Using prescription claims to detect aberrant behaviors with opioids: Comparison and validation of five algorithms

Kathryn Rough[a,b,*], **Krista F. Huybrechts**[a], **Sonia Hernandez-Diaz**[b], **Rishi J. Desai**[a], **Elisabetta Patorno**[a], and **Brian T. Bateman**[b,c]

[a]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts

[b]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

[c]Department of Anesthesiology, Perioperative, and Pain Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

Abstract

Objective—Compare and validate five algorithms to detect aberrant behavior with opioids: Opioid Misuse Score, Controlled Substance-Patterns of Utilization Requiring Evaluation (CS-PURE), Overutilization Monitoring System, Katz, and Cepeda algorithms

Study Design and Setting—We identified new prescription opioid users from two insurance databases: Medicaid (2000-2006) and Clinformatics Data Mart (CDM; 2004-2013). Patients were followed one year, and aberrant opioid behavior was defined according to each algorithm, using Cohen's kappa to assess agreement. Risk differences (RD) were calculated comparing risk of opioid-related adverse events for identified aberrant and non-aberrant users.

Results—3.8 million Medicaid and 4.3 million CDM patients initiated prescription opioid use. Algorithms flagged potential aberrant behavior in 0.02-12.8% of initiators in Medicaid and 0.01-7.9% of initiators in CDM. Cohen's kappa values were poor to moderate (0.00 to 0.50 in Medicaid; 0.00 to 0.30 in CDM). Algorithms varied substantially in their ability to predict opioid-related adverse events; the Overutilization Monitoring System had the highest RD between aberrant and non-aberrant users (14.0% in Medicaid; 13.4% in CDM) and the Katz algorithm had the lowest (0.96% in Medicaid; 0.47% in CDM).

Conclusions—In two large databases, algorithms applied to prescription data had varying accuracy in identifying increased risk of adverse opioid-related events.

Keywords

prescription opioid; misuse; abuse; opioid; validation; claims data

---

[*]Corresponding Author Information: Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Suite 2020, Boston, Massachusetts 02120, Phone: 617-525-9388; ker704@mail.harvard.edu.

INTRODUCTION

Over the past two decades, the misuse and abuse of prescription opioids has become widely recognized as a significant public health crisis in the United States.[1–3] From 1999 to 2009, sales of prescription opioids quadrupled and prescription opioid-related admissions to substance abuse treatment programs increased 6-fold.[4] Since 2000, the rate of fatal overdoses involving opioids tripled, reaching a total of approximately 28,000 in 2014.[5] In 2015, an estimated 12.5 million Americans abused prescription opioids,[6] which costs the United States over $50 billion annually.[7,8]

In response to the prescription opioid epidemic, the US Department of Health and Human Services issued a 2014 directive calling on federal agencies to "expand monitoring of administrative prescription data to identify high-risk prescribing practices and eliminate fraud, waste, and abuse related to opioids."[2] Administrative prescription data can come from several sources, including health insurance claims data and the prescription drug monitoring programs administered by many states to record information on medications scheduled by the United States Drug Enforcement Agency. Health insurance claims databases can have tremendous value in detecting patterns of aberrant behaviors with prescription opioids because they comprehensively capture longitudinal information on the dispensing of prescription medications in addition to diagnoses, procedures, and other healthcare services.

Multiple algorithms have been proposed to detect aberrant behaviors with prescription opioids using insurance claims data. As part of their Overutilization Monitoring System,[9] the Centers for Medicare and Medicaid Services (CMS) created an algorithm to identify 'overuse' of prescription opioids. The algorithm was implemented in Medicare Part D in 2014 to identify beneficiaries who may require individualized safety measures, such as restrictions on prescribers or pharmacies for scheduled drugs. Several academic research groups have also proposed algorithms to detect aberrant behaviors with prescription opioids. [10–13] However, the ability of existing algorithms to successfully identify individuals at risk for adverse events has not been compared and there is a lack of consensus on which criteria should be used to identify aberrant behaviors with prescription opioids from administrative prescription data.[14]

Therefore, the primary objective of our study was to validate and compare the performance of five previously published algorithms. We used administrative claims data from two large health insurance programs in the US, Medicaid and Clinformatics™ Data Mart (CDM; OptumInsight, Eden Prairie, MN), which collectively insure approximately 125 million individuals.

METHODS

Data sources

Data from the Medicaid Analytic eXtract (MAX) spanned the time period from 2000 to 2006 and included 49 U.S. states (excluding Arizona) and Washington D.C., while data from CDM were available from 2004 to 2013 and included all 50 U.S. states and Washington

D.C. Both data sources contain individual-level information on inpatient and outpatient diagnoses and procedures, as well as records of outpatient prescription dispensing. The use of the MAX and CDM de-identified databases for research purposes was approved by the Partners Healthcare Institutional Review Board, and need for informed consent was waived.

Study population

We included all new users of prescription opioids (i.e., no record of previous use for at least 6 months before initiating treatment) who were over 15 years old, had no diagnosed malignancy, and had no history of opioid abuse or dependence in the 6 months prior to the first dispensed opioid prescription. We considered the dispensing of oral and transdermal opioid analgesics; injectable opioid formulations and buprenorphine were excluded (see Appendix Table A1 for a full list of included compounds). We defined the index date as the date of the initial opioid prescription dispensing and required insurance coverage to extend from 6 months before the index date through 12 months after the index date (to allow adequate time for assessment of algorithm criteria). Baseline patient characteristics, including demographics, use of healthcare services, diagnosis of pain conditions, comorbid mental health conditions, and use of psychotropic medications, were measured during the 6 months up to and including the index date.

We deliberately chose to include all new users of prescription opioids in this analysis for several reasons. Most importantly, it reflects how we anticipate these algorithms will be used in practice. Furthermore, even patients with a single opioid prescription are at some risk for opioid-related adverse events. Finally, the primary purpose of this study was to evaluate whether people who met triggering criteria of each of the algorithms had opioid-related adverse events. Each of the algorithms already implicitly requires more than 1 prescription to trigger a positive result.

Algorithms to identify aberrant behaviors with prescription opioids

We consider aberrant behaviors to be "a constellation of behaviors that have grown to be recognized by clinicians as potentially indicative of prescription opioid abuse," as defined by a Tufts Health Care Institute expert panel. [15]

In addition to CMS' Overutilization Monitoring System,[9] we identified four additional algorithms from the peer-reviewed literature: the Opioid Misuse Score,[13] the Controlled Substance–Patterns of Utilization Requiring Evaluation (CS-PURE) algorithm,[12] the Katz et al algorithm,[11] and the Cepeda et al algorithm.[10] We did not employ a systematic search strategy to identify these algorithms.

We used prescription claims information on medication, dose, quantity dispensed, number of days supplied, prescribing physician, pharmacy, and date of dispensing during the 12 months following the index date to identify aberrant behaviors according to each algorithm. Many of the algorithms use a combination of criteria related to the number of prescriptions, number of prescribers, number of pharmacies, total dosage, or duration of supply. Full descriptions of each algorithm are presented in Table 1.

Outcome measures: opioid-related adverse event

We defined opioid-related adverse events as at least one diagnosis of opioid abuse, dependence, or overdose during the 12-month follow-up period. These conditions were identified using the inpatient and outpatient International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9 CM) codes (detailed definitions in Appendix Table A2). Previous validation studies suggest a high concordance between ICD-9 CM codes for opioid abuse and dependence and overdose and diagnoses for adverse opioid events recorded in medical records.[16,17] In sensitivity analyses to ensure high specificity of the outcome measure, we further restricted the definition to opioid poisoning alone, which has a positive predictive value of 96% for opioid-related adverse events [16].

Statistical analysis

Because the MAX and CDM data sets contain different populations over varying time periods, all results are presented stratified by cohort. Baseline characteristics of patients in the MAX and CDM cohorts were evaluated using descriptive statistics. We calculated the proportion of patients meeting each of the five definitions of aberrant behavior. For all pairwise comparisons, we assessed the degree of concordance between algorithms by calculating the percentage agreement (i.e., the proportion of patients two algorithms classify the same way) and Cohen's kappa (i.e., a measure of inter-rater reliability that accounts for chance agreement). A Cohen's kappa value of 0 indicates no agreement and a value of 1 indicates perfect agreement; though there are no accepted thresholds, kappa values between 0 and 0.20 have been interpreted as slight agreement, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1 as near perfect.[18]

The absolute risk of an opioid-related adverse event for patients meeting each definition was calculated, along with unadjusted risk differences (RD) and corresponding 95% confidence intervals (CI) comparing patients who met and did not meet each definition. Sensitivity, specificity, positive predictive value, and negative predictive value were also calculated, comparing all the algorithms to a uniform benchmark of opioid-related adverse event diagnoses.

All analyses were conducted using SAS version 9.4 (SAS Institute, Cary, North Carolina).

