Full title: Social media discourse and internet search queries on cannabis as a medicine: A systematic review

Short title: Social media discourse and search queries on cannabis as medicine

Sedigheh Khademi Habibabadi[1*], Yvonne Ann Bonomo[1,2, #b], Mike Conway[3, #c], and Christine Mary Hallinan[1, #a, *]

[1]Department of General Practice, Melbourne Medical School, Faculty of Medicine, Dentistry & Health Sciences, The University of Melbourne, Victoria, Australia

[2]Centre for Digital Transformation of Health, Victorian Comprehensive Cancer Centre, The University of Melbourne, Victoria, Australia

[3]St Vincent's Health - Department of Addiction Medicine, Melbourne, Victoria, Australia

[#a] Current address: Department of General Practice, Melbourne Medical School, Faculty of Medicine, Dentistry & Health Sciences of the University of Melbourne, Level 3, 780 Elizabeth Street, Carlton, Victoria, 3053

[#b] Current address: St Vincent's Health - Department of Addiction Medicine 38 Fitzroy Street, Fitzroy, Victoria, 3065

[#c] Current address: Centre for Digital Transformation of Health, Victorian Comprehensive Cancer Centre, The University of Melbourne, Victoria, Australia

[*] Corresponding authors

Email: Sedigh.Khademi@unimelb.edu.au; hallinan@unimelb.edu.au

Abstract

Objective: To systematically review the current published literature that uses social media data and internet search engine queries to identify emerging themes on the use of cannabis as a medicine. In this study, the term cannabis as a medicine refers to the use of cannabis with therapeutic intent and includes prescribed cannabis, over-the-counter cannabis, and non-prescribed cannabis used specifically for self-medication (not for recreational purposes).

Materials and Methods: For this systematic review, Medline, Scopus, Web of Science and Embase databases were searched for peer reviewed studies published in English between January 2010 and March 2022. All study types that examined social media data and cannabis as a medicine were included in the review.

Results: Forty studies were included in this review. 21 studies used manually labeled data, 4 studies used existing meta-data, 2 studies used data from social media analytics companies and 13 used computational methods for annotating data. More than half of the studies 22/40 (55%), were published in the last three years. The incidental use of cannabis as a medicine was found in eleven studies that were focused on more general health-related issues, rather than on medicinal cannabis only.

Conclusion: Our systematic review has revealed the growing interest in analyzing user-generated content for studying cannabis as a medicine. This review has also highlighted the need for studies into cannabis use for specific health conditions and for automatic processing of larger datasets using computational methods (machine learning technologies).

Introduction

The plant cannabis sativa was first cultivated for use as a medicine in the Stone Age (1). From the 1800's individuals accessed cannabis as a medicine by either prescription or as an over the counter therapeutic (2). Yet by the mid-20th century, cannabis use was prohibited in many parts of the developed world with the passing of legislation in the USA, the UK and various European countries that proscribed its use (3-6) More recently, new evidence regarding the clinical efficacy of cannabis for some conditions (7) has stimulated public interest in cannabis and cannabis-derived products (8, 9), resulting in a global trend towards public acceptance, and subsequent legalization of cannabis for both medicinal and recreational use.

There is emerging evidence of cannabis efficacy for childhood epilepsy, spasticity and neuropathic pain in multiple sclerosis, acquired immunodeficiency syndrome (AIDS) wasting syndrome, and cancer chemotherapy-induced nausea and vomiting (10-12). Although researchers are investigating cannabis for treating cancer, psychiatric disorders (13), sleep disorders (14), chronic pain (15) and inflammatory conditions such as rheumatoid arthritis (16), there is currently insufficient evidence to support its clinical use. Scientific studies on emerging therapeutics typically exclude vulnerable populations such as pregnant women, young people, the elderly, and those with comorbidities, and those who depend on multiple medications, thus limiting the availability of evidence for cannabis effectiveness in these population groups (17).

Cannabis as medicine is associated with a rapidly expanding industry (18). Patient demand is increasing, as reflected in an increasing number of approvals for prescriptions over time (19), with one study showing that 61% of Australian GPs surveyed reported one or more patient enquiries regarding medical cannabis (20). With this increasing demand, is sophisticated marketing by medicinal cannabis companies that leverages evidence from a small number of studies to promote their products (21, 22). In light of these development, concerns regarding patient safety are warranted especially when marketing for cannabinoid products is associated with inadequate labelling and/or inappropriate dosage recommendations in some cannabis products (23), the provision of over-the-counter cannabis products that do not require a prescription (24) and the illicit drug market (25). Given this dynamic interplay between marketing, product innovation, regulation, and consumer demand new innovative surveillance methods are required to augment existing established monitoring approaches.

Although there is considerable engagement from many stakeholders to improve the scientific evidence regarding the efficacy and safety of cannabis through randomized control trials, many gaps remain in the literature (26). However, data from real-world and patient reported data sources could provide opportunities to address this evidence deficit (27). This real-world data can be captured from a variety of sources such as found in routinely collected health care and health services records that include but is not limited to patient generated data from medical, administrative and claims data as well as patient reported data from surveys, wearable trackers, patient registries, and social media (28-30).

People readily consult the internet when looking for and sharing health information (31, 32). According to 2017 survey of Health Information National Trends, almost 78% of US adults used online searches first to inquire about health or medical information (31). Data resulting from these online activities is labelled ‘user generated’ and is increasingly becoming a component of surveillance systems in the health data domain (33). Monitoring user-generated data on the web can be a timely and inexpensive way to generated population-level insights (34). The collective experiences and opinions shared online are an easily accessible wide-ranging data source for tracking emerging trends – which might be unavailable or less noticeable by other surveillance systems.

The objective of this systematic review is to understand the utility of online user generated text in providing insight into the use of cannabis as a medicine. In this review, we aim to systematically review existing work that utilizes user-generated content to explore cannabis as a medicine. The objective of this systematic review is to synthesise quality primary research that uses social media discourse and internet search engine queries to answer the following questions:

- Does online user-generated text provide a useful data source for studying cannabis as a medicine?
- What are the research questions motivating the studies of online user-generated text that discuss medicinal use of cannabis?
- How can future research leverage user-generated content to study the use of cannabis as a medicine?

Materials and methods

For this systematic review we used a framework for systematic reviews and meta-analyses, the PRISMA guidelines to inform our methods (35). A manual search of Medline, Embase, Web of Science, and Scopus databases was independently conducted by SKH in May 2021 and again by CMH in March 2022. The search was limited to English-language studies that were published between January 1974 and March 2022. Given that much of the computational research in this area is published in peer-reviewed computer science venues that are not necessarily indexed by Medline we also used complementary literature databases (Embase, Web of Science, and Scopus) to generate our initial list of publications for screening.

Literature database queries were developed for four categories of studies. The first three categories used social media text as a data source, the fourth relied on internet search engine query data. For the first category, the database queries combined words used to describe social media forums, and cannabis-related keywords and general medical-related keywords (Table 1 Category 1). The second category also included the social media and cannabis-related keywords, but used keywords specific to psychiatric disorders, for which the use of medical cannabis has been described. Our search terms for this second category were informed by a systematic review of medicinal cannabis for psychiatric disorders (13) (Table 1 Category 2). The third category included social media and cannabis related keywords but focused on non-psychiatric medical conditions for which cannabis is sometimes used (Table 1 Category 3). The fourth category included studies using Internet search engine queries as a data source, there were no medical conditions included in these searches (Table 1 Category 4).

