ADDICTION
SSA SOCIETY FOR THE STUDY OF ADDICTION
RESEARCH REPORT
doi:10.1111/add.15635

New means, new measures: assessing prescription drug-seeking indicators over 10 years of the opioid epidemic

Brea L. Perry[1,2]*, Meltem Odabaş[2]*, Kai-Cheng Yang[3], Byungkyu Lee[2], Patrick Kaminski[2,3], Brian Aronson[2], Yong-Yeol Ahn[1,3], Carrie B. Oser[4], Patricia R. Freeman[5] & Jeffrey C. Talbert[5]

Network Science Institute, Indiana University, Bloomington, IN, USA,[1] Department of Sociology, Indiana University, Bloomington, IN, USA,[2] School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, USA,[3] Department of Sociology, University of Kentucky, Lexington, KY, USA[4] and Department of Pharmacy Practice and Science, University of Kentucky, Lexington, KY, USA[5]

ABSTRACT

Background and aims Prescription drug-seeking (PDS) from multiple prescribers is a primary means of obtaining prescription opioids; however, PDS behavior has probably evolved in response to policy shifts, and there is little agreement about how to operationalize it. We systematically compared the performance of traditional and novel PDS indicators. **Design** Longitudinal study using a de-identified commercial claims database. **Setting** United States, 2009–18. **Participants** A total of 318 million provider visits from 21.5 million opioid-prescribed patients. **Measurements** We applied binary classification and generalized linear models to compare predictive accuracy and average marginal effect size predicting future opioid use disorder (OUD), overdose and high morphine milligram equivalents (MME). We compared traditional indicators of PDS to a network centrality measure, PageRank, that reflects the prominence of patients in a co-prescribing network. Analyses used the same data and adjusted for patient demographics, region, SES, diagnoses and health services. **Findings** The predictive accuracy of a widely used traditional measure (*N* + unique doctors and *N* + unique pharmacies in 90 days) on OUD, overdose and MME decreased between 2009 and 2018, and performed no better than chance (50% accuracy) after 2015. Binarized PageRank measures however exhibited higher predictive accuracy than the traditional binary measures throughout 2009-2018. Continuous indicators of PDS performed better than binary thresholds, with days of Rx performing best overall with 77–93% predictive accuracy. For example, days of Rx had the highest average marginal effects on overdose and OUD: a 1 standard deviation increase in days of Rx was associated with a 6–8% [confidence intervals (CIs) = 0.058–0.061 and 0.078–0.082] increase in the probability of overdose and a 4–5% (CIs = 0.038–0.043 and 0.047–0.053) increase in the probability of OUD. PageRank performed nearly as well or better than traditional indicators of PDS, with predictive performance increasing after 2016. **Conclusions** In the United States, network-based measures appear to have increasing promise for identifying prescription opioid drug-seeking behavior, while indicators based on quantity of providers or pharmacies appear to have decreasing utility.

Keywords Co-prescription networks, drug dependence, opioid use disorder, opiates, overdose, prescription drug-seeking, prescription opioids.

Correspondence to: Meltem Odabaş, Indiana University Schuessler Institute for Social Research 1022 East Third Street Bloomington, IN 47405, USA. E-mail: modabas@iu.edu
Submitted 26 June 2020; initial review completed 2 October 2020; final version accepted 23 June 2021
*Co-authors.

INTRODUCTION

In response to the opioid epidemic, all 50 states and Washington DC have instituted prescription drug monitoring programs (PDMPs) and other supply-side policies to discourage prescription drug-seeking (PDS) for non-medical use and to reduce inappropriate prescribing and dispensing. However, these policies have produced mixed results. Although the number of patients meeting strict criteria for PDS has declined, there has been no simultaneous decrease in non-medical prescription drug use or related negative outcomes [1]. One explanation for these patterns is that both patients and prescribers have adapted their behavior in response to the new regulatory

environment. If true, indicators of PDS that were once useful may be outdated, limiting our ability to 'see' PDS and to accurately evaluate PDMPs. This is consistent with Goodhart's Law [2]: when a measure becomes a target, it ceases to be a good measure. A rapidly evolving regulatory environment demands that researchers continually monitor PDS patterns and revisit how PDS is operationalized to more clearly reflect new phases of the epidemic.

Traditional measures of prescription drug-seeking

There is little agreement in the literature about how to best operationalize PDS (i.e. 'doctor-shopping') [3,4], resulting in a heterogeneous set of measures (see Table 1). One popular approach has been to focus upon the numbers of prescribers, or 'multiple-provider episodes (MPE)'. A binary indicator is defined by setting a threshold number of prescribers in given a time-period (e.g. receiving prescriptions from two or more prescribers within 30 days). Due to the difficulty of obtaining ground-truth, thresholds are often chosen arbitrarily [4–28]. Additionally, classifying PDS behavior using MPEs alone can produce many false negatives and false positives. For example, patients early in drug use trajectories with low levels of PDS or those who engage in heavy PDS for a brief period may not be correctly classified as high-risk (i.e. a false negative). Alternatively, low-risk patients with serious conditions (e.g. cancer) or those who see multiple physicians within the same practice may produce false positives. Other problems with MPE include inability to discriminate between patients with simultaneous versus successive prescriptions [16] and failure to consider quantities of drugs obtained [24,25].

These limitations of MPE have led to other measurement strategies. Most notably, the CDC developed a PDS measure based on daily morphine milligram equivalents (MME) over 6 months [9]. Others have incorporated pharmacy shopping. Katz *et al.* [5] systematically compared the performance of different thresholds for number of prescribers and pharmacies during 1 year. From this, they developed a criterion of four or more different prescribers in addition to four or more pharmacies (the '4 + 4' criterion), arguing that this measure provides the optimum balance between false positives and negatives. However, this analysis was conducted prior to the implementation of many opioid prescribing policies, and has yet to be revisited.