RESULTS

The study sample consisted of 3.8 million new opioid users in MAX and 4.3 million new opioid users in CDM. Of these new users, 18,377 (0.49%) in MAX and 6,356 (0.15%) in CDM had an opioid-related adverse event within 12 months of their first opioid dispensing. For sensitivity analyses restricting outcomes to individuals with codes indicating opioid poisonings, there were 1,915 (0.05%) events in MAX and 1,114 (0.03%) events in CDM.

Compared to the CDM cohort, patients in MAX were younger (median age, 29 years in MAX vs. 40 years in CDM), more likely to be female (72% in MAX vs. 54% in CDM), less likely to have a chronic pain diagnosis (39% in MAX vs. 44% in CDM), more likely to have a mental health disorder (18% in MAX vs. 11% in CDM), and more likely to have a diagnosis of alcohol abuse or non-opioid substance abuse (1.2% and 1.6% respectively in

MAX versus 0.5% and 0.2% respectively in CDM; Table 2). Patients in MAX were also more likely to be dispensed at least one antidepressant (19% in MAX vs. 13% in CDM) or antipsychotic (6.4% in MAX vs. 0.8% in CDM) medication during the study period.

The proportion of new opioid users identified as engaging in aberrant behaviors with prescription opioids over the 12-month follow-up period varied markedly between the five algorithms. The CMS Overutilization Monitoring System classified the smallest proportion of new opioid users as having aberrant behaviors (0.02% in MAX; 0.01% in CDM), while the Katz algorithm criteria were met by 13% of patients in MAX and 8% of patients in CDM. There appeared to be a high percentage of agreement between algorithms (87.3-99.6% in MAX, 92.1-99.9% in CDM), though this is largely due to the vast majority of patients labelled as not having aberrant behaviors. However, Cohen's kappa, which accounts for chance agreement, demonstrates low to modest concurrence between the algorithms (Table 3). The lower values of Cohen's kappa inform us that much of the observed percentage agreement is due to chance because one categorization (i.e., not having aberrant behaviors) is much more likely than the other. The lowest levels of agreement were observed between the CMS Overutilization Monitoring System and the Katz algorithm (Cohen's kappa: 0.00 in MAX; 0.00 in CDM), while the highest levels of agreement were observed between the Cepeda algorithm and the Opioid Misuse Score's 'probable misuse' cutoff (Cohen's kappa: 0.50 in MAX; 0.27 in CDM) and the modified CS-PURE (Cohen's kappa: 0.38 in MAX; 0.30 in CDM).

Within one year of their index date, 0.54% of new opioid users in MAX and 0.15% in CDM were diagnosed with an opioid-related adverse event. Each algorithm found that those who met the criteria had a statistically significant increase in the risk of an event compared to those who did not (Table 4). However, the magnitude of the associations varied; while patients meeting the Katz algorithm had a 0.96% increased risk of an opioid-related adverse event in MAX (95% CI [0.92%, 0.99%]) and a 0.47% increased risk in CDM [0.44%, 0.50%], patients meeting the CMS Overutilization Monitoring System algorithm were at a 14.0% increased risk in MAX [11.7%, 16.4%] and 13.4% increased risk in CDM [9.8%, 16.9%].

Algorithm sensitivity was generally low, and ranged from less than 1% of individuals with an opioid-related adverse event meeting the CMS Overutilization Monitoring System algorithm to approximately one third of the individuals with an opioid-related adverse event meeting Katz algorithm (Table 4; Figure 1). In contrast, specificity was high across all algorithms, ranging from 87.4% to 99.9% in MAX and between 92.1% and 99.9% in CDM, indicating that the vast majority of people who did not experience an opioid-related adverse event were not flagged as having aberrant behaviors. For all algorithms, the probability of a patient who met an algorithm having an opioid-related adverse outcome was relatively low, as reflected by the positive predicted values (1.4% to 15.3% in MAX; 0.6% to 13.5% in CDM), while the negative predictive values (i.e., true negative proportions) were high (99.5% to 99.6% in MAX; 99.9% to 99.9% in CDM).

Sensitivity analyses using only opioid poisoning codes to define opioid-related adverse events did not substantially alter conclusions.

DISCUSSION

Algorithms that can identify potentially aberrant behaviors from electronic databases may be useful tools to combat the prescription opioid epidemic, with possible applications including triggers for early intervention, surveillance, and policy evaluation. We compared the performance of five existing algorithms designed to detect aberrant behaviors with prescription opioids in large cohorts of publicly and commercially insured patients. Each algorithm was able to identify patients who were at a statistically significant increased risk of an adverse opioid-related event. Despite differences in patient characteristics and overall prevalence of opioid-related adverse events, algorithms had similar performance in the MAX and CDM datasets.

