

More accurately assessing opioid risk in Medicare members

Joseph Boschert, MBA, Product Manager
 Jim Hong, MBA, MS, Healthcare Analytics Consultant
 Kyle Ligon, Data Scientist



Both the Centers for Medicare and Medicaid Services (CMS) Minimum criteria and Supplemental criteria were outperformed by the data science model.

According to the Centers for Disease Control and Prevention (CDC), it is estimated that more than 115 people die each day in the United States as a result of opioid overdose and that prescription opioid misuse costs more than \$78.5 billion per year.¹

The latest CDC data shows overall emergency department (ED) visits for suspected opioid overdoses increased 30% from July 2016 to September 2017. As a result, the CDC called for enhanced prevention efforts.²

The uptick in opioid overdose-related deaths and misuse has developed since the late 1990s for a variety of reasons, including:

- An increased number of prescription opioids given to patients for pain management combined with increased quantities
- Increased influence from pharmaceutical companies, including an emphasis on pain as the fifth vital sign and extending marketing from pain specialists to primary care and ED doctors³
- Lack of coordination between physicians and pharmacies
- Lack of insight into patient opioid consumption⁴
- Lack of education regarding alternative treatment modalities for those with non-cancer chronic pain⁵
- A transition to illicit drugs by those who first develop an opioid use disorder (OUD) diagnosis on prescription drugs⁶

Research has shown that the prevalence of opioids in the Medicare population for a given year ranges between 26% and 52% depending on the cohort.⁷ Those who receive an opioid prescription are at higher risk for adverse drug events and opioid-related complications and injury.^{8,9,10}

The opioid epidemic is a complex public health crisis with no simple solution available. All stakeholders need to proactively work to improve the situation. Some of the current efforts by stakeholders include the following:

- Physicians and pharmacists utilizing state electronic Prescription Drug Monitoring Program (ePDMP) systems prior to prescribing and dispensing opioid medications
- State legislators restricting the days' supply for an initial prescription by enacting legislation
- Health plans improving provider education and risk assessment
- Substance Abuse and Mental Health Services Administration (SAMHSA) increasing access to medication-assisted treatment (MAT)
- Healthcare providers improving efforts to integrate information sharing

¹ CDC/NCH (2017). National Vital Statistics System, Mortality. CDC Wonder. Retrieved January 1, 2020, from <https://wonder.cdc.gov>.

² CDC (March 6, 2018). Emergency department data show rapid increases in opioid overdoses. Press release. Retrieved January 1, 2020, from <https://www.cdc.gov/media/releases/2018/p0306-vs-opioids-overdoses.html>.

³ Tompkins D, Andrew J, Hobelmann G, Compton P. (2017). Providing chronic pain management in the "fifth vital sign" era: Historical and treatment perspectives on a modern-day medical dilemma. *Drug & Alcohol Dependence*;173: S11-S21.

⁴ Morone NE, Weiner DK. (2013). Pain as the fifth vital sign: Exposing the vital need for pain education. *Clinical Therapeutics*;35(11):1728-1732.

⁵ Clarke EB, Ellen B., et al. (1996). Pain management knowledge, attitudes and clinical practice: The impact of nurses' characteristics and education. *Journal of Pain and Symptom Management*;1(11): 18-31.

⁶ Compton, Wilson M., Jones, Christopher M., & Baldwin, Grant T. Relationship between nonmedical prescription-opioid use and heroin use. *New England Journal of Medicine* 374.2 (2016): 154-163.

⁷ Jeffery, Molly Moore, et al. (2018). Trends in opioid use in commercially insured and Medicare Advantage populations in 2007-16: Retrospective cohort study. *BMJ* 362: k2833.

⁸ Groeger, J. L., et al. Incidence of long-term opioid use after hospitalization in older adults." *Journal of the American Geriatrics Society* Vol. 66.

⁹ Daoust, Raoul, et al. (2018). Recent opioid use and fall-related injury among older patients with trauma. *CMAJ* 190.16: E500-E506.

¹⁰ Coughlin, Sarah R. (2017). Older Adults' Response to Adverse Opioid Related Events: A Secondary Analysis.

Financial implications

For individuals, opioid dependence can be expensive, harmful to the body, and impactful to the personal relationships of the individual and family. This paper analyzes Medicare Advantage claims and membership data to evaluate the risk and costs associated with OUD. While there are multiple methods to identify OUD, we used a set of common ICD-10 codes (F111, F112, and F119) to identify cases. Medicare Advantage claims and membership data from seven health insurance companies operating throughout the United States were used to analyze risks and costs. As a result of this analysis, we estimate that, for members diagnosed with OUD, their allowed amount costs are on average \$10,000 higher in the 12 months during and after an OUD diagnosis than their costs in the 12 months prior to an OUD diagnosis.

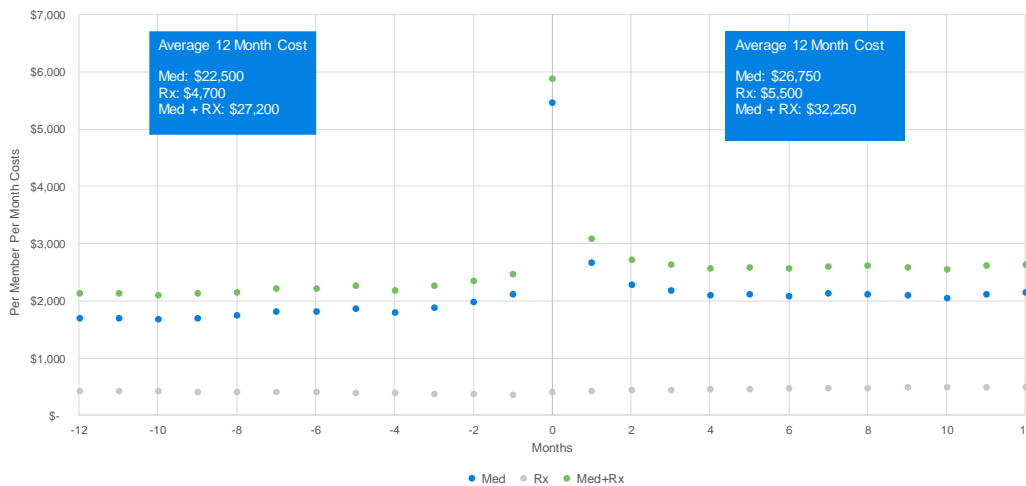
While prevention can help from a cost perspective, so can ensuring the correct accounting of OUD. In some scenarios, individuals may exhibit symptoms of an opioid-related disorder (and may even be treated using a MAT such as buprenorphine), but may not have proper diagnosis codes applied to their medical records. This incomplete accounting has an impact on the revenue of Medicare Advantage health plans, as their reimbursement is tied to properly adjusted risk scores, which require properly coded conditions. In our analysis, a missing Hierarchical Condition Category (HCC) CMS-HCC055 Drug/Alcohol Dependence can result in \$3,700 less annual revenue (assuming a benchmark rate of \$850, an adjusted member bid rate of \$800, and an aged non-dual status).

CMS response

CMS is a major stakeholder in addressing the opioid epidemic. The 2019 Medicare Final Rule enacts new guidelines to reduce the number of beneficiaries who may potentially misuse or overdose on opioids. The 2019 Medicare Program Final Rule made significant changes to the Medicare Advantage (Part C) and Prescription Drug Benefit (Part D) programs to “further reduce the number of beneficiaries who may potentially misuse or overdose on opioids while still having access to important treatment options.”¹¹

The Final Rule aligns with statutory provisions of the Comprehensive Addiction and Recovery Act (CARA) of 2016 and provides clarification on several subjects. It defines at-risk beneficiaries, includes case management and lock-in programs, and specifies requirements in the Part D Opioid Drug Utilization Review (DUR) Policy and Overutilization Monitoring System (OMS). Specifically, new Part D requirements require plans to limit first-time opioid prescriptions for acute pain to seven days and to restrict at-risk beneficiaries to specific prescribers and pharmacies. Essentially, the Final Rule has defined the criteria to identify beneficiaries at risk of opioid addiction.

FIGURE 1: HEALTHCARE COSTS OF INDIVIDUALS WITH AN OUD DIAGNOSIS



¹¹ CMS/HHS (2019). Medicare Program; Contract Year 2019 Policy and Technical Changes to the Medicare Advantage, Medicare Cost Plan, Medicare Fee-for-Service, the Medicare Prescription Drug Benefit Programs, and the PACE Program. Final Rule. Published in the Federal Register 83.73 (2018): 16440.

Risk definition

The Final Rule states that “...it is the Part D sponsor that determines which beneficiaries are at-risk beneficiaries under its drug management program.” CMS created two categories and criteria to identify the requirements,¹¹ which are shown in Figure 2.

FIGURE 2: MINIMUM AND SUPPLEMENTAL CRITERIA

Minimum Criteria	Supplemental Criteria
Use of opioids with an average daily Milligrams Morphine Equivalent (MME) ¹² greater than or equal to 90 mg for any duration during the most recent six months And either: <ul style="list-style-type: none"> ▪ Three or more opioid prescribers and three or more opioid-dispensing pharmacies or <ul style="list-style-type: none"> ▪ Five or more opioid prescribers, regardless of the number of opioid-dispensing pharmacies. 	Use of opioids (regardless of average daily MME) during the most recent six months And either: <ul style="list-style-type: none"> ▪ Seven or more opioid prescribers or <ul style="list-style-type: none"> ▪ Seven or more opioid-dispensing pharmacies

Notes

1. Multiple prescribers associated with the same single Tax Identification Number (TIN) are counted as a single prescriber
2. Pharmacies with multiple locations that share real-time data are counted as one pharmacy
3. Long-term care (LTC) beneficiaries are exempt
4. Benzodiazepines are a frequently abused drug for purposes of Part D drug management programs but are not a factor in these clinical guidelines
5. Buprenorphine is not used to determine the beneficiary's average daily MME

Case management

The new rule also requires clinical staff at the Part D sponsor to engage in case management for each potential at-risk beneficiary. The goal is to clinically engage the prescribers of frequently abused drugs and to assess and verify intent, and whether a potential at-risk beneficiary is indeed at risk.

While many of these efforts will likely have a positive impact on the opioid crisis in the long run, there is also an opportunity to improve and advance the area of identifying those who are at risk of an opioid-related disorder. The Government Accountability Office (GAO) reported on inappropriate prescribing activities and the associated risks, and called for more comprehensive data on Medicare patients and providers to sufficiently monitor and prevent opioid overprescribing.¹³ Additionally, because opioid-related disorders are often at the intersection of a complex web of physical

and mental health conditions, relying on prescription data alone may not provide a full picture.¹⁴ As a result, traditional approaches, such as relying on a Pharmacy Benefits Manager (PBM) and using rules-based methodologies (high MME, prescriber counts, and pharmacy counts), could understate or overstate the individual future risk of an opioid-related disorder.

In a clinical setting, a variety of screening tools are available for clinicians to screen patients for risk of addiction, including, but not limited to, the Opioid Risk Tool (ORT), Screener and Opioid Assessment for Patients with Pain, Revised (SOAPP-R), and Current Opioid Misuse Measure (COMM). Many screening tools today are a set of questions clinicians ask patients, relying on self-reported data. The CDC has called into question the accuracy of these tools and their effectiveness in reducing harm because the evidence and results were inconsistent.¹⁵

Screening and assessing large numbers of either opioid-naïve or high-risk chronic opioid users is made easier with an approach that scales. Data science, specifically predictive modeling and machine learning, may provide one avenue to assist with the opioid epidemic. Machine learning can detect patterns in historical data to identify common criteria that could lead to OUD diagnoses. These patterns may not be easily identifiable through manual reviews of the data.

Predictive analytics: Methodology

To test the applicability of such methods, Milliman developed a model using data science techniques to predict an individual's likelihood of receiving an OUD diagnosis in the next 180 days. This model was developed by training an ensemble tree model on claims and demographic data from an internal Milliman research data set.

We compiled a data set comprised of Medicare Advantage members from seven national health insurance carriers. This data included over 1.2 million distinct individuals' demographic (age, sex, gender), medical (diagnoses and procedures), and pharmacy (drug type, units, days' supply, quantity dispensed) records from June 2016 through November 2018, amounting to over 31 million medical claims records. To address the time lag of claims data, we created features from a two-year window prior to a six-month observation window for events (occurrence of OUD diagnosis). Because most claims data is complete after three months, identifying likelihood within the next 180 days still allows time for intervention.

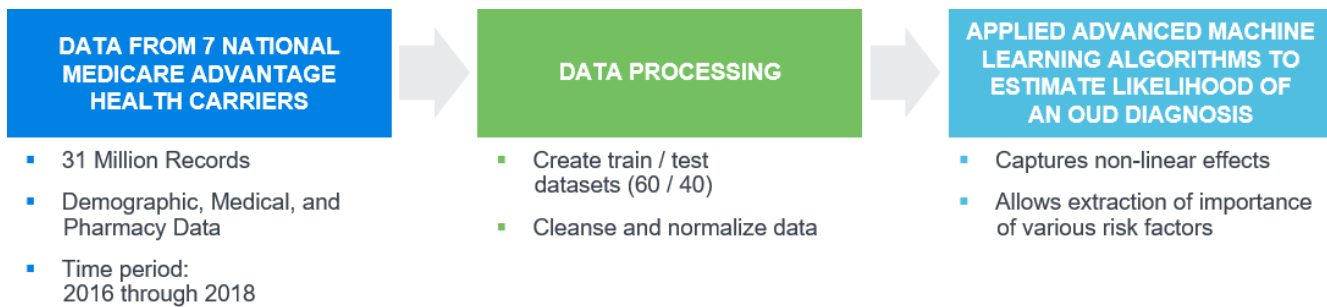
¹² CDC. Calculating Total Daily Dose of Opioids for Safer Dosage. Retrieved January 1, 2020, from https://www.cdc.gov/drugoverdose/pdf/calculating_total_daily_dose-a.pdf.

¹³ GAO (May 29, 2018). Prescription Opioids: Medicare Needs Better Information to Reduce the Risk of Harm to Beneficiaries. Testimony Before the Subcommittee on Health Care, Committee on Finance, U.S. Senate. Retrieved January 1, 2020, from <https://www.gao.gov/assets/700/692085.pdf>.

¹⁴ Rice, J. Bradford, et al. (2012). A model to identify patients at risk for prescription opioid abuse, dependence, and misuse. *Pain Medicine* 13.9: 1162-1173.

¹⁵ Dowell, Deborah, Haegerich, Tamara M. & Chou, Roger (2016). CDC guideline for prescribing opioids for chronic pain—United States, 2016. *JAMA* 315.15: 1624-1645.

FIGURE 3: METHODOLOGY



Finally, to ensure robustness of our model on unseen data and to account for potential overfitting, we split the data: 60% was used for training, while 40% was held out for validation. Typically, the balance of training to holdout may be closer to 80% and 20%, respectively. Given the starting size of our data, we felt there would be sufficient samples in both splits and opted for a smaller training split for efficiency.

RESULTS AND INTERPRETATION

After our initial development and refinement, our model produced promising results, with an area under the curve (AUC) metric of 0.903. AUC is a calculation that is commonly used to demonstrate the accuracy of binary classifiers. Values of AUC closer to 1 indicate that the classifier is more likely to correctly classify whether a person will or will not have the target outcome (in our case who will or will not have an OUD diagnosis). Relatively speaking, a value of 0.903 indicates a strong model, as a value of 1 indicates a perfect model and a value of 0.5 indicates a model that is no better than random chance (e.g., a coin flip). It is important to keep in mind that results are limited to the data sets on which the model has been trained and validated.

Figure 4 displays the top 15 features that are contributing factors to predicting the likelihood of OUD diagnosis, for the Medicare population tested, along with their relative importance.

Figure 4 shows that the features with the strongest influence in determining whether someone has higher or lower likelihood from this model are consistent (from a clinical and public health perspective) with those found in medical literature. For example,

substance-related disorders,¹⁶ age,¹⁷ and geography¹⁸ have all been associated with opioid addiction and have also been identified as having strong influence on the likelihood of an OUD diagnosis. Additionally, polypharmacy features (multiple prescribers and multiple dispensing pharmacies) were important factors, but not the only features to show strong influence. This consistency is encouraging as it suggests that our model bears resemblance to existing determinations of opioid addiction. Yet, beyond simply verifying the associations found in publications, our model produces actual predictions incorporating these demographic variables. Furthermore, it leverages medical and pharmacy claims experience to enhance the determination of the likelihood prediction.

Beyond the overall model features and their consistency with medical literature, our algorithm also allows further examination of contributing features for each individual's risk score. Figure 25 below displays a sample individual's specific contribution factors, with corresponding Shapley Additive exPlanations (SHAP) values, which is an approach to explain the output of any machine learning model. The ranking of the predictors in Figure 4 indicates how key the feature can be in determining an individual's risk of an OUD diagnosis, but does not represent a linear relationship or correlation. For example, the age of an individual is a strong factor in determining the risk of an OUD diagnosis in the next six months, but a person is not necessarily more likely to receive an OUD diagnosis as they get older.

¹⁶ Fleming, Michael F., et al. (2007). Substance use disorders in a primary care sample receiving daily opioid therapy. *The Journal of Pain* 8.7: 573-582.

¹⁷ Edlund, Mark J., et al. (2010). Risks for opioid abuse and dependence among recipients of chronic opioid therapy: Results from the TROUP study. *Drug and Alcohol Dependence* 112.1-2: 90-98.

¹⁸ Keyes, Katherine M., et al. (2014). Understanding the rural-urban differences in nonmedical prescription opioid use and abuse in the United States. *American Journal of Public Health* 104.2: e52-e59.

FIGURE 4: RELATIVE AGGREGATE FEATURE IMPORTANCE USING DEMOGRAPHIC, PHARMACY, AND CLINICAL CLASSIFICATION SOFTWARE (CCS)¹⁹ CATEGORIES

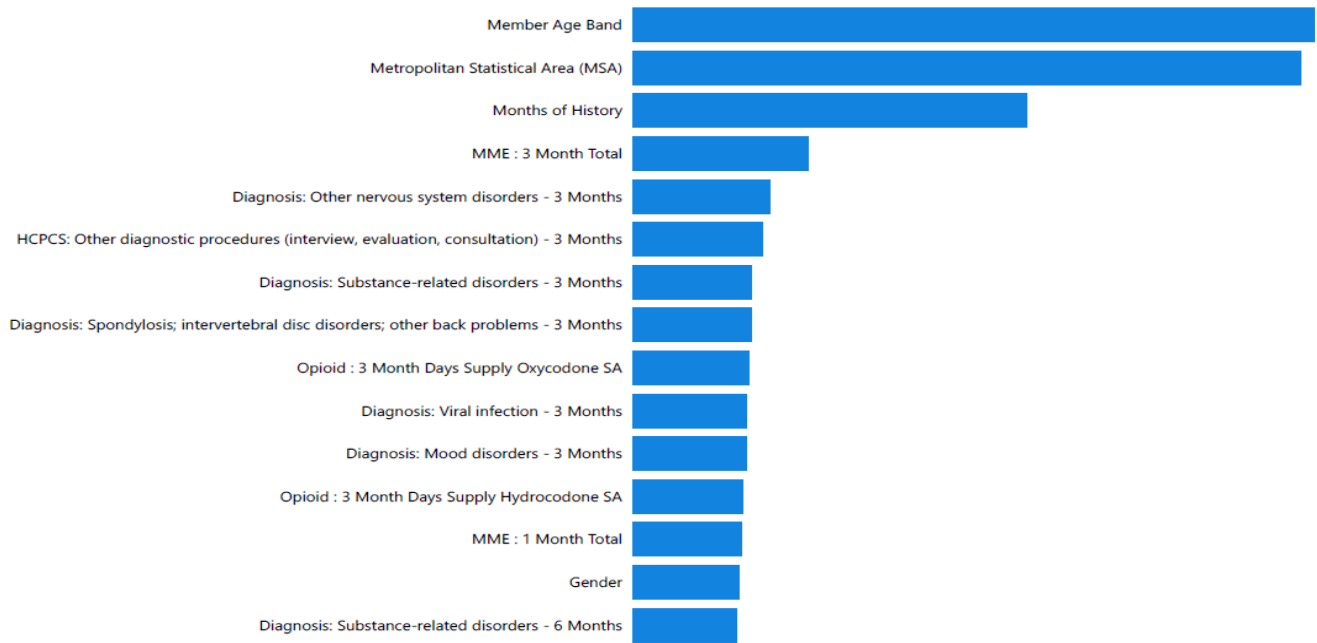
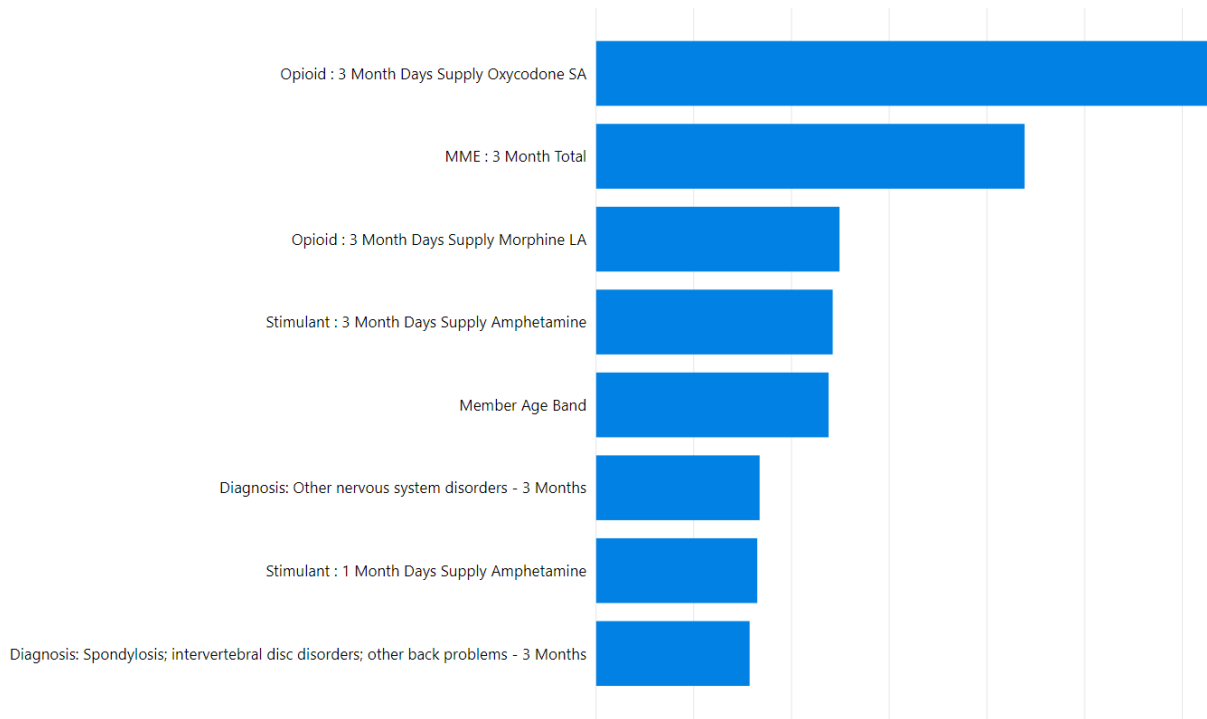


FIGURE 5: SAMPLE SHAP VALUES FOR A RANDOM MEMBER WHO BECAME DIAGNOSED WITH OUD



¹⁹ Healthcare Cost and Utilization Project (March 2017). Clinical Classifications Software (CCS) for ICD-9-CM. Agency for Healthcare Research and Quality. Retrieved January 2, 2020, from <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>.

Comparison of methods of identifying addiction risk

To assess the usefulness of the data science approach for determining OUD diagnosis risk, we compared the performance of our data science method, termed Foresight, against two existing traditional methods (the Minimum and the Supplemental qualifications). Precision is the rate of correct positive predictions; that is, what percentage of the cases predicted as likely to become diagnosed with OUD actually get diagnosed with OUD. True positive rate is the proportion of positive cases identified by the model; in other words, of all cases that are actually diagnosed with OUD, what percentage was identified by the model.

Figure 6 displays the performance results of the three classification methods on the holdout set of members. This test set was 40% of our starting data, containing over 480,000 distinct members and 31 million claim records. Applying the Minimum criteria on the holdout sample of these members, none fit the criteria to be considered at risk for opioid addiction. The Supplemental method was able to identify a few positive cases correctly but still missed a significant number. Both methods were outperformed by the data science approach on both the true positive rate and precision. When predicting that a member will have an OUD diagnosis, the data science approach is correct three times more often than the Supplemental method.

FIGURE 6: QUANTILES FOR THE THREE METHODS

High MME			High Prescriber			Foresight		
Quartile	Count	OUD Cases	Quartile	Count	OUD Cases	Quartile	Count	OUD Cases
1	122,015	1,007	1	122,015	821	1	122,015	1,366
2	122,015	113	2	122,015	113	2	122,015	182
3	122,015	16	3	122,015	16	3	122,015	44
4	122,015	647	4	122,015	833	4	122,015	191

The true positive rate for all methods is pretty tepid as far as percentages go. Even the best performer, our model Foresight, captures only 10% of all addiction diagnosis cases, which means 90% of the cases are unaccounted for. While this is certainly low, our primary intent is to show how predictive modeling and machine learning can outperform existing methods. In this regard, the predictive model beats the Minimum and Supplemental by multiple folds—10% is not very high, but it is many times higher than the existing methods. An actual implementation of the model in the field would be accompanied by evaluations of a cost-benefit analysis to balance the spend on outreach and the costs avoided by intervention.

Extending the concept of looking at individual member performance, we examined the ability of the three methods to segment the whole population based on risk of OUD diagnosis. In Figure 7, we sort the entire validation population three times, once per method, and display the population in ranked quartiles based on the sorting method. We also show the number of members who developed an OUD diagnosis in each quartile of each method. The predictive model has a much higher number of OUD cases in its first quartile, 1, and a much lower number of OUD cases in its last quartile, 4, compared with the first and last quartiles of both the Minimum and Supplemental methods. One interesting thing to note is that all three methods have a decreasing trend in the number of OUD cases when moving from the first quartile to the third quartile, until the fourth quartile where the trend stops. This suggests that there's a component of OUD risk that's still not fully captured by MME, number of prescribers, or our machine learning model.

FIGURE 7: PERFORMANCE METRICS FOR THE THREE METHODS

Metric	Method	Value
True Positive Rate	Minimum	0.0%
	Supplemental	0.2%
	Foresight	10.3%
Precision	Minimum	0.0%
	Supplemental	7.9%
	Foresight	26.0%

All three methods have a decreasing trend in the number of OUD cases when moving from the first quartile to the third quartile, until the fourth quartile where the trend stops.

Conclusion

Properly understanding risk and context are important to assessing whether a Medicare member should be prescribed opioids or continue opioid therapy. There are cases where opioids may be appropriate for specific acute events and chronic pain situations. Pain management professionals, primary care physicians, surgeons, and dentists are best positioned to make these clinical judgments. Yet actively identifying those at greater risk for adverse addiction outcomes may be inefficient and tedious. While existing recommendations on identifying addiction risk exist, more modern and scalable approaches can greatly enhance the precision of this task and identify a greater proportion of patients that may eventually become addicted. Our research has shown that leveraging both machine learning and historical claims data can result in better performance in identifying potential addiction cases. While this approach does not perfectly identify all cases of potential addiction, it does so at rates higher than existing methodologies and can be deployed on a broader scale. As additional data sources become available, the performance of machine learning models will continue to improve. The fact that this method also identifies risk factors that are consistent with medical literature further corroborates its potential clinical relevancy.

Screening and risk assessment are critical to exercising sound clinical judgment and making effective care decisions. OUD diagnoses have continued to increase over the past several years. A multifaceted approach to attacking the problem, including widespread next generation opioid assessment and screening, could play a larger role in reducing societal harm in the future.



Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

[milliman.com](https://www.milliman.com)

CONTACT

Joseph Boschert
joseph.boschert@milliman.com

Jim Hong
jim.hong@milliman.com

Kyle Ligon
kyle.ligon@milliman.com