A social network approach to measuring PDS

Identifying PDS behavior is becoming more difficult as increased regulation has reduced high-volume prescribing and forced patients and prescribers to alter their behavior to avoid sanction. Recently, Perry *et al.* [28] found that using a social network approach is a promising method for characterizing patients who seek out specific prescriber targets, but who do not meet high MPE thresholds. Social network approaches use patterns of co-prescription ties between patients and prescribers to identify irregular and excessive prescribing activity that cluster around central actors. Patients engaging in PDS are likely to visit the same set of prescribers (e.g. those that prescribe high doses of opioids to control patients' pain, those operating pill mills or those that do not monitor PDMP data), causing them to be linked systematically to many other drug-seeking patients through those providers. Network indicators can identify less egregious PDS behavior using information about which prescribers are visited, rather than how many.

Two lines of research support the utility of a social network approach [29–31]. First, PDS is clustered around

Table 1 Measures of prescription drug-seeking (PDS) behavior.

	[4,5]	*[6–15]*	*[16]*	*[17]*	*[18]*	*[19,20]*	*[21]*	*[22]*	*[23]*	*[24,25]*	*[26,27]*	*[28]*
No. Rx[a]									✓			
No. providers[b]		✓								✓		
No. pharmacies[c]								✓				
N + N + 90[d]	✓				✓	✓						
Max. daily MME[e]							✓					
Total MME[f]				✓						✓		
Chronic[g]					✓							
Concomitant[h]					✓							
No. overlapping Rx[i]			✓	✓								
Days of Rx[j]											✓	
PageRank[k]												✓

[a]Total number of prescriptions; [b]total number of unique providers visited; [c]total number of unique pharmacies visited; [d]dichotomous variable: visiting N unique providers and N unique pharmacies in any given 90 days; [e]maximum daily dose of opioid prescription in *x* days (*x* = 90 or 180); converted to morphine milligram equivalent (MME) units; [f]total dose of opioid prescription in *x* days (*x* = 90 or 180); converted to morphine milligram equivalent (MME) units; [g]consuming more than a cutoff level of MME per day for more than *x* days (*x* = 90 or 180); [h]filling more than 30 days of concomitant opioids and benzodiazepines; [i]total number of days in which two or more opioid prescriptions overlap; [j]total number of days with opioid prescriptions; [k]percentile ranking for PageRank centrality measure.

prescribers who are complicit, easily manipulated or unlikely to monitor electronic data [32]. For instance, Cepeda *et al.* [29] estimated that most PDS is concentrated around 13% of providers who prescribed any opioids, and only 2% of prescribers accounting for heavy PDS. Secondly, qualitative research indicates that information about prescriber behavior (e.g. who liberally prescribes opioids) is disseminated through social networks, increasing targeting of a similar set of prescribers by patients engaged in PDS [30]. Thus, together with traditional indicators, we evaluate PageRank, a novel indicator of targeted PDS behavior that uses information regarding links between patients and providers in a prescription drug network [28].

Here, we present the first systematic and longitudinal evaluation of PDS metrics and their effectiveness in predicting adverse outcomes: opioid use disorder (OUD), maximum daily MME (i.e. high overdose risk) and opioid overdose. We also test novel network-based indicators, which we expect to outperform traditional indicators in more recent years due to behavioral adaptation to supply-side prescription drug policies. These indicators reflect patient prominence in a prescription network, which signals the tendency to visit a set of prescribers also visited by others engaging in PDS behavior.

METHODS

Data

Data for these analyses are from an administrative claims database—Optum's de-identified Clinformatics® Data Mart Database. It includes all medical, procedure and pharmacy claims from a large commercially insured population. It contains approximately 21.5 million opioid-prescribed patients, with approximately 318 million unique observations spanning the first quarter of 2009 to the third quarter of 2018. Demographic characteristics of these patients reflect the national commercially insured population.

Dependent variables

We construct three dependent variables to compare the predictive value of PDS indicators over time. All variables are aggregated at the patient level for each quarter. OUD and overdose are generated based on patient diagnostic records using the International Classification of Diseases (ICD). The OUD indicator is inclusive of opioid abuse or dependence, except for remission. Overdose is poisoning due to excessive use of a drug, and is inclusive of prescription and illicit opiates.

Maximum daily MME is the standardized total daily dose of opioids using opioid conversion factors established by the CDC [33]. We use both a binary version (MME $\geq$ 90) and a logged continuous version, depending upon the model. See the Supporting information for additional detail, descriptive statistics, and the ICD and National Drug Codes (NDC) used to generate variables.

Independent variables

Independent variables operationalize prescription drug-seeking behavior for the purpose of comparing how well they predict the above dependent variables over time. Traditional measures of PDS were chosen using a comprehensive literature search (see Table 1). Provider count measures the total number of unique providers from whom a patient received an opioid prescription. Pharmacy count measures the total number of unique pharmacies where an opioid prescription was filled. $N + N + 90$ is a binary indicator that identifies patients who visited at least N providers and N pharmacies within the last 90 days. As studies use different values of N (4 is the most common) we generate three different versions of this binary variable, with values of $N = 3$, 4 and 5. Overlapping Rx measures the number of overlapping opioid prescriptions. Days of Rx aggregates the total number of days of opioid coverage from each opioid prescription the patient received in each quarter.

To operationalize PDS behavior using network-based indicators, we use prescription ties between patients and prescribers. We employ a common measure of network centrality, PageRank, which is described in detail in the Supporting information [34]. In the context of prescription networks, it assigns a high score to a patient if they visit providers who prescribe opioids to many other patients who also have high scores. A higher PageRank score is likely to be indicative of PDS and thus predictive of OUD, overdose and high MME. We use a percentile rank conversion for PageRank (i.e. PageRank percentile, or PRP) due to the right-skewed nature of the variable. We test three thresholds that indicate being at the 99th, 95th and 90th percentiles or above.

We use the following measures for matching procedures described in Analyses: demographic characteristics (age, sex, race, census region), socio-economic status (low-income subsidy; income; household net worth), insurance type, health diagnoses (cancer, palliative care, HIV, hepatitis C, psychological disorder), total number of prescriptions, hospitalization and emergency department visits. Additionally, we control for the presence of medication for OUD (MOUD) consistent with medication assisted therapy, including all buprenorphine products. We remove MOUD prescribed for pain from our analyses; including these cases does not affect substantive results (see Supporting information, Figs S3 and S4 in the Supporting information).

Analyses

We first present descriptive statistics for PDS indicators and dependent variables over time to show trends across the epidemic. We then employ two methods to compare the performance of traditional and network PDS indicators for identifying OUD, overdose and high MME patients.

Comparison of predictive accuracy using the receiver operating characteristic-area under the curve (ROC-AUC)

We measure predictive accuracy of traditional and network indicators of PDS by calculating the area under the ROC-AUC. This method assesses the accuracy of predictions without patient controls and simulates how the indicators may be used in practice (i.e. in prescription monitoring). We predict outcomes at the next quarter (t + 1) and calculate ROC-AUC values longitudinally. Additional information and sensitivity analyses are presented in the Supporting information.

Comparison of average marginal effects using regression

We use survey-weighted generalized linear models (SWGLM) that control for patient-level covariates to compare average marginal effects (AME) for traditional and network measures of PDS. AME is the average change in the probability of an outcome when a predictor increases by one unit. Prognostic score (or outcome-based) matching is used to reduce extensive data, in a process analogous to propensity score matching [35]. Target patients meeting criteria for OUD, overdose or high MME are matched with control patients based on similarity in demographic characteristics, socio-economic status, insurance type and health status. We predict outcomes at the next quarter ($t + 1$) and calculate AME longitudinally

We present average marginal effects from these models for standardized outcomes (i.e. the effect for a 1 standard deviation increase in the PDS indicator) in the text and figures to facilitate comparisons. We also present frequencies and standard errors for binary and continuous measures, respectively, to help highlight the number of patients who meet criteria for a unit increase in each variable. The analysis was not pre-registered, therefore results should be considered exploratory. See Supporting information for details on matching methods and multicollinearity test results.

RESULTS

Trends in OUD, overdose, high MME and prescription drug-seeking, 2009–18

Consistent with expectations, the prevalence of OUD among patients in this claims database increases from 0.1% in 2009 to nearly 0.8% in 2018. Similarly, the prevalence of opioid overdose increases from 0.012% in 2009 to 0.08% in 2018. In contrast, prescription of MME > 90 declines during the period of observation from 2.9% of patients in 2009 to 1.7% in 2018, as does daily MME (See Supporting information, Figs S5 and S6). In sum, OUD and overdose have increased while high doses of opioids have decreased.

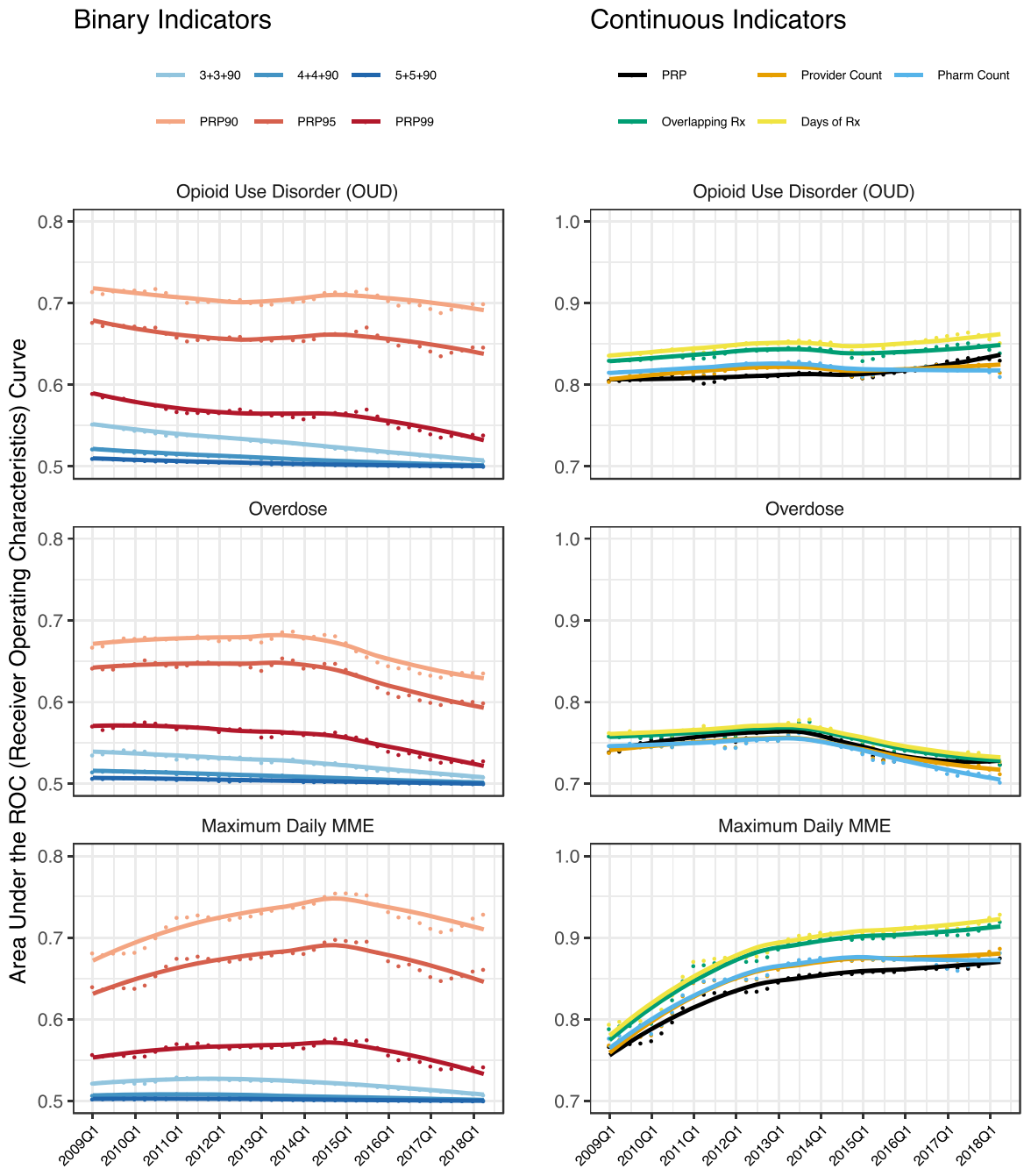
We also examined trends in PDS indicators during the same period which provide insight into the utility of different measures. By design, the network-based PageRank indicators (i.e. PRP90, PRP95 and PRP99) identify 10, 5 and 1% of patients, respectively. As a result, means remain constant during the period of observation. In contrast, as of 2018, only 0.1, 0.02 and 0.003% of patients met criteria for the traditional 3 + 3 + 90, 4 + 4 + 90 and 5 + 5 + 90 multiple provider/pharmacy episode indicators, respectively. This equates to only 10 152, 1246 and 247 patients of approximately 7.4 million. Continuous indicators (both traditional and network) remain stable or increase slightly between 2009 and 2018. The key finding here is that high-threshold binary MPE measures apply to a decreasing percentage of patients over time, with 4 + 4 + 90 and 5 + 5 + 90 identifying so few patients by 2018 that they are essentially obsolete.

Comparison of predictive accuracy using ROC-AUC

As illustrated in Fig. 1, the results for binary indicators show that $N + N + 90$ variables generate high numbers of false-positive and false-negative results. ROC-AUCs for $N + N + 90$ indicators are, at most, 55% for all three outcome variables, indicating only slightly higher predictive accuracy than random guessing. While PageRank at the 99th percentile performs only marginally better with, at most, 60% accuracy, predictive accuracy for PRP90 is in the 69–73% range for OUD, 63–68% for overdose and 68–75% for MME > 90. Network-based indicators of PDS more accurately predict adverse outcomes than traditional $N + N + 90$ thresholds.

Figure 1 also depicts ROC-AUC trends over time. The predictive accuracy of $N + N + 90$ indicators decreases slightly during the study period, from between 51–55 and 50% for all three versions by the end of 2018. For binary PageRank indicators, predictive accuracy remains stable for OUD and overdose from 2009 until 2015, after which point it begins to decline slightly by approximately 3% for OUD and 5% for overdose. Predictive accuracy of binary PageRank indicators for MME > 90 increases from between 56 and 86% in 2009 to a peak of between 57 and 75% in mid-2014. Afterwards, it declines to a range of between 54 and 71% for PRP99 and PRP90, respectively. In sum, all binary thresholds become slightly less predictive over time.

Continuous indicators perform better and are more similar with respect to predictive accuracy, overall, than binary indicators. Overlapping Rx maintains the highest



Notes: To denote potential bias due to the change in ICD code specifications, all quarters of 2015 are shaded gray.

Figure 1 Receiver operating characteristic-area under the curve (ROC-AUC) results for binary classification models predicting opioid use disorder (OUD), overdose and maximum daily morphine milligram equivalents (MME) in the next quarter

