

# **Recharacterizing Medical Marijuana Documentation within Electronic Health Records: Use of a Customized Smart Data Element and Evaluation with a Systematic Chart Review Protocol**

Donielle Beiler, Aanya Chopra, Christina Gregor, Lorraine D. Tusing, Apoorva Pradhan, Katrina M. Romagnoli, Chadd K. Kraus, Brian J. Piper, Eric A. Wright, Vanessa Troiani

Submitted to: JMIR Formative Research  
on: August 30, 2024

**Disclaimer:** © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 23

..... 23

Figures ..... 24

    Figure 1..... 25

Multimedia Appendixes ..... 26

    Multimedia Appendix 1..... 27

# Recharacterizing Medical Marijuana Documentation within Electronic Health Records: Use of a Customized Smart Data Element and Evaluation with a Systematic Chart Review Protocol

Donielle Beiler<sup>1</sup> BS; Aanya Chopra<sup>2</sup> BS; Christina Gregor<sup>2</sup> BS; Lorraine D. Tusing<sup>2</sup> BS; Apoorva Pradhan<sup>2</sup> MD, MPH; Katrina M. Romagnoli<sup>2,3</sup> PhD; Chadd K. Kraus<sup>4</sup> DO, DrPH, MPH; Brian J. Piper<sup>2,5</sup> PhD; Eric A. Wright<sup>2,6</sup> PharmD, MPH; Vanessa Troiani<sup>1</sup> PhD

<sup>1</sup>Autism and Developmental Medicine Institute Geisinger Lewisburg US

<sup>2</sup>Center for Pharmacy Innovation and Outcomes Geisinger DANVILLE US

<sup>3</sup>Department of Population Health Sciences Geisinger DANVILLE US

<sup>4</sup>Department of Emergency and Hospital Medicine Lehigh Valley Health Network Hazelton US

<sup>5</sup>Department of Education Geisinger Commonwealth School of Medicine Scranton US

<sup>6</sup>Department of Bioethics and Decision Sciences Geisinger DANVILLE US

## Corresponding Author:

Donielle Beiler BS

Autism and Developmental Medicine Institute

Geisinger

120 Hamm Drive

Lewisburg

US

## Abstract

**Background:** Medical Marijuana (MMJ) is available in Pennsylvania and participation in the state-regulated program requires a patient to register and receive a certification by an approved physician. There is currently no integration of MMJ certification data in Pennsylvania into health records that would allow for physicians to rapidly identify patients that are using MMJ, as there are with other scheduled drugs. This absence of a formal data sharing structure necessitates tools that aid in consistent documentation practices to enable comprehensive patient care.

**Objective:** Customized smart data elements (SDE) were made available to clinicians at an integrated health system, Geisinger, following MMJ legalization in Pennsylvania. The purpose of this project was to examine and contextualize the use of MMJ SDEs in the Geisinger population. We accomplished this goal by developing a systematic chart review protocol, with the goal of creating a tool that resulted in consistent human data extraction.

**Methods:** We developed a chart review protocol for extracting MMJ-related information. The protocol was developed between August to December of 2022 and focused on a patient group that received one of several MMJ SDE between 1/25/2019 and 5/26/2022. Characteristics were first identified on a small pilot sample of patients (N=5), which were then iteratively reviewed to optimize for consistency. Following the pilot, two reviewers were assigned 200 patient charts, selected randomly from the larger cohort, with a third reviewer examining a subsample to determine reliability. We then summarized the clinician-level and patient-level features from n=156 charts with a table-format SDE that best captured MMJ information.

**Results:** We found the chart review protocol was feasible for those with minimal medical background to complete, with high inter-rater reliability (Kappa = 0.966 (p < 0.001), 95% CI (0.954 - 0.978)). MMJ certification was largely documented by nurses and medical assistants (87.2%) and typically within primary care settings (68.6%). The SDE has 6 pre-set field prompts, including certifying provider, authorized dispensary, certifying conditions, dosage, product, and active ingredient. We found preset fields were overall well-recorded (76.6% across all fields). Individual fields were more heterogeneous in terms of completion, with dispensary specified in 87.8% of documentation, certifying provider specified in 61.5% of documentation, and product dose specified in only 30.8% of documentation.

**Conclusions:** This method of chart review yields high quality data extraction that can serve as a model for other health record inquiries. Our evaluation showed relatively high completeness of SDE fields, primarily by clinical staff responsible for rooming patients. Improving adoption and fidelity of SDE data collection may present a valuable data source for future research on patient

MMJ use and treatment efficacy and outcomes. Clinical Trial: N/A

(JMIR Preprints 30/08/2024:65957)

DOI: <https://doi.org/10.2196/preprints.65957>

## Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org>

## Original Manuscript

# Recharacterizing Medical Marijuana Documentation within Electronic Health Records: Use of a Customized Smart Data Element and Evaluation with a Systematic Chart Review Protocol

Donielle Beiler<sup>1</sup>, Aanya Chopra<sup>2</sup>, Christina Gregor<sup>2</sup>, Lorraine D. Tusing<sup>2</sup>, Apoorva Pradhan<sup>2</sup>, Katrina M. Romagnoli<sup>2,3</sup>, Chadd K. Kraus<sup>4</sup>, Brian J. Piper<sup>2,5</sup>, Eric A. Wright<sup>2,6</sup>, Vanessa Troiani<sup>1</sup>

1. Department of Autism and Developmental Medicine, Geisinger; 2. Center for Pharmacy Innovation and Outcomes, Geisinger; 3. Department of Population Health Sciences, Geisinger; 4. Department of Emergency and Hospital Medicine, Lehigh Valley Health Network; 5. Department of Medical Education, Geisinger Commonwealth School of Medicine; 6. Department of Bioethics and Decision Sciences, Geisinger

## Abstract

**Background:** Medical Marijuana (MMJ) is available in Pennsylvania and participation in the state-regulated program requires a patient to register and receive a certification by an approved physician. There is currently no integration of MMJ certification data in Pennsylvania into health records that would allow for physicians to rapidly identify patients that are using MMJ, as there are with other scheduled drugs. This absence of a formal data sharing structure necessitates tools that aid in consistent documentation practices to enable comprehensive patient care. Customized smart data elements (SDE) were made available to clinicians at an integrated health system, Geisinger, following MMJ legalization in Pennsylvania. The purpose of this project was to examine and contextualize the use of MMJ SDEs in the Geisinger population. We accomplished this goal by developing a systematic chart review protocol, with the goal of creating a tool that resulted in consistent human data extraction.

**Methods:** We developed a chart review protocol for extracting MMJ-related information. The protocol was developed between August to December of 2022 and focused on a patient group that received one of several MMJ SDE between 1/25/2019 and 5/26/2022. Characteristics were first identified on a small pilot sample of patients (N=5), which were then iteratively reviewed to optimize for consistency. Following the pilot, two reviewers were assigned 200 patient charts, selected randomly from the larger cohort, with a third reviewer examining a subsample to determine reliability. We then summarized the clinician-level and patient-level features from n=156 charts with a table-format SDE that best captured MMJ information.

**Results:** We found the chart review protocol was feasible for those with minimal medical background to complete, with high inter-rater reliability (Kappa = 0.966 (p <0.001), 95% CI (0.954 - 0.978)). MMJ certification was largely documented by nurses and medical assistants (87.2%) and typically within primary care settings (68.6%). The SDE has 6 pre-set field prompts, including certifying provider, authorized dispensary, certifying conditions, dosage, product, and active ingredient. We found preset fields were overall well-recorded (76.6% across all fields). Individual fields were more heterogeneous in terms of completion, with dispensary specified in 87.8% of documentation, certifying provider specified in 61.5% of documentation, and product dose specified in only 30.8% of documentation.

**Conclusion:** This method of chart review yields high quality data extraction that can serve as a model for other health record inquiries. Our evaluation showed relatively high completeness of SDE fields, primarily by clinical staff responsible for rooming patients. Improving adoption and fidelity of SDE data collection may present a valuable data source for future research on patient MMJ use and treatment efficacy and outcomes.

**Keywords:** cannabis; learning health system; Epic; prescription drug monitoring program

Preprint  
JMIR Publications

## Introduction

Since 1996, states across the U.S. have legalized cannabis for medical use[1]. The legalization of cannabis or medical marijuana (**MMJ**) adds another layer of complexity to comprehensive patient care, as treatment with MMJ is managed outside of traditional health care management and documentation systems[2]. In Pennsylvania (PA), MMJ is not currently 'prescribed' - rather, a physician with appropriate privileges must certify that a given patient has one of 24 serious medical conditions[3,4]. This certification can then be used to register and obtain an MMJ card for use at dispensaries where a healthcare provider (typically a pharmacist) is available for consultation/advisement and a patient care specialist assists with product selection for a given symptom or diagnosis[5].

Although MMJ use in PA requires certification by a qualified physician, there is not a standardized integration of data from MMJ purchases with health systems, such as those that exist for prescription drugs[6]. While most states with an MMJ program employ a documentation system for the dispensing of MMJ to patients, these systems are not uniformly integrated with state-managed Prescription Drug Monitoring Programs (PDMP) which are often linked to health system records. The integration of PDMP has been shown to improve health outcomes by allowing for comprehensive and coordinated patient care[7], with more direct integration of the databases associated with increased use by clinicians[8].

Electronic health record (**EHR**) systems digitize patient health records to allow for information sharing between providers, institutions, and insurance companies to improve care continuity and track billing based on the standards for meaningful use outlined in the HITECH Act of 2009[9,10]. The records contain diagnoses, encounter documentation, provider notes, scans and lab results, medications, and patient communication as well as demographic details. Most EHR's contain a standard set of input options that have shown utility for patient care and billing purposes; typically focused on capturing details associated with a specific visit (physical exam and current medication review, for example)[11]. Other components of a patient's record are important for longitudinal care - including a social history section where a range of lifestyle and behavior choices can be documented, as well as familial history of disease[12]. Documentation systems can be helpful to care teams because they prompt patient conversations surrounding these topics and encourage consistent documentation to facilitate longitudinal care[13,14].

In addition to standardized data entry and workflows available within EHR's, many customization options for data entry exist[15,16]. One EHR system, *Epic*<sup>®</sup>, allows for smart data elements (**SDE**) which are triggered by built-in 'smart phrases' that load pre-set text with components that can be edited by health care professionals[17,18]. This improves data entry efficiency and creates coded variables from which data can be extracted in an automated way[19]. The SDE (or SmartText) can be linked to other EHR items like a problem list diagnosis or encounter note. For MMJ use documentation, a customized SDE was created, with individual components mimicking those on product labels from MMJ dispensaries[5]. After entering the smart phrase, the *Epic*<sup>®</sup> user will see each of the individual components (in this case - certifying provider, authorized dispensary, certifying condition(s), dosage product provided, dose, active ingredient) and can input text into each section. Even with structured documentation variables in place, implementation and use can be inconsistent[20].

EHR systems also serve as a wealth of information for retrospective research[21,22]. Large databases can be generated from the fields created and entered by clinical providers and hospital staff. Variables within discrete fields are typically easier to extract from the EHR in an automated way, but there is a great deal of potentially useful information within free text fields that require more sophisticated methods of data extraction. One method to extract useful clinical information from EHR notes is to use a chart review protocol in which a human reviewer reads the chart and records relevant information. We have previously developed chart review protocols that enabled the consistent extraction of relevant text that informed opioid use disorder severity[23] and autonomic arousal dimensions[24], but many different approaches for chart review exist and best practice can be determined based on an individual use case and/or purpose for analysis[25-27]. To ensure high fidelity of the extracted data, it is necessary to develop protocols that enable a reviewer to identify specific responses that can be replicated by other reviewers[28,29].

In the current study, we implemented a systematic chart review protocol to document how the MMJ SDE is being utilized within Geisinger. We developed the protocol and implemented it on a random subset (N=200) of the total Geisinger patient population with an existing marijuana SDE (N=2133).



## Methods

### Study Sample Participants

The work described here was approved by our institutional review board. Chart reviews were completed with a waiver of informed consent. Patients who had an *Epic*® SDE coded for marijuana use (N=2133) between 8/1/2017 and 6/29/2022 and were >18 years of age were identified and eligible for chart review as part of a larger study aim. A random number generator was used to assign a value to each record and sorted numerically to obtain a sample cohort (n=5 for the initial development, and n=200 (100 per reviewer) for full chart review). Following chart review, it was realized that the n=200 patients did not all have the same type of marijuana SDE. Rather, n=156 patients in the cohort had the table-format MMJ SDE that listed 6 discrete elements (certifying provider name and location, certifying condition(s), dispensary, dose, product type, and active ingredient) while the remaining n=44 patients had a second type of social history SDE that coded for marijuana use frequency, method of use, and last use date. As our primary goal was to assess whether the table-format SDE was a useful documentation tool for information that is part of patient registration for medical marijuana in PA, we focused our analysis and summary on the 156 patients that contained the table-formatted MMJ SDE.

### Chart Review Process

The primary questions driving this research were (1) whether the MMJ SDE was consistently capturing MMJ related information, (2) where this SDE was being used within the chart, and (3) whether this SDE was serving its intended purpose of making the patient's use of MMJ easily accessible and available to clinicians across our integrated care system. We created a chart review protocol with the goal of extracting information relevant to these primary research questions, with a secondary goal to optimize the protocol to ensure consistent data extraction across chart reviewers. We used an iterative approach to create a list of variables that could easily be extracted from the EHR through manual review. An initial subset of 5 records was comprehensively reviewed by study team members to determine what MMJ-related data could be gleaned from the EHR, and where in the EHR the desired information could be obtained. These records were reviewed by a clinical and research team to determine a consensus of desirable variables. A detailed workflow was created, highlighting specific areas of the patient record to search for the data of interest (Problem List, Encounters, Scanned Documents, Lab Orders, and Search Terms). Explicit wording and directions were created to guide the chart reviewers through a systematic protocol. Final variables included 61 discrete fields with specific expected responses, and 10 descriptive fields for notes and comments as needed. These variables included MMJ SDE documentation location and details, social history documentation, marijuana diagnoses, documenting provider and department information, primary and secondary diagnoses at the time of MMJ documentation, copies of the MMJ certification, relevant toxicological screen results, presence, or absence of MMJ documentation on subsequent encounters, total EHR length, first documentation of both MMJ interest and MMJ use, and any notable side effects (Appendix Textbox 1).

A data capture tool was initially developed in Microsoft Excel to make it easily accessible for team members with limited data collection experience, as it does not require special permissions or training, and template changes can be implemented by someone without specific expertise. A detailed instruction manual with visual aids was also created to be a step-by-step walkthrough of the chart review process (for example, *"Was a marijuana diagnosis present on the current problem list? Yes or No. If yes, please list the diagnosis name(s), ICD Code(s) and date added to the problem list in the subsequent columns. If no, please note 'N/A' in the subsequent columns"*) (Figure 1). Thirteen of the 61 discrete fields were identified as independent variables, some of which branching logic added or skipped whole subsections of the chart review to streamline data collection. In addition to the Microsoft Excel tool, a REDCap (Research Electronic Data Capture)[30,31] instrument that mirrored the Excel data capture tool was created to enter and store data in the fields as described above and to facilitate sharing of the chart review instrument with other institutions. Once the instruction manual and data capture tools were created, the members of the team who would be completing the larger chart review cohort were tasked with testing the process on that same pilot subset. The

results were compared for similarity and in instances of incongruity, edits were made to the instructions to clarify the expected outcome.

The finalized instruction manual was then used by two reviewers to independently review all charts (each patient chart reviewed by one of the two primary reviewers). To determine inter-rater reliability, a third reviewer completed a duplicate review on a random subset of charts (n=30), repeating the full chart review process while being blind to the previous reviewer's documentation.

## Statistical Analyses

Inter-rater reliability was calculated on the 30 charts reviewed by the third reviewer compared with that of the initial reviewer (comparison charts). A standard inter-rater reliability calculation was applied by comparing the results of each of the 61 discrete variables between reviewers and dividing the number of congruent fields by the total number of discrete fields. As discrete responses could be affected by the independent variables, the same calculation process was applied to only those 13 independent variables. In both cases, range, mean, and standard deviation were calculated.

As a measure of inter-rater reliability, Cohen's kappa calculations were also applied[32]. For each of the comparison charts, the discrete fields of both the primary reviewer and the third reviewer were identified as one of three outcomes: the value matched the EHR following the instructions (F), the value did not match the EHR following the instructions (D), or the value was null (N). The outcomes of each paired response were then combined into a two-letter label (FF, FD, FN, DF, DD, DN, NF, ND, NN). The totals of each two-letter combination were summed across all 61 variables and 30 charts for 1,830 points of comparison. These results were analyzed in R statistical analysis software version 3.6.0 by creating a 3x3 table of the points of comparison and applying the Kappa.test package[33]. This yielded values for relative observed agreement (Po), probability of chance agreement (Pe), Cohen's Kappa (**K**) and p-value.

Data recorded in the chart review variables were grouped together by similar features. Descriptive statistics were calculated for all available features using R[33], including means and ranges for numerical data and frequencies and percentages for categorical data. Demographic characteristics of the cohort are summarized in Table 1, features available from SDE documentation encounters are summarized in Table 2, EHR features, including diagnostic information from the patient visit, are summarized in Table 3, and additional chart review information is in Table 4. Because this is the first summary of MMJ SDE use in this population, we also include breakdowns of patient characteristics separately based on biological sex documented in the EHR.

## Results

### Demographic Characteristics

The cohort of 156 patient records containing the table-formatted MMJ SDE consisted of more females (55.1%) than males (44.9%) and were on average 46.1 years of age. The average length of their EHR was 16.0 years (Table 1). Patients were predominantly white and non-Hispanic, consistent with the demographics of this region of Pennsylvania[34]. Demographic characteristics of the chart review cohort were similar to the larger parent SDE population and the general Geisinger patient population (Appendix Table 1).

**Table 1. Demographic characteristics of the chart review cohort (n=156)**

	$\bar{x}$ ( $\pm$ SD)	Range	
		Min	Max
Age (years)	46.1 ( $\pm$ 15.2)	20	83
EHR Length (years)	16.0 ( $\pm$ 7.8)	1.7	33.3
<b>n (%)</b>			

<b>Sex</b>	
Male	70 (44.9%)
Female	86 (55.1%)
<b>Race</b>	
Black or African American	7 (4.5%)
White	145 (92.9%)
Undisclosed or unspecified	4 (2.6%)
<b>Ethnicity</b>	
Hispanic	6 (3.9%)
Not Hispanic	149 (95.5%)
Undisclosed or unspecified	1 (0.6%)

EHR: electronic health record; SD: standard deviation

### Inter-rater Reliability & Chart Review Completion Characteristics

Average completion time for the chart review protocol was 17.7 min ( $\pm 12.4$  min) per chart ( $n=30$ ; the first 15 charts from each reviewer), with a minimum review time of 6 min and a maximum of 75 min. Standard inter-rater reliability calculations yielded a mean reliability percentage of  $98.0\% \pm 1.7\%$  when all 61 variables were compared between the initial reviewer and the third reviewer for each of the 30 comparison charts. When only the 13 independent variables were assessed, there was a mean reliability of  $95.9\% \pm 4.8\%$ . An assessment of those same 30 comparison charts yielded a kappa value of 0.97, with a 95% confidence interval of 0.95-0.98 and a p-value of  $<0.001$ . The kappa value range is graded on a scale from 0 indicating no agreement to 1 indicating perfect agreement. A score of 0.81-0.99 indicates near perfect agreement[32]. The relative observed agreement ( $P_o$ ) was 0.98 with a probability of chance agreement ( $P_e$ ) of 0.52 across 1830 points of comparison.

### Documentation Characteristics

We found the SDE variables were coded in two primary locations: on the active problem list (10.9%), and in encounter/provider notes (89.1%) (Table 2). SDE entry was largely completed by nurses and medical assistants (87.2%) and typically within primary care settings (68.6%). Licensed Practical Nurses (LPNs) were the provider type most frequently documenting SDEs (52.6%), followed by medical assistants (21.2%), Registered Nurses (RNs) and/or nurse practitioners (14.7%), and doctors/physician assistants (9.0%).

Overall, documentation was completed 76.6% ( $\pm 23.7$ ) of the time across all these elements, with certifying provider name and location specified 61.5% of the time, certifying condition(s) listed 93.6%, dispensary listed 87.8%, product type listed 92.9%, dose of product specified 30.8%, and active ingredient specified in 83.3% of patients. Product type was variable with vape/vaporization identified in 45.5% of documentation, dry leaf/flower 38.5%, and tincture/drops/oil 26.3% of the time. A given patient can use several different products and this variability was captured in the SDE documentation, with an average of 2 product types per patient (ranging from 1-6). The active ingredient was also documented, with THC/CBD in combination in the majority (51.3%) of products, THC only in 29.5%, and CBD only in 2.6%. Dose reported was highly variable in terms of the form of measurement reported, with some specifying amount per day, others specifying time to use product (i.e. vape as needed, but tincture before bed), and others listed as PRN (*pro re nata*, meaning 'as necessary').

A physician can list more than one certifying condition as part of their MMJ registration, which would then appear on the patient's card. Of the 156 charts that contained the table-formatted MMJ SDE, 224 total conditions were included, with the number of conditions ranging from 1-5 per person. 'Severe chronic or intractable pain' was the most common certified condition, with 38.5% of patients having this condition listed, followed by anxiety (37.2%) and post-traumatic stress disorder (26.3%) diagnoses. Most patients had one of the 24 qualifying conditions, but many patients also had a condition listed that was not one of the qualifying conditions (Appendix Table 2).

Table 2. Medical marijuana (MMJ) smart data element (SDE) documentation

	All MMJ SDE Charts (N=156)	Female (n=86)	Male (n=70)
	(%): $\bar{x}(\pm SD)$	(%): $\bar{x}(\pm SD)$	(%): $\bar{x}(\pm SD)$
<b>SDE Completion<sup>a</sup></b>	76.6% ( $\pm 23.7\%$ )	79.8% ( $\pm 19.0\%$ )	72.6% ( $\pm 28.2\%$ )
<b>SDE Variable</b>	<b>n (%)</b>	<b>n (%)</b>	<b>n (%)</b>
<b>Location</b>			
Problem list	17 (10.9%)	9 (10.5%)	8 (11.4%)
Encounter notes	139 (89.1%)	77 (89.5%)	62 (88.6%)
<b>Documenting credentials</b>			
Licensed Practical Nurse (LPN)	82 (52.6%)	44 (51.2%)	38 (54.3%)
Medical Assistant	33 (21.2%)	24 (27.9%)	9 (12.9%)
Registered Nurse (RN) or CRNP	23 (14.7%)	13 (15.1%)	10 (14.3%)
<b>Documenting department</b>			
Family Practice/Primary Care	107 (68.6%)	62 (72.1%)	45 (64.3%)
Gastroenterology	17 (10.9%)	8 (9.3%)	9 (12.9%)
Surgery	12 (7.7%)	6 (7%)	6 (8.6%)
<b>Certifying providers or location</b>			
Included	96 (61.5%)	54 (62.8%)	42 (60.0%)
Unique, n	n= 66	n= 38	n= 36
<b>Dispensaries</b>			
Included	137 (87.8%)	81 (94.2%)	56 (80.0%)
Unique, n	n= 43	n= 30	n= 24
<b>Certifying conditions<sup>b</sup></b>			
Severe chronic or intractable pain	60 (38.5%)	36 (41.9%)	24 (34.3%)
Anxiety	58 (37.2%)	33 (38.4%)	25 (35.7%)
Post-traumatic stress disorder	27 (17.3%)	17 (19.8%)	10 (14.3%)
<b>Product type<sup>b</sup></b>			
Vaporization	71 (45.5%)	39 (45.3%)	32 (45.7%)
Dry leaf/flower	60 (38.5%)	34 (39.5%)	26 (37.1%)
Tincture/drops/oil	41 (26.3%)	30 (34.9%)	11 (15.7%)
<b>Dose</b>			
Specified	48 (30.8%)	29 (33.7%)	19 (27.1%)
Unspecified	108 (69.2%)	57 (66.3%)	51 (72.9%)
<b>Active ingredient</b>			
THC/CBD	80 (51.3%)	48 (55.8%)	32 (45.7%)
THC	46 (29.5%)	24 (27.9%)	22 (31.4%)
CBD	4 (2.6%)	4 (4.7%)	0 (0.0%)

<sup>a</sup>Individual SDE component completion that are not listed above include: Certifying condition, 93.6%; product type, 92.9%; and active ingredient, 83.3%

<sup>b</sup>some patients report more than one certifying condition and/or multiple dosage products

CBD: cannabidiol; CRNP: Certified Registered Nurse Practitioner; THC: tetrahydrocannabinol

## Summary of Features of the Encounter on the date of the SDE

One of the anticipated features of the SDE for MMJ is that when paired with a problem list diagnosis, the diagnosis and SDE would then remain on the active problem list, enabling any clinician providing care within all departments to be immediately aware of the MMJ treatment. However, in practice, the SDE can be added separate from a marijuana use diagnosis to other locations in the record, including social history, encounter, and progress notes, or anywhere else in the chart that allows free text entry. We explored documentation and diagnoses present in the record before, during and after the date the SDE was entered (Table 3).

We found that nearly half of records (48.1%; n=75) had some sort of marijuana diagnosis listed on their problem list *prior* to the SDE encounter; this diagnosis sometimes preceded the implementation of the hospital wide SDE and/or could have been entered by a different clinician or at a different clinic. At the SDE encounter, MMJ or other marijuana use was never a primary reason for the visit and was the secondary diagnosis in only 17.9% of the cohort (n=28), although some listed other substance use/abuse related conditions, including opioid use disorder, and/or alcohol-induced cirrhosis/pancreatitis a primary diagnosis. We found that the marijuana use diagnosis remained on the problem list in 74.4% of the next completed encounters.

Most of the primary and secondary diagnoses for the encounter where the SDE's were entered were part of routine care (physical exam, follow-up for history of a given disorder, and screening/lab testing), but some of these encounters included primary diagnoses for pain, diabetes, obesity, digestive and/or cholesterol disorders. Common secondary diagnoses also included psychiatric disorders, endocrine system and/or other signs and symptoms not otherwise classified. Overall, this indicates MMJ use can be brought up in a variety of contexts of a primary care visit.

The social history tab of a patient's chart is supposed to be reviewed/confirmed at every patient visit. Questions include 'Do you drink? Do you smoke? Do you have any history of drug use?'. There is also some additional branching logic that can be utilized to document more specific information in free text. Of the chart review patient cohort, 42.3% of the patients were marked 'yes' for drug use at the SDE encounter, while 39.1% were marked no/never/not currently. Marijuana use, specifically, was noted in the social history of 43.6% of all records in the cohort, marked with either 'yes' or 'not currently'. The information noted in the social history tab of these patients suggests that providers have mixed perspectives on whether MMJ should be characterized as 'drug use' in patient social history.

**Table 3. Encounter characteristics from the SDE documentation date**

Encounter Variable	All Charts (N=156) n (%)	MMJ SDE (n=86) n (%)	Female (n=70) n (%)
<b>Department</b>			
Family Practice/Primary Care	107 (68.6%)	62 (72.1%)	45 (64.3%)
Gastroenterology	17 (10.9%)	8 (9.3%)	9 (12.9%)
Surgery	12 (7.7%)	6 (7.0%)	6 (8.6%)
<b>Primary diagnosis<sup>a</sup></b>			
[Z00-Z99]	34 (21.8%)	16 (18.6%)	18 (25.7%)
[G00-G99; M00-M99]	21 (13.5%)	13 (15.1%)	8 (11.4%)
[R00-R99]	21 (13.5%)	15 (17.4%)	6 (8.6%)
[K00-K95]	17 (10.9%)	8 (9.3%)	9 (12.9%)
<b>Secondary diagnosis<sup>b</sup></b>			
[Z00-Z99]	73 (46.8%)	41 (47.7%)	32 (45.7%)
[F00-F99]	60 (38.5%)	35 (40.7%)	25 (35.7%)
[E00-E89]	46 (29.5%)	28 (32.6%)	18 (25.7%)
[R00-R99]	46 (29.5%)	27 (31.4%)	19 (27.1%)

<b>Marijuana (MJ) diagnosis</b>			
On the problem list prior to the SDE encounter	75 (48.1%)	37 (43.0%)	38 (54.3%)
MMJ in secondary diagnoses of the SDE encounter	28 (17.9%)	10 (11.6%)	18 (25.7%)
On the problem list at the next completed encounter	116 (74.4%)	61 (70.9%)	55 (78.6%)
<b>“Drug Use” documentation</b>			
“Yes”	66 (42.3%)	40 (46.5%)	26 (37.1%)
“Not Currently”	8 (5.1%)	4 (4.7%)	4 (5.7%)
“No” or “Never”	53 (34.0%)	26 (30.2%)	27 (38.6%)
Marijuana specified	68 (43.6%)	41 (47.7%)	27 (38.6%)

<sup>a</sup>[E00-E89] Endocrine, nutritional, and metabolic diseases; [F00-F99] Mental, behavioral and other substance use disorders; [G00-G99; M00-M99] Musculoskeletal and nervous system disorders; [K00-K95] Diseases of the digestive system; [R00-R99] Symptoms, signs, and abnormal findings, not elsewhere classified; [Z00-Z99] Primary care, routine physical exam, or conditions not otherwise noted

<sup>b</sup>Average number of secondary diagnoses per patient when present:  $5.0 \pm 3.9$  (1-18). 30/156 records did not have a secondary diagnosis.

## Summary of Additional Chart Review features in the context of the SDE (before and after SDE documentation)

We explored each patient’s health record for additional context surrounding the SDE documentation, including first mention of MMJ use and/or requests for information in the record (Table 4). The first record of interest in MMJ and MMJ use was most commonly documented in primary care settings (49.4% first interest and 46.2% first use), followed by surgery (7.7% first interest and 7.1% first use). The first record of MMJ interest and/or use was mostly documented by physicians (39.1% in each case) and physician assistants (20.5%/16.7%), whereas actual documentation of SDE components with more specific use information was primarily done by clinical rooming staff. We also assessed whether urine toxicology screens had been completed on the patient at prior visits and whether these were positive for marijuana. We found that 54.5% of the chart review cohort had a drug screening within their health record, with 82.4% of those toxicology screens being positive for marijuana. Urine drug screens could have occurred at any point in time in the patient record. The presence of urine toxicology screens in most patients in this randomly selected portion of the SDE cohort suggests that many patients were using marijuana prior to reporting the use to their doctor. These results also suggest that the EHR may be a useful source for a future retrospective analysis of marijuana use in patients prior to legalization in PA.

**Table 4. Additional chart review data**

	<b>All Charts (N=156)</b>	<b>Female (n=86)</b>	<b>Male (n=70)</b>
	<b>n (%)</b>	<b>n (%)</b>	<b>n (%)</b>
<b>MMJ card scanned</b>	11 (7.1%)	7 (8.1%)	4 (5.7%)
<b>Effect or side effect specified</b>	62 (39.7%)	34 (39.5%)	28 (40.0%)
<b>MJ diagnosis on the current active problem list</b>	127 (81.4%)	68 (79.1%)	59 (84.3%)
MMJ	124 (97.6%) <sup>a</sup>	67 (98.5%) <sup>a</sup>	57 (96.6%) <sup>a</sup>
Other MJ	3 (2.4%) <sup>a</sup>	1 (1.5%) <sup>a</sup>	2 (3.4%) <sup>a</sup>
Dx date different from SDE date	88 (69.3%) <sup>a</sup>	47 (69.1%) <sup>a</sup>	41 (69.5%) <sup>a</sup>
Diagnosing department			
Family Practice/Primary Care	81 (63.8%) <sup>a</sup>	44 (64.7%) <sup>a</sup>	37 (62.7%) <sup>a</sup>

<i>Emergency Medicine</i>	7 (5.5%) <sup>a</sup>	5 (7.4%) <sup>a</sup>	2 (3.4%) <sup>a</sup>
<i>Surgery</i>	3 (2.4%) <sup>a</sup>	2 (2.9%) <sup>a</sup>	1 (1.7%) <sup>a</sup>
<b>Tox Screens</b>			
Tox screens present	85 (54.5%)	36 (41.9%)	49 (70.0%)
Tox screens + for MJ	70 (82.4%) <sup>b</sup>	26 (72.2%) <sup>b</sup>	44 (89.8%) <sup>b</sup>
<b>First mention of interest in MMJ</b>			
Department			
<i>Family Practice/Primary Care</i>	77 (49.4%)	44 (51.2%)	33 (47.1%)
<i>Surgery</i>	12 (7.7%)	5 (5.8%)	7 (10.0%)
<i>Emergency Medicine</i>	7 (4.5%)	4 (4.7%)	3 (4.3%)
<i>General Internal Medicine</i>	7 (4.5%)	5 (5.8%)	2 (2.9%)
Provider type			
<i>Physician (MD/DO)</i>	61 (39.1%)	32 (37.2%)	29 (41.4%)
<i>Physician's Assistant (PA-C)</i>	32 (20.5%)	18 (20.9%)	14 (20.0%)
<i>Licensed Practical Nurse (LPN)</i>	24 (15.4%)	14 (16.3%)	10 (14.3%)
<b>First record of MMJ use</b>			
Department			
<i>Family Practice/Primary Care</i>	72 (46.2%)	41 (47.7%)	31 (44.3%)
<i>Surgery</i>	11 (7.1%)	5 (5.8%)	6 (8.6%)
<i>Emergency Medicine</i>	6 (3.8%)	4 (4.7%)	2 (2.9%)
Provider type			
<i>Physician (MD/DO)</i>	61 (39.1%)	33 (38.4%)	28 (40.0%)
<i>Physician's Assistant (PA-C)</i>	26 (16.7%)	15 (17.4%)	12 (17.1%)
<i>Licensed Practical Nurse (LPN)</i>	24 (15.4%)	14 (16.3%)	9 (12.9%)

<sup>a</sup>Percent is out of those with a marijuana diagnosis present on the current active problem list

<sup>b</sup>Percent is out of those with tox screens present

MJ: marijuana

## Discussion

This research had several primary goals, including to determine (1) whether the SDE was consistently capturing MMJ related information, (2) where this SDE was being used within the chart, and (3) whether the SDE was serving its intended purpose of making the patient's use of MMJ easily accessible and available to clinicians across our integrated care system. We do find that customized SDE's can be utilized for documentation of MMJ use, and that there was reasonably consistent capture of the individual SDE fields when in use. Overall, our results confirm that SDE's have potential to make specific information relatively easy to record and find for future reference. While there was variance in use across clinicians, a concerted effort to educate clinicians and rooming staff on best practices for using the SDE may address this heterogeneity[35]. We found that the SDE was primarily used within an encounter note, rather than as part of a problem list diagnosis. While the presence of the SDE within an encounter note can be helpful for research such as this, the SDE may be more useful clinically if implemented consistently as part of a problem list, so that all clinical providers can access the information quickly as part of a new encounter.

To answer our driving questions, we developed and implemented a systematic chart review protocol. Chart review can be utilized for many purposes, including describing symptoms and prevalence of specific conditions[36,37], risk assessment[38], prediction modeling[39], and as the basis for informing natural language processing and machine learning algorithms[40,41]. We and others have started to make use of more systematic data extraction from patient charts[42]. We have demonstrated here and elsewhere[23,24] that establishing a chart review protocol can result in a highly reliable process that allows for human contextualization of chart information and is also scalable, with average review less than 20 minutes per

patient. By creating the protocol with the input and guidance from clinical stakeholders, we can translate important clinical details into a format that allows for non-experts to reliably perform the review. This point may be important for future use of this type of chart review protocol, particularly for chart review surrounding clinical case/control and natural language processing algorithm development[43,44]. That is, many chart reviews are completed to inform algorithms that are designed to automatically characterize a given patient as a case or control for a specific diagnosis (schizophrenia, for example)[45]. Validation of these algorithms typically involve confirmation as case or control by highly trained clinicians that review the entire patient chart[46]. Typically, additional supporting information is not recorded as part of these reviews and most of the details and clinical expertise required for contextualization are not documented. We show that individuals without clinical expertise can be trained in such a way to search for and identify information within the chart that is relevant to clinical characterization. Future chart reviews that are focusing on diagnostic algorithm evaluation may want to employ a similar process. Beyond use for research algorithms, this procedure may be helpful for training individuals responsible for clinical documentation, such as medical scribes, as previous work indicates wide variability in scribe note structure[47].

This work is not without limitations. We have performed this chart review on a unique population within central and northeast PA that has sought care at an integrated health system, Geisinger. This type of SDE documentation may not be useful if used in EHRs that are not part of integrated care settings. In addition, the data reported here are input into the record by clinical providers and hospital staff and reported by the patient, both of which are prone to human error. While standardization offered by features such as an SDE can guide more consistent data entry, we cannot determine using this analysis whether all the data are accurate, as entered. Further, chart review can be time consuming compared to more automated electronic data extraction. While our chart review protocol attempts to implement a highly reliable process that is also efficient, some of the chart reviews can still take over an hour for a given patient. Finally, this data extraction and chart review focused on individuals >18 and thus does not include documentation of pediatric patients; future work should examine whether the SDE is utilized within pediatric settings with similar fidelity.

The model of chart review described here yielded high-quality data extraction, demonstrating its potential as a prototype for other chart review protocols. We find that the completeness of SDE fields was relatively high and primarily completed by clinical rooming staff. This finding suggests that improving the adoption and fidelity of SDE data collection could provide a valuable data source for consistent documentation of an alternative treatment that is typically not tracked using a formalized drug monitoring system. By leveraging this model of SDE documentation, insights can then be gained into the patterns, trends, and outcomes associated with MMJ use in a clinical setting. This information can also inform the further development of evidence-based guidelines for MMJ use and contribute to a better understanding of its therapeutic potential[48]. Additionally, the model can be adapted to study other areas of healthcare, facilitating the extraction of high-quality data from electronic health records, and contributing to advancements in clinical research more generally.



## Acknowledgements

Special thanks to Julia Soares and Jove Graham for their work on chart review and cohort selection, respectively. Additional thanks to Antionette Sabatino DiCriscio for her assistance in formatting and manuscript review.

The study is funded by Ascend Wellness Holdings, Clinical Registrant through the Geisinger Academic Clinical Research Center IRB# 2022-0498. The funder was not involved in the study or the decision to submit for publication.

## Conflicts of Interest

none declared

## Abbreviations

CBD: cannabidiol

EHR: electronic health record

ICD: International Classification of Diseases

MJ: marijuana

MMJ: medical marijuana

SD: standard deviation

SDE: smart data element

THC: tetrahydrocannabinol

## References

1. California Office of the Attorney General. (2022, December 30). Medicinal Cannabis Guidelines. State of California, Department of Justice, Office of the Attorney General. <https://oag.ca.gov/medicinal-cannabis#:~:text=In%201996%2C%20California%20voters%20approved>
2. Sajdeya, R, Goodin, A J, Tighe, P J. (2021). Cannabis Use Assessment and Documentation in Healthcare: Priorities for Closing the Gap. Preventive Medicine, 153, 106798. DOI:10.1016/j.ypmed.2021.106798
3. Commonwealth of Pennsylvania. (2019). Medical Marijuana Patient and Caregiver Resources. Pennsylvania Department of Health. <https://www.health.pa.gov/topics/programs/Medical%20Marijuana/Pages/Patients.aspx>
4. Commonwealth of Pennsylvania. (2019). Pennsylvania Medical Marijuana Program Guide. Pennsylvania Department of Health. <https://www.health.pa.gov/topics/programs/Medical%20Marijuana/Pages/Medical%20Marijuana.aspx>

5. Pennsylvania General Assembly. (2016, April 17). 2016 Act 16. <https://www.legis.state.pa.us/cfdocs/legis/li/uconsCheck.cfm?yr=2016&sessInd=0&act=16>
6. Commonwealth of Pennsylvania. (2024). Prescription Drug Monitoring Program (PDMP). Pennsylvania Department of Health. <https://www.health.pa.gov/topics/programs/PDMP/pages/pdmp.aspx>
7. Al-Astal, A, Sodhi, K, & Lakhani, H. (2022). Optimization of Prescription Drug Monitoring Program to Overcome Opioid Epidemic in West Virginia. *Cureus*, 14(2). DOI:10.7759/cureus.22434
8. Neprash, H T, Vock, D M, Hanson, A, Elert, B, Short, S, Karaca-Mandic, P, Rothman, A J, Melton, G B, Satin, D, Markowitz, R, Golberstein, E. (2022). Effect of Integrating Access to a Prescription Drug Monitoring Program Within the Electronic Health Record on the Frequency of Queries by Primary Care Clinicians. *JAMA Health Forum*, 3(6), e221852. DOI:10.1001/jamahealthforum.2022.1852
9. Burde, H. (2011). THE HITECH ACT: An Overview. *AMA Journal of Ethics*, 13(3), 172–175. DOI:10.1001/virtualmentor.2011.13.3.hlwa1-1103.
10. HITECH Act of 2009, 42 USC sec 139w-4(0)(2).(2009). Sec 13301, Subtitle A: Promotion of Health Information Technology.
11. Subcommittee E31.25 on Healthcare Data Management, Security, Confidentiality, and Privacy. (2017). Practice for Content and Structure of the Electronic Health Record (EHR). ASTM International, 14.01, 105. DOI:10.1520/e1384-02
12. Wang, M, Pantell, M S, Gottlieb, L M, Adler-Milstein, J. (2021). Documentation and Review of Social Determinants of Health Data in the EHR: Measures and Associated Insights. *Journal of the American Medical Informatics Association*, 28(12), 2608–2616. DOI:10.1093/jamia/ocab194
13. Alanazi, B, Butler-Henderson, K, Alanazi, M R. (2019). The Role of Electronic Health Records in Improving Communication Between Health Professionals in Primary Healthcare Centres in Riyadh: Perception of Health Professionals. *Ebooks.iospress.nl*; IOS Press. <https://ebooks.iospress.nl/volumearticle/52037>
14. Duke, P, Frankel, R M, Reis, S. (2013). How to Integrate the Electronic Health Record and Patient-Centered Communication Into the Medical Visit: A Skills-Based Approach. *Teaching and Learning in Medicine*, 25(4), 358–365. DOI:10.1080/10401334.2013.827981
15. Guo, U, Chen, L, Mehta, P H. (2017). Electronic Health Record Innovations: Helping Physicians – One Less Click at a Time. *Health Information Management Journal*, 46(3), 140–144. DOI:10.1177/1833358316689481
16. Kodama, K, Konishi, S, Manabe, S, Okada, K, Yamaguchi, J, Wada, S, Sugimoto, K, Itoh, S, Takahashi, D, Kawasaki, R, Matsumura, Y, Takeda, T. (2023). Impact of Electronic Medical Record-connected Questionnaire on Efficient Nursing Documentation: A Study of Usability and Efficacy (Preprint). *JMIR Nursing*, 2023(6), e51303–e51303. DOI:10.2196/51303
17. Epic Systems Corporation. EHI Export Specification SmartData. Open.epic.com. Copyright 2024. Retrieved March 14, 2024, from [https://open.epic.com/EHITables/GetTable/CLARITY\\_CONCEPT.htm#:~:text=SmartData%20elements%20are%20discrete%20data](https://open.epic.com/EHITables/GetTable/CLARITY_CONCEPT.htm#:~:text=SmartData%20elements%20are%20discrete%20data)
18. Epic Systems Corporation. (2021, February 16). Smart EHR Design Saves Clinicians 7,000 Hours. <https://www.epic.com/epic/post/smart-ehr-design-saves-clinicians-7000-hours/>
19. Fuller, S. (2014). PS1-45: Get Smart: Finding Structured Data Using SmartTools in Clarity. *Clinical Medicine & Research*, 12(1-2), 95–95. DOI:10.3121/cmr.2014.1250.ps1-45
20. Ebberts, T, Kool, R B, Smeele, L E, Dirven, R, den Besten, C A, Karssemakers, L H E, Verhoeven, T, Herruer, J M, van den Broek, G B, Takes, R P. (2022). The Impact of Structured and Standardized Documentation on Documentation Quality; a Multicenter, Retrospective Study. *Journal of Medical Systems*, 46(7). DOI:10.1007/s10916-022-01837-9
21. Ford, E, Kim, J, Kharrazi, H, Gleason, K, Gumas, D, Decamp, L, Hopkins, J. (2017). A Guide to Using Data from EPIC, MyChart, and Cogito for Behavioral, Social and Systems Science Research

Prepared

for.

[https://ictr.johnshopkins.edu/wp-content/uploads/2018/06/Phase1.Epic\\_Social.Guide\\_2017.06.30\\_ef.pdf](https://ictr.johnshopkins.edu/wp-content/uploads/2018/06/Phase1.Epic_Social.Guide_2017.06.30_ef.pdf)

22. Cowie, M R, Blomster, J I, Curtis, L H, Duclaux, S, Ford, I, Fritz, F, Goldman, S, Janmohamed, S, Kreuzer, J, Leenay, M, Michel, A, Ong, S, Pell, J P, Southworth, M R, Stough, W G, Thoenes, M, Zannad, F, Zalewski, A. (2017). Electronic Health Records to Facilitate Clinical Research. *Clinical Research in Cardiology*, 106(1), 1–9. DOI:10.1007/s00392-016-1025-6
23. Palumbo, S A, Adamson, K M, Krishnamurthy, S, Manoharan, S, Beiler, D, Seiwel, A, Young, C, Metpally, R, Crist, R, Doyle, G, Ferraro, T, Li, M, Berrettini, W, Robishaw, J, Troiani, V. (2020). Assessment of Probable Opioid Use Disorder Using Electronic Health Record Documentation. *JAMA Network Open*, 3(9), e2015909. DOI:10.1001/jamanetworkopen.2020.15909
24. DiCriscio, A S, Beiler, D, Smith, J, Asdell, P, Dickey, S, DiStefano, M, Troiani, V. (2023). Assessment of Autonomic Symptom Scales in Patients with Neurodevelopmental Diagnoses Using Electronic Health Record Data. *Research in Autism Spectrum Disorders*, 108, 102234. DOI:10.1016/j.rasd.2023.102234
25. Rule, A, Chiang, M F, Hribar, M R. (2019). Using Electronic Health Record Audit Logs to Study Clinical Activity: A Systematic Review of Aims, Measures, and Methods. *Journal of the American Medical Informatics Association*, 27(3), 480–490. DOI:10.1093/jamia/ocz196
26. Allison, J J, Wall, T C, Spettell, C M, Calhoun, J, Fargason, C A, Kobylinski, R W, Kiefe, C. (2000). The Art and Science of Chart Review. *The Joint Commission Journal on Quality Improvement*, 26(3), 115–136. DOI:10.1016/S1070-3241(00)26009-4
27. Gilbert, E H, Lowenstein, S R, Koziol-McLain, J, Barta, D C, Steiner, J. (1996). Chart Reviews In Emergency Medicine Research: Where Are The Methods? *Annals of Emergency Medicine*, 27(3), 305–308. DOI:10.1016/s0196-0644(96)70264-0
28. Liddy, C, Wiens, M, Hogg, W. (2011). Methods to Achieve High Interrater Reliability in Data Collection From Primary Care Medical Records. *The Annals of Family Medicine*, 9(1), 57–62. DOI:10.1370/afm.1195
29. Hogg, W, Gyorf-Dyke, E, Johnston, S, Dahrouge, S, Liddy, C, Russell, G, Kristjansson, E. (2010). Conducting Chart Audits in Practice-based Primary Care Research: A User's Guide. *Canadian Family Physician*, 56(5), 495–496. PMC2868621.
30. Harris, PA, Taylor, R, Thielke, R, Payne, J, Gonzalez, N, Conde, JG. (2009). Research Electronic Data Capture (REDCap) – A Metadata-driven Methodology and Workflow Process for Providing Translational Research Informatics Support, *J Biomed Inform.* 2009 Apr;42(2):377-81. DOI:10.1016/j.jbi.2008.08.010
31. Harris, PA, Taylor, R, Minor, BL, Elliott, V, Fernandez, M, O'Neal, L, McLeod, L, Delacqua, G, Delacqua, F, Kirby, J, Duda, SN. (2019). REDCap Consortium, The REDCap Consortium: Building an International Community of Software Partners, *J Biomed Inform.* 2019 May 9. DOI:10.1016/j.jbi.2019.103208
32. McHugh, M L. (2012). Interrater Reliability: the Kappa Statistic. *Biochemia Medica*, 22(3), 276–282. PMC3900052
33. R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
34. U.S. Census Bureau (2020) Pennsylvania Race and Hispanic Origin, Population Census, April 1, 2020. Retrieved from <https://census.gov/>
35. Lee, W W, Alkureishi, M L, Isaacson, J H, Mayer, M, Frankel, R M, London, D A, Wroblewski, K E, Arora, V M. (2018). Impact of a Brief Faculty Training to Improve Patient-centered Communication while using Electronic Health Records. *Patient Education and Counseling*, 101(12), 2156–2161. DOI:10.1016/j.pec.2018.06.020
36. Altman, J, Padilla, C, Merchant, A, Freshwater, K, Weinsztok, S, Clugston, J R, Fournier, K,

- Edenfield, K M. (2021). COVID-19 Prevalence and Presenting Symptoms in a College Student Population: A Retrospective Chart Review. *Journal of American College Health*, 1–5. DOI:10.1080/07448481.2021.1926270
37. Mamman, R, Cheng, A, Tsow, R, Schmidt, J. (2022). Clinician Reports of Self-awareness After Traumatic Brain Injury: a Retrospective Chart Review. *BMC Health Services Research*, 22(1). DOI:10.1186/s12913-022-08444-x
38. Han, C, Redley, B, Lin, C, Chen, L, Ng, C J, Wang, L. (2021). Standard Risk Assessment Forms Help Reduce Preventable Harm in Older Adult Inpatients: A Retrospective Chart Review Study. *Journal of Advanced Nursing*, 77(11). DOI:10.1111/jan.14931
39. Bernstein, S, Gilson, S, Zhu, M, Nathan, A G, Cui, M, Press, V G, Shah, S, Zarei, P, Laiteerapong, N, Huang, E S. (2023). Diabetes Life Expectancy Prediction Model Inputs and Results From Patient Surveys Compared With Electronic Health Record Abstraction: Survey Study. *JMIR Aging*, 6(1), e44037. DOI:10.2196/44037
40. Zheng, C, Duffy, J, Liu, I A, Sy, L S, Navarro, R A, Kim, S S, Ryan, D S, Chen, W, Qian, L, Mercado, C, Jacobsen, S J. (2022). Identifying Cases of Shoulder Injury Related to Vaccine Administration (SIRVA) in the United States: Development and Validation of a Natural Language Processing Method. *JMIR Public Health and Surveillance*, 8(5), e30426. DOI:10.2196/30426
41. Hu, Z, Melton, G B, Moeller, N D, Arsoniadis, E G, Wang, Y, Kwaan, M R, Jensen, E H, Simon, G J. (2017). Accelerating Chart Review Using Automated Methods on Electronic Health Record Data for Postoperative Complications. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, 2016, 1822–1831.
42. Halevy, N, Elias, B, Shilo, S, Nidal Muhanna, Handzel, O, Yahav Oron, Rani Abu Eta, Ungar, O J. (2022). Real life safety of systemic steroids for sudden sensorineural hearing loss: a chart review. *European Archives of Oto-Rhino-Laryngology*, 279(10), 4787–4792. DOI:10.1007/s00405-022-07264-3
43. Castro, V M, Apperson, W K, Gainer, V S, Ananthakrishnan, A N, Goodson, A P, Wang, T D, Herrick, C D, Murphy, S N. (2014). Evaluation of Matched Control Algorithms in EHR-based Phenotyping Studies: A Case Study of Inflammatory Bowel Disease Comorbidities. *Journal of Biomedical Informatics*, 52, 105–111. DOI:10.1016/j.jbi.2014.08.012
44. Noori, A, Magdamo, C, Liu, X, Tyagi, T, Li, Z, Kondepudi, A, Alabsi, H, Rudmann, E, Wilcox, D, Brenner, L, Robbins, G K, Moura, L, Zafar, S, Benson, N M, Hsu, J, R Dickson, J, Serrano-Pozo, A, Hyman, B T, Blacker, D, Westover, M B. (2022). Development and Evaluation of a Natural Language Processing Annotation Tool to Facilitate Phenotyping of Cognitive Status in Electronic Health Records: Diagnostic Study. *Journal of Medical Internet Research*, 24(8), e40384. DOI:10.2196/40384
45. Zheutlin, A B, Dennis, J, Karlsson Linnér, R, Moscati, A, Restrepo, N, Straub, P, Ruderfer, D, Castro, V M, Chen, C, Ge, T, Huckins, L M, Charney, A, Kirchner, H L, Stahl, E A, Chabris, C F, Davis, L K, Smoller, J W. (2019). Penetrance and Pleiotropy of Polygenic Risk Scores for Schizophrenia in 106,160 Patients Across Four Health Care Systems. *American Journal of Psychiatry*, 176(10), 846–855. DOI:10.1176/appi.ajp.2019.18091085
46. Nolan, M, Siwani, R, Helmi, H, Pickering, B, Moreno-Franco, P, Herasevich, V. (2017). Health IT Usability Focus Section: Data Use and Navigation Patterns among Medical ICU Clinicians during Electronic Chart Review. *Applied Clinical Informatics*, 08(04), 1117–1126. DOI:10.4338/aci-2017-06-ra-0110
47. Pranaat, R, Mohan, V, O'Reilly, M, Hirsh, M, McGrath, K, Scholl, G, Woodcock, D, Gold, J A. (2017). Use of Simulation Based on an Electronic Health Records Environment to Evaluate the Structure and Accuracy of Notes Generated by Medical Scribes: Proof-of-Concept Study. *JMIR Medical Informatics*, 5(3), e30. DOI:10.2196/medinform.7883
48. Walsh, Z, Callaway, R, Belle-Isle, L, Capler, R, Kay, R, Lucas, P, Holtzman, S (2013). Cannabis for therapeutic purposes: Patient characteristics, access, and reasons for use. *International Journal of Drug*

Policy, 24(6), 511–516. DOI:10.1016/j.drugpo.2013.08.010

Preprint  
JMIR Publications

## Supplementary Files

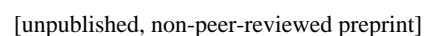
Untitled.

URL: <http://asset.jmir.pub/assets/6419a8bd06092360dc5dfe13d17c9033.docx>

## Figures



<https://preprints.jmir.org/preprint/65957>



## Multimedia Appendixes

Complete list of chart review variables and branching logic.

URL: <http://asset.jmir.pub/assets/44a817404b0e44ab456a3fd560403c8a.docx>