Despite each algorithm's ability to identify patients who were at increased risk, they shared some important limitations; sensitivity was generally low and many patients who would go on to develop an opioid-related adverse event never displayed prescription patterns that were classified as "aberrant." This observation may partially reflect limitations in our data sources, including the inability of prescription claims data to capture illicit acquisition of opioids. For patients primarily acquiring opioids through illicit channels, there may not be any indication of aberrant behaviors in their prescription claims. Low sensitivity and positive predictive values may also highlight an area of potential improvement for future algorithm development. Improved performance may be possible by taking a more empirical approach to algorithm development, especially through the use of machine learning methodologies.

There were also notable differences between the algorithms. The specific criteria and cutoffs employed by each of the five algorithms varied substantially, reflecting a lack of consensus about what characteristics may indicate aberrant behaviors with prescription opioids. It is important to note that algorithms were originally developed in separate populations, using different timeframes for evaluation, and with subtly different goals. As a result, agreement between algorithms was low, with different algorithms capturing different groups of high risk patients. While the Opioid Misuse Score has been demonstrated to have a linear relationship with diagnoses of opioid abuse and dependence,[13] the remaining four algorithms had not previously been evaluated for their ability to predict opioid-related adverse events. We found that measures of performance, including sensitivity, specificity, positive predictive value, and negative predictive value, varied between algorithms.

Given the demonstrated differences in performance, some algorithms may be more appropriate for specific contexts depending on the relative importance of incorrectly classifying a patient as having aberrant patterns (i.e., desiring high specificity) versus failing to detect problematic behaviors (i.e., desiring high sensitivity). If an algorithm were being used to target harm reduction interventions such as naloxone distribution, sensitivity would likely be prioritized over specificity to minimize the number of potentially preventable adverse events that would be overlooked. In contrast, high specificity would be more important for any intervention used to limit access to prescription opioids to avoid creating undue burden on patients experiencing pain who are at low risk of an adverse event. Policy makers must carefully consider individual algorithm strengths and limitations alongside

legal and ethical concerns, including how false positive and false negative algorithm classifications will impact patient safety, privacy, and behavior.

The prevalence of aberrant behaviors with prescription opioids was generally low within the first year of initiating opioid use, except when defined using the Katz algorithm. We also found that the incidence of aberrant behaviors and of opioid-related adverse events was elevated in the MAX cohort compared to the CDM cohort, though these disparities are likely due to underlying differences in each cohort's population. Groups that tend to have elevated risk of opioid abuse had greater representation in MAX, including young people, patients with mental health comorbidities, and individuals with other underlying substance abuse problems.[19,20]

Our study has several limitations. First, records for any opioid prescriptions purchased without using insurance benefits will be missing from the MAX and CDM databases. In one large US-based study, approximately 19% of patients who were dispensed a Schedule II opioid had at least one cash transaction.[21] As a result, some patients in this study who purchased prescription opioids using cash may be misclassified as not meeting a given algorithm. Second, the CMS' introduction of a policy requiring the redaction of substance abuse-related claims [22,23] prevented the use of MAX data beyond 2006 for this analysis. Third, the time period used to evaluate aberrant behaviors was the same time period used to evaluate adverse opioid-related events. Therefore, some outcome events may predate the patient meeting certain algorithm criteria. Fourth, there is likely underreporting of opioid-related adverse events. This may occur if the physician is unaware of the adverse event or because the physician does not code the condition in an insurance claim. Further, some aberrant behaviors, especially diversion and fraud, would never be captured in diagnostic codes. Even some clinical manifestations of aberrant behaviors may not be captured in administrative claims. As a result, individuals who met the algorithm may be incorrectly classified as not having a subsequent opioid-related adverse event, which would result in an underestimation of the true positive predictive value of an algorithm. While there is no single gold standard method for assessing prescription opioid abuse, we have ensured that all algorithms were compared to a uniform benchmark, meaning the magnitude of opioid-related adverse event misclassification will not vary by algorithm. Finally, the generalizability of our findings may be somewhat limited by restricting the sample to new users. However, the large sample, nationwide coverage, and use of both public and private insurance information help to make the results applicable to most new users of prescription opioid medications with insurance coverage in the United States.

Algorithms to detect aberrant opioid use from administrative data will have increasingly important applications in surveillance, research, and the evaluation of policies intended to reduce prescription opioid abuse. There is a growing interest in the use of algorithms to actively monitor prescription drug monitoring program systems,[2] and CMS has already implemented the Overutilization Monitoring System in Medicare Part D.[9] Our study examines the performance of existing algorithms, demonstrating their strengths – high specificity and negative predictive value – and weaknesses – low sensitivity and positive predictive value. These findings have important implications for policy-makers interested in designing targeted interventions based on claims-based algorithms. They also justify further

work to improve the capabilities of these algorithms to detect aberrant behaviors with opioids and identify the patients that may be at increased risk of an adverse opioid-related event.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Manchikanti L, Fellows B, Ailinani H, Pampati V. Therapeutic use, abuse, and nonmedical use of opioids: a ten-year perspective. Pain Physician. 13:401–35.

2. White House Office of National Drug Control Policy. National Drug Control Strategy 2014. 2014. https://www.whitehouse.gov/sites/default/files/ondcp/policy-and-research/ndcs_2014.pdf

3. US Executive Office of the President, Office of National Drug Control Policy. Epidemic: Responding to America's Prescription Drug Abuse Crisis. Washington, D.C.: 2011. https://www.ncjrs.gov/pdffiles1/ondcp/rx_abuse_plan.pdf (accessed 1 Mar2017)

4. Centers for Disease Control and Prevention. Vital signs: overdoses of prescription opioid pain relievers—United States, 1999–2008. MMWR Morb Mortal Wkly Rep. 2011; 60:1487–92. [PubMed: 22048730]

5. Rudd RA, Aleshire N, Zibbell JE, Gladden RM. Increases in Drug and Opioid Overdose Deaths–United States, 2000-2014. MMWR Morb Mortal Wkly Rep. 2016; 64:1378–82. [PubMed: 26720857]

6. Center for Behavioral Health Statistics and Quality. 2015 National Survey on Drug Use and Health: Detailed Tables. Rockville, MD: 2016.

7. Birnbaum HG, White AG, Schiller M, Waldman T, Cleveland JM, Roland CL. Societal costs of prescription opioid abuse, dependence, and misuse in the United States. Pain Med. 2011; 12:657–667. [PubMed: 21392250]

8. Hansen RN, Oster G, Edelsberg J, Woody GE, Sullivan SD. Economic costs of nonmedical use of prescription opioids. Clin J Pain. 2011; 27:194–202. [PubMed: 21178601]

9. Tudor, CG. Memorandum: Medicare Part D Overutilization Monitoring System. 2013. https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/Downloads/HPMS-memo-Medicare-Part-D-Overutilization-Monitoring-System-07-05-13-.pdf

10. Cepeda MS, Fife D, Chow W, Mastrogiovanni G, Henderson SC. Assessing opioid shopping behaviour: a large cohort study from a medication dispensing database in the US. Drug Saf. 2012; 35:325–34. [PubMed: 22339505]

11. Katz N, Panas L, Kim M, Audet AD, Bilansky A, Eadie J, et al. Usefulness of prescription monitoring programs for surveillance–analysis of Schedule II opioid prescription data in Massachusetts, 1996-2006. Pharmacoepidemiol Drug Saf. 2010; 19:115–23. [PubMed: 20014166]

12. Parente ST, Kim SS, Finch MD, Schloff LA, Rector TS, Seifeldin R, et al. Identifying controlled substance patterns of utilization requiring evaluation using administrative claims data. Am J Manag Care. 2004; 10:783–90. [PubMed: 15623267]

13. Sullivan MD, Edlund MJ, Fan M-Y, Devries A, Brennan Braden J, Martin BC. Risks for possible and probable opioid misuse among recipients of chronic opioid therapy in commercial and medicaid insurance plans: The TROUP Study. Pain. 2010; 150:332–9. [PubMed: 20554392]

14. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. National Action Plan for Adverse Drug Event Prevention. Washington, D.C.: 2014. https://health.gov/hcq/pdfs/ade-action-plan-508c.pdf (accessed 1 Mar2017)

15. Katz NP, Adams EH, Chilcoat H, Colucci RD, Comer SD, Goliber P, et al. Challenges in the Development of Prescription Opioid Abuse-deterrent Formulations. Clin J Pain. 2007; 23:648–660. [PubMed: 17885342]

16. Green CA, Perrin NA, Janoff SL, Campbell CI, Chilcoat HD, Coplan PM. Assessing the accuracy of opioid overdose and poisoning codes in diagnostic information from electronic health records, claims data, and death records. Pharmacoepidemiol Drug Saf.

17. Carrell DS, Cronkite D, Palmer RE, Saunders K, Gross DE, Masters ET, et al. Using natural language processing to identify problem usage of prescription opioids. Int J Med Inf. 2015; 84:1057–64.

