American Pain Society
RESEARCH
EDUCATION
TREATMENT
ADVOCACY
PUBLISHED BY ELSEVIER

The Journal of Pain, Vol 16, No 4 (April), 2015: pp 380-387
Available online at www.jpain.org and www.sciencedirect.com

Automated Prediction of Risk for Problem Opioid Use in a Primary Care Setting

Timothy R. Hylan,* Michael Von Korff,† Kathleen Saunders,† Elizabeth Masters,‡ Roy E. Palmer,* David Carrell,† David Cronkite,† Jack Mardekian,§ and David Gross*

**North America Medical Affairs, Global Innovative Pharma, Pfizer Inc, New York, New York.*
†Group Health Research Institute, Seattle, Washington.
‡Outcomes & Evidence, Global Health & Value, Pfizer Inc, New York, New York.
§Statistics, Global Innovative Pharma, Pfizer Inc, New York, New York.

Abstract: **Identification of patients at increased risk for problem opioid use is recommended by chronic opioid therapy (COT) guidelines, but clinical assessment of risks often does not occur on a timely basis. This research assessed whether structured electronic health record (EHR) data could accurately predict subsequent problem opioid use. This research was conducted among 2,752 chronic noncancer pain patients initiating COT (≥70 days' supply of an opioid in a calendar quarter) during 2008 to 2010. Patients were followed through the end of 2012 or until disenrollment from the health plan, whichever was earlier. Baseline risk indicators were derived from structured EHR data for a 2-year period prior to COT initiation. Problem opioid use after COT initiation was assessed by reviewing clinician-documented problem opioid use in EHR clinical notes identified using natural language processing techniques followed by computer-assisted manual review of natural language processing–positive clinical notes. Multivariate analyses in learning and validation samples assessed prediction of subsequent problem opioid use. The area under the receiver operating characteristic curve (c-statistic) for problem opioid use was .739 (95% confidence interval = .688, .790) in the validation sample. A measure of problem opioid use derived from a simple weighted count of risk indicators was found to be comparably predictive of the natural language processing measure of problem opioid use, with 60% sensitivity and 72% specificity for a weighted count of ≥4 risk indicators.**

Perspective: ***An automated surveillance method utilizing baseline risk indicators from structured EHR data was moderately accurate in identifying COT patients who had subsequent problem opioid use.***

Key words: *Chronic opioid therapy, opioid surveillance, primary care, risk for problem opioid use.*

Received October 9, 2014; Revised January 22, 2015; Accepted January 23, 2015.

This study was part of a research collaboration between Group Health Research Institute and Pfizer Inc ("Electronic Research of Opioid Abuse Detection and Surveillance [e-ROADS]"), carried out with financial support from Pfizer. The research concepts and publication plans were agreed upon prior to commencing the study by the Joint Research Governance Committee of the Group Health-Pfizer Research Collaboration, composed of Group Health and Pfizer Inc employees.

Funding for work by Group Health Research Institute from Pfizer Inc included support for the development of this manuscript. M.V.K., K.S., D.Ca., and D.Cr. participated in this work as employees of Group Health Research Institute. T.R.H., E.M., R.E.P., J.M., and D.G. are employees and stockholders of Pfizer Inc. K.S. has stock in Merck.

M.V.K. and D.Ca. are principal investigators of funded and pending grants to Group Health Research Institute from Pfizer Inc for research on use of natural language processing to identify problem opioid use. They are also principal investigators for pending grants to Group Health Research Institute from a consortium of drug companies to conduct FDA-mandated research concerning risks for problem opioid use among patients receiving extended-release opioids.

Address reprint requests to Michael Von Korff, ScD, Group Health Research Institute, 1730 Minor Ave, Suite 1600, Seattle, WA 98101. E-mail: vonkorff.m@ghc.org

1526-5900/$36.00

http://dx.doi.org/10.1016/j.jpain.2015.01.011

Problem use of prescription opioids exacts an increasing clinical, economic, and societal burden within the United States.[1,12,34] Recent calls to action challenge clinicians and health care organizations to ensure patient safety while offering effective care for chronic pain.[8,11,18] Although clinical guidelines recommend a "Universal Precautions" approach to opioid prescribing that includes timely assessment of risks for problem opioid use, available evidence indicates that timely risk assessment usually does not occur.[3,9-11,14,15,20,22,25,26,31] Rigorous longitudinal research evaluating the prediction of subsequent problem opioid use is also lacking.[4]

Accurate identification of patients most likely at risk for problem opioid use is a prerequisite for widely recommended approaches to mitigating these risks. Multiple patient- and clinician-based screeners exist to assess potential problem opioid use risks.[4,13,16,19,29] However,

there are deficiencies in evaluation of the accuracy of these screeners in predicting risks for subsequent onset of problem opioid use, so there is no consensus about the value of screening or which screeners are most useful.[16,19,29] These screeners assess patient risk factors including substance use disorder history, family history of substance use disorder, and significant psychological problems. However, it has proven difficult to implement routine risk assessment in community practice for several reasons. Busy clinicians working in time-constrained care settings may not administer screeners because of competing demands and clinical priorities among often complex chronic pain patients. And it is often unclear when patients transition from short-term to long-term use of opioids, so formal evaluation visits to assess the appropriateness of chronic opioid therapy (COT) do not reliably occur.[14]

Automated assessment utilizing structured electronic health records (EHRs) data could address several of these difficulties, thereby facilitating more timely and consistent opioid risk assessment by clinicians and health care systems. Recently, predictive models of opioid abuse have used insurance claims data to assess the likelihood of problem opioid use, with initially promising results.[6,21,27] However, these studies relied solely upon recorded International Classification of Diseases, Ninth Revision (ICD-9), diagnostic codes to ascertain opioid abuse and were of short duration (≤6 months).

