Comparative Analysis of Patient Distress in Opioid Treatment Programs using Natural Language Processing

Fatemeh Shah-Mohammadi, Wanting Cui, Keren Bachi, Yasmin Hurd and Joseph Finkelstein
Icahn School of Medicine at Mount Sinai, New York, NY, U.S.A.

Keywords: Natural Language Processing (NLP), Opioid Treatment Program (OTP), Determinants of Distress.

Abstract: Psychiatric and medical disorders, social and family environment, and legal distress are important determinants of distress that impact the effectiveness of the treatment in opioid treatment program (OTP). This information is not routinely captured in electronic health record, but may be found in clinical notes. This study aims to explore the feasibility and effectiveness of natural language processing (NLP) strategy for identifying legal, social, mental and medical determinates of distress along with emotional pain rooted in family environment from clinical narratives of patients with opioid addiction, and then using this information to find its impact on OTP outcomes. Analysis in this study showed that mental and legal distress significantly impact the result of the treatment in OTP.

1 INTRODUCTION

In recent years, opioid use disorder has become a significant public health problem in the United States, leading to thousands of death. According to CDC, prescription opioid and heroin overdose deaths have been increasing since 1999 (CDC, 2018). Opioid abusers are exposed to high risk of not only contracting infectious disease, but also development of mental illness (Ehrich, 2015). In addition to health concerns, opioid crisis also creates serious financial costs. In 2009, the annual costs of prescription and illicit opioid abuse, including lost productivity and health care costs, were estimated to be over $55 billion (Birnbaum, H. G, 2011). Methadone and buprenorphine has been reported as effective treatments for opioid dependence, and their widespread use could mitigate the negative health and societal effects of opioid use disorder (Mattick, R. P, 2008). Opioid Treatment Programs (OTPs) are among the licensed providers of medication for opioid abusers, and usually require patients to take medication at a clinic.

While great amount of studies have been conducted on extend, prevention and treatment of the opioid addiction (Han, B, 2021- Malta, M, 2019), research on the effectiveness of OTP and prediction of its outcomes is limited. Among these studies, very sparse numbers have used features extracted from clinical text documentations in their analysis. Clinical notes contain vast amounts of information about the patient such as prescribed medications, detailed physical and mental health conditions, as well as indications of social, legal or family distress. Extraction and analysis of these features may enhance the accuracy of outcome prediction models (Hazlehurst, B, 2019; Green, C. A., 2019). For example, in a recent article (Ettridge, K. A., 2018) it was found that patients with prostate cancer experience social isolations as a side effect of their treatment. Documentation of social determinants of health is promulgated by the National Academy of Medicine however extraction of these parameters from electronic health records (EHR) for systematic research may be challenging due to unstructured nature of clinical notes. Risk factors for opioid misuse such as untreated psychiatric and medical disorders, social or family environment, and legal distress (for example due to incarceration) are not captured routinely and are usually not encoded in EHR. However, this information might be documented in clinical notes in which providers record the information as told by their patients. A potential alternative to identify and extract such information from clinical texts is natural language processing (NLP). Various tools exist to extract information from clinical notes, including Clinical Text Analysis and Knowledge Extraction system (cTAKES) (Savova, G. K., 2010), MataMap/MetaMap Lite (Aronson, A. R., 2010; Demner-Fushman, D., 2017) and Clinical Language

Shah-Mohammadi, F., Cui, W., Bachi, K., Hurd, Y. and Finkelstein, J.
Comparative Analysis of Patient Distress in Opioid Treatment Programs using Natural Language Processing.
DOI: 10.5220/0010976700003123
In *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2022) - Volume 4: BIOSIGNALS*, pages 319-326
ISBN: 978-989-758-552-4; ISSN: 2184-4305
Copyright © 2022 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

Table 1: General statistics of the notes.