18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–74. [PubMed: 843571]

19. Turk DC, Swanson KS, Gatchel RJ. Predicting opioid misuse by chronic pain patients: a systematic review and literature synthesis. Clin J Pain. 2008; 24:497–508. [PubMed: 18574359]

20. Edlund MJ, Martin BC, Russo JE, Devries A, Braden JB, Sullivan MD. The Role of Opioid Prescription in Incident Opioid Abuse and Dependence Among Individuals with Chronic Non-cancer Pain. Clin J Pain. 2014; 30:557–64. [PubMed: 24281273]

21. Cepeda MS, Fife D, Chow W, Mastrogiovanni G, Henderson SC. Opioid shopping behavior: how often, how soon, which drugs, and what payment method. J Clin Pharmacol. 2013; 53:112–7. [PubMed: 23400751]

22. Frakt AB, Bagley N. Protection or Harm? Suppressing Substance-Use Data. N Engl J Med. 2015; 372:1879–1881. [PubMed: 25875196]

23. Rough K, Bateman BT, Patorno E, Desai RJ, Park Y, Hernandez-Diaz S, et al. Suppression of Substance Abuse Claims in Medicaid Data and Rates of Diagnoses for Non–Substance Abuse Conditions. JAMA. 2016; 315:1164. [PubMed: 26978213]

Key points

- With tens of thousands of fatal overdoses involving opioids annually, algorithms to detect aberrant opioid use from administrative data will have increasingly important applications in surveillance, research, and public health interventions.

- Our study examines the performance of existing algorithms, demonstrating their strengths and weaknesses. These algorithms had not been previously compared.

- The agreement of the algorithms was poor to moderate. All five algorithms generally had high specificity and negative predictive value, but low sensitivity and positive predictive value.

- Our findings justify further work to improve the capabilities of algorithms to detect aberrant behaviors with opioids.
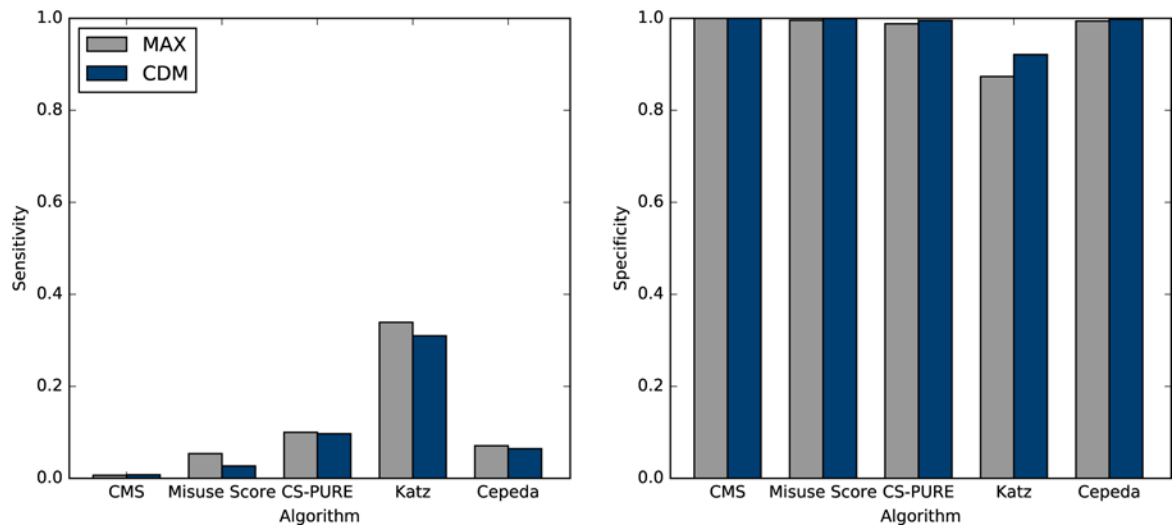
Figure 1.
Sensitivity and specificity of aberrant opioid behavior algorithms' ability to predict opioid related adverse events

Abbreviations: MAX, Medicaid Analytic eXtract; CDM, Clinformatics™ Data Mart; CMS, Centers for Medicaid and Medicare Services; CS-PURE, Controlled Substance Patterns of Utilization Requiring Evaluation

Table 1

Algorithms to identify aberrant behaviors with prescription opioids from insurance claims data

