



Published in final edited form as:

*Subst Abus.* 2022 ; 43(1): 917–924. doi:10.1080/08897077.2021.1986767.

## Clinical documentation of patient-reported medical cannabis use in primary care: Towards scalable extraction using natural language processing methods

David S. Carrell, PhD<sup>1</sup>, David J. Cronkite, MS<sup>1</sup>, Mary Shea, BA<sup>1</sup>, Malia Oliver, BA<sup>1</sup>, Casey Luce, MSPH<sup>1</sup>, Theresa E. Matson, MPH<sup>1</sup>, Jennifer F. Bobb, PhD<sup>1</sup>, Clarissa Hsu, PhD<sup>1</sup>, Ingrid A. Binswanger, MD, MPH, MS<sup>2</sup>, Kendall C. Browne, PhD<sup>3</sup>, Andrew J. Saxon, MD<sup>3</sup>, Jennifer McCormack, MS<sup>4</sup>, Eve Jelstrom, MBA<sup>4</sup>, Udi E. Ghitza, PhD<sup>5</sup>, Cynthia I. Campbell, PhD, MPH<sup>6</sup>, Katharine A. Bradley, MD, MPH<sup>1</sup>, Gwen T. Lapham, PhD, MPH, MSW<sup>1</sup>

<sup>1</sup>Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA

<sup>2</sup>Institute for Health Research, Kaiser Permanente Colorado, Aurora, CO, USA

<sup>3</sup>University of Washington, Psychiatry and Behavioral Sciences, Seattle, WA, USA

<sup>4</sup>The EMMES Corporation, Rockville, MD, USA

<sup>5</sup>National Institutes on Drug Abuse, National Institutes of Health, Rockville, MD, USA

<sup>6</sup>Kaiser Permanente Division of Research, Oakland, CA, USA

### Abstract

**Background**—Most states have legalized medical cannabis, yet little is known about how medical cannabis use is documented in patients' electronic health records (EHRs). We used natural language processing (NLP) to calculate the prevalence of clinician-documented medical cannabis use among adults in an integrated health system in Washington State where medical and recreational use are legal.

**Methods**—We analyzed EHRs of patients 18 years old screened for past-year cannabis use (11/1/2017–10/31/2018), to identify clinician-documented medical cannabis use. We defined medical use as any documentation of cannabis that was recommended by a clinician or described by the clinician or patient as intended to manage health conditions or symptoms. We developed and applied an NLP system that included NLP-assisted manual review to identify such documentation in encounter notes.

**Results**—Medical cannabis use was documented for 16,684 (5.6%) of 299,597 outpatient encounters with routine screening for cannabis use among 203,489 patients seeing 1,274

---

**Corresponding Author:** David S. Carrell, PhD Associate Investigator, Kaiser Permanente Washington Health Research Institute, 1730 Minor Ave Suite 1600, Seattle, WA, 98101 USA, Phone: 206-287-2705, Fax: 206-287-2871.

**Authors' contributions:** All authors contributed to the design of the study. DSC, DJC, MS, MO, CL, TEM, CH, IAB, KAB and GTL contributed to the collection of data. All authors contributed to the analysis and/or interpretation of results. All authors approved the text of the final manuscript. DSC and DJC have expertise in clinical natural language processing (NLP). GTL, KAB, IAB, KCB, AJS, JM, EJ, UEG and CIC have expertise in the field of substance use disorders, assessment of patient medical cannabis use, and/or substance use disorders. KAB has expertise in implementation of screening of patients in primary care settings.

**Competing interests:** We have no competing interests.

clinicians. The validated NLP system identified 54% of documentation and NLP-assisted manual review the remainder. Language documenting reasons for cannabis use included 125 terms indicating medical use, 28 terms indicating non-medical use and 41 ambiguous terms. Implicit documentation of medical use (e.g., “edible THC nightly *for lumbar pain*”) was more common than explicit (e.g., “continues *medical* cannabis use”).

**Conclusions**—Clinicians use diverse and often ambiguous language to document patients’ reasons for cannabis use. Automating extraction of documentation about patients’ cannabis use could facilitate clinical decision support and epidemiological investigation but will require large amounts of gold standard training data.

---

## Introduction

Medical cannabis use is legal in 37 U.S. states. Eighteen states plus the District of Columbia have legalized adult recreational use and another 13 states have decriminalized cannabis use. The National Academy of Sciences has stressed the need for clinical epidemiological research to investigate the safety and effectiveness of medical cannabis use for a variety of health conditions.<sup>1</sup> To enable large-scale observational studies and pragmatic trials to address these knowledge gaps, automated algorithms using electronic health records (EHR) data to define phenotypes of medical cannabis use are needed. Many people may use cannabis to manage medical or behavioral health symptoms, and in some cases to replace prescribed medications or other treatments with proven efficacy—potentially without adequately understanding the trade-offs.<sup>1</sup> Understanding how often that occurs by distinguishing medical from non-medical cannabis use in large patient populations is important for facilitating research to address these knowledge gaps.

Investigations of patient medical cannabis use to date have relied largely on survey data,<sup>2</sup> analysis of online community posts<sup>3</sup>, and information collected from patients who are screened or treated for substance use disorders.<sup>4</sup> Prior work has shown that EHR documentation of patients’ cannabis use entails a diverse language<sup>5 6</sup> and that EHR-documented use corresponds to use self-reported in surveys.<sup>7</sup> However, we are unaware of any prior work attempting to systematically extract on a large-scale evidence of patients’ medical cannabis use recorded in EHRs. Development of such methods was a primary objective of this study.

A unique challenge of medical cannabis research is the lack of a universally accepted definition for what constitutes medical use.<sup>8</sup> The National Institute on Drug Abuse (NIDA) broadly defines medical use as any use intended to treat symptoms of illness or other health conditions.<sup>9</sup> However, the U.S. Food and Drug Administration does not recognize cannabis as a medicine, and federal law acknowledges no accepted medical use.<sup>10</sup> Medicolegal definitions view medical use in terms of treatment of specific diseases (such as cancer pain, seizure disorders and glaucoma), or as recognized in medical cannabis laws,<sup>11</sup> and/or when authorized by a licensed practitioner.<sup>12</sup> Therapeutically-focused definitions encompass any cannabis use intended to treat disease or alleviate medical symptoms, including patient self-defined medical use.<sup>13</sup>

The clinical setting for this study is a large integrated health system in Washington State where medical cannabis use has been legal since 1998 and adult recreational use legal since 2012. The health system, Kaiser Permanente Washington (KPWA), began routine annual screening for the frequency of past-year cannabis use among adult primary care patients in 2015.<sup>14</sup> These screening results are recorded as structured EHR data and clinicians often document reasons for use in clinical notes. Previously published analyses indicate over 80% of KPWA primary care patients receive screening.<sup>15</sup> The primary objective of this paper is to describe free-text documentation of medical cannabis use, as distinct from other reasons for use, among patients receiving care in 25 KPWA primary care clinics during a one-year period. To give a sense for the overall volume of healthcare provided in these 25 clinics, during the 2017 calendar year a total of nearly 280,000 adults had a combined total of more than 750,000 primary care encounters in these clinics. A secondary objective is to describe the natural language processing (NLP) method used to identify descriptions of medical cannabis use in free-text encounter notes and where additional methods development may be needed to facilitate its application in future multi-site studies.

This work is part of an ongoing National Institute on Drug Abuse Clinical Trials Network study investigating patterns of medical cannabis use by primary care patients.

## Materials and Methods

### Definition of medical cannabis use

We defined medical cannabis use as any use recommended by a clinician or characterized by the clinician or patient as use to manage a health condition or symptom, either explicitly (e.g., “continuing medical marijuana most days”) or implicitly (e.g., “cannabis 2–3× weekly for low back pain”). We used this therapeutically-focused definition because it includes both patient and clinician perspectives on medical cannabis use. We did not limit symptoms or health conditions to those recognized by Washington State law as eligible for medical cannabis treatment. We excluded from our definition descriptions of cannabis use that were ambiguous with respect to medical use.

### Setting, data and sample

Data came from KPWA’s Epic® EHR. The patient sample included adults (≥ 18 years old) with KPWA outpatient visits between November 1, 2017 and October 31, 2018 who completed a single-item screen for frequency of past-year cannabis use in connection with a primary care (Family Practice or General Internal Medicine), behavioral health, or urgent care encounter. The cannabis screen is part of KPWA routine care and asks, “*How often in the past year have you used marijuana?*” Possible responses are “*never,*” “*less than monthly,*” “*monthly,*” “*weekly,*” or “*daily/almost daily.*” About 20% of KPWA patients report cannabis use.<sup>16</sup> This screen does not ask patients about their reasons for cannabis use (e.g., medical or recreational), but clinicians may document reasons for use in encounter notes. We limited our sample to patients with cannabis screens so that the data regarding past-year medical cannabis use we generated using NLP could be compared to data obtained by routine screening.

## Overall process design

Following established practices<sup>17–20</sup> we trained an automated NLP system to identify descriptions of medical cannabis use in a sample of EHR notes in which we manually marked all medical cannabis use mentions. We applied this automated system to EHR notes for the entire study cohort; patients with any NLP-identified medical use were classified as such and received no further processing. We applied NLP-assisted manual review to EHR notes of all remaining patients to identify medical use not detected by the automated NLP system. Details of our approach are illustrated in Figure 1 and described in this section.

## Study corpus

Our study corpus consisted of encounter notes from 365 days before through seven days after a patient's cannabis screening date associated with encounters that routinely offer behavioral health screening, including primary care, behavioral health, and urgent care. We obtained these notes from the KPWA Epic® Clarity® database.

## Gold standard corpus

We created our gold standard corpus (also referred to as a “reference standard”) as follows. First, we copied the entire study corpus to a Microsoft SQL Server database, activated SQL Server's built-in full-text indexing service (which allows ad hoc query and rapid retrieval of text containing any term or combination of terms in a corpus), and used standard Microsoft Transact SQL functions<sup>21</sup> to query and display terms in their context. Next, by iteratively querying, reviewing, and quantifying the frequency of candidate cannabis terms in the corpus (an NLP-assisted exploratory review approach we have used previously) we identified six commonly used cannabis terms—marijuana, cannabis, MJ, THC, CBD and weed. We used these terms to query and select from the study corpus a subset of 1,093 notes for gold standard annotation. Based on prior experience we judged a corpus of this size would adequately represent language diversity, provided our sampling scheme over-sampled notes likely to contain content of interest. Accordingly, we selected three-fourths (n=839) of these notes from a group of 293 clinicians with the highest number of patients with cannabis screens (accounting for 75% of all cannabis screens). We selected 254 additional notes from the remaining 636 clinicians. Our random sample included up to two notes per clinician in the first group, and up to one note per clinician in the second group. When a clinician had multiple qualifying notes, we randomly sampled from those containing the highest number of cannabis terms (Appendix B, Table B.2).

Following iterative development of written annotation guidelines, a medical records abstractor with over a decade of medical records experience (MS) and an informaticist (DJC) achieved substantial inter-rater agreement (kappa 0.71) independently annotating all affirmative mentions of cannabis and all novel descriptions of medical cannabis use in a subset of 79 gold standard corpus notes. We assessed inter-rater agreement in this way solely to establish that the medical records abstractor was able to understand and properly implement the annotation guidelines. Thereafter the medical records abstractor, alone, annotated the remainder of this corpus. The abstractor flagged for discussion any relevant content not clearly addressed by the written annotation guidelines and resolved these through discussion with the informaticist and/or subject matter experts (GTL, KAB,

DSC). By “affirmative mentions” we mean mentions not used in a sense that is negated (e.g., “no hx of cannabis use”), hypothetical (e.g., “possible side effects of THC include sleepiness, ...”), uncertain (e.g., “possible marijuana use”), historical (e.g., “+MJ in middle school”), or about someone other than the patient (e.g., “sister using THC for epilepsy”). The abstractor also assembled lists of terms clinicians used to characterize non-medical use and terms that were ambiguous with respect to medical use, though quantifying the frequency of such mentions was beyond the scope of this project.

A random 70% subset of the gold standard corpus was used to train the NLP algorithm and the remaining 30% was reserved for a one-time evaluation of the final NLP algorithm.

### **NLP dictionaries**

Our dictionaries of terms used to describe medical cannabis use, non-medical cannabis use, and use that was ambiguous with respect to therapeutic intent included all such terms appearing in the gold standard corpus. To these terms we added linguistic variations (e.g., for the term “self-treat” we added “self-treats” and “self-treating”). This yielded dictionaries with 21 cannabis terms, 125 medical use terms, 29 non-medical use terms, and 41 ambiguous terms (Appendix A).

### **Automated NLP systems**

We developed an automated NLP algorithm to identify documentation of patients’ medical cannabis use as follows. First, we identified all *affirmative* mentions in the study corpus of current cannabis use by the patient (referred to as Step 2 in Figure 1). We identified mentions of cannabis terms via a commonly used string-matching approach<sup>22</sup> based on our cannabis term dictionary. We considered using but did not use the open source ConText algorithm<sup>23</sup> to identify affirmative mentions because we were concerned it would not perform well due to the absence of proper sentence-level punctuation in our corpus. Instead, we developed and applied a machine-learned algorithm trained on our gold standard data to determine which occurrences were *affirmative* mentions. To isolate affirmative mentions our algorithm used the cannabis term, the three terms preceding it, and the three terms following it as feature inputs to a linear support vector classifier which achieved excellent performance (AUC of 0.93 in validation data; Appendix C). Next, each affirmative mention was processed by a series of NLP rules we manually crafted to reflect patterns of language clinicians used to describe medical cannabis use that appeared in our gold standard training corpus (Appendix C). Medical cannabis use could be described explicitly (e.g., “cannabis use, has a green card for back pain”) or implicitly (e.g., “CBD daily to alleviate anxiety”). The NLP algorithm ignored descriptions of non-medical and ambiguous use.

We considered an NLP algorithm with 90% positive predictive value (PPV, at the mention-level, defined here as the proportion of medical use descriptions identified by the algorithm that truly document medical use according to the gold standard) and 90% sensitivity (the proportion of actual medical use descriptions in the corpus identified by the algorithm) to be suitable for determining which screens describe medical use and which do not because performance at these levels approach that of manual review. We did not know whether developing such a high-performing NLP algorithm was possible. However,

we were confident that a gold standard annotated corpus could be used to calibrate an algorithm with high mention-level PPV—at the expense of sensitivity—and that algorithm performance could be validated in a random subset of the annotated gold standard corpus set aside specifically for such evaluation. An algorithm with high mention-level PPV and low sensitivity would still be useful for identifying positive documentation of medical use, but could not be used to determine when such documentation was absent since low sensitivity yields unacceptably high rates of false negatives (i.e., concluding that descriptions of medical use were absent when in fact they existed). We therefore planned to use the automated NLP algorithm to identify *some* screens with documented medical use and, if the algorithm's sensitivity was below 90%, use NLP-assisted manual review to resolve notes for the remaining screens.

### Steps in identifying EHR-documented medical cannabis use

We generated data about patients' medical cannabis use via the five-step process depicted in Figure 1. **Step 1:** Identify visits with cannabis screening using structured EHR data. **Step 2:** Use NLP methods to identify *affirmative* mentions of (any) cannabis use in patient chart notes. **Step 3:** Identify coded data documenting medical cannabis use based on structured EHR data. **Step 4:** Use a rule-based NLP algorithm to identify mentions of *medical* cannabis use. **Step 5:** Use NLP-assisted manual review to identify mentions of medical cannabis use. For efficiency, as soon as evidence of medical cannabis use was identified for a given screening encounter (in Steps 3, 4, or 5) that encounter was classified as positive for medical use and no remaining notes for that encounter were processed. Additional details for each step follow.

**Step 1:** Patient screenings are documented by structured data in the KPWA Epic® EHR. Using this structured data we determined which patients received cannabis screening during the study period and the screening date. We then assembled all primary care encounter notes for these patients from 365 days before through 7 days after their screening dates.

**Step 2:** Using the NLP system described above and in Appendix C, we identified all affirmative mentions of cannabis use—regardless of the reason for use—in the entire study corpus. If none of the notes for a given screen contained affirmative mentions we classified that encounter as **negative** for medical cannabis use, and no further processing of notes for these screens occurred.

**Step 3:** We determined which of the remaining screens had clinician-recorded diagnosis codes indicating medical cannabis use (defined as the presence of International Classification of Disease, Tenth Revision [ICD-10] code Z79.899 “other long-term current drug therapy”). We classified these screens as **positive** for medical cannabis use and removed their notes from the corpus requiring processing.

**Step 4:** Using the NLP system described above and in Appendix C, we attempted to identify documentation of medical cannabis use in all remaining notes. We classified as positive for medical use each screen having at least one NLP-detected mention of medical cannabis use (and, for efficiency, removed from the corpus to be processed further all notes for these screens).

**Step 5:** To each of the remaining screens we applied NLP-assisted review in an attempt to identify documentation of medical cannabis use not detected by the Step 4 NLP system. **NLP-assisted manual review** is a method that uses NLP to pre-annotate individual clinical notes and present them, sequentially, for human review in a computer interface designed specifically for this purpose. For efficiency the interface prioritizes review of notes containing terms of interest (i.e., terms in our NLP dictionaries [Appendix A]), applies color highlighting to these terms to aid review, and uses a point-and-click interface to simplify recording of reviewer decisions and navigation. Because we were determining whether any medical cannabis use had been documented for a given screen, as soon as the reviewer classified one note as positive for medical use, no further review of notes for that screen were reviewed, and the interface advanced to the next screen. When, during Step 5, the reviewer encountered a reason for use that was not easily classified, it was flagged for discussion and resolution by study investigators (GTL, KAB, DSC). Step 5 continued until there were no notes remaining to be reviewed.

### Analyses and summary measures

In our analyses we used descriptive statistics to summarize the number of cannabis screens per patient, findings regarding medical cannabis use, number of unique patients screened per clinician, and the quantity of clinical note text processed. We do not report the number of patients with cannabis use described as non-medical or ambiguous with respect to therapeutic intent because we did not collect those data for all patients. When a patient was screened more than once during the study period, we randomly selected one screen so that patient-level contributions were balanced.

To illustrate language clinicians used to characterize patient cannabis use we purposively selected and de-identified illustrative examples that were typical yet diverse. Analyses were completed using SAS version 9.4.

This study was approved by the Human Subjects Review Committee of Kaiser Permanente Washington Health Research Institute.

### Results

Characteristics of the patients and cannabis screens included in this study are summarized in Table 1. There were 299,597 cannabis screens during the one-year period, representing 203,489 patients and 1,274 clinicians. The study corpus included 19.4 million chart notes with a combined total of 2.02 billion words, of which 642,376 words were distinct. Sixty percent of patients were female and over half were age 36–65. A large majority of patients (84%) had a single cannabis screen. Most clinicians saw 100 unique screened patients; 1 in 7 clinicians saw >500 screened patients.

“Marijuana” was the most common term clinicians used to document cannabis, appearing in 230,008 notes for 79,332 patients in this corpus. “Marijuana” was more than 10 times more common than the term “cannabis,” which appeared in 21,896 notes for 7,730 patients. However, even less commonly used terms, such as “CBD” (for cannabidiol) and “weed,” appeared in the notes of thousands of patients (see Appendix B Table B.3). Terms clinicians

used most frequently to describe patient cannabis use were “marijuana” (53.5%, including common misspellings), “cannabis” (26.5%), “THC” (8.1%) and “CBD” (6.7%); infrequently used terms were “MJ” (2.1%), “pot” (1.5%), “weed” (0.8%), “indica” (0.5%), “sativa” (0.1%), “cannabinoid” (0.05%) “spice” (0.02%) and tetrahydrocannabinol (0.004%).

Overall, 16,684 (5.6%) of cannabis screens had one or more encounter notes with affirmative mentions of cannabis that also documented medical cannabis use (21 terms). Of these screens, 493 (3%) were classified based on ICD-10 coding (Figure 1 Step 2), 9,077 (54%) by the fully-automated machine-learned NLP algorithm (with specificity of 80%; details in Appendix C) (Figure 1 Step 3), and 7,114 (43%) by NLP-assisted manual review (Figure 1 Step 4).

Clinicians used diverse language to characterize medical cannabis use. These included explicit descriptions of medical use (e.g., “using medical marijuana for ...”) and implicit descriptions (e.g., “... edible THC nightly for lumbar pain”). Implicit descriptions referenced a variety of clinical symptoms including, but not limited to pain, anxiety, insomnia, nausea, and seizures. Language clinicians used to characterize patients’ medical cannabis use, non-medical cannabis use, and ambiguous use are provided in Appendix A and illustrated in Table 2.

## Discussion

In a state where medical and recreational cannabis are legal and a healthcare setting where screening for cannabis use is routine, over twenty different terms were used to reference cannabis in EHR chart notes. Clinical documentation regarding patients’ reasons for cannabis use was often unclear and clinicians used highly diverse language to document patients’ medical use. The diversity was largely driven by a widespread use of *implicit* descriptions of medical cannabis use, typically consisting of a description of cannabis use and a linked statement about one or more clinical symptoms or health conditions that were the reason for the cannabis use. Our automated NLP system was able to accurately identify just over half of medical cannabis use descriptions, necessitating augmentation by NLP-assisted manual review.

Language diversity is a key factor complicating NLP system development; as language diversity increases so does the amount of training data and programming effort needed to develop an NLP system with high accuracy.<sup>24</sup> We may speculate that a training corpus of 1,000 documents may have been sufficient for successful development of an NLP system focused exclusively on explicit descriptions of medical cannabis use (i.e., descriptions involving terms such as “medically”, “clinician-recommended”, or “medical marijuana card”), but it was clearly inadequate for representing the diversity of implicit descriptions we encountered in this corpus (e.g., “edible THC nightly for lumbar pain”). Additional difficulty was introduced by descriptions that were ambiguous with respect to medical use and descriptions of non-medical use, both of which must be distinguished from medical use to achieve an algorithm with high mention-level PPV. Consequently, our automated NLP system was only able to accurately classify documented medical use in about half (54%) of screens.



Because our automated NLP algorithm was developed exclusively using KPWA notes, we would expect it to perform less well in other healthcare settings, due to language usage that is specific to those institutional settings or individual clinicians who practice there—neither of which may be represented in our KPWA corpus.<sup>25</sup> However, the NLP-assisted manual review method described here is directly transferable to other settings and could be used to generate gold standard data in other settings. Development of an automated NLP system achieving high mention-level PPV and high sensitivity when applied in diverse healthcare settings will require the generation of large, representative gold standard corpora from each setting in which it is to be applied. Such corpora have facilitated multi-site development and use of NLP algorithms in other clinical domains.<sup>25–28</sup> An extremely valuable byproduct of the current project is the set of 7,114 descriptions of medical cannabis use identified in KPWA encounter notes—half by our automated NLP system and half by NLP-assisted manual review. These data may be used alone or in combination with comparable data from other settings to advance the development of NLP tools for identifying patients’ medical cannabis use in EHR records.

This study highlights the clinical need for standardized documentation of patients’ reasons for cannabis use—both in support of good clinical care and as a means toward reducing heterogeneity in language clinicians use to document patient cannabis use, as we observed in this study. Standardized assessments have been recommended by experts<sup>1</sup> and the NIDA Clinical Trials Network (CTN) recently included a question about patients’ reasons for cannabis use in an assessment offered to enrollees in studies funded by the NIDA CTN’s “Helping to End Addiction Long-term” (HEAL) initiative.<sup>29</sup> HEAL enrollees are asked to select from a list all their *reasons for cannabis use* (e.g., pain, muscle spasm, sleep, stress, nausea/vomiting, worry or anxiety, and an open-ended “other” category). This HEAL question deliberately avoided the term ‘medical’ out of concern that such language may influence how patients perceive and characterize their own cannabis use. Standardized assessment could be used in conjunction with routine screening to prompt constructive and clear communication between patients and clinicians about patients’ cannabis use,<sup>15</sup> thereby enhancing clinical documentation and, opportunities for improving patient care and facilitating research.

Limitations of this study should be noted. Our results were likely influenced by the design and setting of this study, including 1) the legal status of cannabis use in Washington State, 2) our clinical/therapeutic definition of medical cannabis use, and 3) our health system’s practice of routine screening for past-year cannabis use. Language used to characterize medical cannabis use may vary by state, particularly across the 17 states where cannabis use remains illegal, but also among the 22 states where medical use is legal but recreational use is not. For example, in states where only medical cannabis is legal, we would expect clinician descriptions of patients’ medical use to be skewed toward conditions explicitly permitted under state law, and to include documentation of patients’ eligibility status (e.g., enrollment in a medical cannabis registry or possession of a state-issued “green card”). Similarly, had our definition of medical use required that the description contain an explicit statement regarding *medical* or *clinician-recommended* use, this study would have yielded a much lower prevalence of medical cannabis use. The fact that cannabis use is addressed in routine screening at KPWA may encourage clinicians and patients to discuss

cannabis use more than they otherwise would, and such discussions may increase clinician recognition and documentation of medical use. Though KPWA is a pioneer in conducting annual cannabis screening, trends in state-level legalization and expert recommendations that doctors discuss patient cannabis use—separately from other substance use<sup>30</sup>—may make cannabis screening and identification of patients’ medical cannabis use more common. To ensure representation of the diversity of cannabis-related provider documentation, the corpus was not selected completely randomly. Finally, this study could not address how often clinicians discuss—but fail to document—cannabis use, including medical cannabis use, or when patients were not forthcoming about their cannabis use. We were unable to quantify usage of explicit versus implicit language to document medical use; for efficiency we only ascertained presence/absence of documented medical cannabis use.

## Conclusions

EHR charts of nearly 6% of patients screened for cannabis use contained clinician documentation of medical cannabis use. Clinicians used diverse and ambiguous language to document cannabis use. An automated NLP system was able to accurately classify just over half of this documentation as medical use or not; the remainder was identified by NLP-assisted manual review. Improving automated capture of this information via NLP will require assembling larger training corpora, ideally from multiple and diverse healthcare settings, and may also benefit from application of other modeling approaches including transfer learning. Future improvements in standardization of assessment of cannabis use, including patients’ reasons for use, could improve patient care. Such EHR documentation could be used as well to generate new knowledge about outcomes of medical cannabis use from large prospective cohort studies nested in health systems.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding:

This work was funded by National Institute on Drug Abuse (NIDA) award UG1DA040314 (NIDA Clinical Trials Network Protocol #0077; Lapham, PI). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of NIDA. The NIDA Clinical Trials Network (CTN) Research Development Committee reviewed the study protocol and the NIDA CTN publications committee reviewed and approved the manuscript for publication. The funding organization had no role in the collection, management, analysis, and interpretation of the data or decision to submit the manuscript for publication.

## IRB/Ethics approval:

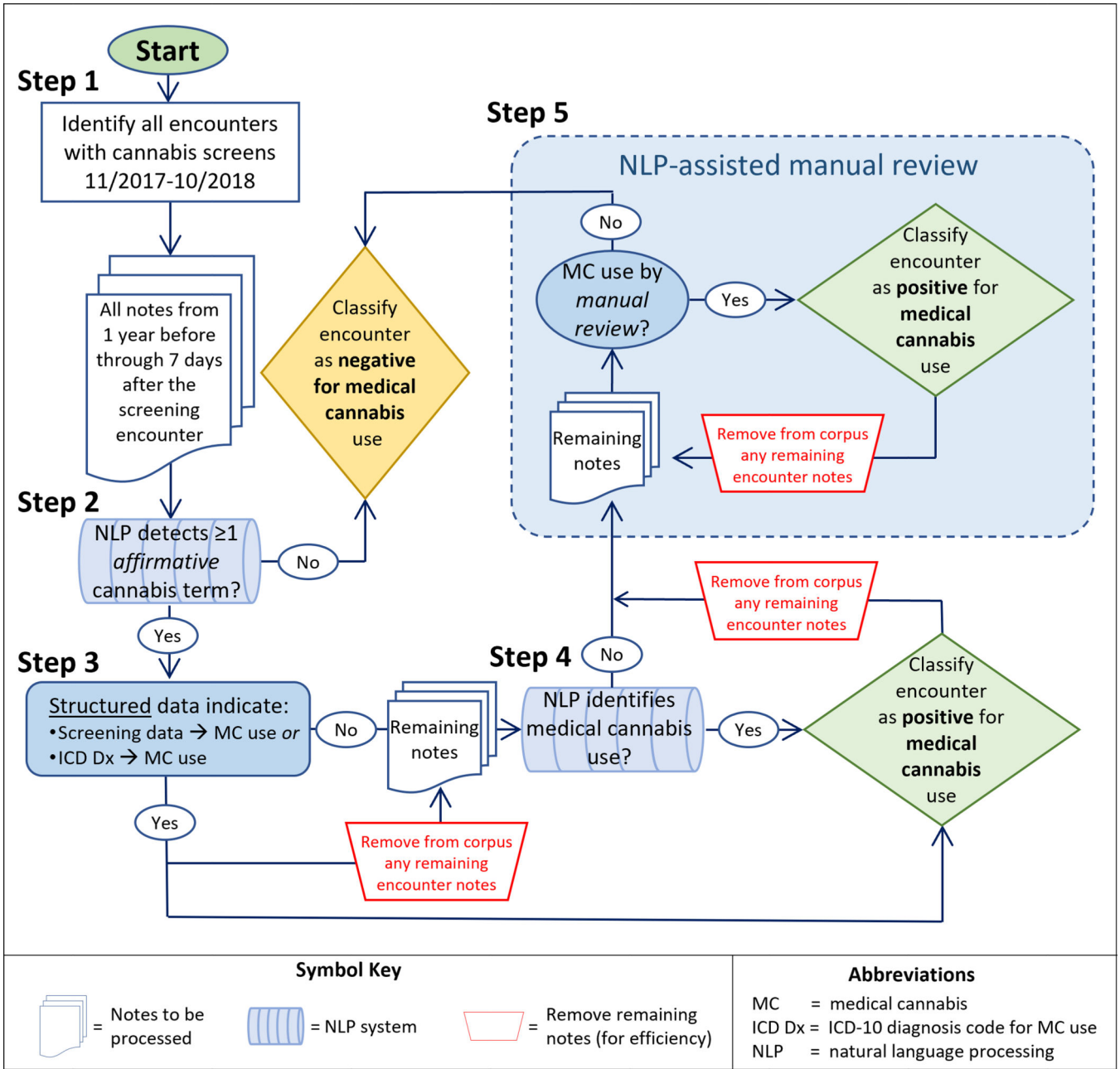
This study, named “Medical cannabis use among primary care patients: using electronic health records to study large populations,” was approved by the Human Subjects Review Committee of Kaiser Permanente Washington Health Research Institute (reference number 1216113–6), which granted the study team a waiver of patient consent to access patient data for the purposes described in the manuscript.

## References

1. National Academies of Sciences Engineering and Medicine. The health effects of cannabis and cannabinoids: the current state of the evidence and recommendations for research. Washington, DC: The National Academies Press, 2017.

2. Ryan-Ibarra S, Induni M, Ewing D. Prevalence of medical marijuana use in California, 2012. *Drug Alcohol Rev* 2015;34(2):141–6. [PubMed: 25255903]
3. Meacham MC, Paul MJ, Ramo DE. Understanding emerging forms of cannabis use through an online cannabis community: An analysis of relative post volume and subjective highness ratings. *Drug Alcohol Depend* 2018;188:364–69. [PubMed: 29883950]
4. Swartz R. Medical marijuana users in substance abuse treatment. *Harm Reduct J* 2010;7:3. [PubMed: 20202221]
5. F BF, W SC, M CD, C C. Identifying Marijuana, Nicotine and Electronic Cigarette References Within a Free-Text Data Field in the Electronic Medical Record Using SAS. *J Patient Cent Res Rev* 2015;2:118.
6. Velupillai S, Mowery DL, Conway M, Hurdle J, Kious B. Vocabulary development to support information extraction of substance abuse from psychiatry notes. Proceedings of the 15th Workshop on Biomedical Natural Language Processing. Berlin, Germany, 2016:92–101.
7. Keyhani S, Vali M, Cohen B, et al. A search algorithm for identifying likely users and non-users of marijuana from the free text of the electronic medical record. *PLoS one* 2018;13(3).
8. Murnion B. Medicinal cannabis. *Aust Prescr* 2015;38(6):212–5. [PubMed: 26843715]
9. Marijuana as Medicine. Secondary Marijuana as Medicine July 2019. [https://www.drugabuse.gov/sites/default/files/marijuanamedicinedrugfacts\\_july2019\\_.pdf](https://www.drugabuse.gov/sites/default/files/marijuanamedicinedrugfacts_july2019_.pdf).
10. Wikipedia. List of Schedule I drugs (US). Secondary List of Schedule I drugs (US) 2020. [https://en.wikipedia.org/wiki/List\\_of\\_Schedule\\_I\\_drugs\\_\(US\)](https://en.wikipedia.org/wiki/List_of_Schedule_I_drugs_(US)).
11. Ebbert JO, Scharf EL, Hurt RT. Medical Cannabis. *Mayo Clin Proc* 2018;93(12):1842–47. [PubMed: 30522595]
12. Health. WSDo. Medical Marijuana: Who Can Authorize. 2020. <https://www.doh.wa.gov/YouandYourFamily/Marijuana/MedicalMarijuana/AuthorizationForm/WhoCanAuthorize>.
13. Whiting PF, Wolff RF, Deshpande S, et al. Cannabinoids for Medical Use: A Systematic Review and Meta-analysis. *JAMA* 2015;313(24):2456–73. [PubMed: 26103030]
14. Lapham GT, Lee AK, Caldeiro RM, et al. Frequency of cannabis use among primary care patients in Washington State. *Journal of the American Board of Family Medicine : JABFM* 2017;30(6):795–805. [PubMed: 29180554]
15. Sayre M, Lapham GT, Lee AK, et al. Routine Assessment of Symptoms of Substance Use Disorders in Primary Care: Prevalence and Severity of Reported Symptoms. *J Gen Intern Med* 2020;35(4):1111–19. [PubMed: 31974903]
16. Matson TE, Carrell DS, Bobb JF, et al. Prevalence of medical cannabis use and associated health conditions documented in electronic health records among primary care patients in Washington State. *JAMA Netw Open* 2021;4(5).
17. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearb Med Inform* 2008:128–44. [PubMed: 18660887]
18. Pestian JP, Deleger L, Savova GK, Dexheimer JW, Solti I. Natural Language Processing – The Basics. In: Hutton JJ, ed. *Pediatric biomedical informatics : computer applications in pediatric research*. New York: Springer, 2012:149–72.
19. Savova GK, Masanz JJ, Ogren PV, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *J Am Med Inform Assoc* 2010;17(5):507–13. [PubMed: 20819853]
20. Savova GK, Deleger L, Solti I, Pestian J, Dexheimer JW. Natural Language Processing: Applications in Pediatric Research. In: Hutton JJ, ed. *Pediatric biomedical informatics : computer applications in pediatric research*. New York: Springer, 2012:173–92.
21. Microsoft. Full-Text Search (SQL Server). Secondary Full-Text Search (SQL Server) 2012. <http://msdn.microsoft.com/en-us/library/ms142571.aspx>.
22. Sun W, Cai Z, Li Y, Liu F, Fang S, Wang G. Data Processing and Text Mining Technologies on Electronic Medical Records: A Review. *J Health Eng* 2018;2018:4302425.
23. Harkema H, Dowling JN, Thornblade T, Chapman WW. ConText: An algorithm for determining negation, experiencer, and temporal status from clinical reports. *J Biomed Inform* 2009;42(5):839–51. [PubMed: 19435614]

24. Juckett D. A method for determining the number of documents needed for a gold standard corpus. *J Biomed Inform* 2012;45(3):460–70. [PubMed: 22245601]
25. Carrell DS, Schoen RE, Leffler DA, et al. Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. *J Am Med Inform Assoc* 2017;24(5):986–91. [PubMed: 28419261]
26. Palmer M, Crosslin D, Carrell DS, Jarvik G, et al. Loci identified by a genome-wide association study of carotid artery stenosis in the eMERGE Network. *Genetic Epidemiology* 2020(In Press).
27. Mehrotra A, Morris M, Gourevitch RA, et al. Physician characteristics associated with higher adenoma detection rate. *Gastrointestinal endoscopy* 2018;87(3):778–86. [PubMed: 28866456]
28. Hazlehurst B, Green CA, Perrin NA, et al. Using natural language processing of clinical text to enhance identification of opioid-related overdoses in electronic health records data. *Pharmacoepidemiol Drug Saf* 2019;28(8):1143–51. [PubMed: 31218780]
29. Developing a standardized MEDical and Recreational Cannabis Use Reasons Evaluation (MERCUR-E) for the National Institute on Drug Abuse Clinical Trials Network. NIDA Clinical Trials Network Steering Committee Meeting; 2019; Bethesda, MD.
30. Bradley KA, Lapham GT, Lee AK. Screening for Drug Use in Primary Care: Practical Implications of the New USPSTF Recommendation. *JAMA Intern Med* 2020;180(8):1050–51. [PubMed: 32515790]



**Figure 1.** Flow diagram of the five steps used to classify notes from screening encounters as positive or negative for medical cannabis use.

**Table 1.**

Characteristics of 203,489 unique patients contributing a total of 299,597 cannabis screens 11/1/2017 through 10/31/2018

Characteristic	Count	Percent
All patients	203,480	100%
Female patients	122,144	60%
Patient age group *		
18–35	52,253	26%
36–65	102,660	50%
66+	48,576	24%
Patient race is white	149,986	80%
Cannabis screens per patient **		
1 during study period	170,990	84%
2+ during study period	32,499	16%
Number of unique screened patients seen per clinician ***		
1–25 unique patients	540	4%
25–100 unique patients	173	14%
101–500 unique patients	388	31%
>500 unique patients	173	14%
Clinical text available per screen ****		
5,000 words	108,688	36%
5,000–20,000 words	121,531	41%
> 20,000 words	64,693	22%
No text available	4,685	2%

\* Age in years at first qualifying screening encounter

\*\* Calendar days with qualifying cannabis screens

\*\*\* For clinicians who saw at least one screened patient, number of *unique patients with any cannabis screens*.

\*\*\*\* Outpatient encounter notes including primary care, behavioral health and urgent care

**Table 2.**

Examples of language clinicians used to characterize medical, non-medical, and ambiguous reasons for cannabis use

Category	Illustrative examples
Medical	<ul style="list-style-type: none"> <li>• She recently began using CBD cannabis oil, daily, about one month ago, and feels it has improved some of her physical symptoms and has alleviated anxiety.</li> <li>• He takes CBD in the form of capsules ongoing for a year to reduce pain and increase his appetite. It has allowed him to reduce morphine dosing.</li> <li>• Hx anxiety, was on Xanax for a couple years, then medical marijuana. Now uses daily</li> <li>• Cannabis ... She has a green card for back pain</li> </ul>
Ambiguous	<ul style="list-style-type: none"> <li>• Occasional recreational marijuana use [is now] daily to try and focus through his headaches and the noise at his work</li> <li>• Smokes marijuana daily, says to “calm down” after work</li> <li>• Reports using marijuana to “keep me solid.”</li> </ul>
Non-medical	<ul style="list-style-type: none"> <li>• He is smoking marijuana occasionally as recreation</li> <li>• She went on a hike (while intoxicated from cannabis) a few days ago</li> <li>• Uses MJ at parties</li> </ul>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript