa followed a variant of the patient- and problem-focused DUR method. The direct cost-savings mix patient-focused profile reviews with problem-focused profile reviews; the total amount saved on drug utilization was calculated as the sum of both reviews. First, the patient-focused DUR consists of an intervention regarding an individual patient requiring a change in therapy. Each template or intervention is then evaluated to determine if the proposed change was implemented, and to calculate the economic implications. The calculation involves a comparison of the member’s (or patient’s) initial profile with his or her re-review profile. Each member profile is a 6-month snapshot of the medications covered by the Medicaid program.

There are 9-month intervals between the initial review and the re-review profiles. For each intervention, the total amount paid on the initial profile for any inter-
vention is noted. According to the intervention, the re-review profile is evaluated for a change in therapy. The amount paid on the re-review profile for the same intervention (as on the initial review) is also noted. The 2 profiles are compared by subtracting the total amount paid in the initial profile from the total amount paid in the re-review profile. This calculation is then annualized by multiplying that number by 2, to get the prerebate annualized savings. A numeric example can be found in Iowa’s FFY 2014 and 2015 Medicaid DUR reports.6,7

Problem-focused reviews highlight specific issues that have been determined during the review to be an area where a targeted educational effort by providers may be valuable (eg, duplicate antidepressants, emergency contraception, ketoconazole utilization). Those topics are selected from the findings of patient-focused reviews or from medical literature reviews.

Although patient-focused reviews may address several clinical situations, a problem-focused review addresses only 1 concern at a time. Review criteria are developed by those doing the review to identify the Medicaid beneficiaries who may benefit from an intervention, and educational materials are then mailed to their providers. The drug utilization cost-savings are calculated based on the difference in costs associated with changes in therapy from the original to a new therapy postintervention that are based on 1 year of therapy.

Direct and Indirect Effects

Five states (ie, Kentucky, Michigan, New Hampshire, Tennessee, and Virginia) used a different methodology, which was developed by the Institute for Health Economics of the University of the Sciences of Philadelphia, to quantify the drug utilization cost-savings that are a direct result of the retrospective DUR letter intervention process, as well as the savings that resulted from any indirect results.

The indirect effects are associated with changes in the prescribing process that are applied by the pharmacist in response to an intervention letter, which implies measuring the effect of 1 patient who had an intervention compared with other patients in the same prescriber’s practice who did not have an intervention.

This model also takes into account the impact of prescription drug inflation; the new drugs introduced into the market; and the changes in drug utilization rates, number of recipients, and demographics. This methodology is not explicitly described in any report from the states.

The drug utilization cost-savings are tracked month by month over a 12-month period. Similar to the pre- and postimplementation gross difference methodology, changes in prescription drug costs are totaled, to yield the overall cost-savings for the specific review period. In addition, the calculations are estimated for specific drug classes.

Varied Methodologies Used in 2015

In 2015, 36 states conducted a Medicaid retrospective DUR program evaluation that included the estimated prescription drug cost-savings and cost-avoidance; of these, 30 states reported information on the method used.7 In most states, the estimation was done by mostly private, third-party companies, which improves the accountability of the estimations.7 In total, according to CMS’s consolidated report, states saved an average of $1,122,094.7 However, as shown in this current analysis, at least 6 different methodologies are used by states across the country, thus making comparisons and descriptive statistics unreliable. In addition, some states did not report their methods, or they reported them in general terms only.

Despite being amply accepted that quasi-experimental designs with nonequivalent control groups are superior to quasi-experimental designs without control groups in accounting for time-variant observed and unobserved factors that could bias estimates,8 at least 11 states were still using such a method in FFY 2015 in the retrospective DUR program reports.

The analysis presented here is limited by the amount of information included by the states in their annual DUR reporting to CMS, and by the information available from Medicaid on its website.

Conclusion

There is great variation among states in the methods they use to measure drug utilization cost-savings and cost-avoidance. At the same time, the lack of published details on the methods used makes it difficult to compare savings among states and analyses of best practices. This deficiency is significant, because it hinders states from learning from each other, and, ultimately, the federal government is unable to evaluate and improve the Medicaid DUR program. In addition, the national and international academic communities could benefit greatly if states converge and use a common methodology to calculate prescription drugs savings. There is also a need for more transparency in creating publicly available methods for measuring cost-savings using retrospective DUR programs.

Acknowledgments

The author acknowledges Andrés Aguirre, Johan Loaiza, and Andrea Arenas for their excellent research assistance.

Author Disclosure Statement

Dr Prada reported no conflicts of interest.
The concerns regarding drug utilization review (DUR) raised by Prada could be the same as with dozens of other federal programs. The article’s key messages about these shortcomings come through loud and clear—there are no benchmarks for DUR; no benefit for completing a DUR, and no real penalty for failure to do so; no requirement for a savings level that a state can compare itself with others (either the results or the attempted programs), and no requirement that states disclose their DUR methods.

Legislation or regulation that requires action needs to consider the type of action that is necessary to achieve it, and the appropriate measurement rather than just hope for smart behaviors. The requirement for DURs leaves the decision of what type of review to conduct to the states and assumes that they will make good choices. Prada demonstrates the range of responses provided by various states, which allows me to make a set of observations and conclusions.

Aside from the ambiguous requirements for DURs, and the chaotic variation across the states, the study relays the findings of the Centers for Medicare & Medicaid Services (CMS) of an average savings of just slightly more than $1 million with this program. It’s not clear what the costs of these studies were, but the mention of third-party analytics and outsourcing of the studies, as well as common knowledge of most program costs, suggest that all or substantially all the savings were used to fund those third parties. That the savings accrue to CMS, and that each state is required to conduct the review but receives no real benefit for doing so, is noteworthy, until you realize how many states are paying lip service to this requirement.

The approaches to the required analysis are many, and fewer than 50% of the states used the most vigorous analytic methods, which is a useful analog for many future scenarios regarding the use of burgeoning real-world healthcare data. It is a laudable approach to ask recipients of funding to make sure that they are using the funds appropriately, and even to offer them a share of the savings if they improve, as in the accountable care organization model. With that in mind, drafters of future federal legislation and regulation could learn from this example what it achieves, and what it does not.

The DUR requirement is either red tape that allows states to pay lip service to it, or an enlightened state of flexibility that leaves room for local adaptation. The challenge is that without clear benchmarks and more detailed and harmonized requirements, it is impossible to discern which it is.