Algorithm name	Description	Original algorithm timeframe
CMS Overutilization Monitoring System	120mg of average morphine equivalents for 90 consecutive days AND >3 prescribers for any opioid drug AND >3 pharmacies for any opioid drug	3 months
Opioid Misuse Score	For each 6-month (180 day) period: Days supplied of IR/SA opioids: '0' points if 185 days, '1' point if 186-240 days, '2' points if > 240 days Days supplied of ER/LA opioids: '0' points if 185 days, '1' point if 186-240 days, '2' points if > 240 days Number of pharmacies for opioids: '0' points if 0-2 pharmacies, '1' point if 3-4 pharmacies, '2' points if 5 pharmacies Number of prescribers of opioids: '0' points if 0-2 prescribes, '1' point if 3-4 prescribers, '2' points if 5 prescribers Scores for both periods are summed and classified as: no misuse (0-1 points), possible misuse (2-4 points), probable misuse (5 points)	12 months (two 6 month periods)
Modified *CS-PURE	6 prescribers for the same opioid drug OR 4 pharmacies for the same opioid drug OR 4 prescriptions for butorphanol tartrate OR 2 prescriptions resulting in an overlapping supply of ER/LA opioids for 90 consecutive days	12 months
Katz et al	2 prescribers for the same opioid drug AND 2 pharmacies for the same opioid drug	12 months
Cepeda et al	2 overlapping opioid prescription episodes (1 day of overlap for 2 opioid drugs prescribed by different physicians) AND 3 pharmacies used for opioid drugs during overlapping episodes	18 months

Abbreviations: CMS, Centers for Medicaid and Medicare Services; CS-PURE, Controlled Substance Patterns of Utilization Requiring Evaluation; ER/LA, extended release/long acting; IR/SA, immediate release/short acting

*
The original CS-PURE contains ten items to identify possible aberrant behaviors related to multiple controlled substances. Only CS-PURE criteria related to prescription opioid compounds were retained for this analysis.

Table 2

Baseline characteristics of new opioid users in MAX (2000-2006) and CDM (2004-2013)

	MAX (N = 3,752,066)		CDM (N = 4,298,537)	
	n	%	n	%
Age (years)				
15 to 19	904,185	24.1	422,796	9.8
20 to 34	1,389,628	37.0	1,161,905	27.0
35 to 49	928,393	24.7	1,509,819	35.1
50 to 64	509,941	13.6	1,083,216	25.2
Over 65	19,919	0.5	120,801	2.8
Sex				
Male	1,039,029	27.7	1,964,498	45.7
Female	2,713,008	72.3	2,333,700	54.3
Unknown	29	0.0	339	0.0
Race and ethnicity				
Black/African American	1,036,256	27.6	n/a	–
White	1,810,815	48.3	n/a	–
Hispanic/Latino	477,973	12.7	n/a	–
Asian/Pacific Islander	141,757	3.8	n/a	–
American Indian	44,826	1.2	n/a	–
Other/unknown	240,439	6.4	n/a	–
Region				
Northeast	689,409	18.4	383,799	8.9
Midwest	849,110	22.6	1,160,360	27.0
South	1,289,613	34.4	2,085,504	48.5
West	923,934	24.6	667,098	15.5
Baseline hospitalizations				
0	3,007,229	80.1	3,844,627	89.4
1	604,253	16.1	421,633	9.8
2 or more	140,584	3.7	32,277	0.8
Baseline prescriptions				
0	817,490	21.8	1,151,378	26.8
1 to 3	1,325,455	35.3	1,840,991	42.8
4 to 6	786,773	21.0	819,451	19.1
7 or more	822,348	21.9	486,717	11.3
Chronic pain diagnosis	1,447,047	38.6	1,905,528	44.3
Mental health diagnosis	656,637	17.5	468,462	10.9
Antidepressant use	692,644	18.5	561,144	13.1
Antipsychotic use	241,023	6.4	32,302	0.8
Benzodiazepine use	272,110	7.3	283,960	6.6
Tobacco use	114,949	3.1	179,142	4.2
Alcohol abuse	44,410	1.2	19,615	0.5

	MAX (N = 3,752,066)		CDM (N = 4,298,537)	
	n	%	n	%
Non-opioid substance abuse	58,290	1.6	9,083	0.2

Abbreviations: MAX, Medicaid Analytic eXtract; CDM, Clinformatics™ Data Mart

Table 3

Agreement between algorithm classifications for aberrant behaviors: Cohen's kappa