Table 1. Search Terms

Category 1 - Social media, cannabis, and medical terms as keywords		
Social media related Keywords	*Cannabis keywords*	*Medical keywords*
"Social media" OR twitter OR reddit OR instagram OR youtube OR pinterest OR facebook OR "social network forum" OR "Online health community" OR "message board"	cannabis OR cannabis OR cannabinoids OR delta-9-tetrahydrocannabinol OR cannabidiol OR cbd OR cbg OR cbn OR thc OR weed	medical OR medicinal OR patient OR patients OR medicine OR doctor OR position OR care OR therapy OR therapeutic
Category 2 - Social media, cannabis, and psychiatric disorders keywords		
Social media related Keywords	*Cannabis keywords*	*Psychiatric disorders*
"Social media" OR twitter OR reddit OR instagram OR youtube OR pinterest OR facebook OR "social network forum" OR "Online health community" OR "message board"	cannabis OR cannabis OR cannabinoids OR delta-9-tetrahydrocannabinol OR cannabidiol OR cbd OR cbg OR cbn OR thc OR weed	depression OR depressive OR "mental illness*" OR "mental disorder*" OR "mental health" OR "mood disorder*" OR "affective disorder*" OR anxi* OR "panic disorder" OR "obsessive compulsive" OR adhd OR "attention deficit" OR phobi* OR bipolar OR psychiat* OR psychological OR psychosis OR psychotic OR schizophr* OR "severe mental*" OR "serious mental*" OR antidepress* OR antipsychotic* OR "post traumatic*" OR "personality disorder*" OR stress
Category 3 - Social media, cannabis and various medical (non-psychiatric) conditions/ illnesses keywords		
Social media related Keywords	*Cannabis keywords*	*Medical conditions*
"Social media" OR twitter OR reddit OR instagram OR youtube OR pinterest OR facebook OR "social network forum" OR "Online health community" OR "message board"	cannabis OR cannabis OR cannabinoids OR delta-9-tetrahydrocannabinol OR cannabidiol OR cbd OR cbg OR cbn OR thc OR weed	Pain, Opioid, Alzheimer, sleep OR insomnia, inflammatory, arthritis, Multiple Sclerosis, Endometriosis
Category 4 - Search engine queries and cannabis keywords		
Search Engine keywords	*Cannabis keywords*	
"Search engine" OR "search log" OR "search queries" OR "online search" OR "internet Search" OR "web search"	cannabis OR Cannabis OR Cannabinoids OR Delta-9-Tetrahydrocannabinol OR Cannabidiol OR CBD OR CBG OR CBN OR THC OR weed	

The inclusion criteria for this systematic review were: (i) primary research studies, (ii) studies which used online user-generated text as a data source, and (iii) research that was either directly focused on cannabis and cannabis products that have an impact on health, or were health-related studies that found medicinal use of cannabis.

Exclusion criteria comprised published: (i) editorials, letters, commentaries, and book chapters; (ii) studies that used social media for recruiting participants; (iii) studies where the full text of the publication was not available; (iv) studies primarily focused on electronic nicotine delivery systems adapted to deliver cannabinoids and (v) studies that used bots or autonomous systems as the main data source and (vi) studies that focused exclusively on synthetic cannabis.

Included studies were reviewed by CMH, YB, MC, and SKH. Where initial disagreement existed between reviewers regarding the inclusion eligibility of a study, team members met to discuss the disputed article's status until consensus was achieved. For all the articles that met the inclusion criteria, pairs of reviewers independently critiqued the articles using a checklist developed for this study. The purpose of the checklist was to provide an overall assessment of quality rather than generate a specific score, a summary version of this checklist is presented (S1). Assessments of quality were based on evidence of weakness in aims or objectives, main findings, data collection method, analytic methods, data source, and evaluation and interpretations of the study.

Studies were excluded if they had: (i) poorly defined objectives; (ii) poorly described methodology; (iii) inappropriate analytic approach; (iv) small volume of data and/or (v) results that are not clearly reported).

Of the 1,271 titles identified in the electronic database searches, 781 were duplicates and 441 were excluded based on lack of relevance. This screening provided 49 potentially relevant articles for inclusion. Of these, five were excluded due to full text inaccessibility, and four were excluded based on reasons listed earlier. This provided 40 papers for inclusion in the review. The PRISMA flow diagram is presented in Fig 1. The PRISAM checklist is reported in S2 Checklist.

Fig 1. The PRISMA flow diagram

Results

Of the 40 papers synthesized for this review most were journal articles 38/40 (95%), and two were conference-based publications 2/40 (5%). Table 2 provides a summary of each paper that includes author names, publication year, data source, duration of the study, number of collected posts, number of analyzed posts, and the coding/labelling approach used.

Data collection and annotation

Table 2 shows that the largest dataset labelled by the researchers belongs to one of the earliest studies in this domain, where around 47,000 tweets are manually labelled (36). This paper was one of 21/40 (53%) of the papers which have either collected a limited number of data points or have sampled their collected data and manually coded the data to gain an in-depth understanding of the domain. 4/40 (10%) have used existing meta-data including geo-location data. 2/40 (5%) have used social media analytics companies' data (37, 38). 13/40 (32%) have used an automated method for labelling data - these include machine learning, lexicon or rule-based algorithms. The earliest study that employed an automated labelling method was performed in 2017 (39), which used machine learning, thereafter the papers show a trend for using automation. The ability of such automated data-centric (40) approaches to handle more data means they can be utilized to broaden the available analyses, as more data enables the use of more sophisticated techniques such as machine learning models. Insights gained from the smaller qualitative studies can be used to inform the design of automated approaches, and even the manually labelled data can be used for training these models, so in this way automated methods can be built upon and leverage initial manual methods.

Table 2. Articles included in the review

#	Study	Source	Duration	Collected data	Annotated/analyzed	Coding/labelling approach
1	McGregor et al., 2014 (41)	Online forums, Facebook, Twitter, YouTube	Not Available	3,785 items	All data	Manual coding
2	Cavazos-Rehg et al., 2015 (42)	Twitter	Feb 2014 – Mar 2014	7,653,738 tweets	7,000 tweets	Manual coding using crowdsource services
3	Daniulaityte et al., 2015 (43)	Twitter	Oct 2014– Dec 2014	125,255 tweets 27,018 geolocated	All data	No coding - used existing geographical fields
4	Gonzalez-Estrada et al., 2015(44)	YouTube	4-8 June 2014	200 most viewed videos	All data	Manual coding
5	Krauss et al., 2015(45)	YouTube	22nd of Jan 2015	116 Videos	All data	Manual coding
6	Thompson et al., 2015 (36)	Twitter	Mar 2012 – July 2013	36,939 original tweets 10,000 retweets	~47,000	Manual coding
7	Cavazos-Rehg et al., 2016 (46)	Twitter	Jan 2015	206,854 tweets	5,000	Manual coding using crowdsource services
8	Lamy et al., 2016 (47)	Twitter	May 2015 – Jul 2015	100,182 tweets 26,975 geolocated	3,000	Manual coding
9	Mitchell et al., 2016 (48)	Online Forums	Oct 2014	268 forum threads	46 threads 880 posts	Manual coding
10	Andersson et al., 2017 (49)	Online Forums	18-19 Apr 2016	32 topics	All data by researchers	Manual coding
11	Dai & Hao, 2017 (39)	Twitter	Aug 2015 – Apr 2016	1,253,872 post traumatic stress disorder (PTSD) tweets 66,000 cannabis PTSD tweets	2,000 labelled, remaining data by machine learning	Automated coding using machine learning
12	Greiner et al., 2017 (50)	Online Forums	Nov 2014 – Mar 2015	717 posts	All data	Manual coding
13	Turner & Kantardzic, 2017 (51)	Twitter	Aug 2015 – Apr 2016	40,509 geolocated tweets	2,000 labelled, remaining data by machine learning	Automated coding using machine learning
14	Westmaas et al., 2017 (52)	Online Forums	Jan 2000 – Dec 2013	468,000 posts	All data	Automated coding using topic modelling
15	Yom-Tov & Lev-Ran, 2017 (53)	Bing search engine	Nov 2016 – Apr 2017	Not available	All data	Automated coding using lexicons
16	Cavazos-Rehg et al., 2018 (54)	YouTube	10-11 June 2015	83 videos	All data	Manual coding
17	Glowacki et al., 2018 (55)	Twitter	Aug 2016 – Oct 2016	73,235 tweets	All data	Automated coding using topic modelling
18	Meacham et al., 2018 (56)	Reddit	Jan 2010 – Dec 2016	~400,000 posts	All data	Automated coding using lexicons and pattern matching

19	Leas et al., 2019 (57)	Google Trends	Jan 2004 – April 2019	Not available	Summary data	No coding – used Google trends data
20	Meacham et al., 2019 (58)	Reddit	Jan 2017– Dec 2017	193 dabbing questions and/or posts	All data	Manual coding
21	Nasralah et al., 2019 (37)	Twitter	Jan 2015 – Feb 2019	20,609 tweets	All data	Automated coding supplied by an analytics company
22	Pérez-Pérez et al., 2019 (59)	Twitter	Feb 2018– Aug 2018	24,634 tweets	All data	Automated coding using lexicons
23	Shi et al., 2019 (60)	Google Trends, Buzzsumo	Jan 2011 – Jul 2018	Not available	Summary data	No coding – used Google trends data
24	Allem et al., 2020 (61)	Twitter	May 2018– Dec 2018	19,081,081 tweets	60,861 non-bots 8,874 bots	Automated coding using rule-based methods
25	Janmohamed et al., 2020 (62)	Blogs, news, forums, others	Aug 2019 – Apr 2021	4,027,172 documents	All data	Automated coding using topic modelling
26	Jia et al., 2020 (63)	Google, YouTube Facebook	Sep 2019	51 Google websites 126 Facebook posts 37 YouTube videos	All data	Manual coding
27	Leas et al., 2020 (64)	Reddit	Jan 2014 – Aug 2019	104,917	3,000 initial data 376 as testimonials	Manual coding
28	Merten et al., 2020 (65)	Pinterest	31 Jul, 18 Aug, 1 Sep 2018	1,280	226	Manual coding
29	Mullins et al., 2020 (38)	Twitter	22 Jun – 5 Jul 2017	941 tweets	All data	Automated coding supplied by an analytics company
30	Saposnik & Huber, 2020 (66)	Google Trends	Jan 2004 – Dec 2019	Not available	Summary data	No coding – used Google trends data
31	Song et al., 2020 (67)	GoFundMe	Jan 2012 – Dec 2019	1,474 campaigns	500 campaigns	Manual coding
32	Tran & Kavuluru, 2020 (68)	Reddit & FDA comments	Jan 2019– April 2019	64,099 reddit comments 3,832 FDA (U.S. Food & Drug Administration) comments	All data	Automated coding using rule-based methods
33	van Draanen et al., 2020 (69)	Twitter	Jan 2017 – June 2019	1,200,127 tweets	All data	Automated coding using Topic modelling
34	Zenone et al., 2020 (70)	GoFundMe	Jan 2017 – May 2019	155 campaigns	All data	Manual coding
35	Pang et al., 2021 (71)	Twitter	Dec 2019 – Dec 2020	17,238 tweets	1,000 tweets	Manual coding
36	Rhidenour et al., 2021 (72)	Reddit	Jan 2008 – Dec 2018	974 posts	All data	Manual coding
37	Smolev et al., 2021 (73)	Facebook	Nov 2018 – Nov 2019	7,694 posts	All data	Manual coding
38	Zenone et al., 2021 (74)	GoFundMe	Jun 2017 – May 2019	164 campaigns	All data	Manual coding
39	Soleymanpour et al., 2021 (75)	Twitter	July 2019	2,200,000 tweets	2,000 labelled, remaining data by machine learning	Automated coding using machine learning
40	Allem et al. 2022 (76)	Twitter	Jan – Sep 2020	353,353 tweets	1,092 labelled, remaining data by classifiers	Automated coding using rule-based methods

Data analysis

Most of the manually labelled studies focused on manual or statistical analysis (36, 41, 42, 46-50, 58, 63-65, 67, 70-74), while three analyzed online videos (44, 45, 54). Studies that utilized a large volume of data allowed for the use of computational methods, including sentiment analysis, topic modeling, and rule-based text mining. Sentiment analysis aims to analyse people's sentiments, opinions, and attitudes (75, 77). Topic modelling is a machine learning method for automatically discovering common themes in a collection of documents (75, 78). Rule-based text mining involves classification of posts into pre-existing health-related categories (76). Studies also explored social network analysis, which examines how social media users connect through friendship, sending links and information, or tagging and following (79).

Research Themes

In this review, we categorized 40 research articles related to user-generated online text (social media discourse and internet search engine queries) into six broad themes. The themes are centered around the research questions motivating the studies.

General cannabis-related

Nine studies were included in this theme (Table 3). The main keywords used in these studies included general terms such as 'cannabis', 'marijuana', 'pot' and 'weed'. The major aim of these studies was to either identify topics of conversations regarding cannabis, or to examine their sentiments. These studies are included because they reported on conversations around cannabis use for medical purposes, or sentiment associated with perceptions of health benefits of cannabis and reports of adverse effects. For example, a study on veterans use of cannabis found that cannabis is used to self-medicate a number of health issues, including Post-Traumatic Stress

Disorder (PTSD), anxiety and sleep disorder (72). Six of the studies used Twitter as a data source (36, 42, 51, 61, 69, 76) , one studied content of YouTube videos about cannabis (54), one investigated online self-help forums (50) and another used Reddit (72).

Table 3. Theme of papers studying general conversations

Study	Aim	Health-related effects/claims	Data identification
Cavazos-Rehg et al., 2015 (42)	To examine the sentiment and themes of cannabis-related tweets from influential users and to describe the users' demographics.	A common theme of pro tweets was that cannabis has health benefits. Anti-cannabis posts spoke of the harm experienced in using cannabis. 77% of posts had positive sentiments, with 12 times higher reach than other posts.	Cannabis-related keyword
Thompson et al., 2015 (36)	To examine cannabis- related content in Twitter, especially content tweeted by adolescent users, and to examine any differences in message content before and after the legalization of recreational cannabis in two US states.	More tweets described perceived positive benefits of cannabis use, including relaxation and escaping life problems. Tweets described cannabis as less harmful than other drugs or as not harmful at all, and suggested its medical role for conditions such as depression and cancer. Less than 1% of tweets expressed a concern about cannabis use.	Cannabis-related keyword
Greiner et al., 2017 (50)	To investigate online content of cannabis use/addiction self-help forums.	Self-help forums on cannabis share a theme around cannabis users seeking help for addiction and withdrawal issues.	keywords "cannabis" and "forum" and "help"
Turner et al., 2017 (51)	To examine if cannabis legalization policies impact Twitter conversations and the social networks of users contributing to cannabis conversations.	Medical cannabis was a major topic in the conversations.	Cannabis-related keywords
Cavazos-Rehg et al., 2018 (54)	To investigate cannabis product reviews and the relationship between exposure to product reviews and cannabis users' demographics and characteristics.	Product reviews promoted cannabis for helping with relaxation, pain relief, sleep, improving emotional well-being. Medical cannabis users are more likely to be exposed to cannabis product reviews.	"weed review", "marijuana review", "cannabis review"
Allem et al., 2020 (80)	To identify and describe cannabis-related topics of conversation on Twitter, and the public health implications of these.	Health and medical was the third most prevalent topic of the 12 topics identified in the data. Posts suggested that cannabis could help with cancer, sleep, pain, anxiety, depression, trauma, and posttraumatic stress disorder. Health-related posts from social bots were almost double of genuine posts.	Cannabis-related keywords

Van Draanen et al., 2020 (69)	To examine differences in the sentiment and content of cannabis-related tweets in the US (by state cannabis laws) and Canada.	Medical cannabis use was one of the main topics of conversations in cannabis-related tweets from both countries.	Tweets filtered on US and Canada geolocation and then further filtered on cannabis-related keywords
Rhidenour et al., 2021 (72)	To explore Veterans' Reddit discussions regarding their cannabis use.	Over a third of the Reddit posts described the use of medical cannabis as an aid for psychological and physical ailments. Overall, veterans discussed how the use of medical cannabis reduced PTSD symptoms, anxiety, and helped with their sleep.	The veteran subreddit
Allem et al., 2022 (76)	To determine the extent to which a medical dictionary could identify cannabis-related motivations for use and health consequences of cannabis use.	There were posts related to both health motivations and consequences of cannabis use. The health-related posts included issues with the respiratory system, stress to the immune system, and gastrointestinal issues, among others.	Cannabis-related keywords

Cannabis mode of use

Seven studies reported on the use of cannabis as a medicine in relation to its mode of use (Table 4). These studies collected data using keywords such as 'vape', 'vaping', 'dabbing' and 'edibles'. Conversations around modes of use revealed a theme about lacking, seeking, or sharing knowledge about possible health consequences of the modes of use. Another theme was around the perceived health benefits of various modes of use including sleep improvement and relaxation resulting from dabbing oils (46) or consuming 'edibles' (47). The findings suggest that for emerging modes of use such as dabbing, where the availability of evidence-based information is limited, people seek information from others' experiences.