Discipline	Number of patients	Frequency
Counsellor	6674	70%
Nurse	6339	67%
Physician assistant	6019	63%
Medical doctor	4353	46%
Social worker	2767	29%
Vocational rehab counsellor	1221	13%
Assistant supervisor	59	0.6%
Clinic manager	91	0.9%
Financial counsellor	1	0.01%
Social work intern	91	0.9%

Annotation, Modeling and Processing (CLAMP) (Soysal, E., 2018). While MetaMap and cTAKES are general purpose NLP systems, CLAMP provides an integrated development environment with GUIs for the users who need to build customized NLP pipelines for their individual applications. In this study we used CLAMP to extract information from the clinical notes.

Mental, social and legal distresses along with family environment and physical health conditions are important determinants that encourage opioid misuse and impact the effectiveness of the treatment in OTP. The aim of this study is to explore the feasibility and effectiveness of NLP strategy for identifying legal, social, mental and medical determinates of health along with emotional pain rooted in family environment from clinical narratives of patients with opioid addiction, and then using this information to find its impact on OTP outcomes.

2 METHOD

2.1 Dataset

A dataset used in this study contained a harmonized aggregation of several relevant sources including notes taken by different providers (such as nurses, medical doctors and counselors) in OTP, information on admission, transfer, follow-up and discharge records from OTP in the New York City area since 1960s. Data was collected from the New York State Office of Addiction Service and Support's (OASAS) opioid treatment program for patients who received treatment at Mount Sinai Health System (MSHS) in New York City. This data includes admission records from May 1965 to March 2021. Out of 31,685 unique patients enrolled in OTP, 9511 have records of the notes coming from 10 general disciplines as follow: assistant supervisor, clinic manager, counsellor, financial counsellor, medical doctor, nurse, physician assistant, social worker, vocational rehab counsellor and social work intern. Table.1 lists for each discipline the percentage of patients who have records of notes from the corresponding discipline. The most common author types were counsellor, nurse and physician assistant. The average number of documents per patient was 54 (maximum: 780; minimum: 1).

2.2 NLP Tool

As mentioned earlier, we used CLAMP as entity extraction tool. This tool follows pipeline-based architecture composed of multiple NLP components. Various machine learning-based methods and rule-based methods have been used in developing these components. The list of CLAMP's components are as follow: sentence boundary detection, tokenizer, part-of-speech tagger, section header identifier, abbreviation reorganization and disambiguation, named entity recognition, UMLS encoder and rule engine. CLAMP is currently available in two versions: CLAMP-CMD (a command line NLP system) and CLAMP-GUI that provides a GUI for building customized NLP pipelines. The output of CLAMP contains the start and the end point of the word (or sequence of words) detected as entity within the text, the semantic tag associated with it, Concept Unique Identifier (CUI) number (along with RX-Norm code for the entities tagged as drug), assertion, and the actual text extracted as entity (Table. 2 shows a screenshot from CLAMP output). The semantic tag is divided into 20 categories as follow: 'temporal', 'treatment', 'problem', 'history', 'drug', 'test', 'strength', 'route', 'frequency', 'body location', 'course', 'duration', 'subject', 'condition', 'generic', 'lab value', 'form', 'dosage', 'severity'. The assertion can be "present" or "absent" (in case of negation). Information regarding each category can be found in CLAMP official site (https://clamp.uth.edu).

According to Table 1, less than 1% of the patients have notes taken by the last four disciplines. This encouraged us to check how informative the notes from these disciplines are and how many informative terms they contain. To find the answer, we randomly selected 500 patients. For each patient we extracted the notes from all disciplines. For each discipline, we integrated the notes from all patients to a single large documents. We then used *CountVectorizer* from

Table 2: Screenshot of CLAMP output.

Start	End	Semantic	CUI	Assertion	Entity
138	143	temporal	null	null	daily
224	232	drug	C0028040, RxNorm=[7407]	present	nicotine
475	479	subject	null	null	wife
494	497	problem	null	present	ill
503	516	problem	null	absent	throat cancer

scikit-learn library in python to extract for each discipline the list of tokens and to count the frequency of each token. The input parameters *min_df* and *ngram_range* are set to be 2 and (1, 2), respectively. To prevent the *CountVectorizer* tool from considering the stop words, we constructed a list of stop words and passed it as the next input parameter to the tool (https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html). The tokens were converted to lowercase and low frequency tokens and punctuation marks were discarded. Figures. 1 and 2 show the appearance frequency of the most frequent tokens (only the first 20) in notes taken by medical doctor and assistant supervisor. According to these figures, compared to the medical doctor the notes from assistant supervisor contain only 8 tokens (after deleting the stop words) which none of them are informative. The distribution for financial counsellor, clinic manager and social network intern were the same. As a result, in the next analyses we only considered the notes from the first six authors in Table. 1. For each patient, the notes from the first six disciplines were aggregated into a single large document and then fed as an input to the CLAMP.