The increasing availability of data from patients' EHRs including clinicians' notes offers new opportunities to develop predictive models for assessing potential risks for problem opioid use. In this research, we generated and evaluated a measure of problem opioid use based on natural language processing (NLP) techniques utilizing information contained within clinicians' notes in the EHR pertaining to COT patients.

The aim of this article is to report on a predictive model developed to assess the likelihood of problem opioid use over a 2- to 5-year period following initiation of COT within a large health plan. We assessed whether readily accessible baseline risk indicators obtained from structured EHR data at initiation of long-term opioid therapy were able to accurately predict subsequent problem opioid use.

Methods

Setting and Data

This study was conducted at Group Health Cooperative (GHC), a large, mixed-model health plan established in 1945 that now serves approximately 600,000 patients in urban, suburban, and rural areas of Washington State. Group Health is composed of both integrated group practice (IGP; salaried physicians working in GHC clinics) and network (fee-for-service physicians working in community settings) segments. This research is limited to members in the IGP because the study relies on clinical electronic medical records data that are available only in the IGP. The GHC uses the Epic EHR system (http://www.epic.com/; Epic, Verona, WI) to document care delivery, including primary, specialist, and emergency care encounters and hospital discharge summaries. The GHC Epic EHR system is the digital chart used for patient care within the IGP, having replaced the paper chart in 2005. The GHC maintains structured data, which are data elements entered into the GHC Epic system by health care staff using predefined options (eg, diagnostic codes), including information on most aspects of care delivery such as patient demographics, diagnostic codes, procedures, medications, and laboratory results. In addition, the GHC Epic EHR system includes free text clinical notes entered by health care staff during patient encounters and other contacts with patients. Approval for this study was granted by the Group Health Human Subjects institutional review board.

The sample consisted of GHC members aged 18 years or older who initiated COT for noncancer pain between 2008 and 2010. COT was defined as receipt of ≥70 days' supply of transdermal or oral opioids (except buprenorphine) in a calendar quarter, which corresponded to >75% of the days in the quarter covered by an opioid prescription. We employed this definition because it corresponded to the operational definition of COT employed by GHC for a COT risk reduction initiative implemented in 2010. Because analyses examined opioid use over multiple years, we counted the number of quarters that study patients met this criterion for receiving COT. To ensure sustained use of opioids and a sufficient COT duration to permit clinician identification of problem opioid use, we required that study patients receive at least 2 quarters of COT within a 1-year period. We identified the first calendar quarter between 2008 and 2010 in which a subject received COT (index quarter). If the subject received at least 1 more quarter of COT in the 3 quarters following the index quarter, she or he was provisionally eligible for the study. To restrict the sample to those initiating COT between 2008 and 2010, we excluded subjects who received COT in any quarter in 2006 and 2007. Further, subjects were required to be enrolled in the health plan at least 6 of the 8 calendar quarters before the index quarter and 6 of the 8 quarters following the index quarter (inclusive) so each subject had a 2-year "pre-period" and at least a 2-year follow-up period. These enrollment requirements ensured the availability of EHR data to capture utilization prior to the initiation of COT use and the potential availability of clinical notes data to document problem opioid use after COT initiation. To restrict the study to patients receiving COT for noncancer pain, patients were excluded if they had ≥2 visits with cancer diagnoses (excluding nonmelanoma skin cancer) during any calendar 1-year period between 2006 and 2012 or had received an opioid prescription from an oncologist or were admitted to hospice during the study period. Based on a random assignment performed on the entire pool of subjects potentially eligible for the study (eg, before application of study inclusion/exclusion criteria), patients were grouped into "learning" and "validation" samples.

NLP Measure of Problem Opioid Use

To construct a measure of clinician-documented problem opioid use, we began with a definition of opioid misuse or abuse using the National Institute of Drug Abuse definition: "Prescription drug abuse is the use of a medication without a prescription, in a way other than as prescribed, or for the experience or feelings elicited."[17] We defined clinician-documented problem opioid use as a practitioner statement in the clinical record of a diagnosis or assessment that the subject is overusing, is misusing, is abusing, or has become addicted to prescription opioids. Relevant documentation was in the form of an explicit description in the narrative of a clinical note or a coded entry into the subject's active problem list. In this prospective study, problem opioid use had to be documented during the individual's follow-up period. A person's follow-up period extended from the index quarter through December 31, 2012, or his or her disenrollment from the health plan, whichever was earlier.

