




Sentiment analysis of social media posts on pharmacotherapy: A scoping review

Chanakya Sharma¹  | Samuel Whittle² | Pari D. Haghghi³ | Frada Burstein³ | Helen Keen¹

¹Rheumatology, Fiona Stanley Hospital, Perth, Australia

²Rheumatology, The Queen Elizabeth Hospital, Adelaide, Australia

³Faculty of Information Technology, Monash University, Melbourne, Australia

Correspondence

Chanakya Sharma, Department of Rheumatology, Fiona Stanley Hospital, 11 Robin Warren Drive, Murdoch 6150, Australia.

Email: chanakya_s@hotmail.com

Abstract

Social media is playing an increasingly central role in patient's decision-making process. Advances in technology have enabled meaningful interpretation of discussions on social media. We conducted a scoping review to assess whether Sentiment Analysis (SA), a big data analytic tool, could be used to extract meaningful themes from social media discussions on pharmacotherapy. A keyword search strategy was used on the following databases: OneSearch, PubMed, Medline, EMBASE, and Cochrane. One hundred and ninety-four titles were identified of which 10 studies were included. We extracted themes about uses and implications of SA of social media discussions on pharmacotherapy. Twitter was the most frequently analyzed platform. Assessment of public sentiment about a particular medication was the most common use of SA followed by detection of adverse drug reactions. Studies also revealed a significant impact of news media on public sentiment. Implications for real world practice include identifying reasons for a negative sentiment, detecting adverse drug reactions and using the impact of news media on social media sentiment to drive public health initiatives. The lack of a consistent approach to SA between the studies reflects the lack of a gold standard for the technology and consequently the need for future research. Sentiment Analysis is a promising technology that can allow us to better understand patient opinion regarding pharmacotherapy. This knowledge can be used to improve patient safety, patient-physician interaction, and also enhance the delivery of public health measures.

KEYWORDS

data mining, pharmacotherapy, sentiment analysis, social media

Abbreviations: LB, Lexicon based; ML, machine learning; SA, sentiment analysis; SVM, support vector machine.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Pharmacology Research & Perspectives* published by British Pharmacological Society and American Society for Pharmacology and Experimental Therapeutics and John Wiley & Sons Ltd

1 | INTRODUCTION

The development of Web 2.0 has allowed the internet to become a more interactive platform for its users, thus allowing social media to flourish.¹ Social media growth has been explosive, and its power to shape opinion is demonstrated by its impact on mass political movements including the Arab Spring, Brexit, and the American presidential election of 2016.^{2,3} The powerful impact social media can have on users and their friends and family is being explored by more and more industries (health care included) to gain insights into their user base and consequently drive change.^{4,5}

While there are different methods of conducting social media content analysis, one way to detect the aggregate opinion held towards a particular treatment is to analyze the sentiment expressed in social media posts. This can be done via a technique known as Sentiment Analysis (SA); also termed "opinion mining".⁶ Sentiment Analysis involves assigning an integer value to each word in a corpus of text, depending on the sentiment being expressed in that text. Words with negative sentiment get negative scores and vice versa.⁷ For example, the term "painful" might receive a negative score, whereas "beautiful" will usually receive a positive score.

Sentiment Analysis is usually conducted by one of two methods: Lexicon Based (LB) or Machine Learning (ML). The LB method requires the development of a "lexicon" or collection of words or phrases with their sentiment polarity mapped and scored. These words are then searched for in the target document and their scores are aggregated to obtain an overall sentiment score for the document.⁸ The ML methods use computer programs that allow classification of text, requiring the development of a program to detect sentiment, then training that program on a labelled, representative corpus of text, to assess and enhance its accuracy, prior to running it on the target document.⁹ However, before these methods are used, the data being analyzed is transformed from its raw format to one that is more readable by the software. This technique is known as "data pre-processing" and has been shown to improve the accuracy of data analysis.¹⁰

Each method has its own advantages and disadvantages. Lexicon based approaches do not require labeling and training of a classifier for each task, however, they are completely dependent on the lexica being used, which might have been generically designed and not specific for the topic being researched, thus impacting accuracy. Machine learning approaches on the other hand might demonstrate a high level of accuracy but require training of classifiers which can be time and cost consuming.¹¹