2.3 Development of Lexicon

This stage involved generating five different lexicons of the terms that are indication of social, legal, mental and medical distress along with the distress coming from the family environment. Since documented standard and data collection strategies for theses determinant of health in the EHR are in an early stage, generating lexicon for each determinant was challenging. The initial list of term were collected using standard terminologies in SNOMED-CT. Since clinical notes are often documented as natural language, examining standard terminologies alone may lead to miss important information embedded in clinical notes. Therefore we then queried every term in every lexicons against CLAMP's extracted entities (for all notes of 500 patients) to incorporate any spelling variant of each term and find any relevant lexical representations. For example, the terms "low income" and "loss of income" are indication of social distress but former is a standard term in SNOMED-CT while the latter has been extracted through NLP analysis and search into CLAMP output. So, finding variant of any standard terms resulted in forming an enhanced and refined list for all lexicons. Then we used two domain expert's assessment to construct the final lexicons that appropriately represents social, legal, mental, medical and family-based distress. The terms indicating mental distress included "anxiety", "depression", "adhd", "insomnia", "psychiatric disorders", "borderline personality disorder", "ptsd disorders", "substance induced psychological disorders", "dissociative identity disorder", "multiple personality disorder", "panic disorder". The terms indicating social distress included "low income", "loss of income", "financial issues", "immigration issues", "loss of housing", "homelessness", "homelessness issues", "job loss", and "work related issues". The terms that are indication of medical problem included "diabetes", "uncontrolled high blood pressure", "heart attack", "surgery in left knee", "obesity" and "asthma". Regarding the terms indicating distresses rooted in family environment we assumed that the appearance of the terms in the note is a direct indication of family-related distress in the patient's life and the validity of this assumption were manually assessed by two domain experts. The terms included "wife", "dead wife", "brother", "husband", "step son", "interpersonal relationship issues", "family issues", "marital issues", "family stress", "increased family stress". Lastly, the terms indication legal distress included "legal problems", "legal related issues", "criminal legal issues", "second arrest", "multiple arrests", "recent release from prison", "drug related arrest", "multiple drug related incarcerations", "recent incarceration". The size of lexicons is as follow: 143, 38, 236, 66 and 31 for respectively mental, legal, health, family and social distress lexicon. Our considered process of lexicon development is consistent with previous research (Zhu, V. J., 2019; Bejan, C. A., 2018). Programming language for all analysis was Python 3.8.

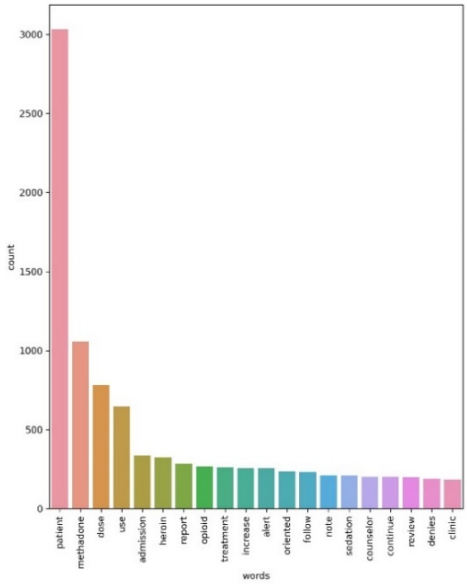

Figure 1: Medical doctor.