Because of the large number of clinical notes available for each subject (on average, 355 clinical notes per patient), we used NLP techniques to identify subjects who had text in the clinician's notes that was potentially indicative of problem opioid use. NLP uses computing techniques to automate identification of meaningful content in unstructured text, facilitating discovery of such content within large numbers of documents. NLP methods can address linguistic variation (different words with the same meaning), polysemy (single words with several meanings), negation (eg, "reports no pain" vs "reports severe pain"), ambiguity, and temporality. NLP methods can be particularly useful in searching the free text clinical notes of the EHR.

We performed 3 phases of NLP searching through all the relevant clinical notes for each subject (2008 through 2012). First, the algorithm identified subjects who had clinical notes that indicated opioid overuse, misuse, abuse, or addiction. Second, the algorithm identified subjects who had clinical notes that indicated opioid "dependence," which sometimes indicated only physiologic dependence on opioids. Third, we identified any additional subjects who had an ICD-9 code for opioid abuse or dependence [304.0, .00, .01, .02; 304.7, .70, .71, .72; 305.5, .50, .51, .52] in structured EHR data who were not identified in the prior 2 phases as having clinician-documented opioid overuse, misuse, or abuse. Complete details of the development of computer-assisted NLP techniques applied in this study are described elsewhere.[5]

Two pairs (M.V.K. and a research specialist, and M.V.K. and K.S.) of Group Health Research Institute chart reviewers then used a customized computer program that highlighted relevant text in NLP-identified clinical notes to conduct a manual validation of all subjects in the study sample for whom the NLP algorithm indicated positive for problem opioid use (NLP-identified positive), reviewing up to 4 NLP-identified clinical notes per subject. Before the manual validation was begun, interrater reliability between the pairs of reviewers was tested and was found to be high (kappa = .86, 97% agreement for the first pair of raters [M.V.K. and a research specialist] on a random sample of 50 NLP-identified positive subjects; kappa = .71, 88% agreement for the second pair of raters [M.V.K. and K.S.] on a random sample of 137 NLP-identified positive subjects).

Baseline Risk Indicators

A set of baseline risk indicators was constructed from the structured data in the EHR to include in the multivariate analyses used to evaluate each measure of problem opioid use. These baseline risk indicators were taken from the GHC Epic EHR system and developed from structured data (not the clinical notes) entered by the GHC health care staff during the patient's visits occurring in the 2-year "pre-period" and are likely to be correlated with problem opioid use based on earlier literature.[6]

Baseline risk indicators included in the model were as follows: age (18–44, 45–64, and ≥65 [as of the index quarter]), gender (female/male), ethnicity (white, nonwhite, unknown), smoking status (equal to 1 as a current smoker if indicated at any time during the study period, 0 otherwise), ICD-9 diagnosis of opioid abuse/dependence (equal to 1 if indicated in any quarter during the 2-year "pre-period" and 0 otherwise; this variable was obtained from the structured EHR data available on the patient and is different from the ICD-9 code "mention" in the clinician's notes used for the NLP-derived measure of problem opioid use), ICD-9 diagnosis of nonopioid prescription drug abuse/dependence (equal to 1 if indicated in any quarter during the 2-year "pre-period" and 0 otherwise), ICD-9 diagnosis of a mental disorder (equal to 1 if indicated in any quarter during the 2-year "pre-period" and 0 otherwise; mental disorders included depression, bipolar disorder, anxiety, autism, schizophrenia, dementia, other psychosis, and self-inflicted injury [including "possible"]), ICD-9 diagnosis of alcohol abuse/dependence (equal to 1 if indicated in any quarter during the 2-year "pre-period" and 0 otherwise), and ICD-9 diagnosis of hepatitis C (equal to 1 if indicated in any quarter during the 2-year "pre-period" and 0 otherwise).

Weighted Count of Risk Indicators

We also developed a simple, weighted count of risk indicators to determine if it predicted NLP-ascertained problem opioid use. We did this because the use of a logistic regression model to predict the likelihood of future problem opioid use would be difficult in many community practice settings. The weighted count of risk indicators could be readily implemented in any setting able to determine the status of patients on each of the risk indicators from medical records data.

To construct the weighted count of risk indicators, we preidentified relevant risk indicators, including age, smoking status, and ICD-9 codes in the structured data (not the clinical notes) that appeared in the 2-year period prior to COT initiation. These diagnoses included opioid abuse/dependence, nonopioid drug abuse/dependence,

alcohol abuse/dependence, mental disorders, and hepatitis C. For each risk indicator, we estimated a separate logistic regression model, with each predictor entered individually as the independent variable and problem opioid use as the dependent variable. We assigned weights to each of the risk indicators based on the odds ratio (OR) from the logistic regression models according to the following criteria: for $1 < OR \leq 2.5$, weight = 1; for $2.5 < OR \leq 5$, weight = 2; and for $OR > 5$, weight = 3. The following weights were employed: for age 18 to 44, weight = 3; for age 45 to 64, weight = 1; for prior opioid abuse/dependence diagnosis, weight = 3; for prior nonopioid drug abuse/dependence, weight = 3; for prior alcohol abuse/dependence diagnosis, weight = 1; for prior mental disorder diagnosis, weight = 2; for current smoking, weight = 2; and for prior hepatitis C diagnosis, weight = 2. We then summed these weights for each subject. The weighted count of risk indicators was tested in the learning, validation, and combined samples. This weighted count was also used to assess whether a prediction model estimated in the learning sample achieved comparable prediction in the validation sample.