	MAX			
	Opioid Misuse Score (Probable)	**Modified * CS-PURE**	**Katz et al**	**Cepeda et al**
CMS Overutilization Monitoring System	0.07	0.03	0.00	0.06
Opioid Misuse Score (Probable)	–	0.35	0.01	0.50
Modified * CS-PURE	–	–	0.13	0.38
Katz et al	–	–	–	0.08
	CDM			
	Opioid Misuse Score (Probable)	Modified * CS-PURE	Katz et al	Cepeda et al
CMS Overutilization Monitoring System	0.10	0.03	0.00	0.06
Opioid Misuse Score (Probable)	–	0.18	0.01	0.27
Modified * CS-PURE	–	–	0.09	0.30
Katz et al	–	–	–	0.05

Abbreviations: MAX, Medicaid Analytic eXtract; CMS, Centers for Medicaid and Medicare Services; CS-PURE, Controlled Substance Patterns of Utilization Requiring Evaluation; CDM, Clinformatics™ Data Mart

*
The original CS-PURE contains ten items to identify possible aberrant behaviors related to multiple controlled substances. Only CS-PURE criteria related to prescription opioid compounds were retained for this analysis.

Table 4

Opioid-related adverse events according to aberrant behavior algorithm classifications: Risk differences, sensitivity, specificity, positive predictive value, and negative predictive value

MAX cohort

	Total sample: met algorithm n (%)	Met algorithm: had outcome n (%)	Risk difference RD (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	Positive predictive value % (95% CI)	Negative predictive value % (95% CI)
CMS Overutilization Monitoring System	881 (0.02)	135 (15.32)	14.04 (11.72, 16.37)	0.67 (0.55, 0.78)	99.98 (99.98, 99.98)	15.32 (12.94, 17.70)	99.46 (99.46, 99.47)
Opioid Misuse Score (Probable)	16,810 (0.45)	1087 (6.47)	5.64 (5.28, 6.01)	5.36 (5.05, 5.67)	99.58 (99.57, 99.59)	6.47 (6.09, 6.84)	99.49 (99.48, 99.49)
Modified *CS-PURE	46,711 (1.24)	2025 (4.34)	3.64 (3.46, 3.82)	9.98 (9.57, 10.39)	98.80 (98.79, 98.81)	4.34 (4.15, 4.52)	99.51 (99.50, 99.51)
Katz et al	478,882 (12.76)	6878 (1.44)	0.96 (0.92, 0.99)	33.91 (33.25, 34.56)	87.35 (87.32, 87.39)	1.44 (1.40, 1.47)	99.59 (99.58, 99.60)
Cepeda et al	24,215 (0.65)	1432 (5.91)	5.10 (4.81, 5.39)	7.06 (6.71, 7.41)	99.39 (99.38, 99.40)	5.91 (5.62, 6.21)	99.49 (99.49, 99.50)

CDM cohort

	Total sample: met algorithm n (%)	Met algorithm: had outcome n (%)	Risk difference RD (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	Positive predictive value % (95% CI)	Negative predictive value % (95% CI)
CMS Overutilization Monitoring System	363 (0.01)	49 (13.50)	13.35 (9.84, 16.87)	0.77 (0.56, 0.99)	99.99 (99.99, 99.99)	13.50 (9.98, 17.01)	99.85 (99.85, 99.86)
Opioid Misuse Score (Probable)	2,154 (0.05)	169 (7.85)	7.84 (6.70, 8.99)	2.67 (2.27, 3.06)	99.95 (99.95, 99.96)	7.85 (6.71, 8.98)	99.86 (99.85, 99.86)
Modified *CS-PURE	20,619 (0.48)	614 (2.98)	2.89 (2.66, 3.13)	9.68 (8.96, 10.41)	99.53 (99.53, 99.54)	2.98 (2.75, 3.21)	99.87 (99.86, 99.87)
Katz et al	340,569 (7.92)	1965 (0.58)	0.47 (0.44, 0.50)	30.99 (29.86, 32.13)	92.11 (92.09, 92.14)	0.58 (0.55, 0.60)	99.89 (99.89, 99.89)
Cepeda et al	8,993 (0.21)	408 (4.54)	4.41 (3.98, 4.84)	6.44 (5.83, 7.04)	99.80 (99.80, 99.80)	4.54 (4.11, 4.97)	99.86 (99.86, 99.87)

Abbreviations: MAX, Medicaid Analytic eXtract; CMS, Centers for Medicaid and Medicare Services; CS-PURE, Controlled Substance Patterns of Utilization Requiring Evaluation; RD, risk difference; CI, confidence interval; CDM, Clinformatics™ Data Mart

*
The original CS-PURE contains ten items to identify possible aberrant behaviors related to multiple controlled substances. Only CS-PURE criteria related to prescription opioid compounds were retained for this analysis