Healthcare is a frequently discussed topic in the online community, with patients using social media not only as a platform to discuss their medical conditions and treatment, but also to seek support.¹² In 2012, 26% of internet users were using social media for health issues, making it a rich source of information about patient beliefs.¹³ A key aspect of healthcare is pharmacotherapy, with adherence to prescribed medications being a pre-requisite for good health. However, studies show that majority of patients with chronic conditions are non-compliant with their prescribed

medications with up to 69% of hospital admissions being caused by this non-compliance.¹⁴ Patient's personal belief play a significant role in medication compliance, and recent studies have shown that online content strongly influences health these related beliefs and attitudes.¹⁵ These beliefs have traditionally been studied through qualitative, labour intensive methods; social media content analysis technique represents a novel approach to improving our understanding of patient beliefs.^{16,17}

The aim of this scoping review was to describe the available evidence as it pertains to SA of Social Media specifically about pharmacotherapy. Themes will be generated about the published uses of SA and the real-world implications of the knowledge generated.

2 | MATERIALS AND METHODS

Due to the novelty of the topic, we used a scoping review methodology to summarize all available information from a variety of sources. The framework outlined by Arksey and O'Malley was followed.¹⁸

The research question was identified as "*Can sentiment analysis be conducted on social media platforms to understand public sentiment held towards pharmacotherapy?*"

Social media is defined as "*a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content*".¹⁹ Pharmacotherapy was defined as the use of pharmaceutical drugs to treat or prevent medical conditions.

Literature published between 2002 (inception of web 2.0) and 2019 was collected from OneSearch, PubMed, Medline, EMBASE, and Cochrane. A keyword search strategy was employed using the words Sentiment Analysis, Opinion mining, Social Media, Medication, Pharmacotherapy, Drugs, Pharmaceutical, Medicine, Facebook, Twitter.

Articles were eligible for inclusion in this review if their primary aim was to conduct SA of social media posts regarding pharmacotherapy. Only articles published in English were included in this study. We also excluded articles that did not contain original data (eg letters to editor, opinion pieces). Reviews and Meta-analyses were excluded but manually searched for potential studies.

From all the included studies, information was collected on the following aspects on a predesigned template: authorship, year and journal published, social media platform(s) mined, medical condition(s), pharmacotherapy, type of SA used, outcomes generated, and potential use in clinical settings as described in the study.

3 | RESULTS

Our search strategy revealed 194 articles, 95 of which were excluded after title and abstract review for not meeting inclusion criteria. Of the remaining 99, 89 were excluded as they were not analysing at

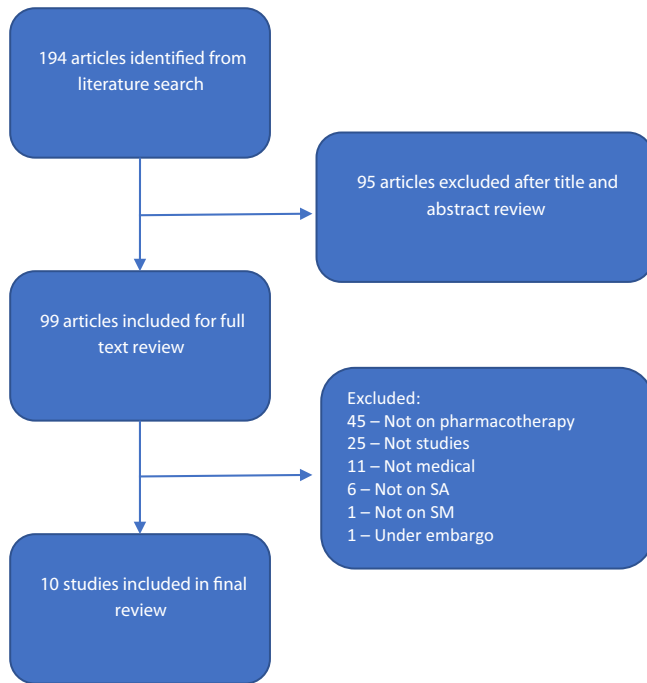


FIGURE 1 Study flow diagram

least one of the required topics of pharmacotherapy, medicine, or social media. A total of 10 studies were finally included (Figure 1)²⁰⁻²⁹

All the studies found were published after 2013. Eight of the ten included studies performed data mining on a single forum. Twitter was the most common platform mined (50%). Majority of the studies aimed to understand the sentiment being expressed towards a particular treatment, some of them also used this to explore other avenues such as adverse drug reaction detection, the role of new media in influencing social media sentiment and the sentiment dynamics on social media forums (Table 1).

3.1 | Sentiment analysis techniques and accuracy

Seven of the studies used a LB approach, two used ML and one used both methods. Most of the studies used a different lexicon for their analysis, with none of them being specifically geared for medical terminology. The studies that used ML algorithms also utilized different algorithms, namely AdaBoost Classifier in one and Support Vector Machine (SVM) in the other two. Both these are types of ML algorithms that allow stratification of data into different categories. While AdaBoost does this by sequentially weighting the results of weak classifiers to form a strong classifier, SVM finds the ideal margin to separate the dataset into desired categories.^{30,31}

The study by Ebrahimi et al was the only one that compared ML techniques to LB and also against manually classified sentiment. They used SVM to create a ML based algorithm and compared that to a LB algorithm. The ML algorithm outperformed the LB algorithm on both the primary (identifying forum posts mentioning drug side effects) and secondary objectives (identifying posts mentioning disease symptoms).²⁰

Data pre-processing was employed by five of the studies.^{20,21,23,24,27} The methods used by the studies varied, with tokenization (breaking sentences into small word groups or phrases that are more easily read by a program) being common. The other studies did not explicitly state whether they conducted data pre-processing, and if so then what techniques were used.

The study by Rocchetti et al compared the performance of its lexical SA technique to that of manual (human) coding of sentiment and found that there was a high degree of correlation for the extremes of sentiment (positive and negative), and less so for the neutral sentiments.²² Du et al conducted a manual analysis of a small corpus of tweets classified by their ML algorithm and found the overall accuracy to be acceptable.²⁴

3.2 | Sentiment analysis use

The most common application of SA (seven studies) was to analyze opinion regarding a particular medication.^{22-24,26,27,29} Six of these used LB approaches and one used ML. While majority of these studies directly analyzed the cumulative polarity of the posts for each medication, the study by Rocchetti et al reversed the process to analyze which therapy generated the strongest sentiment (positive or negative).

The next most common application of SA (three studies) was to identify adverse drug reactions (ADR) from social media chatter.^{20,21,28} The studies differed in both the platforms that they mined and the approach to SA. Ebrahimi et al mined an online forum (www.drugratingz.com) using both ML and LB algorithms to assess whether sentiment expressed in forum posts can be used to identify drug side effects from disease symptoms. Korkontzelos et al mined forums and tweets using five different LB methods to assess whether the addition of a SA feature to a pre-existing adverse drug reaction detection algorithm would improve its efficacy. Liu et al mined www.webmd.com, specifically reviewing diabetic medication forums. Their aim was to see if the addition of SA to pre-existing ADR detection algorithms would enhance detection. All three studies provided evidence that SA can be used to detect ADR mentions from social media posts.

One study also explored the interaction between news media and social media through the lens of sentiment.²⁴ Du et al analyzed the impact of sentiment towards Human Papilloma Virus (HPV) vaccination, as expressed by tweets, before and after publication of a positive New York Times article.²⁴ While the average number of tweets (positive, negative and neutral) pertaining to the topic was 1245 per day, the immediate period after publication of a New York Times article on HPV saw this number jump to 16,000 with the proportion of positive sentiment tweets rising from 35% to 66%. This was a remarkable demonstration of the impact of real-world events on social media sentiment.

Three studies analyzed the sentiment dynamics in cancer forums.²⁵⁻²⁷ The study by Portier et al looked at how the sentiment expressed by users in each thread influences the sentiment of the

TABLE 1 Summary of studies

Authors	Title, Journal and year	Data source and quality assessment (QA)	Type of SA and data pre-processing	Outcome of interest	Result	Significance
Ramagopalan et al ²³	Using Twitter to investigate opinions about multiple sclerosis treatments: a descriptive, exploratory study, F1000Research. 2014	Twitter QA not stated	LB - Hu & Liu's opinion lexicon Data pre-processing - Yes	The Sentiment Score (mean and summed) for each treatment	Overall positive sentiment scores for all drugs apart from Novantrone and Tysabri	Oral treatments had the highest mean summed scores which showing that patients prefer oral medications as opposed to injections
Portier et al ²⁵	Understanding Topics and Sentiment in an Online Cancer Survivor Community, Journal of the National Cancer Institute Monographs. 2013	Cancer survivors network QA not stated	ML using Adaboost classifier Data pre-processing - Not explicitly stated	Does the sentiment of the person making a post change with regards to responses received for that post?	Thread about treatment side effects had the lowest initial sentiment score, but also the greatest shift in sentiment (towards positive).	Treatment and side effect related posts are usually highly negative but are associated with the most shift in sentiment polarity, thus showing the positive support that is provided in the community
Rocchetti et al ²²	Attitudes of Crohn's Disease Patients: Infodemiology Case Study and Sentiment Analysis of Facebook and Twitter Posts, Journal of Medical Internet Research Public Health and Surveillance. 2017	Facebook and twitter QA: Used a "HoneyPot" approach to identify social spammers and to ensure that data being gathered is from patients.	LB using OpinionFinder Data pre-processing - Not explicitly stated	What topic within Crohn's disease generates that strongest sentiment from patients? Correlation between SA and human scores	Infliximab (an antibody used to treat Crohn's disease) was the most sentiment related term for both positive and negative sentiment High degree of correlation between positive and negative scores, less so for neutral score	This study showed that a data mining approach provided material of simple interpretation, regardless of the analysts' scientific and professional background. This shows that the analysis of such data can be completely automated with significant accuracy
Du et al ²⁴	Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data BioMed Central Medical Informatics and Decision Making. 2017	Twitter QA not stated	ML using SVM Data pre-processing - Yes	Sentiment toward HPV vaccination. Also looked at the impact of new media on sentiment and change in sentiment as it relates to the day of the week	35.8% were "Positive"; 32.1% were "Neutral"; and 32.0% tweets were "Negative". Safety was the biggest factor in negative tweets. They also found that mainstream media can have a significant influence on public opinion with 66.21% positive rate on the day a favorable news article was published compared to the previous positive rate of 35.8%	This study revealed the significant impact of mainstream media articles on public sentiment, a fact that can be used to promote public health

(Continues)

TABLE 1 (Continued)

Authors	Title, Journal and year	Data source and quality assessment (QA)	Type of SA and data pre-processing	Outcome of interest	Result	Significance
Cobb et al ²⁶	<i>Sentiment Analysis to Determine the Impact of Online Messages on Smokers' Choices to Use Varenicline</i> , Journal of the National Cancer Institute Monographs. 2013	QuitNet QA not stated	LB (Saliency Engine 4.1) Data pre-processing - No	Whether exposure to positive messages re: varenicline resulted in more people switching to it and sticking with it	Registrants who started or continued with varenicline were exposed to a statistically significantly greater number of positive-sentiment varenicline messages than negative-sentiment messages	While they cannot draw conclusions about causality, emotional content of online communications about health behavior intervention is associated with decision making around pharmaceutical choices
Korkontzelos et al ²¹	<i>Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts</i> , Journal of Biomedical Informatics. 2016	DailyStrength forum and Twitter	LB, 5 lexica used - the Hu&Liu Lexicon of Opinion Words (H&L), the Subjectivity Lexicon (SL), the NRC WordEmotion Association Lexicon (NRC), the NRC HashTag Sentiment Lexicon (NRC#), and the Sentiment 140 Lexicon (S140)	Whether the addition of sentiment analysis feature to ADRMine (a software already designed to pick up ADR mentions) would increase accuracy of picking up ADRs	There was an increase in pick up rate of ADRs for posts taken from twitter but not for posts from daily strength	Thus, there is potential for sentiment analysis to be used to pick up ADRs
Ebrahimi et al ²⁰	<i>Recognition of side effects as implicit-opinion words in drug reviews</i> Emerald Insight. 2016	www.drugratingz.com QA Not stated	ML using SVM and a Rule based version of lexicon based Data pre-processing - Yes	To evaluate if implicit sentiment can be used to identify drug side effects from disease symptom. These were tested against the manual annotation of the same drug reviews by a pharmacist	Experimental results show that ML outperforms the rule-based algorithm significantly for both disease symptom and especially side effect detection where it was almost two-fold better	The main finding was that drug review side effect recognition can be handled by using the ML algorithm, which significantly outperforms the regular expression-based algorithm

(Continues)

TABLE 1 (Continued)

Authors	Title, Journal and year	Data source and quality assessment (QA)	Type of SA and data pre-processing	Outcome of interest	Result	Significance
Liu et al ²⁶	<i>Adverse drug reaction related post detection using sentiment features</i>	Webmd.com; Manual annotation of posts done	LB - SentiWordNet Data pre-processing - Not stated	To use sentiment features to detect and identify if a post was related to an ADR. They compared the accuracy of detecting ADRs using three approaches; 1. Using N-gram and domain features 2. Adding sentiment to the above, 3. Using CHI statistic to select posts with high correlation between sentiment, n-gram and domain features	This method was very efficient in picking up ADR related posts. Compared to similar studies (which had use some of the methods but not all three) it had the highest F-measure (81.4%)	The addition of sentiment analysis to detect ADRs from social media forums results in greater accuracy than seen in previous methods
Cabling et al ²⁷	<i>Sentiment Analysis of an Online Breast Cancer Support Group: Communicating about Tamoxifen</i>	Breastcancer.org QA not stated	LB; Liu's dictionary Date pre-processing - yes	What is the sentiment expressed towards Tamoxifen	Most active users were 80% more positive than least active users, while the least active users were 48% more negative than the most active ones	Online support groups allow for stronger ties to be created around a specific sentiment, with less connection from those with dissimilar sentiments to the dominant group
Zhang et al ²⁹	<i>Utilizing twitter data for analysis of chemotherapy</i>	Twitter QA not stated	LB - using TextBlob Data pre-processing - Not explicitly stated	To assess and compare perceptions about chemotherapy of patients and healthcare providers through analysis of chemo-related tweets	Individuals are more likely to post emotional tweets about side effects than organizations	Twitter data can be used to understand behavioral patterns associated with treatments for cancer and for understanding how individuals and organizations communicate about health care concerns and discovering cancer patients' need, which could aid in developing personalized therapy plans

person who started the thread. They were able to show that discussions especially about pain and chemotherapy side effects typically started with a negative sentiment but gradually underwent a positive sentiment shift, reflecting the power of community support in improving sentiment.²⁵ The study by Cabling et al looked at the sentiment of the posters in a breast cancer forum on tamoxifen and found that the most active posters were more likely to have a positive sentiment than those who posted less frequently.²⁷ The study by Cobb et al was interesting as it was perhaps the only one to assess the direct impact of sentiment on compliance. It studied whether.

4 | DISCUSSION

This scoping review shows that SA can be used to gauge public perceptions regarding pharmacotherapy as expressed on social media. The most common application that emerged was of using SA to assess patient opinion regarding pharmacotherapy. While there was some consistency with regards to the platform being mined (Twitter being the most common), there was no consistent “gold standard” approach used by the authors to conduct SA. This likely reflects the fact that SA is still in its early stages of development, with various methods currently being explored in order to establish a standard.³²

Lexicon based approaches were more popular than ML based approaches, especially when the aim was to detect sentiment toward a particular treatment, with all of them being successful in detecting the sentiment expressed. The accuracy of this sentiment, as judged by a manual review, however, was infrequently done.^{22,24} Rocchetti et al conducted a manual analysis of a small corpus of tweets to judge the accuracy of their SA. This analysis was conducted by medical specialist and a software engineer who individually reviewed the posts and assigned a sentiment to each one. It was interesting to note that while the agreement between the two manual observers was good (kappa 0.647) it was not perfect, thus showing that even amongst human reviewers there can be disagreement about the underlying sentiment of the text being analyzed. While their algorithm had adequate accuracy in detecting positive and negative sentiment, it was more likely to classify those posts with less obvious sentiment as neutral. The one study (Du et al) that used a ML algorithm to analyze sentiment also conducted a manual comparison of a small corpus of tweets which suggested acceptable accuracy. It appears that SA might be unable to detect the polarity of posts with subtle sentiment and tends to classify them as neutral. This is a reassuring finding for two reasons, firstly, it would be better to classify a post with subtle positive or negative emotion as neutral than the opposite category (as was seen with the human reviewers where the computer scientist assigned more posts as either positive or negative than the gastroenterologist), thus highlighting that SA can negate some of the inherent experiential biases that come with human sentiment coding. Secondly, posts that describe significant ADRs are unlikely to have subtle emotion, thus more likely to be picked up by SA.

Three studies applied SA to improve the detection of ADRs, an important cause of morbidity and mortality.³³ While some ADRs are detected during clinical trials, a large number only become obvious during the post marketing surveillance phase.³⁴ There were significant differences between the studies in terms of both the platforms being mined (DailyStrength forum and Twitter, www.drugratingz.com and webmd.com) and the technique used (LB by two and both ML and LB by the other). The study by Korkontzelos et al added different types of lexicon-based SA to an existing adverse drug reaction detection program (ADRMine—an algorithm-based software designed to detect adverse drug reaction mentions in social media posts) to assess whether identification of negative sentiment would increase the detection rate. While ADRMine is designed to be highly sensitive, the addition of SA slightly improved the rate of detection of ADRs. The most successful lexica employed in this analysis were developed from Twitter, suggesting that SA is highly domain specific.³⁵ A similar study was conducted by Liu et al who added SA to pre-existing ADR detection processes such as N-gram and domain features and demonstrated that this resulted in increased detection of ADRs. In contrast, the study by Ebrahimi et al applied both LB and machine learning SA directly to the mined data and successfully detected ADRs from the forum posts. This was the only study that compared ML to LB algorithms, using manual review of the ADRs identified. While ML based approaches were superior at picking up ADR mentions and detection of disease effects, the authors concluded that both approaches were promising and that in future perhaps a hybrid of the two could be used for even more accuracy.²⁰

Another potential application of SA is understanding the interaction between news media and social media through the sentiment expressed. The article by Du et al showed the remarkable positive impact a positive news media publication can have on social media sentiment, thus demonstrating its potential use in public health. This is an exciting area deserving of further analysis as the relationship between News media and social media would provide a powerful tool to help promote and assess the efficacy of public health initiatives, especially relevant in the current pandemic.

Perhaps more important is the potential impact of social media sentiment on real-world behavior. This has already been demonstrated in other fields such as movies and stock markets, with positive sentiment resulting in positive box-office and market returns.^{36,37} Thus, the question arises whether social media sentiment might influence individual decisions related to pharmacotherapy. This concept was evaluated by Cobb et al who used SA to evaluate the impact of online messages on a smoker's decision to use a particular medication (Varenicline) to help them quit smoking.²⁶ They analyzed smokers who posted information about their pharmacotherapy use on QuitNet, a forum for smokers. Users who started or continued with varenicline were exposed to a greater number of positive sentiment varenicline messages and had a significantly higher ratio of positive to negative messages. While the authors refrained from drawing concrete conclusions on causality of sentiment on medication preference and compliance, the results certainly warrant further scrutiny with targeted studies.

Cabling et al also looked at the sentiment dynamics on medical forums (specifically Tamoxifen related posts on Breastcancer.org) and found that the most active posters were much more likely to express positive sentiment, thus perhaps explaining the positive sentiment that persistent users from Cobb et al study were exposed to.

The specifics of negative sentiment associated with certain medications and side effects suggests SA could be used to identify specific issues which could be addressed by individual clinicians with their patients, to allay their fears and improve adherence. This was demonstrated in the study by Ramagopalan et al on Multiple Sclerosis medications. This study revealed that patients preferred oral medications to injections and were more concerned about some side effects (eg infections) than others. Similarly, the study by Zhang et al was also able to demonstrate user sentiment towards specific side effects of chemotherapy, showing some side effects generate less negative sentiment ("nausea," "hair loss") as opposed to others ("Fatigue," "neuropathy"), which generated much more negative sentiment. This knowledge can be used by clinicians and pharmacists to better target medication related counselling, thus potentially improving adherence.

While this review does provide preliminary evidence that SA can be used to understand mass opinion about pharmacotherapy, several questions remain about the overall process and the technique of SA. We found heterogeneity between the studies at several stages of the analytic process, especially at the key stage of conducting the analysis but also at the earlier stage of data pre-processing and the subsequent stage of accuracy analysis. These different approaches are however not specific to SA of medical texts and reflect the ongoing development and evolution of the technology itself.³⁸ There is presently no universally accepted gold standard approach. Current evidence suggests that the choice of method may be domain-specific (depend on the condition/therapy being analyzed, the platform being mined and the outcome that is sought). The few studies that have compared the different approaches have generally failed to establish a gold standard, with each approach having its own set of advantages and disadvantages.^{39,40}

As the technology is further refined, standardization of methodology and the establishment of healthcare specific SA methods (either ML algorithms or a medical-sentiment lexicon) may facilitate the development of further validity regarding the application of this technology to the health care sector.^{41,42}

This review has a few limitations. Sentiment analysis is dependent on the domain or topic being studied, thus the lack of validated lexica or ML algorithms of conducting SA specific to the field of healthcare meant that the quality of the SA would be limited. Future work in this field to establish either a standardized medical lexicon or appropriate classifier would enhance the quality of the SA being conducted.

Our inclusion criteria were intentionally specific, thereby limiting the focus of SA just to the realm of pharmacotherapy; however, there are other applications of SA in the field of healthcare including (but not limited to) mining opinions regarding healthcare received,

determining clinical outcomes and understanding emotions of being unwell.⁴²

5 | CONCLUSION

This scoping review provides an overview of current evidence on the multifaceted applicability of SA. While the most obvious utilization is in the assessment of public sentiment about particular medications, the fact that SA is also being used for other tasks such as adverse drug reaction detection is a promising glimpse into the hitherto untapped potential of this technology. The heterogeneity of approach to SA across the studies reflects the rapid pace at which this technology continues to evolve. While it has already found use in the fields of commerce and marketing, its current state of clinical equipoise may be resolved if a universally agreed standardized approach is established. This will have far reaching consequences across various domains of healthcare, including but not limited to patient safety and public health initiatives.

DATA SHARING STATEMENT

Data are available upon reasonable request by contacting the corresponding author.

CONFLICT OF INTEREST

Authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Dr Chanakya Sharma – Data collection, analysis and writing the manuscript, study design. Dr Helen Keen – Analysis of data, editing of manuscript, study design. Dr Samuel Whittle – Analysis of data, editing of manuscript, study design. Dr Frada Burstein – Data collection, editing manuscript, Sentiment Analysis expert, study design. Dr Pari Delir Haghighi – Data collection, editing manuscript, Sentiment Analysis expert, study design.

ORCID

Chanakya Sharma  <https://orcid.org/0000-0002-4830-9547>

REFERENCES

1. Murugesan S. Understanding Web 2.0. *IT Prof Magazine*. 2007;9:34-41.
2. Wilson MI, Corey KE. The role of ICT in Arab spring movements. *Netcom*. 2014;2012:343-356.
3. Hall W, Tinati R, Jennings W. From brexit to trump: social media's role in democracy. *Computer*. 2018;51:18-27.
4. Bond RM, Fariss CJ, Jones JJ, et al. A 61-million-person experiment in social influence and political mobilization. *Nature*. 2012;489:295-298.
5. Kurniawati K, Nargiza B, Graeme S. *The Business Impact of Social Media Analytics*. ECIS 2013 - Proceedings of the 21st European Conference on Information Systems. 2013.
6. Pang B, Lee L. Opinion mining and sentiment analysis. *Found Trends Inf Ret*. 2008;2:1-135.
7. Bali R. In: Sarkar D, Sharma T, eds. *Sentiment Analysis. Learning Social Media Analytics with R*. Birmingham: Birmingham: Packt Publishing; 2017.

8. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M. Lexicon-based methods for sentiment analysis. *Comput Linguist.* 2011;37:267-307.
9. Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques. 2002.
10. Haddi E, Liu X, Shi Y. The role of text pre-processing in Sentiment analysis. *Procedia Comput Sci.* 2013;17:26-32.
11. Vohra SM, Teraiya JB. A comparative study of sentiment analysis techniques. 2013:313-337.
12. Benetoli A, Chen TF, Aslani P. How patients' use of social media impacts their interactions with healthcare professionals. *Patient Educ Couns.* 2018;101:439-444.
13. Hamm MP, Chisholm A, Shulhan J, et al. Social media use among patients and caregivers: a scoping review. *BMJ Open.* 2013;3:e002819.
14. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med.* 2005;4:487-497.
15. Hu Y, Shyam SS. Effects of online health sources on credibility and behavioral intentions. *Commun Res.* 2009;25:105-132.
16. Jelin E, Granum V, Eide H. Experiences of a web-based nursing intervention-interviews with women with chronic musculoskeletal pain. *Pain Manag Nurs.* 2012;13:2-10.
17. Gaspar R, Pedro C, Panagiotopoulos P, Seibt B. Beyond positive or negative: qualitative sentiment analysis of social media reactions to unexpected stressful events. *Comput Hum Behav.* 2016;1:179-191.
18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;1:19-32.
19. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus Horiz.* 2010;53:59-68.
20. Ebrahimi M, Yazdavar AH, Salim N, Eltyeb S. Recognition of side effects as implicit-opinion words in drug reviews. *Online Inform Rev.* 2016;4:1018-1032.
21. Korkontzelos I, Nikfarjam A, Shardlow M, Sarker A, Ananiadou S, Gonzalez GH. Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. *J Biomed Inform.* 2016;62:148-158.
22. Rocchetti M, Marfia G, Salomoni P, et al. Attitudes of Crohn's disease patients: Infodemiology case study and sentiment analysis of Facebook and Twitter posts. *JMIR Public Health Surveill.* 2017;3:e51.
23. Ramagopalan S, Wasiak R, Cox AP. Using Twitter to investigate opinions about multiple sclerosis treatments: a descriptive, exploratory study. *F1000Res.* 2014;3:216.
24. Du J, Xu J, Song H-Y, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. *BMC Med Inform Decis Mak.* 2017;17(Suppl 2):69.
25. Portier K, Greer GE, Rokach L, et al. Understanding topics and sentiment in an online cancer survivor community. *J Natl Cancer Inst Monogr.* 2013;2013:195-198.
26. Cobb NK, Mays D, Graham AL. Sentiment analysis to determine the impact of online messages on smokers' choices to use varenicline. *J Natl Cancer Inst Monogr.* 2013;2013:224-230.
27. Cabling ML, Turner JW, Hurtado-de-Mendoza A, et al. Sentiment Analysis of an online breast cancer support group: communicating about tamoxifen. *Health Commun.* 2018;33:1158-1165.
28. Liu J, Jiang X, Chen Q, Song M, Li J. Adverse drug reaction related post detecting using sentiment feature. *Iran J Public Health.* 2018;47:861-867.
29. Zhang L, Hall M, Bastola D. Utilizing Twitter data for analysis of chemotherapy. *Int J Med Informatics.* 2018;120:92-100.
30. Freund Y, Schapire R. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci.* 1997;55:119-139.
31. Devroye L, Györfi L, Lugosi G. *A Probabilistic Theory of Pattern Recognition. Stochastic Modelling and Applied Probability.* New York, NY: Springer; 1996:187-213.
32. Seerat B, Azam F. Opinion mining: Issues and challenges (a survey). *Int J Comput Appl.* 2012;49:42-51.
33. Sloane R, Osanlou O, Lewis D, Bollegala D, Maskell S, Pirmohamed M. Social media and pharmacovigilance: a review of the opportunities and challenges. *Br J Clin Pharmacol.* 2015;80:910-920.
34. Sultana J, Cutroneo P, Trifiro G. Clinical and economic burden of adverse drug reactions. *J Pharmacol Pharmacother.* 2013;4(Suppl 1):S73-S77.
35. Aue A, Gamon M. Customizing sentiment classifiers to new domains: a case study. In: Submitted to RANLP-05, the International conference on recent advances in natural language processing [Internet]. Borovets, BG; 2005.
36. Yu Y, Duan W, Cao Q. The impact of social and conventional media on firm equity value: a sentiment analysis approach. *Decis Support Syst.* 2013;55:919-926.
37. Rui H, Liu Y, Whinston A. Whose and what chatter matters? The effect of tweets on movie sales. *Decis Support Syst.* 2013;55:863-870.
38. Collomb A, Costea C, Joyeux D, Hasan O, Brunie L. A study and comparison of sentiment analysis methods for reputation evaluation. Rapport de recherche RR-LIRIS-2014-002. 2014.
39. Devika MD, Sunitha C, Ganesh A. Sentiment analysis: a comparative study on different approaches. *Procedia Comput Sci.* 2016;87:44-49.
40. Gonçalves P, Araújo M, Benevenuto F, Cha M. Comparing and combining sentiment analysis methods. 2014.
41. Balahur A, Jacquet G. Sentiment analysis meets social media – Challenges and solutions of the field in view of the current information sharing context. *Inf Process Manage.* 2015;51:428-432.
42. Denecke K, Deng Y. Sentiment analysis in medical settings: New opportunities and challenges. *Artif Intell Med.* 2015;64:17-27.

How to cite this article: Sharma C, Whittle S, Haghghi PD, Burstein F, Keen H. Sentiment analysis of social media posts on pharmacotherapy: A scoping review. *Pharmacol Res Perspect.* 2020;e00640. <https://doi.org/10.1002/prp2.640>