Table 4. Theme of papers studying mode of use

Study	Aim	Health-related effects/claims	Data identification
Daniulaityte et al., 2015 (43)	To explore Twitter data on concentrate ('dabs') use and examine the impact of cannabis legalization policies on concentrate use conversations.	Twitter data suggest popularity of dabs in the US states with legalized recreational/medical use of cannabis. Dabbing as an emerging mode of use could carry significant health risks.	dab-related keywords for U.S. location

Krauss et al., 2015 (45)	To explore the content of cannabis dabbing-related videos on YouTube.	Only 21% of videos contained some warning about dabbing, such as preventing explosions, injury, or negative side effects. 22% of videos specifically mentioned medical cannabis or getting "medicated", either in the video itself or in the accompanying text description.	Dabbing related keywords
Cavazos-Rehg et al., 2016 (46)	To study themes of dabbing conversations and to investigate the consequences of high-potency cannabis consumption.	4th theme (of 7 themes) was about cannabis helping with relaxation, sleep or solving problems. Extreme effects were both physiological and psychological. The most common physiologic effects were passing out and respiratory, with coughing the most common respiratory effect.	Dabbing related keywords
Lamy et al., 2016 (47)	To study themes of edibles conversations and examine legalization policies' impact on cannabis-related tweeting activity.	Twitter data suggest mostly positive attitudes toward cannabis edibles. Positive tweets describe the quality of the "high" experienced and how cannabis edibles facilitate falling asleep. Negative tweets discuss the unreliability of edibles' THC dosage and delayed effects that were linked to over-consumption, which could lead to potential harmful consequences.	Cannabis edible-related keywords
Meacham et al., 2018 (56)	To analyse discussions of emerging and traditional forms of cannabis use.	Less than 2% of conversations described adverse effects. The most mentioned adverse effects were anxiety-related in the context of smoking, edibles, and butane hash oil, and "cough" for vaping and dabbing.	A cannabis specific subreddit on various modes of use
Meacham et al., 2019 (58)	To study themes of dabbing-related questions and responses.	Health concerns are the 5th category of dabbing questions - including respiratory effects, anxiety, and vomiting. Respondents in these conversations usually spoke from personal experience.	Search for "Dab" and "question" on cannabis subreddits
Janmohamed et al.,2020 (62)	To map temporal trends in the web-based vaping narrative, to indicate how the narrative changed from before to during the COVID-19 pandemic.	The emergence of a vape-administered CBD treatment narrative around the COVID-19 pandemic.	Vape-related keywords

Cannabis as a medicine for a specific health issue

Six studies were included in this theme (Table 5). These studies investigated conversations around the use of cannabis or cannabidiol for a specific health issue. The health conditions included glaucoma (63), PTSD (39), cancer (60, 70), Attention Deficit Hyperactivity Disorder

(ADHD) (48) and pregnancy (71). These studies mostly discovered that conversations claimed benefits of cannabis as an alternative treatment for these health conditions, although mentions of harm, and both harm and therapeutic effects, were also present (48).

Table 5. Theme of papers studying specific health issues

Study	Aim	Health-related effects/claims	Data identification
Mitchell et al., 2016 (48)	To examine the content of online forum threads on ADHD and cannabis use to identify trends about their relation, particularly regarding therapeutic and adverse effects of cannabis on ADHD.	25% of individual posts indicated that cannabis is therapeutic for ADHD, as opposed to 8% that claim it is harmful, 5% that it is both therapeutic and harmful, and 2% that it has no effect on ADHD.	Cannabis and ADHD keywords
Dai & Hao, 2017 (39)	To evaluate factors that could impact public attitudes to PTSD related cannabis use.	5.3% of all PTSD tweets were related to cannabis use and these tweets predominantly supported cannabis use for PTSD.	Cannabis and PTSD keywords
Shi et al., 2019 (60)	To characterize trends in use of cannabis for cancer and analysis of content and impact of popular news about cannabis for cancer.	Between 2011-2018, the relative google search volume of 'cannabis cancer' queries increased at a rate 10 times faster than 'standard cancer therapies' queries. Popular 'false news' stories had a much higher engagement than contrary 'accurate' news stories.	Cannabis vs standard therapies for cancer
Jia et al., 2020 (63)	To analyze the content quality and risk of readily available online information regarding cannabis and glaucoma.	While the American Glaucoma Society recommends against cannabis use for glaucoma treatment, 21% of Facebook, 24% of Google, and 59% of YouTube search results were pro cannabis use for glaucoma treatment.	Cannabis and Glaucoma keywords
Zenone et al., 2020 (70)	To use crowdfunding campaigns to understand how cannabidiol is represented/misrepresented as a cancer-related care.	CBD use was reported to reduce the side effects of conventional treatments or can be used with other complementary cancer treatments. Reported uses included stimulating appetite, general pain relief, assisting with sleep, countering nausea, or general recovery purposes. Most campaigners presented definite efficacy of CBD for pain or symptom management.	CBD term variants and "cancer"
Pang et al., 2021 (71)	To examine cannabis and pregnancy-related tweets over a 12-month period.	36% mentioned safety during pregnancy, 2.3% of posts asked about safety during postpartum, and 2.7% of posts expressed use of cannabis during pregnancy to help with pregnancy symptom i.e. to help with morning sickness, nausea, vomiting, headaches, pain, stress, and fatigue. The authors conclude that health providers discuss risks and provide official information about cannabis use in pregnancy.	Cannabis and pregnancy related keywords

Cannabis as a medicine as part of discourse on illness and disease

Eleven studies were included in this theme. In this category, the research focus was on social media topics relating to management and treatment options for a range of health conditions rather than on medicinal cannabis per se (Table 6). There were eleven studies, and the health conditions included inflammatory and irritable bowel disease(59), opioid use disorder (37, 55), pain (38), ophthalmic disease (41), cluster headache and migraine(49), asthma (44), cancer (52, 67), autism disorder (66) and brachial plexus injury (73).

Table 6. Theme of papers studying illness and disease discourse

Study	Aim	Health-related effects/claims	Data identification
McGregor et al., 2014 (41)	To analyse the ophthalmic content of social media platforms.	Treatment was one of the main themes, with complementary therapy featuring most prominently on Twitter, where 87% of posts on complementary therapy described the use of medical cannabis for glaucoma.	Glaucoma patient forums and glaucoma keywords
Gonzalez-Estrada et al., 2015 (44)	To determine the educational quality of YouTube videos for asthma.	The most common video content was regarding alternative medicine (38%), and included cannabis as well as live fish ingestion; salt inhalers; raw food, vegan, gluten-free diets; yoga; Ayurveda; reflexology; acupressure; and acupuncture; and Buteyko breathing.	Asthma related videos
Andersson et al., 2017 (49)	To understand the use of non-established or alternative pharmacological treatments used to alleviate cluster headaches and migraines.	Cannabis was discussed for its potential to alleviate symptoms or reduce the frequency of migraine attacks. Some had used cannabis for other purposes but experienced additional benefits for headaches. The effects of self-treatment with cannabis appeared more contradictory and complex than treatment with other substances.	Search for “treatment migraines” on three alternative treatment online forums
Westmaas et al., 2017 (52)	To investigate contexts in which smoking or quitting is discussed in a cancer survivor network.	Use of cannabis (primarily for nausea), was the 4th topic.	Smoking/cessation-related keywords from the Cancer Survivors Network (CSN)

Glowacki et al.,2018 (55)	To identify public reactions to the opioid epidemic by identifying the most popular topics.	Mentions of cannabis as an effective alternate to opioids for managing pain.	Opioid-related keywords
Nasralah et al., 2019 (37)	To understand the concerns of opioid-addicted users.	Cannabis was found in two of five main themes: "In recovery" and "taking illicit drugs" for pain management.	Users who self-identified as addicted to, or previously addicted to, opioids
Pérez-Pérez et al., 2019 (59)	To characterize the bowel disease community on Twitter.	Medical cannabis was the 4th most mentioned term in the bowel disease (BD) community. Medical cannabis and its components was the most discussed drug, with mentions of its benefits in mitigating common BD symptoms.	Inflammatory bowel disease, Irritable bowel disease keywords
Mullins et al., 2020 (38)	To examine pain-related tweets in Ireland over a 2-week period.	The 4th most occurring keyword was cannabis. 90% of Cannabis related tweets were non-personal, with highly positive sentiment and highest number of impressions per tweet. Cannabis had by the largest number of tweets aimed at generating awareness.	Pain-related keywords
Saposnik & Huber, 2020 (66)	To analyse of trends in web searches for the cause and treatments of autism spectrum disorder (ASD).	ASD and cannabis web searches have continued to rise since 2009. Apart from searches on Applied Behavioral Analysis and Autism, cannabis and ASD have been searched more than other ASD interventions since 2013.	"Autism" and key search terms for causes and treatments of autism
Song et al., 2020 (67)	To understand the cancer patient's perspective for using complementary and alternative medicine, or for declining traditional cancer therapy.	Cannabidiol oil was 10th amongst the most used alternative treatments.	20 most prevalent cancers in the U.S. and a list of top most utilized complementary and alternative medicine including yoga, herbal, meditation, etc.
Smolev et al., 2021 (73)	To analyse themes of brachial plexus injury Facebook conversations.	There were 313 posts regarding cannabinoids as a preferred alternative pain management medication.	"traumatic brachial plexus injury" keyword

Cannabidiol (CBD)

There were six studies in the cannabidiol (CBD) category (57, 64, 65, 68, 74, 75). These studies concentrated on conversations related to the benefits of CBD products, product sentiment (positive, negative, or neutral), the factors that impact on a person's decision to use CBD products, and the trends in therapeutic use of CBD.

Table 7. Theme of papers studying CBD

Study	Aim	Health-related effects/claims	Data identification
Merten et al., 2020 (65)	To analyse how CBD is portrayed on Pinterest.	Most pins (57.5%) did not make a specific health benefit claim yet 42.5% claimed mental, physical, or both mental and physical health benefits.	'cannabidiol' or 'CBD'
Zenone et al., 2021 (74)	To analyse the CBD informational pathways which bring consumers to CBD for medical purposes.	Self-directed research was the most common pathway to CBD. The proposed uses of CBD were for cancer, seizure-inducing diseases/conditions, joint/inflammatory diseases, mental health disorders, nervous system diseases, and autoimmune diseases.	'cannabidiol' or 'CBD'
Leas et al., 2019 (57)	To analyse public interest trends in CBD using Google Trends.	Searches for CBD exceed searches for yoga and around half as much as searches for dieting.	'cannabidiol' or 'CBD' vs other alternative medicine including diet, yoga, etc.
Leas et al., 2020 (64)	To assess if individuals are using CBD for diagnosable conditions which have evidence-based therapies.	Psychiatric conditions were the most commonly cited diagnosable condition, mentioned in 63.9% of testimonials. The second most cited subcategory was orthopedic conditions (26.4%), followed by sleep (14.6%), neurological (6.9%), and gastroenterological (3.9%) conditions.	CBD subreddit posts
Tran & Kavuluru, 2020 (68)	To examine social media data to determine perceived remedial effects and usage patterns for CBD.	Anxiety disorders and pain were the two conditions dominating much of the discussion surrounding CBD, both in terms of general discussion and for CBD as a perceived therapeutic treatment. CBD is mentioned as a treatment for mental issues (anxiety, depression, stress) and physiological issues (pain, inflammation, headache, sleep disorder, seizure disorders, nausea, and cancer).	CBD subreddit posts
Soleymanpour et al., 2021 (75)	To perform content analysis of marketing claims for CBD in Twitter.	Over 50% of CBD tweets appear to be marketing related chatter. Pain and anxiety are the most popular conditions mentioned in marketing messages. Edibles are the most popular product type being advertised, followed by oils.	"cbd", "cbdoil" and "cannabidiol"

Adverse drug reactions and adverse effects

One study focused explicitly on adverse reaction detection (Table 8) This study explored the prevalence of internet search engine queries relating to the topic of adverse reactions and cannabis use. Mentions of adverse effects of cannabis consumption were found in eight studies, these included respiratory effects and anxiety (56, 58, 76), addiction and withdrawal syndromes (50), neurological harm (76), harm in pregnancy (71, 76) and general harm (45, 46, 48, 76).

Table 8. Theme of papers studying reactions and adverse effects

Study	Aim	Health-related effects/claims	Data identification
Yom-Tov & Lev-Ran, 2017 (53)	To check if search engine queries can be used to detect adverse reactions of cannabis use.	A high correlation between the side effects recorded on established reporting systems and those found in the search engine queries. These side effects included anxiety, depression-related symptoms, psychotic symptoms such as paranoia and hallucinations, cough and other symptoms.	Cannabis related keywords

Discussion

Currently, there exist systematic reviews of cannabis and cannabinoids for medical use based on clinical efficacy outcomes from randomized clinical trials (15) and reviews on the use of social media for illicit drug surveillance (81). However, to our knowledge, this paper constitutes the first systematic review examining studies that used user-generated online text to understand the use of cannabis as a medicine in the global community.

Our systematic review found that the use of social media and internet search queries to investigate cannabis as a medicine is a rapidly emerging area of research. Over half of the studies included in this review were published within the last three years, this reflects not only increase community interest in the therapeutic potential of cannabinoids, but also world-wide trends towards cannabis legalization (4, 82-86). Regarding social media platforms, Twitter was the data source in seventeen (42.5%) of the 40 studies, almost three and a half times the number of studies using Reddit (5, 12.5%) and just under three times the number of studies using data from

Online forums (6, 15%). Three (7.5%) GoFundMe studies and three (7.5%) Google Trends studies were also included in the review. Hence, much of the data in this systematic review comprised posts from the Twitter platform. Several factors may explain this finding, firstly Twitter is real-time in nature, it has a high volume of messages, and it is publicly accessible. These factors makes it a useful data source for public health surveillance (87).

Regarding the subjects of the studies, eleven (29%) focused on general user-generated content regarding the treatment of health conditions (glaucoma, autism, asthma, cancer, bowel disease, brachial plexus injury, cluster headaches, opioid disorder). These studies were not explicitly designed to investigate cannabis as a medicine, yet they generated results that incidentally found cannabis mentioned as an alternative or complimentary treatment, either formally prescribed or via self-medication.

Qualitative studies feature prominently in the research, but while their contribution is valuable, especially in the context of hypothesis generation, they tend to be limited by their smaller datasets, which frequently comprised manually annotated samples. The recent emergence of powerful machine learning-based natural language processing (NLP) models suggests that it should be possible to automate the continuous processing of far larger datasets using NLP technologies, built upon the insights gained from initial qualitative studies, and even leveraging their annotated data for training purposes. Recent trends in the social science data landscape have shown a convergence between social science and computer science expertise, where the ability to use computational methods has greatly assisted the collection and validation of robust datasets that can form the basis of deeper social science research (88).

We found much heterogeneity in approaches applied to analyse user-related content, and inconsistent quality in the methodologies adopted. While we endeavored to include as many studies as possible, some of the publications initially identified as suitable for inclusion were not suitable based on a minimum quality requirements checklist (S1). This checklist was designed to ensure that selection of data source, choice of platform, data acquisition and preparation, analysis and evaluation delivers data and conclusions that are appropriate for answering the research questions.

The utilization of user-generated content for health research is subject to several inherent limitations which include the lack of control that researchers have in relation to the credibility of information, the frequently unknown demographic characteristics and geographical location of individuals generating content, and the fact that social media users are not necessarily representative of the wider community (89). Furthermore, the uniqueness, volume, and salience of social media data has implications that need to be considered when used for health information analysis (90). Volume is usually inversely related to salience: a platform such as Twitter has a very high volume of information, but, apart from its frequency, much of which is not highly pertinent for analyzing an effect; whereas the volume of information contained in a blog will much less but likely more salient for analysis. Notwithstanding these limitations, user-generated content comprises large-scale data that provides access to the unprompted organic opinions and attitudes of cannabis users in their own words, and is an effective medium through which to gauge public sentiment. To date, insights regarding cannabis as medicine have gained primarily through surveys or focus groups which have their own limitations regarding the format of data collection and potential bias in participant recruitment.

Conclusion

This systematic review has shown that user-generated content as a data source for studying cannabis as a medicine is a growing area worthy of investigation given that it provides another means to understand how cannabis is being used and perceived by the community. As such, it is another potential 'tool' with which to engage in pharmacovigilance of, not only cannabis as a medicine, but also other novel therapeutics as they enter the market.

Acknowledgments

This systematic review was supported by the Australian Centre for Cannabinoid Clinical and Research Excellence (ACRE), funded by the National Health and Medical Research Council (NHMRC) through the Centre of Research Excellence scheme (NHMRC CRE APP1135054).

References

1. Russo EB. History of cannabis and its preparations in saga, science, and sobriquet. Chemistry and Biodiversity. 2007;4(8):1614-48.
2. Bridgeman MB, Abazia DT. Medicinal cannabis: history, pharmacology, and implications for the acute care setting. Pharmacy and therapeutics. 2017;42(3):180-.
3. Grinspoon L, Bakalar JB. Marihuana: the forbidden medicine. New Haven, USA: Yale University Press; 1993.
4. Kalant H. Medicinal use of cannabis: history and current status. Pain Research and Management. 2001;6(2):80-91.
5. Taylor S. Medicalizing cannabis—Science, medicine and policy, 1950–2004: An overview of a work in progress. Drugs: education, prevention and policy. 2008;15(5):462-74.
6. Musto DF. The marihuana tax act of 1937. Archives of General Psychiatry. 1972;26(2):101-8.
7. National Academies of S, Medicine E. The health effects of cannabis and cannabinoids: the current state of evidence and recommendations for research. 2017.
8. Hall W. Medical use of cannabis and cannabinoids: questions and answers for policymaking. 2018.
9. Aguilar S, Gutiérrez V, Sánchez L, Nougier M. Medicinal cannabis policies and practices around the world. International Drug Policy Consortium. 2018(April):32-.

10. TGA. Access to medicinal cannabis products Canberra, Australia: Therapeutic Goods Administration2021 [Available from: Retrieved September, 2021, from https://www.tga.gov.au/access-medicinal-cannabis-products-1

11. National Academies of Sciences EaM. The health effects of cannabis and cannabinoids: the current state of evidence and recommendations for research. 2017.

12. Pearlson G. Medical marijuana and clinical trials. Weed Science: Cannabis Controversies and Challenges Academic Press; 2020. p. 243-60.

13. Sarris J, Sinclair J, Karamacoska D, Davidson M, Firth J. Medicinal cannabis for psychiatric disorders: A clinically-focused systematic review. BMC Psychiatry. 2020;20(1):1-14.

14. Suraev AS, Marshall NS, Vandrey R, McCartney D, Benson MJ, McGregor IS, et al. Cannabinoid therapies in the management of sleep disorders: A systematic review of preclinical and clinical studies. Sleep Medicine Reviews. 2020;53:101339-.

15. Whiting PF, Wolff RF, Deshpande S, Di Nisio M, Duffy S, Hernandez AV, et al. Cannabinoids for medical use: A systematic review and meta-analysis. JAMA - Journal of the American Medical Association. 2015;313(24):2456-73.

16. Schulze-Schiappacasse C, Durán J, Bravo-Jeria R, Verdugo-Paiva F, Morel M, Rada G. Are Cannabis, Cannabis-Derived Products, and Synthetic Cannabinoids a Therapeutic Tool for Rheumatoid Arthritis? A Friendly Summary of the Body of Evidence. JCR: Journal of Clinical Rheumatology. 2021;Publish Ah(00):1-5.

17. Bonomo Y, Souza JDS, Jackson A, Crippa JAS, Solowij N. Clinical issues in cannabis use. British Journal of Clinical Pharmacology. 2018;84(11):2495-8.

18. Owens B. The professionalization of cannabis growing. Nature. 2019;572(7771).

19. Hallinan CM, Gunn JM, Bonomo YA. The Rise and Rise of Medicinal Cannabis: An Interrupted Segmented Regression Time Series Analysis of Australian Medicinal Cannabis Prescribing Data 2017-2021. The Medical Journal of Australia 2021;mja21.00965.

20. Karanges EA, Suraev A, Elias N, Manocha R, McGregor IS. Knowledge and attitudes of Australian general practitioners towards medicinal cannabis: a cross-sectional survey. British Medical Journal Open. 2018;8(7):e022101.

21. Caputi TL. The medical marijuana industry and the use of “research as marketing”. American Journal of Public Health. 2020;110(2):174-5.

22. Ayers JW, Caputi TL, Leas EC. The need for federal regulation of marijuana marketing. Jama. 2019;321(22):2163-4.

23. Bonn-Miller MO, Loflin MJE, Thomas BF, Marcu JP, Hyke T, Vandrey R. Labeling accuracy of cannabidiol extracts sold online. JAMA - Journal of the American Medical Association. 2017;318(17):1708-9.

24. Hallinan CM, Eden E, Graham M, Greenwood LM, Mills J, Popat A, et al. Over the counter low-dose cannabidiol: A viewpoint from the ACRE Capacity Building Group. Journal of Psychopharmacology. 2021.

25. Fitzcharles MA, Shir Y, Häuser W. Medical cannabis: Strengthening evidence in the face of hype and public pressure. Cmaj. 2019;191(33):E907-E8.

26. Freeman TP, Hindocha C, Green SF, Bloomfield MAP. Medicinal use of cannabis based products and cannabinoids. BMJ (Online). 2019;365(April):1-7.

27. Graham M, Martin JH, Lucas CJ, Hall W. Translational hurdles with cannabis medicines. 2020(December 2019):1325-30.

28. US Food and Drug Administration. Framework for FDA's real world evidence program. 2018.
29. Nabhan C, Klink A, Prasad V. Real-world Evidence - What Does It Really Mean? JAMA Oncology. 2019;5(6):2019-21.
30. Hallinan CM, Gunn JM, Bonomo YA. Implementation of medicinal cannabis in Australia: Innovation or upheaval? Perspectives from physicians as key informants, a qualitative analysis. BMJ Open. 2021;11(10):1-12.
31. Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, Allen SV, Moser RP, Hesse BW. Online Health Information Seeking Among US Adults: Measuring Progress Toward a Healthy People 2020 Objective. Public Health Reports. 2019;134(6):617-25.
32. Gualtieri LN. The doctor as the second opinion and the internet as the first. 2009. p. 2489-98.
33. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaeen M, Nurmambetova E, Lucini F, et al. Digital public health surveillance: a systematic scoping review. npj Digital Medicine. 2021;4(1):41.
34. Paul MJ, Sarker A, Brownstein JS, Nikfarjam A, Scotch M, Smith KL, et al. Social Media Mining for Public Health Monitoring and Surveillance. Pacific Symposium on Biocomputing. 2016;21(January):468-79.
35. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. PLOS Medicine. 2021;18.
36. Thompson L, Rivara FP, Whitehill JM. Prevalence of Marijuana-Related Traffic on Twitter, 2012-2013: A Content Analysis. Cyberpsychology, Behavior, and Social Networking. 2015;18(6):311-9.
37. Nasralah T, El-gayar OF, Wang Y. What Social Media Can Tell Us About Opioid Addicts : Twitter Data Case Analysis. 2019.
38. Mullins CF, Ffrench-O'Carroll R, Lane J, O'Connor T. Sharing the pain: An observational analysis of Twitter and pain in Ireland. Regional Anesthesia and Pain Medicine. 2020;45(8):597-602.
39. Dai H, Hao J. Mining social media data on marijuana use for Post Traumatic Stress Disorder. Computers in Human Behavior. 2017;70:282-90.
40. Sarker A, Deroos A, Perrone J. Mining social media for prescription medication abuse monitoring: A review and proposal for a data-centric framework. Journal of the American Medical Informatics Association. 2020;27(2):315-29.
41. McGregor F, Somner JEA, Bourne RR, Munn-Giddings C, Shah P, Cross V. Social media use by patients with glaucoma: What can we learn? Ophthalmic and Physiological Optics. 2014;34(1):46-52.
42. Cavazos-Rehg PA, Krauss M, Fisher SL, Salyer P, Grucza RA, Bierut LJ. Twitter chatter about marijuana. Journal of Adolescent Health. 2015;56(2):139-45.
43. Daniulaityte R, Nahhas RW, Wijeratne S, Carlson RG, Lamy FR, Martins SS, et al. "Time for dabs": Analyzing Twitter data on marijuana concentrates across the U.S. Drug and Alcohol Dependence. 2015;155:307-11.

44. Gonzalez-Estrada A, Cuervo-Pardo L, Ghosh B, Smith M, Pazheri F, Zell K, et al. Popular on YouTube: A critical appraisal of the educational quality of information regarding asthma. Allergy and Asthma Proceedings. 2015;36(6):e121-e6.
45. Krauss MJ, Sowles SJ, Mylvaganam S, Zewdie K, Bierut LJ, Cavazos-Rehg PA. Displays of dabbing marijuana extracts on YouTube. Drug and Alcohol Dependence. 2015;155:45-51.
46. Cavazos-Rehg PA, Sowles SJ, Krauss MJ, Agbonavbare V, Grucza R, Bierut L. A content analysis of tweets about high-potency marijuana. Drug and Alcohol Dependence. 2016;166:100-8.
47. Lamy FR, Daniulaityte R, Sheth A, Nahhas RW, Martins SS, Boyer EW, et al. "Those edibles hit hard": Exploration of Twitter data on cannabis edibles in the U.S. Drug and Alcohol Dependence. 2016;164:64-70.
48. Mitchell JT, Sweitzer MM, Tunno AM, Kollins SH, Joseph McClernon F. "I use weed for my ADHD": A qualitative analysis of online forum discussions on cannabis use and ADHD. PLoS ONE. 2016;11(5):1-13.
49. Andersson M, Persson M, Kjellgren A. Psychoactive substances as a last resort-a qualitative study of self-treatment of migraine and cluster headaches. Harm Reduction Journal. 2017;14(1):1-10.
50. Greiner C, Chatton A, Khazaal Y. Online self-help forums on cannabis: A content assessment. Patient Education and Counseling. 2017;100(10):1943-50.
51. Turner J, Kantardzic M. Geo-social analytics based on spatio-temporal dynamics of marijuana-related tweets. ACM International Conference Proceeding Series. 2017;Part F1282:28-38.
52. Westmaas JL, McDonald BR, Portier KM. Topic modeling of smoking- and cessation-related posts to the American Cancer Society's Cancer Survivor Network (CSN): Implications for cessation treatment for cancer survivors who smoke. Nicotine and Tobacco Research. 2017;19(8):952-9.
53. Yom-Tov E, Lev-Ran S. Adverse Reactions Associated With Cannabis Consumption as Evident From Search Engine Queries. JMIR Public Health and Surveillance. 2017;3(4):e77-e.
54. Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Murphy GM, Bierut LJ. Exposure to and Content of Marijuana Product Reviews. Prevention Science. 2018;19(2):127-37.
55. Glowacki EM, Glowacki JB, Wilcox GB. A text-mining analysis of the public's reactions to the opioid crisis. Substance Abuse. 2018;39(2):129-33.
56. 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 and Alcohol Dependence. 2018;188(March):364-9.
57. Leas EC, Nobles AL, Caputi TL, Dredze M, Smith DM, Ayers JW. Trends in Internet Searches for Cannabidiol (CBD) in the United States. JAMA network open. 2019;2(10):e1913853-e.
58. Meacham MC, Roh S, Chang JS, Ramo DE. Frequently asked questions about dabbing concentrates in online cannabis community discussion forums. International Journal of Drug Policy. 2019;74:11-7.
59. Pérez-Pérez M, Pérez-Rodríguez G, Fdez-Riverola F, Lourenço A. Using twitter to understand the human bowel disease community: Exploratory analysis of key topics. Journal of Medical Internet Research. 2019;21(8).

60. Shi S, Brant AR, Sabolch A, Pollom E. False News of a Cannabis Cancer Cure. Cureus. 2019;11(1):1-11.
61. Allem J-P, Escobedo P, Dharmapuri L. Cannabis Surveillance With Twitter Data: Emerging Topics and Social Bots. American Journal of Public Health. 2020;110(3):357-62.
62. Janmohamed K, Soale AN, Forastiere L, Tang W, Sha Y, Demant J, et al. Intersection of the Web-Based Vaping Narrative with COVID-19: Topic Modeling Study. Journal of Medical Internet Research. 2020;22(10).
63. Jia J, Mehran N, Purgert R, Zhang Q, Lee D, Myers JS, et al. Marijuana and Glaucoma: A Social Media Content Analysis. Ophthalmology Glaucoma. 2020.
64. Leas EC, Hendrickson EM, Nobles AL, Todd R, Smith DM, Dredze M, et al. Self-reported Cannabidiol (CBD) Use for Conditions With Proven Therapies. JAMA network open. 2020;3(10):e2020977-e.
65. Merten JW, Gordon BT, King JL, Pappas C. Cannabidiol (CBD): Perspectives from Pinterest. Substance Use and Misuse. 2020;55(13):2213-20.
66. Saposnik FE, Huber JF. Trends in web searches about the causes and treatments of autism over the past 15 years: Exploratory infodemiology study. JMIR Pediatrics and Parenting. 2020;3(2).
67. Song S, Cohen AJ, Lui H, Mmonu NA, Brody H, Patino G, et al. Use of GoFundMe® to crowdfund complementary and alternative medicine treatments for cancer. Journal of Cancer Research and Clinical Oncology. 2020;146(7):1857-65.
68. Tran T, Kavuluru R. Social media surveillance for perceived therapeutic effects of cannabidiol (CBD) products. International Journal of Drug Policy. 2020;77:102688-.
69. van Draanen J, Tao H, Gupta S, Liu S. Geographic Differences in Cannabis Conversations on Twitter: Infodemiology Study. JMIR Public Health and Surveillance. 2020;6(4):e18540-e.
70. Zenone M, Snyder J, Caulfield T. Crowdfunding cannabidiol (CBD) for cancer: Hype and misinformation on gofundme. American Journal of Public Health. 2020;110:S294-S9.
71. Pang RD, Dormanesh A, Hoang Y, Chu M, Allem JP. Twitter Posts About Cannabis Use During Pregnancy and Postpartum:A Content Analysis. Substance Use and Misuse. 2021;0(0):1-4.
72. Rhidenour KB, Blackburn K, Barrett AK, Taylor S. Mediating Medical Marijuana: Exploring How Veterans Discuss Their Stigmatized Substance Use on Reddit. Health Communication. 2021;00(00):1-11.
73. Smolev ET, Rolf L, Zhu E, Buday SK, Brody M, Brogan DM, et al. “Pill Pushers and CBD Oil”—A Thematic Analysis of Social Media Interactions About Pain After Traumatic Brachial Plexus Injury. Journal of Hand Surgery Global Online. 2021;3(1):36-40.
74. Zenone M, Snyder J, Crooks VA. What are the informational pathways that shape people’s use of cannabidiol for medical purposes? Journal of Cannabis Research. 2021;3(1).
75. Soleymanpour M, Saderholm S, Kavuluru R. Therapeutic Claims in Cannabidiol (CBD) Marketing Messages on Twitter. 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Bioinformatics and Biomedicine (BIBM), 2021 IEEE International Conference on: IEEE; 2021. p. 3083-8.
76. Allem J-P, Majmundar A, Dormanesh A, Donaldson SI. Identifying Health-Related Discussions of Cannabis Use on Twitter by Using a Medical Dictionary: Content Analysis of Tweets. JMIR Form Res. 2022;6(2):e35027.

77. Yue L, Chen W, Li X, Zuo W, Yin M. A survey of sentiment analysis in social media. Knowledge and Information Systems. 2019;60(2):617-63.
78. Blei DM, Lafferty JD. Topic models. Chapman and Hall/CRC; 2009. p. 101-24.
79. Himelboim I. Social Network Analysis (Social Media). The International Encyclopedia of Communication Research Methods. 2017:1-15.
80. Allem JP, Escobedo P, Dharmapuri L. Cannabis surveillance with twitter data: Emerging topics and social bots. American Journal of Public Health. 2020;110(3):357-62.
81. Kazemi DM, Borsari B, Levine MJ, Dooley B. Systematic review of surveillance by social media platforms for illicit drug use. Journal of Public Health (United Kingdom). 2017;39(4):763-76.
82. Grinspoon L, Bakalar JB. Marihuana, the forbidden medicine: Yale University Press; 1997.
83. Belackova V, Ritter A, Shanahan M, Chalmers J, Hughes C, Barratt M, et al. Medicinal cannabis in Australia–Framing the regulatory options. Sydney: Drug Policy Modelling Program. 2015.
84. CPFG. Home Office Circular 1 November 2018: Rescheduling of cannabis-based products for medicinal use in humans. In: Unit CDaA, editor. London, United Kingdom: Crime, Policing, and Fire Group (CPFG); 2018.
85. Shover CL, Humphreys K. Six policy lessons relevant to cannabis legalization. Am J Drug Alcohol Abuse. 2019;45(6):698-706.
86. EMCDDA. Cannabis policy: status and recent developments Lisbon, Portugal: The European Monitoring Centre for Drugs and Drug Addiction; 2021 [cited 2021 August]. Available from: https://www.emcdda.europa.eu/publications/topic-overviews/cannabis-policy/html_en.
87. Lardon J, Bellet F, Aboukhamis R, Asfari H, Jaulent M, Beyens M, et al. Evaluating Twitter as a complementary data source for pharmacovigilance Expert Opinion on Drug Safety. 2018;17(8):763–74.
88. Bode L, Davis-Kean P, Singh L, Berger-Wolf T, Budak C, Chi G, et al. Study designs for quantitative social science research using social media. 2020.
89. Chancellor S, De Choudhury M. Methods in predictive techniques for mental health status on social media: a critical review. npj Digital Medicine. 2020;3(1).
90. Abbasi A, Sarker A, Ginn R, Smit K, Oconnor K, Abbasi A. Social Media Analytics for Smart Health.

Supporting information

S1. Quality assessment checklist (Doc)

S2. PRISMA checklist (Doc)

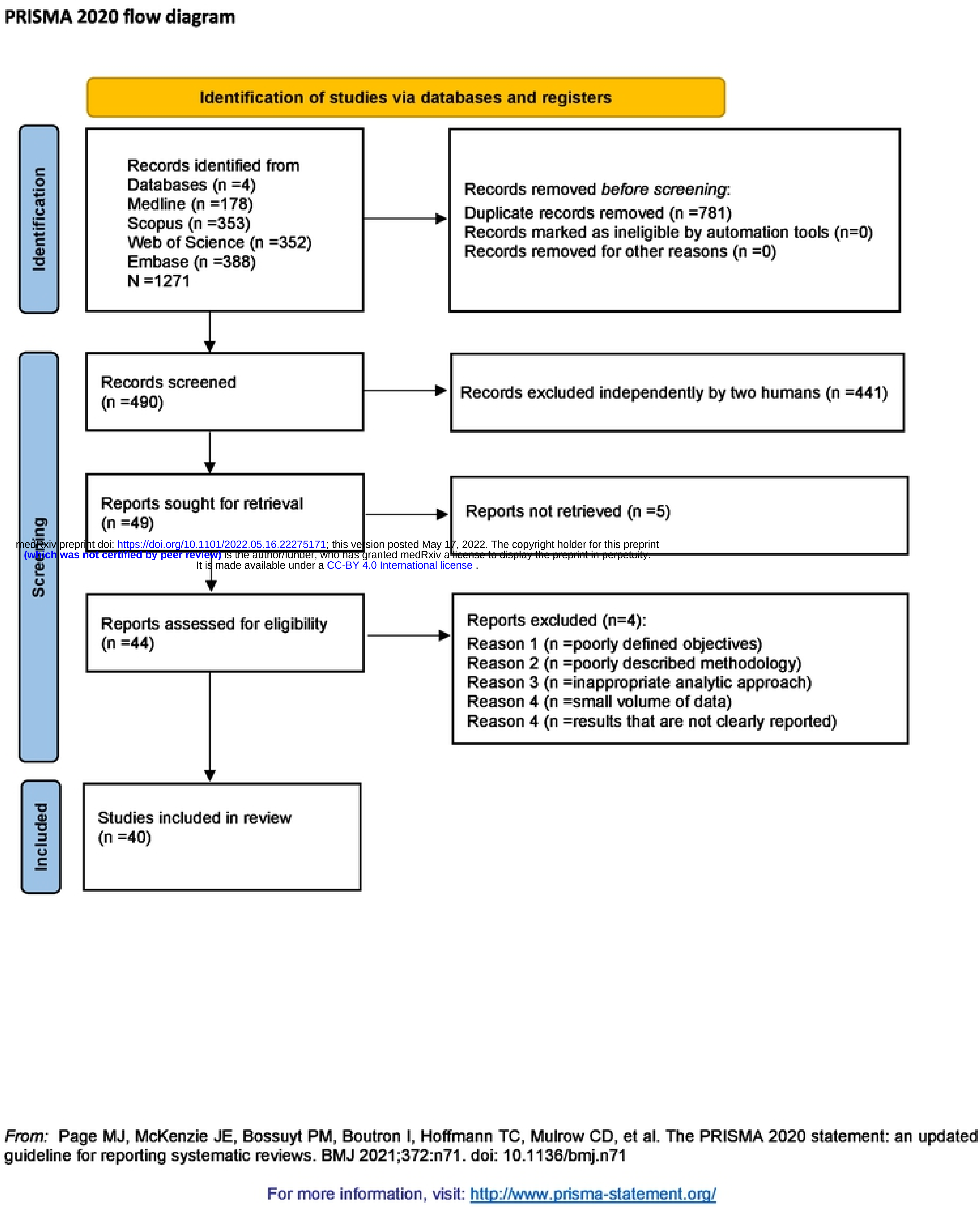

Fig1 - PRISMA flow diagram