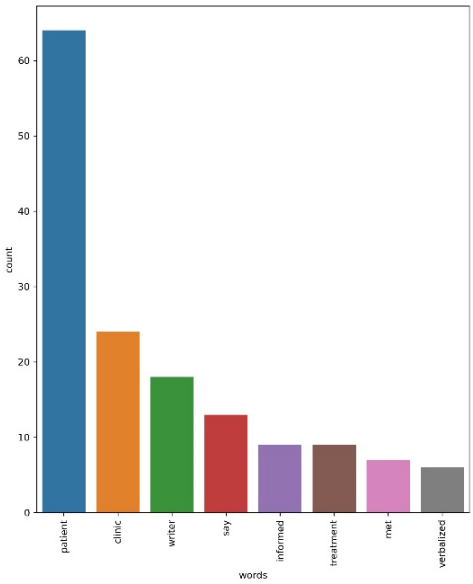

Figure 2: Assistant supervisor.

Table 3: Frequency count for determinants of distress.

Patient ID	Mental distress	Social distress	Legal distress	Medical distress	Family distress	Geometric mean
Patient 1	21	1	1	9	1	1.55
Patient 2	5	1	1	5	1	1.90
Patient 3	69	5	1	77	1	7.67
…	…	…	…	…	…	…
Patient 51	1	1	1	29	9	3.04

2.4 Study Design

To determine patients's treatment effectiveness, we focused on discharge records and the variable named as "discharge status". This variable recorded different unique reasons of discharge from the program among which we focused only on two extreme cases as follow: "completed treatment: all treatment goals met" and "treatment not complete: no goals met". The patients with the former discharge status were considered as the patients whom treatment was successful, while the patients with the latter status identified as failed treatment cases. Out of 31,065 unique patients discharged from the program, 2,939 patients met all treatment goals while 9524 patients met no goals. Out of 2,939 patients who met all the treatment goals, only 51 had notes available. We selected notes from those patients and put them in a folder called as "succeeded". We then extracted the notes from randomly selected 51 patients who failed all treatment's goals and put them in a separate folder called as "failed" group. For each patient in each group, the notes from the first 6 disciplines in Table. 1 were aggregated into a single large note. We then extracted entities for each patient's large note in each group using CLAMP. Next, we counted the number of words out of three lexicons that appeared as extracted entity for each patient across both groups. Table. 3 shows a screenshot of this analysis for "succeeded" group. It should be noted that two domain expert reviewers manually assessed the notes and confirmed that the existence of any term from the five lexicons is indication that the patient actually experiences that specific distress in his/her life. We further added a new feature as "geometric mean" which calculated by raising the product of word counts for all determinants of distress to the inverse of the total length of determinants (i.e. 5). Since different patients may have different dimensions of distress affected, we use geometric mean to compare overall distress level between different patients.

Table 4: Descriptive statistics of the groups.

Group	S	F
Age		
Mean	40.23	40.60
Std	12.20	11.57
Min	23.00	21.00
25%	28.50	31.25
50%	38.00	39.00
75%	51.50	51.00
Max	64.00	64.00
Sex		
Female	25.00%	25.4%
Male	75.00%	74.60%
Race		
Black / African American	20.00%	35.00%
White / Non-Hispanic	37.00%	27.00%
Hispanic	35.00%	38.00%
Asian	7.00%	0.00%
Education		
Not finished high school	33.00%	47.00%
High school	39.00%	27.00%
Above high school	28.00%	28.00%
Employment		
Employed	26.00%	12.00%
Not in labor force	15.00%	23.00%
Unemployed	59.00%	65.00%
Primary Substance		
Heroin	96.00%	96.00%
Other	4.00%	4.00%
Secondary substance		
None	49.00%	43.00%
Cocaine	16.00%	22.00%
Crack	6.00%	6.00%
other	29.00%	29.00%

The aim of generation of this table for both groups is to recognize the risk factors for opioid misuse and find out whether specific risk factors (of being mental, social, medical, legal and family distresses) are significant factors to the success of the treatment. Next section contains the result of this assessment.

3 RESULTS

In both group, the average age of patients at admission was 40 years old, but the standard deviation in patients with unsuccessful treatment is slightly less than the other group. The number of males is around 3 times more than females in both groups. Around 96% of patients in both group would consume heroin as their primary abuse substance and around 50% were addicted to more than one substance. Around 35% of patients failed in the treatment was Black or African American, 27% White and the rest Hispanic. While the distribution of race among patients with successful treatment was more diverse: 20% Black or African American, 37% White, 7% Asian and 35% Hispanic. Just around 47% of patients in failed group have not received their diploma, 27% have diploma and 28% have some graduate degree. Only around 12% of these patients are employed. The scenario in group with successful treatment is as follow: around 33% without diploma, 39% with diploma, and 28% with some graduate degree. The employment rate for this group is 26%.