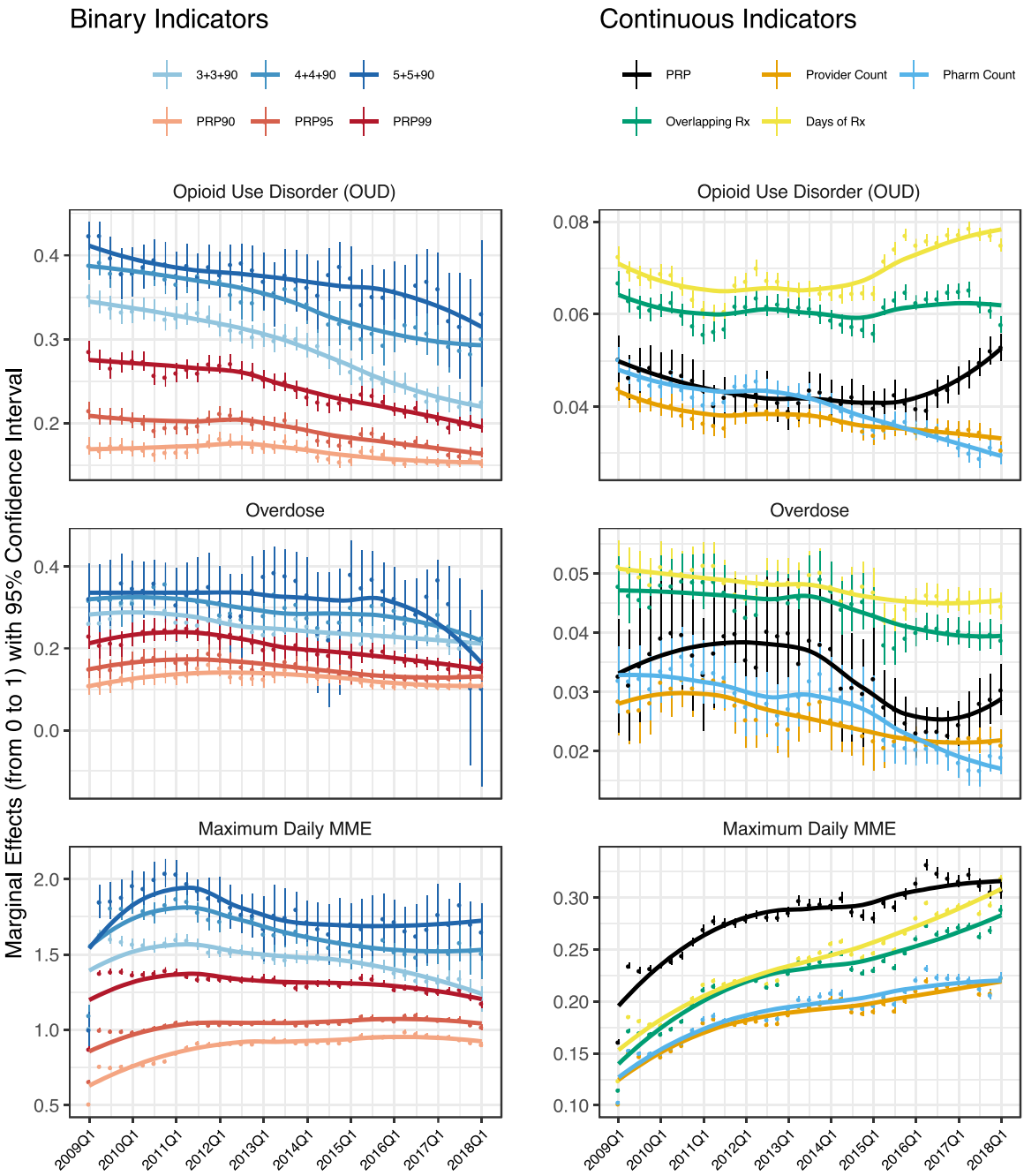
predictive accuracy over the period of observation, ranging from 84 to 86% for OUD, 73 to 76% for overdose and 73 to 93% for MME > 90. All other continuous indicators have similar ROC-AUC values, ranging from only 1 to 2% lower than overlapping Rx (for days of Rx) to 4 to 6% lower (for pharmacy count), with the other indicators in between.

Over time, the continuous indicators exhibit similar trends in predictive accuracy. For OUD these measures exhibit relatively stable accuracy, increasing very slightly until 2015. PageRank percentile shows the largest improvement in recent years, increasing in predictive accuracy from 81 to 84%. For overdose, continuous indicators remained stable until mid-2013, but declined substantially

from a peak of between 75 and 77% to a low of between 70 and 78% thereafter. Predictive accuracy for MME > 90 increased rapidly from between 76 and 78% in 2009 to between 85 and 89% in 2013, and then began to improve more slowly to a peak of between 87 and 93% in the final quarter of 2018. In short, our ability to accurately predict overdose using continuous PDS measures of any kind has declined. For OUD it has remained stable, and for high MME it has improved.

Comparison of average marginal effect size

Figure 2 presents AMEs, which can be interpreted as the increase in the predicted probability or predicted value of



Notes: To denote potential bias due to the change in ICD code specifications, all quarters of 2015 are shaded gray.

Figure 2 Average marginal effects for survey-weighted generalized linear models predicting opioid use disorder (OUD), overdose and maximum daily morphine milligram equivalents (MME) in the next quarter

an outcomes for each one-unit increase in a predictor. For all outcome variables, $N + N + 90$ indicators have higher average marginal effect sizes than binarized PageRank percentile (PRP) indicators throughout 2009–18. This is the inverse of the patterns observed for predictive accuracy, reflecting a trade-off between accuracy and effect size for these binary indicators. This can also be seen in confidence intervals, which are very large for higher threshold $N + N + 90$ indicators, especially after 2014, but much smaller for network-based PRP measures. The 5 + 5 + 90 indicator has the largest AMEs for most of the study period (between 0.30 and 0.43 for OUD, 0.10 and 0.27 for overdose and 1.0 and 2.1 for logged maximum daily MME). The lowest AMEs are observed for PageRank at the 90th percentile threshold (between 0.15 and 0.17 for OUD, 0.15 and 0.17 for overdose and 0.5 and 0.9 for logged maximum daily MME). For overdose, AMEs are more similar across indicators, especially in 2018, when confidence intervals for 5 + 5 + 90 and 4 + 4 + 90 overlap with those of all PRP indicators.

Figure 2 also illustrates that AMEs for $N + N + 90$ indicators decrease during the observation period, while binary PRP measures are more stable. For example, for OUD, the 3 + 3 + 90 indicator decreases from a high AME of 0.35 in 2009 to a low of 0.22 in 2018. Decreases are also observed in the AMEs for 5 + 5 + 90, for instance, for overdose from 0.37 in 2009 to 0.10 in 2018, and for MME from 2.1 in 2010 to 1.4 in 2018. In contrast, PRP90 is stable or increases in AME over time, and PRP99 shows smaller decreases, especially for OUD and overdose. Again, results suggest that $N + N + 90$ thresholds, in particular, are declining in significance.

The trade-off between accuracy and effect size does not apply to continuous indicators: those with higher predictive accuracy also have higher AME, in general. In predicting OUD and overdose, a 1 standard deviation increase in days of Rx has the highest AMEs, ranging from 0.06 to 0.08 and 0.04 to 0.05, respectively. In contrast, PageRank percentile results in the highest AMEs for maximum daily MME, ranging from 0.23 to 0.31. The lowest AMEs are observed for provider count (0.04 for OUD, 0.02 to 0.03 for overdose and 0.1 to 0.2 for MME).

With respect to trends over time, patterns differ across outcomes. For OUD, AMEs are relatively stable for all indicators until 2014, at which point AMEs increase for days of Rx (from 0.06 to 0.08) and PageRank percentile (from 0.04 to 0.06), but decline for other indicators. By the end of the study period, the AME gap between overlapping Rx and PRP has closed. Trends for overdose reflect slight decreases in AME for all indicators, although PRP is increasing between 2016 and 2018. Finally, AMEs for maximum daily MME increased sharply from 2009 to 2011 for all indicators, then continued to increase more slowly for the remainder of the observation period. In general, continuous indicators, and especially those based on prescription patterns and network ties between patients and prescribers, are increasing in effect size over the period from 2009 to 2018.

DISCUSSION

In the current study, we compare social network-based indicators of prescription drug-seeking behavior to traditional indicators during 11 years of the opioid epidemic. We consider indicator performance with respect to predictive accuracy (i.e. does the indicator produce few or many false-positives and false-negatives) and average marginal effect size (i.e. how much does risk for adverse outcomes change with a 1-unit increase in a predictor). This exercise is significant because, to date, there have been few attempts to compare measurement strategies for operationalizing PDS behavior in secondary data sets, and none have assessed the predictive value of PDS measures over time.

Review of main findings

Examining longitudinal trends reveals that threshold indicators based on the number of prescribers and pharmacies (i.e. $N + N + 90$) have declined in relevance during the past decade of the opioid epidemic, and currently perform no better than chance. While meeting these thresholds is associated with strong effects on OUD, overdose or high MME, they produce the 'wrong' prediction approximately half the time, and identify only between 0.01 and 0.003% of patients. This suggests that they produce very high rates of false negatives. A simple count of prescribers and pharmacies performs better than do thresholds, but worse than Rx and network-based measures. Overall, days of Rx is the best performing predictor of adverse outcomes, with the fewest prediction errors and the highest average marginal effects overall. However, the continuous PageRank indicator performs nearly as well or better (for MME) than continuous Rx indicators, and is increasing in performance across all outcomes after 2016.

Implications for behavioral dynamics across the opioid epidemic

Changes in the performance of different indicators over time may reflect key phases or shifts in the nature of the opioid epidemic. One contributing factor may be the implementation of PDMPs and similar policies that have limited the supply of prescription opioids by constraining provider behavior. Research suggests that the number of patients meeting stringent criteria for PDS has decreased with increased regulation, with no evidence of a decline in prescription drug use or related adverse outcomes such as overdose [1]. Our findings bear this out, with MME > 90

decreasing by 57% between 2009 and 2018, and fewer patients meeting $N + N + 90$ criteria during the course of the epidemic. During the same period, we find that meeting a high threshold for PDS behavior and simply counting the number of prescribers or pharmacies has become less predictive of a future OUD diagnosis or overdose event.

Similarly, trends regarding network-based and Rx indicators may reflect behavioral dynamics. That is, in response to increase supply-side regulation, patients and prescribers have probably changed their behavior to avoid outlier prescribing or physician visitation patterns. As prescribers at large have become less willing or less able to provide access to prescription opioids, it is plausible that patients have had to establish provider–patient relationships with those that adhere less stringently to new prescribing guidelines or do not access PDMP data [28]. Alternatively, patients might maintain provider–patient relationships with only those who are more liberal or sympathetic prescribers, foregoing visits to those who respond more conservatively to new regulation. Long-time pain patients with high opioid tolerance, in particular, may seek out or preserve patient–provider relationships that allow them to maintain their established medication regimens. Unlike traditional measures of multiple provider episodes, indicators of network prominence are sensitive to patterns of PDS behavior that reflect fewer but more targeted provider visits. Days of Rx and overlapping Rx have maintained their predictive performance, by and large, because obtaining consistently high quantities of MME is still associated with risk for OUD and overdose. It is how and from whom those prescriptions are obtained that has probably changed.

Another important implication of our findings relates to our declining ability to predict overdose with any measure of PDS. Concurrent with increasing regulation of opioid prescribing, there has been a rise in the supply of heroin during the second wave of the epidemic and synthetic opioids during the third wave [36]. Consequently, some individuals who engage in non-medical drug use may supplement with and/or convert to using illicit opioids or other controlled substances. As high-volume PDS and prescribing has become more constrained in the contemporary regulatory environment, and many have turned to illicit opioids, PDS indicators have become less predictive of opioid overdose.

Limitations and future directions

Our analysis has several notable limitations. Because we use secondary administrative data, social mechanisms underlying network structure are not directly observable. However, by using a matching strategy, we are able to rule out many alternative explanations for associations between network centrality and adverse outcomes. Also, because our data are derived from claims billed through commercial insurance, we are unable to observe cash transactions and Medicaid claims. As self-payment is a strategy for avoiding detection of drug-seeking behavior [29], our findings may underestimate the effects of PDS indicators. This research should be replicated using PDMP data. Finally, it is possible that different indicators of PDS perform more or less well in different states due to distinct regulatory environments. While we are unable to test this with our current data, where the most granular geographic unit is the census region, future research should examine this possibility.

CONCLUSION

In sum, these results have critical implications for research and policy. First, regarding measurement, researchers and state and federal regulators should consider avoiding $N + N + 90$ thresholds for identifying probable PDS behavior. Alternatively, the indicator of PDS that best optimizes practical utility (because it is easy to calculate) and predictive performance in the current regulatory environment is continuity of opioid prescriptions (days of Rx). For researchers with more advanced quantitative analysis skills, PageRank or other measures of co-prescription network centrality are a promising indicator of adverse outcomes that may continue to increase in relevance in coming years.

Secondly, these findings may be reflective of the shift in the opioid epidemic towards illicit drugs and black-market supply chains that has been identified in the literature. While it is relatively convenient and economical to use health-care administrative data sets to research and monitor drug-seeking behavior and the epidemiology of OUD and overdose, this strategy in isolation is doomed to fail. To reverse the opioid epidemic, future studies must focus on identifying social, economic and psychological determinants of drug-seeking behavior, identifying and addressing demand-side predictors while maintaining reasonable regulatory constraints on prescribing.

Declaration of interests

None.

Acknowledgements

This research was funded by the National Institute on Drug Abuse (R01 DA039928), an Indiana University Addictions Grand Challenge Grant, and a University of Kentucky Interdisciplinary Research Collaborations Seed Grant. We would like to thank the information technology team at the Indiana University Network Science Institute for their technical support.

Author Contributions

Brea Perry: Conceptualization; funding acquisition; methodology; project administration; resources; supervision. **Meltem Odabas:** Formal analysis; methodology; validation; visualization. **Kai-Cheng Yang:** Conceptualization; data curation; formal analysis; investigation; methodology; software; validation. **Byungkyu Lee:** Data curation; formal analysis; methodology; validation; visualization. **Patrick Kaminski:** Formal analysis; validation. **Brian Aronson:** Validation. **Yong Yeol Ahn:** Conceptualization; data curation; funding acquisition; project administration; resources; software; supervision. **Carrie Oser:** Conceptualization; funding acquisition. **Patricia Freeman:** Funding acquisition. **Jeffery Talbert:** Writing-review & editing-Equal

References

1. Lee B., Zhao W., Yang K. C., Ahn Y. Y., Perry B. L. Systematic evaluation of state policy interventions targeting the US opioid epidemic, 2007–2018. *JAMA Netw Open* 2021; **4**: e2036687–e2036687.
2. Goodhart C. A. Problems of monetary management: the UK experience. In: Goodhart C. A., editor. *Monetary Theory and Practice*. Palgrave, London; 1984, pp. 91–121.
3. Sansone R. A., Sansone L. A. Doctor shopping: a phenomenon of many themes. *Innov Clin Neurosci* 2012; **9**: 42.
4. Simeone R. Doctor shopping behavior and the diversion of prescription opioids. *Subst Abuse Res Treat* 2017; **11**; https://doi.org/10.1177/1178221817696077
5. Katz N., Panas L., Kim M., Audet A. D., 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.
6. Jena A. B., Goldman D., Weaver L., Karaca-Mandic P. Opioid prescribing by multiple providers in Medicare: retrospective observational study of insurance claims. *BMJ* 2014; **348**: g1393.
7. Hall A. J., Logan J. E., Toblin R. L., Kaplan J. A., Kraner J. C., Bixler D., *et al*. Patterns of abuse among unintentional pharmaceutical overdose fatalities. *JAMA* 2008; **300**: 2613–20.
8. Maxwell J. C. The prescription drug epidemic in the United States: a perfect storm. *Drug Alcohol Rev* 2011; **30**: 264–70.
9. Morris B. J., Zumsteg J. W., Archer K. R., Cash B., Mir H. R. Narcotic use and postoperative doctor shopping in the orthopaedic trauma population. *J Bone Joint Surg* 2014; **96**: 1257–62.
10. Peirce G. L., Smith M. J., Abate M. A., Halverson J. Doctor and pharmacy shopping for controlled substances. *Med Care* 2012; **50**: 494–500.
11. Hagihara A., Tarumi K., Odamaki M., Nobutomo K. A signal detection approach to patient–doctor communication and doctor-shopping behaviour among Japanese patients. *J Eval Clin Pract* 2005; **11**: 556–67.
12. Rouby F., Pradel V., Frauger E., Pauly V., Natali F., Reggio P., *et al*. Assessment of abuse of tianeptine from a reimbursement database using 'doctor-shopping' as an indicator. *Fundam Clin Pharmacol* 2012; **26**: 286–94.
13. Lineberry T. W., Bostwick J. M. Taking the physician out of 'physician shopping': a case series of clinical problems associated with Internet purchases of medication. *Mayo Clin Proc* 2004; **79**: 1031–4.
14. Shaffer E. G., Moss A. H. Physicians' perceptions of doctor shopping in West Virginia. *WV Med J* 2010; **106**: 10–4.
15. Wang M.-J., Lin S.-P. Study on doctor shopping behavior: insight from patients with upper respiratory tract infection in Taiwan. *Health Policy* 2010; **94**: 61–7.
16. Cepeda M. S., Fife D., Chow W., Mastrogiovanni G., Henderson S. C. Assessing opioid shopping behaviour. *Drug Saf* 2012; **35**: 325–34.
17. Pradel V., Delga C., Rouby F., Micallef J., Lapeyre-Mestre M. Assessment of abuse potential of benzodiazepines from a prescription database using 'doctor shopping' as an indicator. *CNS Drugs* 2010; **24**: 611–20.
18. Chang H.-Y., Murimi I. B., Jones C. M., Alexander G. C. Relationship between high-risk patients receiving prescription opioids and high-volume opioid prescribers. *Addiction* 2018; **113**: 677–86.
19. Walker A. M., Weatherby L. B., Cepeda M. S., Bradford D., Yuan Y. Possible opioid shopping and its correlates. *Clin J Pain* 2017; **33**: 976–82.
20. Young S. G., Hayes C. J., Aram J., Tait M. A. Doctor hopping and doctor shopping for prescription opioids associated with increased odds of high-risk use. *Pharmacoepidemiol Drug Saf* 2019; **28**: 1117–24.
21. Bohnert A. S., Valenstein M., Bair M. J., Ganoczy D., McCarthy J. F., Ilgen M. A., *et al*. Association between opioid prescribing patterns and opioid overdose-related deaths. *JAMA* 2011; **305**: 1315–21.
22. Buurma H., Bouvy M. L., De Smet P., Floor-Schreudering A., Leufkens H. G. M., Egberts A. C. G. Prevalence and determinants of pharmacy shopping behaviour. *J Clin Pharm Ther* 2008; **33**: 17–23.
23. Neutel C. I., Skurtveit S., Berg C., Sakshaug S. Multiple prescribers in older frequent opioid users–does it mean abuse? *J Popul Ther Clin Pharmacol* 2013; **20**: e397–e405.
24. Pradel V., Frauger E., Thirion X., Ronfle E., Lapierre V., Masut A., *et al*. Impact of a prescription monitoring program on doctor-shopping for high dosage buprenorphine. *Pharmacoepidemiol Drug Saf* 2009; **18**: 36–43.
25. Pradel V., Thirion X., Ronfle E., Masut A., Micallef J., Bégaud B. Assessment of doctor-shopping for high dosage buprenorphine maintenance treatment in a French region: development of a new method for prescription database. *Pharmacoepidemiol Drug Saf* 2004; **13**: 473–81.
26. Meara E., Horwitz J. R., Powell W., McClelland L., Zhou W., O'Malley A. J., *et al*. State legal restrictions and prescription-opioid use among disabled adults. *N Engl J Med* 2016; **375**: 44–53.
27. Buchmueller T. C., Carey C. The effect of prescription drug monitoring programs on opioid utilization in Medicare. *Am Econ J Econ Policy* 2018; **10**: 77–112.
28. Perry B. L., Yang K. C., Kaminski P., Odabas M., Park J., Martel M., *et al*. Co-prescription network reveals social dynamics of opioid doctor shopping. *PLOS ONE* 2019; **14**: e0223849.
29. Cepeda M. S., Fife D., Yuan Y., Mastrogiovanni G. Distance traveled and frequency of interstate opioid dispensing in opioid shoppers and nonshoppers. *J Pain* 2013; **14**: 1158–61.
30. Worley J., Thomas S. P. Women who doctor shop for prescription drugs. *West J Nurs Res* 2014; **36**: 456–74.
31. Ong M.-S., Olson K. L., Cami A., Liu C., Tian F., Selvam N., *et al*. Provider patient-sharing networks and multiple-provider prescribing of benzodiazepines. *J Gen Intern Med* 2016; **31**: 164–71.

32. Worley J. What prescribers can learn from doctor shoppers. *J Nurse Pract* 2014; **10**: 75–82.
33. Dowell D., Haegerich T. M., Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. *JAMA* 2016; **315**: 1624–45.41.
34. Page L., Brin S., Motwani R. & Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab, Stanford, CA; 1999. http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf. Accessed on 8 July 2019.
35. Hansen B. B. The prognostic analogue of the propensity score. *Biometrika* 2008; **95**: 481–8.
36. Roehler D. R., Hoots B. E., Olsen E. O., Kariisa M., Wilson N. O., Rudd R. A. *et al.* Annual Surveillance Report of Drug-Related Risks and Outcomes. Atlanta, GA: National Center for Injury Prevention and Control (NCIPC), Centers for Disease Control and Prevention (CDC), US Department of Health and Human Services; 2019.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1 Number of observations and patients included in the matching datasets for OUD, Overdose, and Maximum Daily MME.
Table S2 Descriptive Statistics for Continuous and Binary Variables based on OUD Matching Dataset.
Table S3 Descriptive Statistics for categorical variables based on OUD, Overdose and Maximum Daily MME Matching Datasets.
Table S4 Descriptive Statistics for Continuous and Binary Variables based on Overdose Matching Dataset.
Table S5 Descriptive Statistics for Continuous and Binary Variables based on Maximum Daily MME Matching Dataset.
Table S6 ICD Codes Used for Diagnosis Classification.
Fig S1 Clustered ROC-AUC results for binary classification models predicting OUD, overdose, and maximum daily MME in the next quarter (clustering by census division region).
Fig S2 Correlation Matrix for PDS Indicators at Quarter 2013Q1.
Fig S3 ROC-AUC results for binary classification models predicting OUD, overdose, and maximum daily MME in the next quarter (alternative specification of medication for OUD).
Fig S4 Average marginal effects for survey-weighted generalized linear models predicting OUD, overdose, and maximum daily MME in the next quarter (alternative specification of medication for OUD).
Fig S5 Trends of prescription drug seeking outcomes, 2009–2018.
Fig S6 Trends in prescription drug seeking indicators, 2009–2018.
Fig S7 Trends in prescribing behavior, 2009–2018.