Multivariate Analyses

We performed logistic regression analysis on the learning, validation, and combined samples to evaluate the ability to predict subsequent problem opioid use using the NLP classification of problem opioid use. We also assessed the ability of the weighted count of risk indicators to predict the NLP classification of problem opioid use utilizing logistic regression. Receiver operating characteristic (ROC) curves were generated, and corresponding c-statistics and 95% confidence intervals (CIs) were calculated for each logistic regression. The c-statistic measures the discriminatory ability of the model, with a greater area under the curve representing better prediction, where .50 indicates prediction no better than chance and 1.00 indicates perfect prediction. We also estimated the weighted count measure model's sensitivity and specificity and the positive predictive value (PPV, defined as the number of patients predicted positive for problem opioid use by the weighted count measure who were also positive on the NLP classification of problem opioid use, divided by the total number of patients) and negative predictive value (NPV, defined as the number of patients predicted to not have problem opioid use by both the weighted count measure and NLP classification, divided by the total number of patients).

Software

We developed a regular expression-based NLP system in the Python programming language, version 2.7 (Python Software Foundation; http:www.python.org), modeled after the named entity recognition in the widely used Apache Clinical Text Analysis and Knowledge Extraction System (cTAKES; The Apache Software Foundation; http:www.apache.org).[24] The system uses a list of phrases to identify mentions of problem opioid use in clinical notes, accounting for negated, uncertain, hypothetical, and historical language qualifiers, references to persons other than the patient, and spelling errors. Additionally, the Python GUI library wxPython 2.8 was used to facilitate the manual validation of clinician-documented problem opioid use with highlighting of key terms. All analytic file construction,

Table 1. Sample Characteristics (N = 2,752)

Variable	% or Mean (SD)	N
Age*		
18–44	21.2	583
45–64	49.9	1,373
≥65	28.9	796
Gender		
Male	37.7	1,039
Female	62.3	1,713
Ethnicity		
Non-Hispanic white	77.8	2,141
Nonwhite	13.3	365
Unknown	8.9	246
Current smoker		
Yes	33.3	916
No	33.7	1,836
ICD-9 diagnosis of opioid abuse/dependence†		
Yes	.9	25
No	99.1	2,727
ICD-9 diagnosis of drug abuse/dependence (excluding opioids)†		
Yes	2.1	58
No	97.9	2,694
ICD-9 diagnosis of alcohol abuse/dependence†		
Yes	4.9	135
No	95.1	2,617
ICD-9 diagnosis of mental health disorder†		
Yes	43.2	1,189
No	56.8	1,563
ICD-9 diagnosis of hepatitis C†		
Yes	4.8	132
No	95.2	2,620
Mean no. of quarters received ≥30 days' supply‡	12.2 (5.0)	—
Mean no. of quarters received ≥70 days' supply‡	8.8 (4.8)	—
Mean (median) number of quarters in follow-up period	13.4 (13)	—
Proportion (number) clinician-documented problem opioid use§		
Yes	5.7	158
No	94.3	2,594
Weighted count of risk indicators		
0	15.1	415
1	18.1	497
2	10.9	299
3	25.8	709
4	3.2	89
5	17.3	476
6	1.5	40
7	5.9	161
≥8	2.4	66
Mean (median) weighted count of risk indicators	2.90 (3.00)	—
Range of weighted count of risk indicators	0–16	—

*Assessed at start of study period.
†Assessed in the 2-year pre-period prior to the index quarter.
‡Assessed between January 1, 2008, and the end of follow-up period.
§Assessed between the index quarter and the end of follow-up period.

descriptive statistics, and multivariate analyses (including statistical calculations) were performed using SAS, version 9.3 (SAS Institute, Cary, NC).

Results

Overall, the study had 24,716 potential subjects. Of these, 10.4% (n = 2,574) had 1 of the cancer exclusions, and 43% (n = 9,475) of the remaining 22,142 subjects were excluded from this analysis because they had received at least 1 quarter of COT in 2006 or 2007. Of the remaining 12,667 subjects, 57.9% (n = 7,339) were excluded because they did not meet our definition of COT between 2008 and 2010, and an additional 2,576 were excluded because of failure to satisfy the pre- and/or post-index enrollment criteria, resulting in a total of 2,752 COT subjects who were eligible for this study. The original pool of 24,716 subjects had been assigned at random to a learning sample or to a validation sample. Among the 2,752 subjects eligible for the current analysis, 1,449 had been randomly assigned to the learning sample and 1,303 to the validation sample.

Approximately half of the 2,752 study subjects were 45 to 64 years of age (see Table 1). Close to two-thirds were women, whereas one-third were smokers. Of note, less than 1% had a structured ICD-9 diagnosis of opioid abuse/dependence recorded in structured EHR data in the 2 years prior to COT initiation. Among the 2,752 COT initiators, 158 (5.7%) were identified as positive for subsequent problem opioid use over the follow-up period based on the text contained within the clinicians' notes portion of the EHR, using the NLP classification and computer-assisted manual validation. The median weighted count of risk indicators was equal to 3.

Table 2 reports the multivariate model results for the NLP measure of problem opioid use in the learning (n = 1,449), validation (n = 1,303), and combined (N = 2,752) samples. In the combined sample, variables indicating age 18 to 44, opioid abuse diagnosis, positive smoking status, and mental disorder diagnoses during the 2-year prior period were significant predictors. It should be noted that probably because of small numbers with the risk indicator, the results for opioid abuse diagnosis were not significant in the learning sample. ROC analysis yielded an area under the curve (c-statistic) of .754 (CI = .689, .819), .739 (CI = .688, .790), and .745 (CI = .705, .785) for the learning, validation, and combined samples, respectively.

We then assessed the ability of the weighted count of risk indicators to predict the NLP classification of problem opioid use. Fig 1 reports the ROC curves of this model in both the learning and validation samples. The ROC area under the curve (c-statistic) was .733 (CI = .666, .800), .718 (CI = .668, .768), and .724 (CI = .684, .764) for the learning, validation, and combined samples, respectively.

We also calculated the overall sensitivity and specificity of the weighted count of risk indicators model by applying the estimates from the logistic regression model that predicted problem opioid use in the learning sample to the validation sample. The sensitivity of the regression model prediction of problem opioid use was 58.3%, and specificity was 71.2% at a predicted probability cutoff value equal to .046. This predicted probability corresponded to a count of ≥4 risk indicators.

We also estimated sensitivity, specificity, PPV, and NPV at varying cut points for the weighted count of risk indicators. These estimates are reported in Table 3 for the

Table 2. Multivariate Logistic Regression of NLP Measure of Problem Opioid Use

Baseline Variable	*Learning Sample (n = 1,449)*	*Validation Sample (n = 1,303)*	*Combined Sample (N = 2,752)*
Age			
18–44	4.58 (1.99, 10.53)	4.41 (2.22, 8.74)	4.54 (2.68, 7.69)
45–64	1.20 (.52, 2.80)	1.42 (.72, 2.82)	1.33 (.78, 2.26)
≥65 (referent)	1.0	1.0	1.0
Smoking status			
Current	1.41 (.81, 2.46)	1.57 (.99, 2.49)	1.47 (1.03, 2.08)
Past/never (referent)	1.0	1.0	
ICD-9 diagnosis of opioid abuse/dependence			
Yes	2.90 (.59, 14.16)	16.33 (4.12, 64.73)	7.44 (2.79, 19.80)
No (referent)	1.0	1.0	1.0
ICD-9 diagnosis of drug abuse/dependence (excluding opioids)			
Yes	2.12 (.69, 6.55)	1.75 (.55, 5.58)	1.63 (.73, 3.65)
No (referent)	1.0	1.0	1.0
ICD-9 diagnosis of alcohol abuse/dependence			
Yes	.87 (.30, 2.53)	1.32 (.53, 3.25)	1.02 (.51, 2.03)
No (referent)	1.0	1.0	1.0
ICD-9 diagnosis of mental health disorder			
Yes	1.84 (1.05, 3.21)	1.61 (1.02, 2.53)	1.73 (1.22, 2.45)
No (referent)	1.0	1.0	1.0
ICD-9 diagnosis of hepatitis C			
Yes	2.55 (1.00, 6.45)	1.10 (.42, 2.92)	1.56 (.80, 3.05)
No (referent)	1.0	1.0	1.0

NOTE. Values are OR (95% CI).

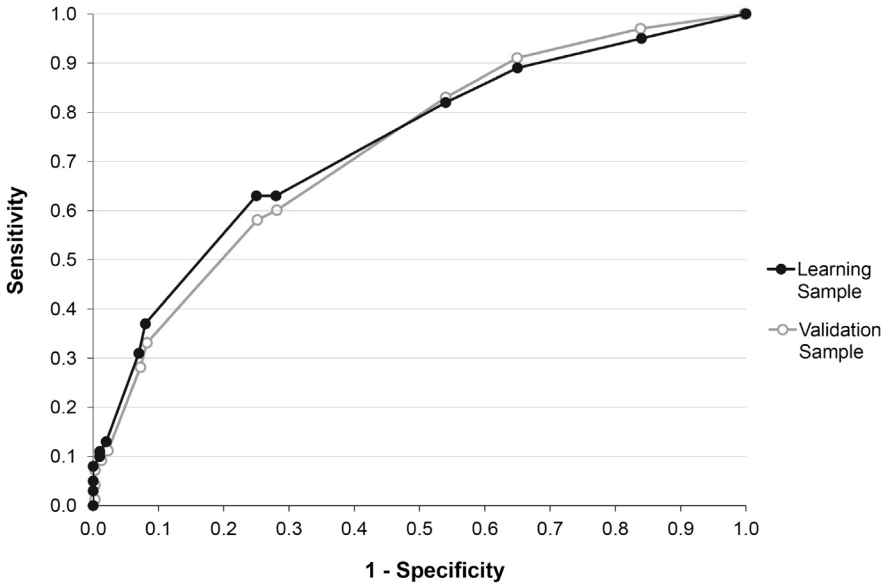

Figure 1. ROC curves for learning (n = 1,449) and validation (n = 1,303) samples for predicting problem opioid use using the weighted count measure.

combined study sample. A weighted count of risk indicators of ≥3 provided good sensitivity (82.8%) but poor specificity (45.6%). A weighted count of risk indicators of ≥4 provided moderate predictive ability, with fair specificity (71.6%) balanced with lower sensitivity (60.1%). Table 4 provides more detail on how sensitivity, specificity, PPV, and NPV were calculated for a weighted count of risk indicators of ≥4. Using a cutoff of ≥4, the PPV was 11.4% (95/832), and the NPV was 96.7% (1,857/1,920).

Discussion

Our results suggest the potential utility of automated surveillance to assess risks of problem opioid use utilizing EHR data as predictors. Performance of both the NLP measure model and the weighted count of risk indicators was similar to that of widely used risk screeners in other disease states (eg, cardiovascular disease predictive models have reported area under the curve c-statistics in the range of .68–.77).[35] The observed sensitivities and specificities using the weighted count of risk indicators reflect only fair prediction of the risk for subsequent problem opioid use as assessed by the NLP measure.

Table 3. Prediction of Clinician-Documented Problem Opioid Use at Various Cut Points (N = 2,752)

Weighted Count of Risk Indicators Cutoff	PPV	NPV	Sensitivity	Specificity
≥1	6.6	98.8	96.8	15.8
≥2	7.8	98.4	90.5	34.6
≥3	8.5	97.8	82.9	45.6
≥4	11.4	96.7	60.1	71.6
≥6	19.5	95.7	32.9	91.7
≥8	27.3	94.8	11.4	98.2

NOTE. Values are percentages.

If we view the NLP classification as a "gold standard" for problem opioid use, the 60% sensitivity means that about 4 of 10 patients who have subsequent problem opioid use would be predicted as negative for subsequent problem opioid use when applying the NLP classification approach. The 72% specificity means that 28% of patients who are negative for subsequent problem opioid use in clinical notes would be predicted as positive when applying the NLP classification approach. Although the performance of the NLP classification approach is similar to that of tests widely used in clinical practice, the attenuated sensitivities and specificities mean that initial screening cannot accurately predict who will and who will not subsequently develop problem opioid use.[13,16]

These results suggest that it may be feasible to develop simple risk stratification algorithms to initially alert clinicians to potential patients at higher risk for problem opioid use, with the caveat that such algorithms may have far from optimal sensitivity and specificity for predicting who will develop subsequent problem opioid use. Because a variety of factors often prevent timely

Table 4. Agreement of Classification of Problem Opioid Use: NLP-Based Classification and Weighted Count Risk Screener

	NLP-Based Classification of Problem Opioid Use		
Weighted Count Risk Screener of Problem Opioid Use*	Absent	Present	Total
0–3 risk indicators	1,857	63	1,920
≥4 risk indicators	737	95	832
Total	2,594	158	2,752

NOTE. Sensitivity: 95/158 = 60.1%; specificity: 1,857/2,594 = 71.6%; PPV: 95/832 = 11.4%; NPV: 1,857/1,920 = 96.7%.
*With classification of problem opioid use based on the weighted count measure with a cut point of ≥4.

assessment of risk factors for problem opioid use among patients initiating COT in community practice settings, automated screening may help to initially identify some patients at elevated risk for subsequent problem opioid use for whom further evaluation may be warranted.[14] Although the risk indicators from EHR data tested in this research are similar to those included in patient-administered risk assessment tools, it is not possible to determine the performance of the automated tool evaluated here relative to the patient-administered scales tested in prior research.[4,29] It remains unclear whether it is possible to identify truly low-risk patients among those considering long-term use of opioids.

Limitations of our methods include the use of data from only 1 health plan. Previous research on GHC COT patients has generalized to other settings[7,23,32,33]; however, future research should replicate the prediction methods developed in this paper in other settings. With the growing use of EHR systems in diverse community practice settings, replication of this research in other clinical settings is possible. GHC instituted additional care processes for the management of prescription opioid use in its population toward the latter part of 2010.[28] Whether these changes influenced how clinicians observed and documented problem opioid use in the EHR after 2010 is not known at this time. We did not include information from urine drug screens in the analysis because this information was not systematically collected for a large percentage of COT patients within the GHC system until after 2010.

Our NLP classification of problem opioid use included instances of prescription opioid addiction, abuse, misuse, diversion, and/or overuse of opioids. At this time, it is not known whether reliable differentiation of different types of problem opioid use (eg, addiction vs overuse) is possible. This limitation does not diminish the potential value of improving problem opioid use risk assessment, regardless of type, in community practice.

For practical reasons, notably the large number of clinical notes per patient, we were not able to manually validate the absence of problem opioid use in the entire electronic chart of study subjects. It is possible that clinical notes not manually reviewed contained information pertaining to problem opioid use. To the extent that problem opioid use was unrecognized and/or undocumented by clinicians in this setting, ascertainment of problem opioid use was incomplete. Previous literature has also shown that problem opioid use is often underrecognized and underdiagnosed.[14] As such, our estimates of the risk for developing problem opioid use likely underestimate risks for problem opioid use in clinical practice.

In terms of strengths, this study is one of the few longitudinal analyses of problem opioid use risk assessment, with by far the longest duration of follow-up. We assessed study subjects for 2 years prior to initiation of chronic opioid use and up to 5 years thereafter. Second, our results complement those of earlier predictive models that used only ICD-9 diagnoses from medical claims data to identify problem opioid use.[6,21,27] Third, our method offers busy, time- and resource-constrained health plans an opportunity to more efficiently implement automated surveillance of risk indicators for problem opioid use within their patient populations. The weighted count of risk indicators may be more readily implemented in community practice settings than use of a regression-based prediction model. This can be done by applying the weights we generated from our analysis to the values of the respective set of baseline risk indicators for each patient to generate a weighted count measure for each patient. Our results suggest that it may be possible to improve the efficiency of developing information relevant to assessing risk for problem opioid use among patients initiating COT.

Clinical guidelines for chronic pain recommend that opioids be considered only after an adequate trial of nonopioid options, yet accurately assessing risk for problem opioid use remains a challenge.[2,10,30] Steps to improve timely identification of risk indicators for problem opioid use could play a role in reducing adverse selection of higher risk patients into COT; however, the effectiveness of recommended risk mitigation strategies has not been determined. Because predictive models for problem opioid use are only moderately accurate, at best, there is always a need for the clinician's vigilance to ensure safe and appropriate opioid use for long-term management of chronic musculoskeletal pain.

Acknowledgment

The authors would like to recognize the contributions of Patty Yarbro and Megan Addis and research specialist Chrystel Kratohvil of Group Health Research Institute and Annmarie Gillespie of Pfizer Inc for administrative support and of Sean Donevan of Pfizer Inc for medical and scientific input during the project.

References

1. 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 12:657-667, 2011

2. Brown J, Setnik B, Lee K, Wase L, Roland CL, Cleveland JM, Siegal S, Katz N: Assessment, stratification, and monitoring of the risk for prescription opioid misuse and abuse in the primary care setting. J Opioid Manag 7:467-483, 2011

3. Chou R, Fanciullo GJ, Fine PG, Adler JA, Ballantyne JC, Davies P, Donovan MI, Fishbain DA, Foley KM, Fudin J, Gilson AM, Kelter A, Mauskop A, O'Connor PG, Passik SD, Pasternak GW, Portenoy RK, Rich BA, Roberts RG, Todd KH, Miaskowski C, American Pain Society-American Academy of Pain Medicine Opioids Guidelines Panel. Opioid Treatment Guidelines: Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. J Pain 10:113-130, 2009

4. Chou R, Fanciullo GJ, Fine PG, Miaskowski C, Passik SD, Portenoy RK: Opioids for chronic noncancer pain: Prediction and identification of aberrant drug-related behaviors: A

review of the evidence for an American Pain Society and American Academy of Pain Medicine clinical practice guideline. J Pain 10:131-146, 2009

5. Cronkite D: Estimating Prevalence of Clinician-Labeled Opioid Abuse: Structured vs. Unstructured Clinical Data. HMO Research Network Annual Meeting Presentation; 2014. Phoenix, AZ

6. Dufour R, Joshi AV, Pasquale MK, Schaaf D, Mardekian J, Andrews GA, Patel N: The prevalence of diagnosed opioid abuse in commercial and Medicare managed care populations. Pain Pract 14:E106-E115, 2014

7. Dunn KM, Saunders KW, Rutter CM, Banta-Green CJ, Merrill JO, Sullivan MD, Weisner CM, Silverberg MJ, Campbell CI, Psaty B, Von Korff M: Opioid prescriptions for chronic pain and overdose: A cohort study. Ann Intern Med 152:85-92, 2010

8. Frakt A: Painkiller abuse, a cyclical challenge. New York Times. Available at: http://www.nytimes.com/2014/12/23/upshot/painkiller-abuse-a-cyclical-challenge.html?_r=0&abt=0002&abg=0. Accessed December 29, 2014

9. Gourlay DL, Heit HA, Almahrezi A: Universal precautions in pain medicine: A rational approach to the treatment of chronic pain. Pain Med 6:107-112, 2005

10. Washington State Agency Medical Directors' Group: Interagency Guideline on Opioid Dosing for Chronic Noncancer Pain: An Educational Aid to Improve Care and Safety With Opioid Treatment. Olympia (WA), Washington State Department of Labor and Industries, 2010. Available at: http://www.agencymeddirectors.wa.gov/Files/OpioidGdline.pdf. Accessed December 28, 2014

11. FDA Statement. FDA Commissioner Margaret A Hamburg Statement on Prescription Opioid Abuse. Available at: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm391590.htm. Accessed April 28, 2014

12. Institute of Medicine of the National Academies: Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. Washington, DC, National Academies Press, 2011

13. Jones T, Moore T, Levy JL, Daffron S, Browder JH, Allen L, Passik SD: A comparison of various risk screening methods in predicting discharge from opioid treatment. Clin J Pain 28:93-100, 2012

14. Krebs EE, Ramsey D, Miloshoff JM, Bair MJ: Primary care monitoring of long-term opioid therapy among veterans with chronic pain. Pain Med 12:740-746, 2011

15. Meier B: FDA urging a tighter rein on painkillers. Available at: http://www.nytimes.com/2013/10/25/business/fda-seeks-tighter-control-on-prescriptions-for-class-of-painkillers.html. Accessed March 30, 2014

16. Moore TM, Jones T, Browder JH, Daffron S, Passik SD: A comparison of common screening methods for predicting aberrant drug-related behavior among patients receiving opioids for chronic pain management. Pain Med 10:1426-1433, 2009

17. National Institute of Drug Abuse: Prescription Drug Abuse. Available at: http://www.drugabuse.gov/publications/research-reports/prescription-drugs/what-prescription-drug-abuse. Accessed March 31, 2014

18. Office of National Drug Control Policy: Epidemic: Responding to America's Prescription Drug Abuse Crisis [updated 2011 Apr 11; cited 2011 Jun 10]. Available at: http://www.whitehousedrugpolicy.gov/publications/pdf/rx_abuse_plan.pdf. Accessed December 28, 2014

19. Passik SD, Kirsh KL, Casper D: Addiction-related assessment tools and pain management: Instruments for screening, treatment planning, and monitoring compliance. Pain Med 9:S145-S166, 2008

20. Prescription Drug Monitoring Program (PDMP) Center of Excellence, Brandeis University. National Rx Drug Abuse Summit Meeting, 22 April 2014

21. Rice JB, White AG, Birnbaum HG, Schiller M, Brown DA, Roland CL: A model to identify patients at risk for prescription opioid abuse, dependence, and misuse. Pain Med 13:1162-1173, 2012

22. Rolfs RT, Johnson E, Williams NJ, Sundwall DN: Utah clinical guidelines on prescribing opioids for treatment of pain. J Pain Pall Care Pharm 24:219-235, 2010

23. Saunders KW, Von Korff M, Campbell CI, Banta-Green CJ, Sullivan MD, Merrill JO, Weisner C: Concurrent use of alcohol and sedatives among persons prescribed chronic opioid therapy: Prevalence and risk factors. J Pain 13:266-275, 2012

24. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, Chute CG: Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. J Am Med Inform Assoc 17:507-513, 2010

25. Sehgal N, Manchikanti L, Smith HS: Prescription opioid abuse in chronic pain: A review of opioid abuse predictors and strategies to curb opioid abuse. Pain Physician 15(Suppl 3):ES67-ES92, 2012

26. Starrels JL, Becker WC, Alford DP, Kapoor A, Williams AR, Turner B: Systematic review: Treatment agreements and urine drug testing to reduce opioid misuse in patients with chronic pain. Ann Intern Med 152:712-720, 2010

27. Sullivan MD, Edlund MJ, Fan MY, 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 50:332-339, 2010

28. Trescott CE, Beck RM, Seelig MD, Von Korff MR: Group Health's initiative to avert opioid misuse and overdose among patients with chronic noncancer pain. Health Aff 30:1420-1424, 2011

29. Turk DC, Swanson KS, Gatchel RJ: Predicting opioid misuse by chronic pain patients: A systematic review and literature synthesis. Clin J Pain 24:497-508, 2008

30. US Food and Drug Administration: letter to Extended Release/Long Acting opioid application holders. Available at: www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM367697.pdf. Accessed July 10, 2014

31. Von Korff M, Kolodny A, Deyo RA, Chou R: Long-term opioid therapy reconsidered. Ann Intern Med 155:325-328, 2011

32. Von Korff M, Merrill JO, Rutter CM, Sullivan M, Campbell CI, Weisner C: Time-scheduled vs. pain-contingent opioid dosing in chronic opioid therapy. Pain 152:1256-1262, 2011

33. Von Korff M, Ormel J, Keefe FJ, Dworkin SF: Grading the severity of chronic pain. Pain 50:133-149, 1992

34. White AG, Birnbaum HG, Mareva MN, Daher M, Vallow S, Schein J, Katz N: Direct costs of opioid abuse in an insured population in the United States. J Manag Care Pharm 11:469-479, 2005

35. Wilson P, D'Agostino R, Levy D, Belanger A, Silbershatz H, Kannel W: Prediction of coronary heart disease using risk factor categories. Circulation 97:1837-1847, 1998