Figures. 3 to 8 shows the distribution of word counts among 51 patients for both succeeded and failed groups. It can be seen in all figures that the underlying distribution for all determinates of distress across both groups is not normal. As a result, non-parametric statistical test is used to recognize risk factors to the success of treatment in OTP. To devise this, we ran Mann Whitney U test between the two groups for all five determinants of distress. The testing is two sided. We considered that p value higher than 0.05 is not statistically significant. Table. 4 shows the general statistics for the both groups and for all determinants, along with the result for the statistical test (i.e. p-value). The letters “S” and “F” in this table indicate the succeeded and failed group, respectively. The p value for mental and legal distress along with geometric mean is less than 0.05.

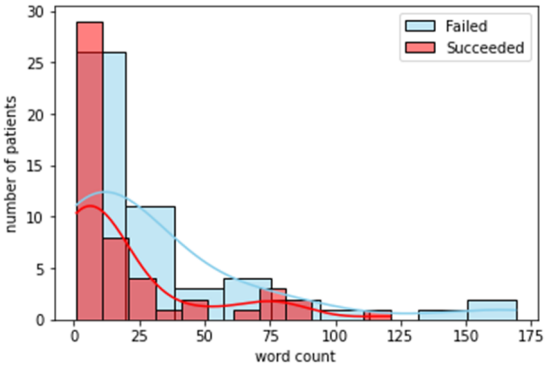

Figure 3: Mental distress.

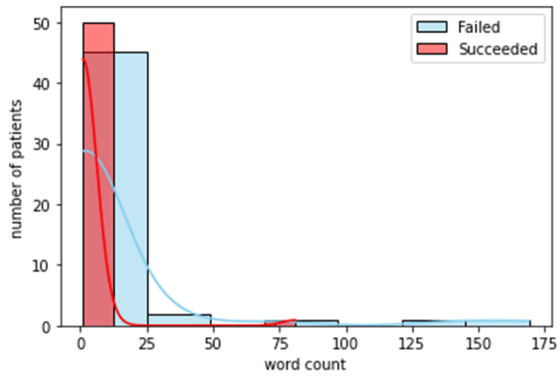

Figure 4: Legal distress.

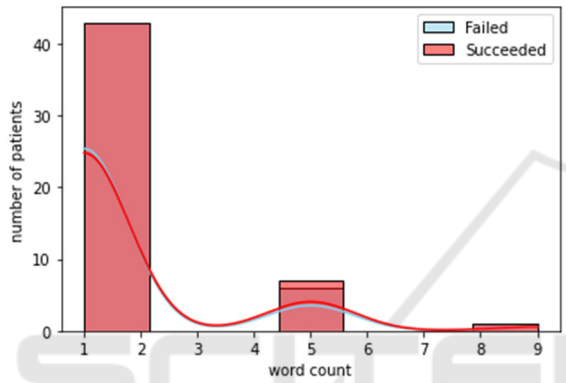

Figure 5: Social distress.

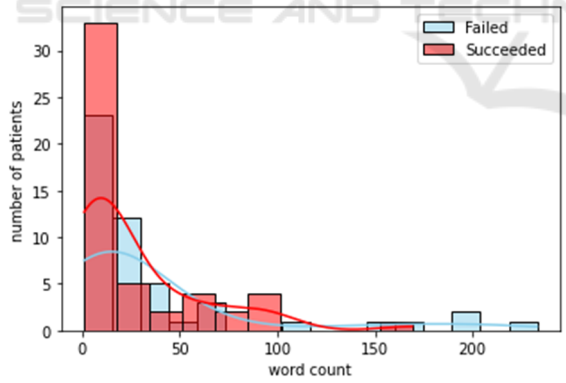

Figure 6: Health distress.

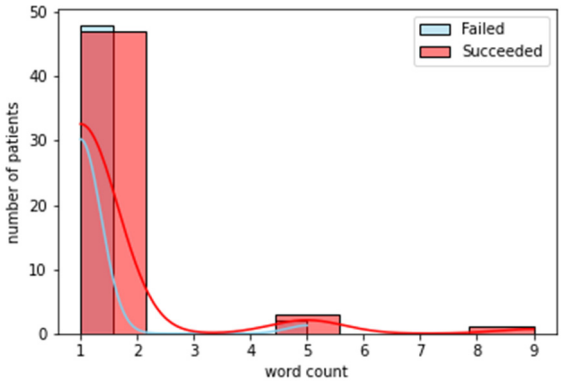

Figure 7: Family distress.

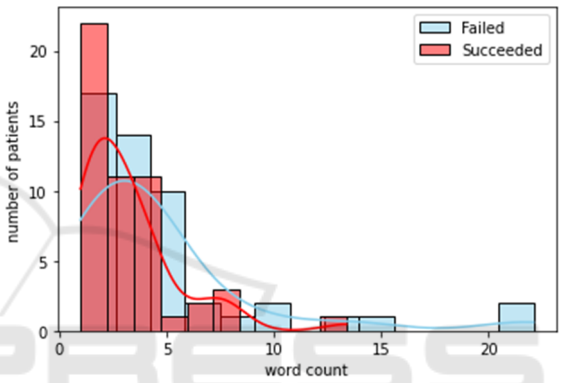

Figure 8: Geometric mean.

4 DISCUSSION

According to table 4, the patients with unsuccessful treatment are younger, mostly Black African American and Hispanic, with higher rate of unemployment and lower education. The age, employment and education status can be considered as factors that may contribute to opioid abuse. According to table. 5, on the other hand, for mental distress the mean and median values of succeeded and failed group are significantly different with a p value equal to 0.03. This is a strong indication that existence of mental distress in the life of patients has a significant effect on their success in the treatment for opioid abuse. The same scenario is valid for the legal distress. For this determinant of distress the p value is 0.003. Accordingly, mental and legal distress can be considered as a risk factors to the effectiveness of the treatment in OTP. Based on the data in the same for the social, family and medical source of distresses, the p value is more than 0.05 which shows that these determinant of distress are not significant factors to the effectiveness of the treatment. There is also statistically significant difference in geometric mean between two groups demonstrating that the two study groups differ in overall level of distress.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

Table 5: Statistics for the determinant of distress.

	Mental distress		Social distress		Legal distress		Medical distress		Family distress		Geometric mean	
	S	F	S	F	S	F	S	F	S	F	S	F
Mean	19.8	32.2	1.7	1.64	2.73	10.84	27.43	37.48	1.39	1.16	3.23	4.63
Std	28.6	41.4	1.73	1.68	11.24	32.2	34.9	55.67	1.44	0.79	2.36	4.5
Min	1	1	1	1	1	1	1	1	1	1	1	1
25%	1	5	1	1	1	1	5	9	1	1	1.61	1.94
50%	9	17	1	1	1	1	13	17	1	1	2.43	3.65
75%	21	40	1	1	1	1	37	37	1	1	3.92	5.12
Max	121	169	9	9	81	169	169	233	9	5	13.37	22.08
P-value	0.03		0.4		0.003		0.2		0.2		0.03	

Our study is consistent with previous works but provides more comprehensive review of different interventions affecting the result of the treatment in OTP (Malta, M., 2019; Carrell, D. S., 2019). In the future, we'll increase the sample size for the groups to develop predictive models using the identified risk factors.

5 CONCLUSIONS

In this study we explored the feasibility and effectiveness of NLP strategy for identifying legal, social, mental and medical determinates of health along with source of distress rooted in family environment from clinical narratives of patients with opioid addiction, and then used this information to find its impact on OTP outcomes. Five lexicons were generated for all five determinants of distress. The lexicons generated in this study combine standard concepts and domain expert knowledge. To the best of our knowledge this paper is the first one that studies a comprehensive review of dimensions of information extracted through NLP from the notes taken by different authors in OTP and explores their impact on the effectiveness of the treatment in OTP. Our preliminary analysis showed that mental and legal distress significantly impact the result of the treatment.

