Determining Unfairness in Medicaid Reimbursement Rates

Between Idaho and the Nation - Conclusion

By Robert Barnett

# A. Project Highlights

## A1. Research Question

Is there a significant difference in Medicaid reimbursement rates between Idaho and the nation?

## A2. Context & Background

Healthcare organizations share a common view that Medicaid often pays below costs. While Medicaid is overseen by the Centers for Medicare and Medicaid Services (CMS) on a federal level, much of a Medicaid program’s operation is left to the authority of each state, including pricing methodologies. Medicaid programs subscribe to a national database, such as Medispan or First Databank, for a standardized drug information catalog. However, due to each state’s ability to manage their own Medicaid programs, each program’s pricing methodologies can give rise to different reimbursement rates. Some programs attempt to augment the burden by obtaining kickbacks from manufacturers, incentivizing providers to adhere to therapy regimens, and other tactics. Some tactics are executed successfully while others create gaps in coverage and reimbursements. Atypical reimbursement rates would be a useful indicator to justify an investigation into suboptimal tactics employed by a Medicaid program’s pricing methodologies. Consistently low reimbursement rates can adversely affect a healthcare organization’s viability to provide care to Medicaid beneficiaries.

## A3. Scope

The scope of the project included analysis and visualizations utilizing Python in a Jupyter Notebook with two CSV datasets as input. The input was to calculate statistical significance using a Z-score test for the state of Idaho. The datasets were aggregated from sources in 2022 and were used to compare 1st quarter data. This project did not aim to find a statistical significance for other states or other years as that fell outside this scope.

Specifically, this project included an evaluation of the dataset filtered to explore the reimbursement rates of countable drugs within the 1st quarter of 2022 for the state of Idaho and the nation, subset by branded and generic drugs to find a statistical significance. This project did not explore any statistical significance for other states or any other subsets of drugs.

## A4. Overview of the Solution

The project analyzed the drug utilization data of Medicaid reimbursements in the 1st quarter of 2022. The analysis was conducted using Python within a Jupyter Notebook environment using the pandas, NumPy, MatplotLib, and SciPy libraries. A null hypothesis was tested using a Z-score statistical test. The rates of reimbursement were calculated for each NDC as total amount reimbursed per unit dispensed. Histograms of the top 5 brand and top 5 generic drugs, by prescription count, were created to visualize distributions of reimbursement rates across the country. The distributions appeared Gaussian which indicated that the application of a Z-test was appropriate. The rate of reimbursement feature was grouped by state and a Z-score was calculated. A histogram was created to visualize the distribution of Z-scores across the country. Box plots were created for brand and generic subsets to visualize the distribution of Z-scores for each drug, by state. Idaho’s average Z-score was compared against the national average Z-score and assessed against the Z-critical value of 1.96 for an α of 0.05. The null hypothesis was accepted. There were no statistically significant differences in reimbursement rates between Idaho and the nation for branded and generic drugs.

The solution of this analysis can be beneficial in many ways:

* Pharmacies can use this information to identify which drugs are underpaid and apply it to their inventory management and billing practices.
* Other states may be able to utilize the same study to identify unfairness in reimbursement.
* Organizations can use this information to take legal action against their state’s Medicaid program if any laws or rules are broken.
* Organizations may be able to identify opportunities for improving billing best practices by comparing this data against their dispensing data.

The benefits can directly assist in decision-making by creating access to a list of low-revenue drugs and searching for NDCs with better reimbursements.

An example benefit could be that a pharmacy identifies some of their dispensings of a specific NDC that was underpaid yet did not show as underpaid in the analysis. This would indicate that they most likely billed the item incorrectly, creating a systematic training opportunity.

# B. Project Execution

## B1. Goals, Objectives, & Deliverables

The goal of this project was to evaluate if a significant difference exists between Idaho and the nation with Medicaid reimbursements.

The following objectives were set to meet this goal:

* Acquire the datasets.
* Clean the datasets.
* Explore the datasets.
* Compare the Z-scores of reimbursements among the states.
* Finalize a report of the findings.

## B2. Project Planning Methodology

This project used the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. It’s useful for analytics projects and the following phases were utilized:

* Business understanding: Define the objectives, scope, and criteria for success concerning the business needs. In this step, I identified the need to address underpayments.
* Data understanding: Collect, explore, and clean the data while selecting relevant features for the analysis. I found datasets available from a payor with features that would be appropriate for analysis.
* Data preparation: Transform, clean, and integrate the data. This step will be performed within a Jupyter Notebook using Python and various libraries.
* Evaluation: This step will evaluate the analysis for statistical significance with a Z-score test
* Deployment: This step will ensure that the functions can be useful on similar datasets provided by Medicaid to evaluate other years.

## B3. Timeline for Milestones

|  |  |  |  |
| --- | --- | --- | --- |
| Milestone | Duration | Start | End |
| Acquire Data | 1 day | 11/23 | 11/24 |
| Clean Data | 1 day | 11/24 | 11/25 |
| Explore Data | 1 day | 11/25 | 11/26 |
| Analyze Data | 2 days | 11/26 | 11/27 |
| Finalize Report | 1 day | 11/27 | 11/28 |

The timeline for each milestone was met.

## B4. Resources and Associated Costs

1. PC: $1000
2. Python 3: $0
3. Anaconda: $0
4. Datasets: $0
5. 40 work hours: $2000 @ $50 per hour

## B5. Criteria for Success

The criteria for success were:

* A completed Z-score test
* Acceptance or rejection of the null hypothesis
* Data visualizations to represent distributions.

## B6. Variances in Execution

The project was executed as planned without any deviations. The datasets were not difficult to wrangle and were straightforward to work with. The resources used were appropriate for the tasks and the conclusion was easy to reach.

# C. Data Collection Process

## C1. Data Collection Methods

The DrugUtilization2022\_updated.csv file was downloaded from <https://data.medicaid.gov/dataset/200c2cba-e58d-4a95-aa60-14b99736808d/data>. The nadac-national-average-drug-acquisition-cost-2022.csv file was downloaded from <https://data.medicaid.gov/dataset/dfa2ab14-06c2-457a-9e36-5cb6d80f8d93>. This selection and collection process went according to the project plan.

No significant obstacles were encountered while collecting the data. I ensured that all data governance issues were amenable when I chose the data for acquisition, and no data governance issues were encountered during the completion of this project.

An advantage of the main dataset used is that it’s continually updated by Medicaid on an annual basis. This allows reuse of the analysis on other datasets as input which can provide value. A disadvantage of the dataset was that it lacks granularity of drug information. However, since it contains a column for NDC, other datasets can easily be used to supplement the missing information as needed, such as the NADAC dataset that was used for that purpose.

# D. Data Extraction and Preparation

## D1. The Data Extraction and Preparation Process

The data was downloaded from data.medicaid.gov as a CSV file. It was then loaded into pandas dataframes and evaluated for datatypes. Additionally, since the NDC was to be used as a key to join the datasets, it was evaluated for proper format in both datasets prior to merging. The NDC columns were converted to a numeric datatype when read into the dataframes which required casting them back to string datatypes and then filling in the leading zeros that were dropped, back to 11 digits. A few columns were renamed for ease of identification. The reimbursement rate column was created as a calculation between the existing total reimbursement and units reimbursed columns.

Duplicate rows for the NDC column were dropped in the NADAC dataset so that it could be used as a primary key for merging. The duplicate rows held no information that was important to this project.

Unnecessary columns were dropped as well as any rows with NaN values. These rows likely resulted with missing data that was not submitted in time from various states or for NDCs that were simply not covered in those states. By dropping the rows with NaN values, a fair comparison among the common data elements was ensured.

The datasets were merged by joining the drug utilization and NADAC data as a left join on the NDC column as a key using a many to one relationship.

After the merge, the resulting dataframe was filtered to omit over-the-counter medications, drugs with ML or GM units of measurement, and other quarters outside of the scope of this project. Over-the-counter medications were eliminated because many Medicaid programs do not cover them, so the data would be unusually skewed. ML and GM units of measure were omitted since different payors in different regions have various policies regarding paying for partial and full packages for partial dispensings, which would unusually skew the data. Additionally, some software vendors submit the dispensed quantities with or without decimals which can affect the actual reimbursement rates when compared between states. The majority of dispensings are for the EA type of unit of measure (such as tablets, capsules, etc.). Eliminating the EA and ML units of measure was appropriate to ensure appropriateness of comparisons. If EA and ML were measured, they would need a separate subset and that comparison falls outside the scope of this project.

A column was then created for the Z-scores using SciPy’s zscore function on the reimbursement rate, grouped by NDC. This tool allowed efficient use of resources to calculate a Z-score grouped on a column rather than requiring further manipulation of the dataframe for unnecessary calculated columns.

The data was then ready for use within the scope of this project. It was split into a subset for Generic drugs and another for Brand drugs. This is because they are reimbursed on different a different cost-basis. Histograms were created using MatplotLib to view the distributions and assess appropriateness for applying a Z-score test. Those subsets were split once more by Idaho and the Nation to apply the Z-score statistical test. Box plots were created using MatplotLib to visually assess the distributions.

# E. Data Analysis Process

## E1. Data Analysis Methods

A Z-score test was used to evaluate the Z-score of Idaho’s reimbursement rates against the population’s (Nation’s) reimbursement rates. An α of 0.05 was used with a Z-critical of 1.96. The null hypothesis was that there is no significant difference in reimbursement rates between Idaho and the nation. This test fits since the analysis compares Idaho’s sample against a population mean.

The pandas library was used to shape the data in dataframes. The MatplotLib library was used to create histogram and box plot visualizations. The SciPy library was used to calculate Z-scores for the dataframe.

The prepared data was preliminarily viewed by evaluating a sample of NDCs. These samples were 5 generic and 5 brand drugs of the highest prescribed count. The dataframes were sorted by highest prescribed count to obtain a list. A histogram was plotted for each list to observe the shapes of distributions. One pair of histograms viewed all 5 drugs each in a plot. Then a series of 5 histograms were viewed per subset to view the distributions shapes for each drug. The distribution appeared Gaussian.

A box plot of the Z-scores was created with the generic subset, grouped by state to observe the distributions nationwide. Another box plot was created for the brand subset.

A final pair of box plots was created for brand and generic to specifically view Idaho vs The Nation Z-score distributions.

The mean Z-score was calculated for brand and generic reimbursement rates for Idaho and the nation. Idaho’s Z-scores were evaluated to be below the Z-critical value and the null hypothesis was accepted for both brand and generic tests. Idaho’s generic Z-score was 0.433 and the brand Z-score was -0.003.

## E2. Advantages and Limitations

Advantages and limitations of tools used:

|  |  |  |
| --- | --- | --- |
| Tool | Advantage | Limitation |
| Python | It is a very capable programming language for data analysis. | The lack of a GUI requires a strong understanding of programming concepts. |
| Jupyter Notebook | It can execute blocks of code within cells. | It can only run one programming language as a kernel at a time. |
| Pandas | It can handle large amounts of data. | The lack of a GUI requires a strong understanding of programming concepts. |

Jupyter Notebook was an excellent tool for exploring this data in an organized manner. Pandas facilitated the use of dataframes to work with the large datasets in a quick manner. NumPy allowed the ability to set an array for a histogram plot function. MatplotLib was useful for creating appropriate visuals using the data.

## E3. Application of Analytical Methods

The analytical method was descriptive. The analysis observed a past dataset to describe what happened to reimbursement rates in Idaho compared to the nation. Descriptive analysis is used to transform past data into information that can be acted upon.

The Z-score test was implemented by evaluating the Z-scores for reimbursement rates. A Z-score test requires a population which a mean can be known and a population greater than 30 samples. The dataset contained 50 states to group by for each drug.

The Z-scores were created by using the SciPy library. Then the Z-scores were evaluated against a Z-critical value of 1.96 based on an α of 0.05.

At the end of this process, the descriptive analytical method explored the data of reimbursement rates in 2022 to provide insights that can be acted upon by pharmacies and other organizations.

# F. Data Analysis Results

## F1. Statistical Significance

The statistical test used was a Z-score test. Two tests were conducted.

For an α of 0.05, the Z-critical was 1.96.   
Null Hypothesis: There is no significant difference in reimbursement rates for **generic drugs** between Idaho and the nation.  
The Z-score for Idaho’s generic drug reimbursement rates was 0.433.   
There was sufficient evidence to accept the null hypothesis that there is no significant difference in reimbursement rates between Idaho and the nation for generic drugs.

Null Hypothesis: There is no significant difference in reimbursement rates for **brand drugs** between Idaho and the nation.  
The Z-score for Idaho’s brand drug reimbursement rates was -0.003.  
There was sufficient evidence to accept the null hypothesis that there is no significant difference in reimbursement rates between Idaho and the nation for brand drugs.

## F2. Practical Significance

This analysis provided insights into the distribution of reimbursement rates across the nation. In the box plot titled "Z-Scores for all Brand Medications in Each State", Kansas appears to have alarmingly low reimbursement rates. Other states may be able to benefit from similar analyses even though Idaho may not have directly benefited from it. Evidence exists to pursue this study for other states.

The practical significance of this project is that pharmacies will be able to identify items that they may be underpaid for. If there is a systematic underpayment, then it can lead to legislative action. On a granular level, a pharmacy can compare their Medicaid reimbursement data against this project’s data to identify items that do not match and investigate the cause which may be a billing error by the pharmacy. Such billing errors could be a systemic training issue in their organization which may affect their reimbursements from other payors. Larger organizations, like state boards of pharmacy or a pharmacy services administrative organization (PSAO), may look at region-wide data to find legally actionable issues, such as the case in a Washington lawsuit.

Furthermore, the analysis could be useful for internal evaluation of a pharmacy's reimbursements compared to their state's average reimbursements. This could help a pharmacy root out procedural opportunities to optimize their billing practices to match their colleagues' best practices.

## F3. Overall Success

The criteria for success were projected to be:

* **A completed Z-score test**. The Z-score will be calculated and compared to a Z-critical score of 1.96 based on an α of 0.05. If the absolute value of the Z-score is less than the absolute value of the Z-critical value, then the null hypothesis will be accepted. Otherwise, the alternate hypothesis is accepted.
* **Acceptance or rejection of the null hypothesis** based on the calculated Z-score. The acceptance of either the null hypothesis or alternate hypothesis shall be stated clear and concise.
* **Data visualizations to represent distributions**. Successful visualizations will display the data with easily observable inferences.

This project was successful. Z-scores were and compared against the Z-critical score, resulting in an acceptance of the null hypothesis. Histograms and box plots were created to visualize the distributions of the data as appropriate.

This project provided additional insights into possibly other studies to be evaluated in a similar way for other states, such as Kansas. As it is, this project can be utilized by various organizations to improve their operations. Pharmacies can identify drugs to mitigate low reimbursements and legislative actors can identify opportunities for improvement in various state Medicaid programs’ pricing methodologies.

# G. Conclusions

A statistical test was used to explore if unfairness existed in Idaho’s Medicaid reimbursement rates for generic and brand medications. The project resulted in accepting that there is no significant difference in reimbursement rates between Idaho and the nation for generic and brand drugs. However, some insights were gained during this exploration. While Idaho may not be a victim to low reimbursement rates, it became evident that other states were. This project can provide an opportunity for other states to explore that information. Furthermore, this project can provide actionable information for pharmacies, regulatory bodies, and other organizations to improve their outcomes.

The project created insightful visualizations. The histograms provided a look at the distributions of reimbursement rates nationwide on both a granular and broad level. While it’s reasonable to expect that the rates should be a single point in an ideal world, it’s not surprising that a normal distribution could be observed. These distributions helped validate the appropriateness of the statistical test performed. Box plots were also created to demonstrate the extent to which the variation in reimbursements is spread. They were able to show how each state fared in comparison to each other. It made the similarity of Idaho’s data to the national data easier to conceptualize.

A further course of action could be pursued by Kansas. Kansas showed an alarmingly low rate of reimbursement for both brand and generic drugs. They may be able to pursue similar courses of action that Washington’s pharmacies are undergoing in their lawsuit against their Medicaid administration (National Association of Chain Drug Stores v. Xavier Becerra, 2021).

Another course of action could be pursued by legislators. While Idaho may not have shown evidence of unfair reimbursements overall, the visualizations in this project have made it apparent that unfairness may exist in other states. If legislators do not act to rectify the underpayments, then other groups can pressure legislators into creating fair policies that empower healthcare providers to serve Medicaid beneficiaries better.

# H. Panopto Link

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=3a8c3c93-b5bf-4d21-b9df-b0c800c2d17d>

# I. Appendices

Medicaid\_Reimbursement\_2022\_Study.ipynb  
DrugUtilization2022\_updated.csv  
nadac-national-average-drug-acquisition-cost-2022.csv

## Sources

Ford, T., & Michener, J. (2022, June 16). Medicaid Reimbursement Rates Are a Racial Justice Issue. *The Commonwealth Fund*. Retrieved from [https://www.commonwealthfund.org/blog/2022/medicaid-reimbursement-rates-are-racial-justice-issue](https://www.commonwealthfund.org/blog/2022/medicaid-reimbursement-rates-are-racial-justice-issue%20)

“Pharmacy Groups Lobby for Revisions to Proposed Medicaid Rule” (2009, June 11). *Kaiser Health News*. Retrieved from <https://kffhealthnews.org/morning-breakout/dr00045610/>

National Association of Chain Drug Stores v. Xavier Becerra, 2:21-cv-00576-RSM (W.D. Wash. filed April 29, 2021)