REFERENCES

Centers for Disease Control and Prevention, National Center for Health Statistics (2018). Multiple Cause of Death 1999-2017 on CDC WONDER Online Database.

Ehrich, E., Turncliff, R., Du, Y., Leigh-Pemberton, R., Fernandez, E., Jones, R., & Fava, M. (2015). Evaluation of opioid modulation in major depressive disorder. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 40(6), 1448–1455.

Birnbaum, H. G., White, A. G., Schiller, M., Waldman, T., Cleveland, J. M., & Roland, C. L. (2011). Societal costs of prescription opioid abuse, dependence, and misuse in the United States. *Pain medicine (Malden, Mass.)*, 12(4), 657–667.

Mattick, R. P., Kimber, J., Breen, C., & Davoli, M. (2008). Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *The Cochrane database of systematic reviews*, (2), CD002207.

Han, B., Jones, C. M., Einstein, E. B., & Compton, W. M. (2021). Trends in and Characteristics of Buprenorphine Misuse Among Adults in the US. *JAMA network open*, 4(10), e2129409.

Malta, M., Varatharajan, T., Russell, C., Pang, M., Bonato, S., & Fischer, B. (2019). Opioid-related treatment, interventions, and outcomes among incarcerated persons: A systematic review. *PLoS medicine*, 16(12), e1003002.

Hazlehurst, B., Green, C. A., Perrin, N. A., Brandes, J., Carrell, D. S., Baer, A., DeVeaugh-Geiss, A., & Coplan, P. M. (2019). Using natural language processing of clinical text to enhance identification of opioid-related overdoses in electronic health records data. *Pharmacoepidemiology and drug safety*, 28(8), 1143–1151.

Green, C. A., Perrin, N. A., Hazlehurst, B., Janoff, S. L., DeVeaugh-Geiss, A., Carrell, D. S., Grijalva, C. G., Liang, C., Enger, C. L., & Coplan, P. M. (2019). Identifying and classifying opioid-related overdoses: A validation study. Pharmacoepidemiology and drug safety, 28(8), 1127–1137.

Ettridge, K. A., Bowden, J. A., Chambers, S. K., Smith, D. P., Murphy, M., Evans, S. M., Roder, D., & Miller, C. L. (2018). "Prostate cancer is far more hidden…":

Perceptions of stigma, social isolation and help-seeking among men with prostate cancer. *European journal of cancer care*, 27(2), e12790.

Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., & Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *Journal of the American Medical Informatics Association : JAMIA*, 17(5), 507–513.

Aronson, A. R., & Lang, F. M. (2010). An overview of MetaMap: historical perspective and recent advances. *Journal of the American Medical Informatics Association : JAMIA*, 17(3), 229–236.

Demner-Fushman, D., Rogers, W. J., & Aronson, A. R. (2017). MetaMap Lite: an evaluation of a new Java implementation of MetaMap. *Journal of the American Medical Informatics Association : JAMIA*, 24(4), 841–844.

Soysal, E., Wang, J., Jiang, M., Wu, Y., Pakhomov, S., Liu, H., & Xu, H. (2018). CLAMP - a toolkit for efficiently building customized clinical natural language processing pipelines. *Journal of the American Medical Informatics Association : JAMIA*, 25(3), 331–336.

Zhu, V. J., Lenert, L. A., Bunnell, B. E., Obeid, J. S., Jefferson, M., & Halbert, C. H. (2019). Automatically identifying social isolation from clinical narratives for patients with prostate Cancer. BMC medical informatics and decision making, 19(1), 43.

Bejan, C. A., Angiolillo, J., Conway, D., Nash, R., Shirey-Rice, J. K., Lipworth, L., Cronin, R. M., Pulley, J., Kripalani, S., Barkin, S., Johnson, K. B., & Denny, J. C. (2018). Mining 100 million notes to find homelessness and adverse childhood experiences: 2 case studies of rare and severe social determinants of health in electronic health records. Journal of the American Medical Informatics Association : JAMIA, 25(1), 61–71.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS