Journal of the American Medical Informatics Association, 26(6), 2019, 561–576
doi: 10.1093/jamia/ocz009
Advance Access Publication Date: 25 March 2019
Review

Review

A systematic literature review of machine learning in online personal health data

Zhijun Yin,[1] Lina M Sulieman,[1] and Bradley A Malin[1,2,3]

[1]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA, [2]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA, and [3]Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, Tennessee, USA

Corresponding Author: Zhijun Yin, PhD, Department of Biomedical Informatics, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 1412A, Nashville, TN 37203, USA (zhijun.yin@vanderbilt.edu)

Received 23 September 2018; Revised 6 January 2019; Editorial Decision 9 January 2019; Accepted 11 January 2019

ABSTRACT

Objective: User-generated content (UGC) in online environments provides opportunities to learn an individual's health status outside of clinical settings. However, the nature of UGC brings challenges in both data collecting and processing. The purpose of this study is to systematically review the effectiveness of applying machine learning (ML) methodologies to UGC for personal health investigations.

Materials and Methods: We searched PubMed, Web of Science, IEEE Library, ACM library, AAAI library, and the ACL anthology. We focused on research articles that were published in English and in peer-reviewed journals or conference proceedings between 2010 and 2018. Publications that applied ML to UGC with a focus on personal health were identified for further systematic review.

Results: We identified 103 eligible studies which we summarized with respect to 5 research categories, 3 data collection strategies, 3 gold standard dataset creation methods, and 4 types of features applied in ML models. Popular off-the-shelf ML models were logistic regression ($n = 22$), support vector machines ($n = 18$), naive Bayes ($n = 17$), ensemble learning ($n = 12$), and deep learning ($n = 11$). The most investigated problems were mental health ($n = 39$) and cancer ($n = 15$). Common health-related aspects extracted from UGC were treatment experience, sentiments and emotions, coping strategies, and social support.

Conclusions: The systematic review indicated that ML can be effectively applied to UGC in facilitating the description and inference of personal health. Future research needs to focus on mitigating bias introduced when building study cohorts, creating features from free text, improving clinical creditability of UGC, and model interpretability.

Key words: systematic review, machine learning, online environment, online health community, social media, patient portal, personal health

INTRODUCTION

Over the past decades, the structured data in electronic medical records (EMRs) have become critical resources for medical informatics research.[1–3] However, this clinically centric data often lack a patient's self-reported experiences and attitudes, as well as a characterization of their feelings and emotional states, thus providing only a partial view of a patient's health and wellness status.[4,5] As the Internet continues to permeate every aspect of daily life,[6,7] platforms that support anytime, anywhere communications have gained in popularity, such that individuals are increasingly sharing highly detailed information regarding many aspects of their life in online environments (eg, via social media platforms like Twitter and online health communities [OHCs] like PatientsLikeMe),[8,9] including their health and wellness.[10,11] This provides opportunities for healthcare providers and researchers to learn about an individual's health status and treatment experiences outside of clinical settings. This notion is supported by a recent systematic review,[12] which found that the benefits induced by incorporating online environments into health care included peer emotional support, public health surveillance, and potential to influence health policy.

However, there are various challenges associated with the collection and application of user-generated content (UGC) in online environments for healthcare research.[13] First, in social medical platforms, discussions can wander over a wide variety of topics, many of which are not necessarily pertinent to personal health.[14–16] Second, unlike the structured information in EMRs or clinical notes composed by healthcare providers, UGC generated by patients is often expressed in un- or semistructured text with layman words, such that interpretable factors need to be detected and extracted to gain intuitions into an individual's health status.[4,17] Third, in many situations, the health status or outcomes of the users of such environments need to be inferred from their discussion.[18] While manual review can be applied to tackle these challenges, such methods are often time consuming and lack scalability.[4] Crowdsourcing may speed up the process,[19] but it can be quite expensive (eg, domain experts are costly in their expertise and time, while the number of tasks could be on the scale of millions, leading to rapid cost escalation) and, in certain instances, privacy concerns limit the ability to share such data to crowd workers.[20] To address these challenges, automated techniques, often based on machine learning (ML), are increasingly adopted to process UGC in online environments.[21–24]

It should be noted that systematic reviews in this research area have been conducted, but they mainly focused on public health surveillance,[25] adverse drug reaction (ADR) detection,[26] and interventions on health-related behaviors through online environments.[27] While surveillance and ADR detection using massive online UGC can potentially track public health emergencies and enable drug safety, few reviews examine literature with a focus on personal health, which is important given the emerging era of precision medicine.[28] Online interventions can improve personal health-related behaviors, but most of them are experimental or behavioral studies. Considering that most patients, especially those with chronic diseases or assigned to long-term treatments, spend most of their time outside of formal clinical environments, UGC can provide additional resources to assist healthcare providers to learn about a patient's condition and treatment experience, and possibly predict their health status. In this systematic review, we investigated the effectiveness of applying UGC in online environments to study personal health though ML methods. Specifically, we summarized the personal health problems, the types of data, the ML methodologies, the scientific findings, and the challenges investigators encountered in this research area. In doing so, we provided intuitions into best practices for processing such data, as well as future challenges and opportunities.

MATERIALS AND METHODS

This investigation followed the guidelines of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework for preparation and reporting.[29]

Eligibility criteria

This study focused on peer-reviewed publications that applied ML to UGC to predict or infer factors related to personal health outcomes or behaviors. We used 3 criteria to search and screen publications that focused on (1) UGC in online environments, (2) quantitative analysis of UGC, and (3) personal health.

In particular, by invoking the first criterion, we concentrated on the free text that people openly expressed in online environments. This criterion leads to studies that relied on data from social media platforms (eg, Twitter), OHCs (eg, breastcancer.org), and patient portals (eg, MyHealthAtVanderbilt.com). It excludes the data generated by individuals through surveys or interviews, whereby respondents are required to answer predefined questionnaires in an online environment. The second criterion ensures that the selected publications applied techniques based on ML or statistical inference. It excludes publications that are based solely on a qualitative analysis. The third criterion indicates that the data under investigation was applied to learn about an individual's (or their relatives') health outcomes or behaviors. It excludes publications that focused on population-based phenomena, such as public health surveillance, drug interactions, and ADRs that were not focused on individual-level results.

Information sources and search

We searched for peer-reviewed publications in 5 resources: PubMed, Web of Science, ACM Library, IEEE Xplore, AAAI Library, and the ACL anthology. We restricted our search to research articles published in English and in peer-reviewed journals or conference proceedings (excluding conference affiliated and standalone workshops) between January 1, 2010, and June 30, 2018. We grouped the query keywords into 2 sets, which were combined through an AND operator. The first set of keywords correspond to online social media platforms, OHCs, and patient portals. We added an additional term, *health*, to confine the publications about online social media platforms to those that focused on health-related topics. The second set of keywords correspond to ML techniques, such as regression, classification, and prediction. We applied each query to the titles and abstracts of the publications. The queries for this study are available in Supplementary Appendix A.

Study selection

Two team members (LMS and ZY) independently evaluated the eligibility of publications. They first screened the publications by examining the titles, abstracts and methods, and then obtained eligible publications through reading full text. Disagreement was resolved by discussion with the third team member (BAM).

Data collection and analysis

The data that were documented for each eligible publication included the objectives, methods, environments, problems investigated, language, and dataset (see Table 1). A narrative synthesis of all eligible studies was conducted and organized with respect to (1) research questions, (2) ML methods, and (3) scientific findings. While the first perspective showed applications of UGC in healthcare research, the second perspective provided insights into current research methods and challenges when processing and analyzing UGC using ML. The third perspective demonstrated what can be learned from UGC regarding personal health issues.

RESULTS

Figure 1 illustrates the process of identifying eligible publications. Initially, our queries returned 3315 publications. We removed 534 duplicate publications and entries that were either workshop articles or not original studies in their own right, such as letters to editors, proceeding summaries, and descriptions of keynotes. Two team members read the abstracts and titles separately, leading to the removal of 2228 publications and retention of 553 publications for a

Table 1. Summary of the 103 eligible studies

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
Aramaki et al (2011)[30]	Classification (SVM)	Identifying tweets that mention actual influenza patients and detecting influenza epidemics	Twitter	Influenza	English	0.4 million
Qiu et al (2011)[31]	Classification (AdaBoost, LogitBoost, Bagging, SVM, Logistic Regression, Neural Networks, BayesNet, and Decision Tree)	Analyzing the sentiments of posts and the changes in the sentiments in the same thread	American Cancer Society Cancer Survivors Network	Breast Cancer	English	468 000 posts, 27 173 participants
Jamison-Powell et al (2012)[32]	Content Analysis (Keyword Search)	Exploring the discussion of mental health issues and coping with insomnia	Twitter	Insomnia	English	18 901 tweets
Wen and Rose (2012)[33]	Classification (Logistic Regression), Analysis	Examining the behavior and the disease trajectories overtime	Multiple Platforms: online breast cancer support groups	Breast Cancer	English	2145 users
Biyani et al (2013)[34]	Classification (Co-training with Logistic Regression)	Classifying the sentiment content in cancer survivor network	American Cancer Society Cancer Survivors Network	Cancer Survivor	English	786 000 posts
De Choudhury et al (2013)[35]	Classification, Prediction (SVM)	Identifying users with depression in Twitter	Twitter	Depression	English	69 514 tweets, 489 users
De Choudhury et al (2013)[36]	Classification (SVM)	Identifying mothers at risk of postpartum depression	Twitter	Postpartum Depression	English	376 mothers
De Choudhury et al (2013)[37]	Classification (SVM)	Detecting and diagnosing major depressive disorder in individuals	Twitter	Depression	English	476 users
Greenwood et al (2013)[38]	Classification (Naive Bayes)	Extracting the patient experience about chronic obstructive pulmonary disease	Multiple Platforms: 17 active blogs discussing chronic obstructive pulmonary disease	Chronic Obstructive Pulmonary Disease	English	100 posts
Lamb et al (2013)[39]	Classification (Log-Linear)	Differentiating the concerned awareness vs Infection influenza tweets	Twitter	Influenza	English	11 990 tweets
Lu (2013)[40]	Classification (SVM, C4.5, and Naive Bayes)	Classifying topics in online posts	Komen.org	General	English	4041 messages
Lu et al (2013)[41]	Content Analysis (Clustering)	Identifying health-related hot-topics in online forums (lung cancer, breast cancer, and diabetes)	MedHelp	General	English	100 000 messages
North et al (2013)[42]	Content Analysis	Evaluating the risk of death and emergent hospitalization using portal messages or eVisits	Patient Portal	General	English	7322 messages
Ofek et al (2013)[43]	Classification (Logistic Regression, Random Forest, Rotation Forest, and Adaboost), Sentiment Analysis	Understanding the sentiment of posts content	American Cancer Society Cancer Survivors Network	Cancer Survivor	English	468 000 posts
Sokolova et al (2013)[44]	Classification (semantic-based methods)	Identifying personal health information	Twitter	General	English	3017 tweets
Beykikhoshk et al (2014)[45]	Content Analysis, Classification (Naive Bayes and Logistic Regression)	Examining linguistic and semantic aspects of tweets about autism	Twitter	Autism	English	944 568 tweets

(continued)

Table 1. continued

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
Bodnar et al (2014)[46]	Classification (AdaBoost, Bayesian, Decision Tree, Logit Boost, Weighted Voting)	Identifying influenza diagnosis on Twitter	Twitter	Influenza	English	30 988 557 tweets, 913 186 users
Biyani et al (2014)[47]	Classification (Multinomial Naive Bayes)	Identifying emotional support and informational support in cancer patients posts	American Cancer Society Cancer Survivors Network	Cancer	English	240 posts
Chomutare et al (2014)[48]	Analysis, Recommendation	Modeling a recommendation system for threads in online patient communities	Diabetes, Obesity forums, and nonhealthcare data obtained from *Yahoo!* Research	Diabetes and Obesity	English	26 083 diabetes, 1436 obesity, 2360 yahoo posts
De Choudhury et al (2014)[21]	Classification (Logistic Regression)	Detecting users with postpartum depression	Facebook	Postpartum	English	578 220 posts, 165 mothers' accounts
De Choudhury et al (2014)[49]	Prediction (Negative Binomial Regression), Content Analysis	Studying the linguistic characteristics of mental health disclosure on Reddit	Reddit	Mental Health	English	20 411 posts, 27 102 users
Lin et al (2014)[50]	Classification (CNN)	Detecting individual's psychological stress using content and behavior patterns	Sina Weibo, Tencent Weibo, Twitter	Stress	English, Chinese	492 676 tweets, 23 304 users
Nguyen et al (2014)[51]	Classification (Logistic Regression With Lasso), Content Analysis, Sentiment Analysis	Classifying users in depression communities vs others	Multiple Platforms: Online Health Communities(Depression)	Depression	English	267 964 posts
Opitz et al (2014)[52]	Content Analysis	Extracting breast cancer topics from an online breast cancer forum	CancerDuSein.org	Breast Cancer	French	16 961 posts, 675 users
Paul and Dredze (2014)[53]	Content Analysis	Identifying general public health topics	Twitter	General	English	144 million tweets
Tuarob et al (2014)[54]	Classification (Random Forest, SVM, Naive Bayes)	Classifying social media posts and discovering health-related information	Twitter, Facebook	General	English	15 128 tweets, 10 000 status
Wilson et al (2014)[55]	Content Analysis	Analyzing the linguistic contents and the topic themes about depression on Twitter	Twitter	Depression	English	13 279 tweets
Adrover et al (2015)[56]	Analysis, Classification (AdaBoost, SVM, Bagging, Neural Network)	Determining the adverse effects of HIV drug treatment and associated sentiments	Twitter	Adverse Effects of HIV Drugs	English	1642 tweets
Beykikhoshk et al (2015)[57]	Classification (Naive Bayes, Logistic Regression with Lasso), Content Analysis (Word Cloud on pattern regarding hashtag, word frequency, and part of speech)	Analyzing the content of tweets about autism	Twitter	Autism Spectrum Disorder	English	11 million tweets
Burnap et al (2015)[58]	Classification (Ensemble Learning)	Identifying tweets that are related to suicide and containing worrying contents	Twitter	Suicide	English	2000 tweets
Chomutare et al (2015)[48]	Classification (Naive Bayes, SVM, and Decision Trees)	Discovering mood disorder cues from Internet chat messages	Multiple Platforms: Trained and evaluated model on Chat Lingo, evaluated on diabetes obesity and a non-health community	Severe Mood Disorders and Depression	English	200 messages

(continued)

Table 1. continued

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
Davis et al (2015)[59]	Content Analysis (Latent Semantic Analysis), Statistical Inference (Linear Mixed-Effect Model)	Examining the association between characteristics of the Facebook user and the likelihood of receiving responses after the announcement of a surgery	Facebook	Surgery	English	3899 users
De Choudhury (2015)[60]	Classification (SVM), Statistical Inference	Characterizing behavioral characteristics of anorexia users and detecting anorexia content	Tumblr	Anorexia	English	55 334 posts
Guan et al (2015)[61]	Classification (Logistic Regression and Random Forest)	Identifying users with high suicide possibility	Sina Weibo	Suicide	Chinese	909 users
Hu et al (2015)[62]	Classification, Prediction (Logistic Regression)	Classifying and predicting the depression scores	Sina Weibo	Depression	Chinese	10 102 questionnaires
Huang et al (2015)[63]	Classification (SVM)	Identifying suicide ideation	Sina Weibo	Suicide Ideation	Chinese	7314 users
Jimeno-Yepes et al (2015)[64]	Entity Recognition (CRF)	Evaluating state-of-the-art approaches to reproduce the manual annotations	Twitter	General (Knowledge Extraction)	English	1000 tweets
Kanouchi et al (2015)[65]	Classification (Logistic Regression)	Identifying the subject of a disease or symptom in a tweet	Twitter	Cold	English	3000 tweets
Kumar et al (2015)[22]	Content Analysis	Detecting the rate of suicide and the changes in posts after a celebrity's suicide	Reddit	Suicide	English	19 159 unique users
Nie et al (2015)[24]	Classification (Deep Learning)	Inferring the possible diseases automatically using questions in community-based health services	Multiple Platforms: EveryoneHealthy, WebMD, HealthTap, and MedlinePlus	General	English	220 000 1uestions and answers
Tamersoy et al (2015)[66]	Classification (Logistic Regression with Ridge)	Distinguishing individuals with long-term from short-term for smoking or drinking abstinence	Reddit	Smoking and Drinking Abstinence	English	146 036 posts
Tuarob et al (2015)[67]	Classification (Customized method)	Modeling the dynamics of infectious diseases and identifying individuals with infection	Facebook	Infection (SIRS and Influenza)	English	264 students
Yang and Mu (2015)[68]	Classification (Customized method, using context and *Diagnostic and* Statistical Manual of Mental Disorders)	Detecting depressed users in Twitter and analyzing spatial patterns	Twitter	Depression	English	286 users
Yin et al (2015)[14]	Classification (Naive Bayes)	Detecting the disclosing of personal health status and information	Twitter	General: 34 health issues	English	3400 tweets
Zhou et al (2015)[69]	Classification, Sentiment Analysis (SVM)	Identifying anti-vaccine opinions on Twitter	Twitter	HPV vaccine	English	42 533 tweets, 21 166 users
Ben-Sasson and Yom-Tov (2016)[70]	Classification (Decision Tree)	Classifying child's risk of autism using parents' narrative	*Yahoo!* Answers	Autism Spectrum Disorder	English	1081 queries
Braithwaite et al (2016)[71]	Classification (Decision Tree)	Identifying individuals with suicide risk on social media	Twitter	Suicide	English	135 users
Bui et al (2016)[72]	Classification (Adaboost), Sentiment Analysis (Causal Inference)	Analyzing the temporal causality of sentiment change	American Cancer Society Cancer Survivors Network	Cancer Survivor	English	468 000 posts

(continued)

Table 1. continued

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
Chancellor et al (2016)[23]	Survival Analysis (Cox Proportional Hazards Model)	Predicting the likelihood of pro-anorexia recovery	Tumblr	Anorexia	English	68 380 375 posts, 13 317 users
Chancellor et al (2016)[73]	Prediction (Regularized Multinomial Logistic Regression)	Quantifying, and forecasting mental illness severity in pro-eating disorder posts	Instagram	Mental illness severity for pro-eating disorder	English	434 000 posts
Daniulaityte et al (2016)[74]	Classification (Logistic Regression, Naive Bayes, and SVM)	Identifying the trends of cannabis- and synthetic cannabinoid–related drugs using tweets	Twitter	General (Drugs)	English	4000 tweets
Dao et al (2016)[75]	Sentiment Analysis	Extracting the factors of affective sentiment, mood, and emotional transitions	LiveJournal	Depression, Autism, and General Mental-Related Conditions	English	28 235 posts, 2000 users
De Choudhury et al (2016)[76]	Classification, Prediction (Regularized Logistic Regression)	Discovering shifts to suicidal ideation from mental health content in social media	Reddit	Suicidal Ideation	English	79 833 posts
He and Luo (2016)[77]	Classification (CMAR)	Identifying Tumblr and Twitter posts that encourage eating disorder	Tumblr, Twitter	Eating Disorder	English	5965 posts
Kavuluru et al (2016)[78]	Classification (SVM)	Identifying helpful comments of posts in the subreddit SuicideWatch community	Reddit	Suicide	English	3000 comments
Krishnamurthy et al (2016)[79]	Classification, Prediction (based on similarity between labeled and unlabled instance)	Identifying individuals with mental health issues	PatientsLikeMe, GoodNightJournal	Psychiatric Disorders and Behavioral Addiction	English	100 users
Lee et al (2016)[80]	Prediction (Odds Ratio)	Predicting individuals at risk of having back pain using tweets	Twitter	Back pain	English	742 028 tweets.
Marshall et al (2016)[81]	Analysis	Comparing and contrasting symptom clusters learned from messages on a breast cancer forum	MedHelp: breast cancer forum	Breast cancer	English	50 426 messages, 12 991 users
Niederkrotenthaler et al (2016)[82]	Content Analysis	Identifying the difference in message and communication style in suicide boards	Forum: 7 suicide message boards	Suicide	German	21 681 threads
Ping et al (2016)[83]	Content Analysis (K-Medoid Clustering)	Identifying cancer symptom clusters	Medhelp.com and Questionnaires	Breast Cancer Survivors	English	50 426 messages
Rus and Cameron (2016)[84]	Statistical Inference (Multi-level, Negative Binomial Regression)	Identifying the predictive features of user engagement in diabetes-related Facebook page	Facebook	Diabetes (Engagement)	English	500 posts
Saha et al (2016)[85]	Classification (Customized Method)	Classifying co-occurring mental health–related issues in online communities	Multiple Platforms: Online Health Communities	General	English	620 000 posts
Sarker et al (2016)[86]	Classification (Naive Bayes, SVM, Maximum Entropy, and Decision Tree), Content Analysis	Monitoring the abuse of prescription medication automatically	Twitter	Medication Abuse (Adderall, Oxycodone, and Quetiapine)	English	119 809 tweets
Yang et al (2016)[87]	Content Analysis (Modified LDA)	Analyzing the user-generated contents and sentiment in a health community	MedHelp	Breast Cancer	English	1568 threads

(continued)

Table 1. continued

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
Yin et al (2016)[15]	Classification (SVM, Logistic Regression, and Random Forest), Content analysis (NMF)	Detecting and learning the semantics patterns of health status disclosure on Twitter	Twitter	34 Health Issues	English	277 957 tweets
De Quincey et al (2016)[88]	Correlation Study	Examining correlation between fever trends on Twitter and reports from authorities	Twitter	Fever	English	512 000 tweets
Alimova et al (2017)[89]	Classification (SVM)	Extracting drugs' side effects and reactions from users' reviews	Online Health Community: Otzovik	Beneficial Effects, Adverse Effects, Symptoms of Drugs	Russian	580 reviews
Alnashwan et al (2017)[90]	Sentiment Analysis, Classification (Random Forest, Logistic Regression, and Neural Network)	Identifying the emotions categories and classifying users posts into the discovered emotions groups	Online Health Forums	Lyme Disease	English	1491 posts
Benton et al (2017)[91]	Predication (Multitask Learning, Deep Learning)	Predicting mental health	Twitter	Mental Health	English	9611 users
Birnbaum et al (2017)[92]	Classification	Combining algorithm and domain experts to classify schizophrenia users from the control	Twitter	Schizophrenia	English	671 users
Cheng et al (2017)[93]	Analysis (Logistic Regression)	Assessing one's suicide risk and emotional distress using onlitpdel.ne discussion	Weibo	Mental Health	Chinese	974 users
Cohan et al (2017)[94]	Classification (Boost Tree)	Classifying the severity of users' posts based on the indications of self-harm ideation	ReachOut.com	Mental Health: Self-harm	English	1188 posts
Cronin et al (2017)[95]	Classification (Logistic Regression, Random Forest, and Naive Bayes)	Classifying patient portal message and identifying the need communicated in the messages	Patient portals	Patients' needs	English	3253 messages
De Choudhury et al (2017)[96]	Analysis	Examining the gender-based and cross-cultural dimensions of mental health content on social media	Twitter	Mental Health	English	470 471 tweets
Du et al (2017)[97]	Classification (Naive Bayes, Random Forest, and SVMs)	Extracting opinions about HPV vaccines on Twitter using sentiment analysis	Twitter	HPV Vaccine	English	6000 tweets
Gkotsis et al (2017)[5]	Classification (CNN)	Classifying posts related to mental illness on Reddit	Reddit	Mental Health	English	900 000 subreddits
Huang et al (2017)[98]	Analysis	Detecting changing points in users committed suicide in weibo.	Weibo	Suicide	Chinese	130 users
Lim et al (2017)[99]	Classification	Detecting latent infectious disease communicated in Twitter	Twitter	Infectious Diseases	English	37 599 tweets
Luciana et al (2017)[100]	Classification (Deep Learning)	Classifying users with depression from users without it	Reddit	Depression	English	486 articles
Mowery et al (2017)[101]	Analysis	Developing annotated corpus to encode and analyze depressive symptoms and psychosocial stressors	Twitter	Behavioral Depression	English	9300 tweets
Nguyen et al (2017)[102]	Classification (Logistic Regression)		Live Journal	Mental Health	English	38 041 posts

(continued)

Table 1. continued

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
		Characterizing the online discussion between users in an online depression forum				
Nzali et al (2017)[103]	Content Analysis	Detecting discussed topics about breast cancer on social media and compare to Quality of Life Questionnaire Core topics	Facebook and other forums	Breast Cancer	French	86 960 messages
Oscar et al (2017)[104]	Analysis, Classification	Analyzing the content of Alzheimer's disease and dementia portrayal tweets	Twitter	Alzheimer	English	311 tweets
Roccetti et al (2017)[105]	Content Analysis and Sentiment Analysis	Extracting the topics and sentiments in Crohn's disease posts	Facebook	Crohn's disease	English	261 posts
Salas-Zarate et al (2017)[106]	Classification (based on the scores from SentiWordNet)	Analyzing the sentiment in diabetes related topics on Twitter	Twitter	Diabetes	English	900 tweets
Simms et al (2017)[107]	Classification (Decision Tree, Logistic Regression, Naive Bayes, Multilayer Perceptron, and K-Nearest-Neighbors)	Detecting cognitive distortion	Tumblr	Cognitive Distortion	English	459 posts
Smith et al (2017)[108]	Analysis	Evaluating posting patterns grouped by diseases	Facebook and EMR System	General	English	695 patients
Stanovsky et al (2017)[109]	Classification (RNN, Active learning)	Identifying mentions of ADRs	AskaPatient	General	English	1244 posts
Stewart and Abidi (2017)[110]	Content Analysis (Knowledge Map)	Applying semantic mapping technologies	Medical Mail List	General (Knowledge Extraction)	English	317 000 messages
Strapparava and Mihalcea (2017)[111]	Classification (Multinomial Naive Bayes)	Identifying the drugs behind experience	Erowid.org	General	English	4636 documents
Sulieman et al (2017)[112]	Classification (CNN)	Classifying patient portal messages into different groups	Patient Portals	Patient Needs	English	3000 messages
Vedula et al (2017)[113]	Content Analysis, Classification (Gradient-Boosted Decision Trees)	Detecting symptomatic cues of depression using linguistic and emotional signals	Twitter	Depression	English	150 users, 15 530 tweets
Wang et al (2017)[114]	Classification (Logistic Regression, CNN, and Customized Regularization Algorithm)	Detecting posts containing self-harming	Flickr	Self-harm	English	850 000 posts
Wang et al (2017)[115]	Analysis, Classification (Naive Bayes, Support Vector Machine and K-Nearest-Neighbors)	Detecting eating disorder communities and characterizing interactions among individuals	Twitter	Eating Disorder	English	1000 tweets.
Workewych et al (2017)[116]	Content Analysis and Sentiment Analysis	Characterizing the content of traumatic brain injury–related tweets	Twitter	Brain Injury	English	7483 tweets
Yazdavar et al (2017)[117]	Customized Method	Applying semisupervised method to evaluate how the duration of symp-	Twitter	Depressive Symptoms	English	7046 users

(continued)

Table 1. continued

Study	Methods	Objective	Environment	Health Issue	Language	Dataset Size
		toms mentioned on Twitter align with medical findings reported via Patient Health Questionnaire-9				
Zhang et al (2017)[118]	Classification (Conditional Random Field), Longitudinal Analysis	Identifying treatment mentions in social media	Autism-pdd.net	Autism Spectrum Disorder	English	500 posts
Zhang et al (2017)[119]	Analysis, Classification (LDA classifier, SVM, and CNN)	Classifying health topics and analyzing the changes in topics in online health communities	Breastcancer.org	Breast Cancer	English	1008 posts
Zhu et al (2017)[120]	Analysis and Classification (Word Embedding, Conditional Random Field)	Extracting medical events and temporal relations between them	Online Patient Consultation	Medical Events: Problem, Treatment, and Test	Chinese	8600 posts
Abdellaoui et al (2018)[121]	Content Analysis	Detecting the messages with noncompliant behavior for depression and psychotic drugs	French forums	Drug Compliance	French	5814 posts
Bryan et al (2018)[122]	Multilevel Model	Identifying and comparing temporal changes in suicide vs nonsuicide deaths for users who served in the military	Multiple social media sources	Suicide	English	315 users
Karisani and Agichtein (2018)[123]	Logistic Regression, Deep Learning	Detecting personal health problems in social media posts	Twitter	Influenza, Alzheimer's, Heart Attack, Parkinson's, Cancer, Depression, and Stroke	English	11 422 posts
Yadav et al (2018)[124]	Sentiment analysis (CNN)	Analyzing the sentiment about medical conditions reported by users	Patient.info	General	English	7490 posts

CMAR: classification based on multiple association rules; CNN: convolutional neural network; CRF: conditional random field; HPV: human papillomavirus; LDA: latent dirichlet allocation; NMF: non-negative matrix factorization; RNN: recurrent neural network; SIRS: systemic inflammatory response syndrome; SVM: support vector machine.

more in-depth review. After examining methods and accounting for the aforementioned inclusion and exclusion criteria, 173 publications were retained for further inspection. We excluded 2 studies for which we could not obtain access to the manuscripts. After the full article review, 66 additional publications were excluded because either (1) they failed to perform content analysis or (2) health care was not their primary focus. During this process, the 2 team members disagreed on 10 publications, for which the third team member broke the ties and recommended inclusion of 6 of them. Finally, 103 publications were included in the systematic review. Table 1 summarizes the publications with respect to their objectives, methods, dataset sizes, environments and posting languages, and investigated health issues.

Online platforms and languages

The choice of the social media platforms varied but was mostly dominated by Twitter (38 studies), Facebook (8 studies), Reddit (7 studies), and other OHCs (33 studies). There are 3 studies that examined the messages generated in patient portals. Most posts were published in English (89 studies). Other posting languages included Chinese (7 studies), French (3 studies), German (1 study), and Russian (1 study).

Personal health discussed in UGC

We summarized the research problems along 5 categories. The first is characterizing health issues and patients. These studies aimed to identify health problems, symptoms, and treatments,[14,22,50,61,63–65,68,70,78,94,101,107,116,119,123] as well as classify users into treatment vs control groups (eg, users with or without mental illness)[5,48,58,62,66,70,71,79,85,92,93,100,102,114,119,122,125]. The second is predicting the occurrence of a health issue (eg, suicide),[23,35,36,62,76,80,91,98] including learning posting patterns (eg, language and writing styles) and their capability of predicting a health issue or event (eg, anorexia, depression, undergoing surgery).[21,33,36,37,59,60] The third is investigating the correlation between posts about a health issue on social media and reports from authorities (eg, infectious disease).[46,88,99,117]; The fourth is characterizing pharmaceutical usage, including drug identification, ADRs, trends of drug usage, and drug abuse or addiction.[56,74,86,89,109,111] The fifth is detecting sentiments or emotions. These studies focused on sentiment classification, characterizing emotions when coping with a major health event (eg, postpartum depression)[31,56,74,75,87,90,105,106,124,126] and their impact on users' online posting behaviors.[72,113]

UGC processing and analysis using ML

Data collection. The datasets applied in these studies were mainly created through 3 methods: (1) snowball, (2) funnel, and (3) random sampling. In the snowball method, a small study cohort is carefully constructed following certain criteria and then is expanded based on their online social connections (eg, followers in Twitter, or post responding relationship in OHCs).[49,127] By contrast, the funnel method begins with a large dataset and excludes samples based on criteria defined by investigators or domain experts.[23,35,36,57,62,76] In random sampling methods, a dataset is randomly selected from an initial collected dataset.[14,15,70]

Creating a gold standard dataset. Lacking explicit clinical knowledge (eg, health status or treatment history), the data collected from online environments have to be annotated before further analysis. There are 3 strategies that were commonly adopted in these studies: (1) manual annotation, which promises a certain degree of accuracy, but has limited scalability[35–37]; (2) using keywords or patterns to filter dataset, which is fast, but occasionally inaccurate and biased to the predefined rules[30,32,33,57,121]; and (3) data-driven methods, in which ML models are first trained on a small number of labeled records and then, subsequently, applied to label an unannotated dataset in a scalable manner.[4,14–16,18,33,69,96,104,127]

Feature engineering. There were 4 main types of high-level features that were applied in these studies: (1) *post* summary statistics (eg, post length, time of publication, number or frequency of posts)[31,38,54,57]; (2) linguistic features, which can be further characterized into 4 subtypes—term-based features (eg, a bag of words or n-grams [of words or characters]),[15,39,45,47,65] grammar-related features (eg, parts of speech or dependency structure),[14,39,47] features based on dictionaries (eg, synsets, drug-slang lexicons, ontologies),[43,44,54,67,86,93,120] and topics (eg, word clusters based on either clustering algorithm or meaning extraction method, topics extracted using latent topic modeling methods [eg, Latent Dirichlet Allocation], and predefined semantic vocabularies [eg, Linguistic Inquiry and Word Count)[24,33,54,60,62,81,87,101,108,127]; (3) sentiments and emotions, including scores calculated through applying ML models, and emojis[14,31,34,43,57,127]; and (4) geographic location.[74]

Due to the high dimensionality of natural language, various processing techniques were applied to detect signals or reduce noise. This included using the odds ratio of features in treatment vs control groups, frequency analysis, and feature selection.[5,33,38,49,56,69,70] Recently, dense dimensional representation of words (eg, word2vec) has proven effective in building classifiers for short text.[4,15,16,18,124] In addition, many studies combined different types of features to improve model performance.[15,31,40,46,51,60,86]

Models. The ML methodologies in these studies were summarized based on their purpose into classification (67 studies), prediction (8 studies), content analysis (23 studies), sentiment analysis (9 studies), and other analysis (12 studies). It should be noted that we use the term *prediction* to refer to the models that applied information in the past to predict health outcomes or events in the future. Most reviewed studies adopted off-the-shelf models for either classification or prediction tasks. Logistic regression (with lasso or ridge regularization, 22 studies), support vector machines (18 studies), naive Bayes (17 studies), and ensemble learning (eg, random forests and adaboost; 12 studies) were the 4 most common models. Deep learning, though recently introduced, is rapidly becoming a popular technique in this domain (11 studies). In addition, 5 investigations proposed customized models to solve their specific research problems. Understanding and analyzing UGC often relied on unsupervised learning and analytical techniques. Clustering and topic modeling (23 studies) were the most common content analysis methods in these studies. Other analytical techniques included negative binomial regression,[49,84] survival analysis,[23] linear mixed-effect models,[59] and causal inference.[72]

Research findings about personal health

We grouped the findings of these studies by the investigated health issues: cancer, eating disorder and sleep issues, mental health, vaccines, and others. It should be noted that mental health and cancer were the 2 most popular studied health issues, and were investigated in 39 and 15 studies, respectively.

Mental health. Relying on predefined semantic vocabularies, studies showed that users with mental health problems (eg, postpartum depression) in online environments (eg, Twitter,

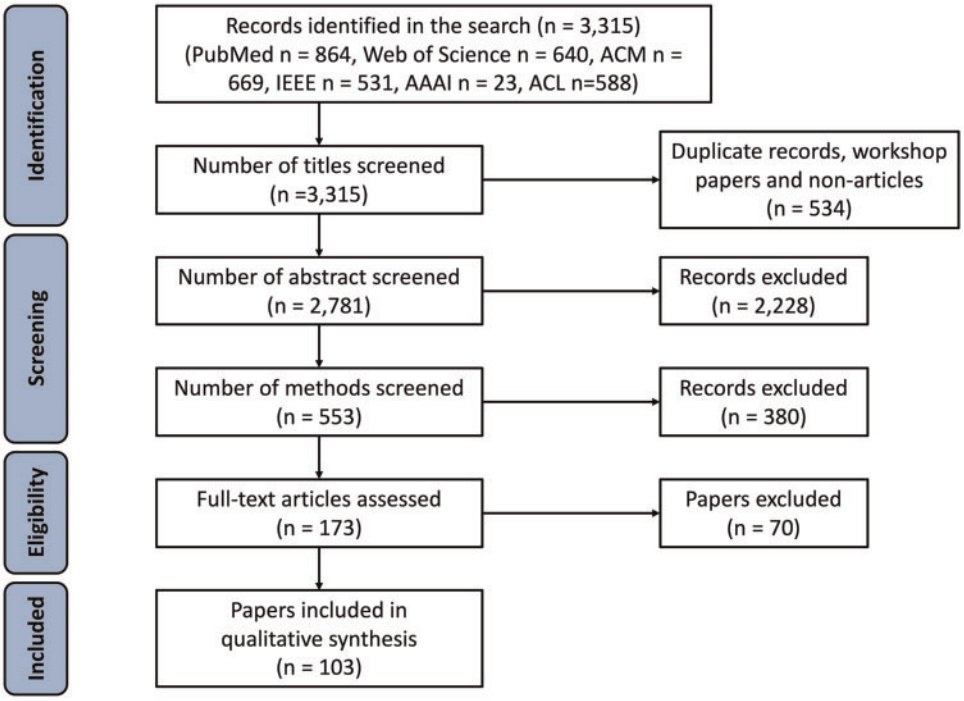

Figure 1. Illustration of the steps used in the literature search.

Facebook, Reddit) often expressed negative feelings and emotions, such as hopefulness and anxiety,[36,49,55,96] and exhibited lower social engagement and activity.[21,36,37] The posts of these users contained different psycholinguistics, writing and linguistic style, and poor linguistic coherence.[51,55,76] A general analysis on posting frequencies showed that social media has witnessed an increasing proportion of posts of medium and high mental illness severity.[73] Additionally, such changes, or the increasing number of posts about mental health problems, might be triggered by a major event. For instance, a study showed that the number of posts about suicide in Reddit increased after reporting a celebrity suicide.[22] Finally, a study showed that communications between users in suicide message boards, including active listening, sympathy, and provision of constructive advice, could improve the psychological content of UGC about their mental health.[82]

Cancer. Patients with cancer wrote about their symptoms, including pain, fatigue, sleep, weight change, and loss of appetite.[81,87] Moreover, patients talked about medications, daily matters, personal lives, and nutrition, as well as the complications they experienced after a procedure or taking medication.[52,83,87,119] The topics varied by cancer stages. For example, breast cancer patients in early stages tended to discuss their diagnosis, while patients in late cancer stages tended to establish online connections with others.[119] Further, cancer patients expressed their emotions in OHCs, which associated with different health-related behaviors (eg, stopping or completing hormonal therapy).[4] Other studies analyzed replies to posts of cancer patients and found that the sentiment of replies can influence others' sentiment, particularly the originators of discussion threads.[31,43,72,128,129]

Eating disorder and sleep issues: Social media analysis helped identify the characteristics of users with eating disorder or anorexia, including young age, high social anxiety, self-focused attention, deep negative emotions, and increased mental instability.[60,77,115] Some Reddit or Tumblr users with anorexia showed signs of recovery in their posts,[23,60] while others who exhibited eating disorder showed signs associated with their body image using #hashtags in Twitter.[77,115] Additionally, Twitter users with insomnia or sleeping problems made fewer connections to others and were less active in general, but were relatively more active at some specific times (eg, during sleep hours).[32,130]

Vaccines: Social media users expressed their opinions and attitudes toward vaccines. Studies found differences in the posts published by users who were anti-vaccine and users who were pro-vaccine. For example, through examining the language, studies found that anti-vaccinators on Twitter tended to use more direct language and exhibited more negative opinion and anger compared with the pro-vaccine posts.[69,97,127] A study suggested that surveillance of anti-vaccine opinion may help understand the driving factors of negative attitudes toward vaccines.[127]

Other health issues: Studying trends and detecting infectious diseases, such as fever, influenza, and systemic inflammatory response syndrome (SIRS), have been investigated extensively.[46,54,88,99] While some studies investigated mentions of medications on social media and found that users talked about medication abuse and outcomes,[56,74,86] other studies found general health topics that users posted about themselves.[53,103] For instance, through a content analysis, studies found that users in WebMD, PatientsLikeMe, YouTube, and Twitter talked about diagnoses, symptoms, feelings, and emotions, and the related therapeutic techniques.[41,42,108,110] Further, it is showed that users in Twitter and OHCs discussed chemicals, drugs and their efficiency, complications, and ADRs.[41,56] In autism OHCs, users who were mainly parents wrote about behaviors, needs, concerns, and treatments that their kids or themselves experienced in daily life.[57,70,118] Finally, a study using deep learning models showed that UGC in Flickr can be applied to detect attempts of self-harm by inspecting changes in patterns of language, platform usage, activity, and visual content.[114]

DISCUSSION

The majority of the reviewed studies demonstrated that UGC in online environments can be effectively applied to learn about personal health via ML. Our investigation suggested that UGC can be utilized to learn factors related to personal health that are rarely recorded in EMR systems. For example, with the help of ML techniques, UGC can be a useful data source to extract people's opinions, sentiments and emotions, coping strategies, and social support regarding a broad range of health issues (eg, cancer and mental health). This is significant because these factors can potentially influence a person's health-related behavior, confirming the importance of UGC in describing an individual's health. However, despite its notable advantages, UGC in online environments also brings challenges.

Clinical creditability of UGC: First, a majority of the studies assumed that what people claimed about their health status is credible. Yet this assumption might not be true. For example, 5 studies indicated that their findings did not correspond to a clinical diagnosis and the credibility of the findings depended on the reliability of the information posted by users.[23,48,81,96,130] Only 1 study investigated a cohort with health status confirmed with medical records.[46] Two other studies applied a Patient Health Questionnaire Depression Screening tool to detect users with depression.[21,37] By contrast, studies that investigated UGC in patient portals were discovered in the initial search, but only 3 met our inclusion criteria.[42,95,112] This suggests that there is a need for research that bridges the gap between the EMR information collected during clinical encounters and the patients' health information outside the clinical environment. Analyzing UGC and aligning the findings with EMR data may empower patients and provide clinicians with a more complete version of their health and life.

Challenges in Processing UGC: Second, the ML models applied to UGC have to deal with linguistic complications in the analysis of natural language text (eg, misspellings, jokes, humor, metaphors, ambiguity, sarcasm, grammar errors and emotions). For example, Twitter, as an all-purpose social media platform, is used to communicate various topics beyond one's personal health. Many studies applied keyword filtering or ML based methods to filter tweets for further analysis. While OHCs contain more health-focused discussion, these methods still need to be applied to extract particular health outcomes or events. For example, certain studies applied a combination of keyword searching and ML models to extract medication discontinuation events from the online discussion board in breastcancer.org.[4,16,18]

Further, 37 of the reviewed studies suffered from selection bias caused by the 3 summarized dataset creation methods. For instance, it is not uncommon for Twitter users to misspell a complex medical keyword (eg, writing *tamoxfen* instead of *tamoxifen*) or to use layman terms to describe health conditions (eg, using high blood pressure to represent hypertension). Keyword filtering could hardly capture all of these variations. The ML-based method exhibited a high precision, but it missed mentions that failed to follow the patterns incorporated into the models. Additionally, manual annotation was often applied to identify a small study cohort from Twitter or OHCs, which might not be able to represent the study population. Other biases caused by the nature of online environments include selection bias due to an individual's willingness to share online information and sampling bias (eg, analyzing only active users or specific users such as adult patients and healthcare providers).

Interpretability of models: Third, while there is an increasing body of research in this area that applied deep learning models to improve model performance instead of interpretability, most of the reviewed studies applied classical off-the-shelf models (eg, linear and logistic regression) and dedicated more effort on feature engineering. There were 3 major types of content-related features in these studies: (1) term-level features (eg, n-gram characters or words), (2) topic-level features (eg, topics, word clusters or semantic groups), and (3) sentiments or emotions. Interpreting sentiment or emotion features is straightforward because their values represent the magnitude of positivity in UGC. However, interpreting values of term-level or topic-level features is more challenging. For example, the coefficient of a topic in a logistic regression represents the rate of change in the log-odds when the distribution of the topic changes 1 unit. However, the 1-unit increase of a topic is often problematic in its interpretation for several reasons. First, it is difficult to explain what is 1-unit increase of a topic. Second, and perhaps most importantly, when a patient mentions a notable issue, such as a side effect, they might merely be indicating that there were no side effects experienced when taking a medication. Hence, there is a need to establish a more interpretable feature representation for training interpretable models, such as accounting for negation, building more robust topic models, and directly measuring health severity or emergence from text.

Application in Practice: Fourth, it is challenging to apply ML results in practice to benefit both patients and healthcare providers. For example, who is responsible for continuously monitoring a patient's posting behavior? Some studies suggested that on Reddit, the moderators can monitor such behavior and help direct in-time psychological services for users with potential mental health problem.[49,76] We believe that it will be fruitful for health care once an effective way is established to connect patients, platform moderators, and healthcare providers to solve this issue collaboratively. Finally, it should be noted that it is worth investigating the discrepancy between the information that a patient receives from their healthcare professionals and that from the online environment. Doing so provides insight into the extent to which the information patients receive in online settings reinforces or conflicts with their doctor's guidance.

Limitations

There are several limitations in this systematic review that we wish to acknowledge. First, we included many ML-related keywords in the search queries to cover as many related publications as possible. However, this process might miss some studies that failed to mention such terms. Second, we removed 345 workshop articles before screening eligible publications and 2 after full article review, which could be considered in future review. Third, we removed many studies that focused on public health and ADRs but neglected to investigate or discuss personal health. In addition, some studies that investigated immunizations and performed opinion mining were excluded because there was no further investigation on their impact on personal health. Finally, the ethical and privacy concerns of using ML methods to UGC is an important consideration,[131] but was beyond the focus of this systematic review.

CONCLUSION

This systematic review summarized how ML has been applied to UGC in online settings to study personal health issues. We specifically focused on the information that social media users shared about their health to seek information and support and to express opinions. While the findings of the reviewed studies (eg, creating study cohort, extracting sentiments and emotions, predicting depression, learning about cancer treatment experience) suggested that ML for the analysis of health information in online environments has advanced and achieved certain benefits, there remains a variety of challenges that need further investigation. These include, but are not limited to, the ethical aspects of analyzing personally contributed data, bias induced when building study cohorts and dealing with natural language, interpretation of modeling results, and reliability of the findings.

FUNDING

This work was supported by the National Science Foundation grant number IIS1418504.

AUTHOR CONTRIBUTORS

ZY and BAM contributed to the idea of the work. ZY and LMS performed article collection, screening, full article examination and summarization. BAM resolved the disagreement during the article screening process. ZY and LMS composed and revised the manuscript. BAM edited, commented and approved the final manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

Conflict of interest statement. None declared.

REFERENCES

1. Collen M, Ball M. *The History of Medical Informatics in the United States*. New York: Springer; 2015.
2. King J, Patel V, Jamoom EW, *et al*. Clinical benefits of electronic health record use: national findings. *Health Serv Res* 2014; 49 (1pt2): 392–404.
3. Bowton E, Field JR, Wang S, *et al*. Biobanks and electronic medical records: enabling cost-effective research. *Sci Transl Med* 2014; 6: 234cm3. doi:10.1126/scitranslmed.3008604. http://stm.sciencemag.org/content/6/234/234cm3/tab-pdf
4. Yin Z, Malin B, Warner J, *et al*. The power of the patient voice: learning indicators of treatment adherence from an online breast cancer forum. In: Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017), 2017: 337–46.
5. Gkotsis G, Oellrich A, Velupillai S, *et al*. Characterisation of mental health conditions in social media using informed deep learning. *Sci Rep* 2017; 7: 45141. doi:10.1038/srep45141. https://www.nature.com/articles/srep45141
6. Rahimi MR, Ren J, Liu CH, *et al*. Mobile cloud computing: a survey, state of art and future directions. *Mobile Netw Appl* 2014; 19 (2): 133–43.
7. Botta A, De Donato W, Persico V, *et al*. Integration of cloud computing and internet of things: a survey. *Future Gener Comput Syst* 2016; 56: 684–700.
8. Ma WWK, Chan A. Knowledge sharing and social media: altruism, perceived online attachment motivation, and perceived online relationship commitment. *Comput Human Behav* 2014; 39: 51–8.
9. Perrin A.Social media usage: 2005–2015. *Pew Res Cent* 2015; 2005–15. doi:202.419.4372. http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/. Last accessed on January 1, 2019.
10. Pittman M, Reich B. Social media and loneliness: why an instagram picture may be worth more than a thousand Twitter words. *Comput Human Behav* 2016; 62: 155–67.
11. Cookingham LM, Ryan GL. The impact of social media on the sexual and social wellness of adolescents. *J Pediatr Adolesc Gynecol* 2015; 28 (1): 2–5.
12. Moorhead SA, Hazlett DE, Harrison L, *et al*. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *J Med Internet Res* 2013; 15 (4): e85.
13. Househ M, Borycki E, Kushniruk A. Empowering patients through social media: the benefits and challenges. *Health Informatics J* 2014; 20 (1): 50–8.
14. Yin Z, Fabbri D, Rosenbloom ST, *et al*. A scalable framework to detect personal health mentions on Twitter. *J Med Internet Res* 2015; 17 (6): e138.
15. Yin Z, Chen Y, Fabbri D, *et al*. # Prayfordad: learning the semantics behind why social media users disclose health information. In: Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016), 2016: 456–65.
16. Yin Z, Xie W, Malin BA. Talking about my care: detecting mentions of hormonal therapy adherence behavior in an online breast cancer community. *AMIA Annu Symp Proc* 2017; 2017: 1868–77.
17. Gao J, Liu N, Lawley M, *et al*. An interpretable classification framework for information extraction from online healthcare forums. *J Healthc Eng* 2017; 2017: 2460174.
18. Yin Z, Song L, Malin B. Reciprocity and its association with treatment adherence in an online breast cancer forum. In: 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), IEEE, 2017: 618–23.
19. Ye C, Coco J, Malin B, *et al*. A crowdsourcing framework for sensitive and complex data sets. In: Proceedings of the American Medical Informatics Association Informatics Summit 2018, 2017: 273–80.
20. Khare R, Good BM, Leaman R, *et al*. Crowdsourcing in biomedicine: challenges and opportunities. *Brief Bioinform* 2016; 17 (1): 23–32.
21. De Choudhury M, Counts S, Horvitz EJ, *et al*. Characterizing and predicting postpartum depression from shared Facebook data. In: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social computing—CSCW '14, 2014: 626–38. doi:10.1145/2531602.2531675.
22. Kumar M, Dredze M, Coppersmith G, *et al*. Detecting changes in suicide content manifested in social media following celebrity suicides. In: Proceedings of the 26th ACM Conference on Hypertext & Social Media—HT '15, 2015: 85–94. doi:10.1145/2700171.2791026.
23. Chancellor S, Mitra T, De Choudhury M. Recovery amid pro-anorexia: analysis of recovery in social media. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016: 2111–23.
24. Nie L, Wang M, Zhang L, *et al*. Disease inference from health-related questions via sparse deep learning. *IEEE Trans Knowl Data Eng* 2015; 27 (8): 2107–19.
25. Velasco E, Agheneza T, Denecke K, *et al*. Social media and internet-based data in global systems for public health surveillance: a systematic review. *Milbank Q* 2014; 92 (1): 7–33.
26. Sarker A, Ginn R, Nikfarjam A, *et al*. Utilizing social media data for pharmacovigilance: a review. *J Biomed Inform* 2015; 54: 202–12.
27. Maher CA, Lewis LK, Ferrar K, *et al*. Are health behavior change interventions that use online social networks effective? A systematic review. *J Med Internet Res* 2014; 16 (2): e40.
28. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015; 372 (9): 793–5.
29. Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; 4: 1. doi:10.1186/2046-4053-4-1. https://systematicreviewsjournal.biomedcentral.com/track/pdf/10.1186/2046-4053-4-1
30. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using Twitter. In: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 2011: 1568–76. doi:10.1136/emermed-2011-200617.7.
31. Qiu B, Zhao K, Mitra P, *et al*. Get online support, feel better – sentiment analysis and dynamics in an online cancer survivor community. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and

Trust and Third International Conference on Social Computing, 2011: 274–81. doi:10.1109/PASSAT/SocialCom.2011.127.

32. Jamison-Powell S, Linehan C, Daley L, *et al*. "I can't get no sleep": discussing #insomnia on Twitter. In: Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems—CHI '12, 2012: 1501–10. doi:10.1145/2207676.2208612.
33. Wen M, Rose CP. Understanding participant behavior trajectories in online health support groups using automatic extraction methods. In: Proceedings of the 17th ACM International Conference on Supporting Group work - GROUP '12, 2012: 179–88. doi:10.1145/2389176.2389205.
34. Biyani P, Caragea C, Mitra P, *et al*. Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community. In: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining—ASONAM '13, 2013: 413–7. doi:10.1145/2492517.2492606.
35. De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. In: Proceedings of the 5th Annual ACM Web Science Conference—WebSci '13, 2013: 47–56. doi:10.1145/2464464.2464480.
36. De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13, 2013: 3267–76. doi:10.1145/2470654.2466447.
37. De Choudhury M, Gamon M, Counts S, *et al*. Predicting depression via social media. In: Proceedings of the 7th International AAAI Conference on Weblogs and Social Media, 2013: 128–38. doi:10.1109/IRI.2012.6302998.
38. Greenwood M, Elwyn G, Francis N, *et al*. Automatic extraction of personal experiences from patients' blogs: a case study in chronic obstructive pulmonary disease. In: 2013 International Conference on Cloud and Green Computing, 2013: 377–82. doi:10.1109/CGC.2013.66.
39. Lamb A, Paul M, Dredze M. Separating fact from fear: tracking flu infections on Twitter. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2013: 789–95. doi:10.1.1.378.2542.
40. Lu Y. Automatic topic identification of health-related messages in online health community using text classification. *Springerplus* 2013; 2: 309. doi:10.1186/2193-1801-2-309. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3736074/pdf/40064_2012_Article_423.pdf
41. Lu Y, Zhang P, Liu J, *et al*. Health-related hot topic detection in online communities using text clustering. *PLoS One* 2013; 8: e56221. doi:10.1371/journal.pone.0056221. https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0056221&type=printable
42. North F, Crane SJ, Stroebel RJ, *et al*. Patient-generated secure messages and eVisits on a patient portal: are patients at risk? *J Am Med Inform Assoc* 2013; 20 (6): 1143–9.
43. Ofek N, Caragea C, Biyani P, *et al*. Improving sentiment analysis in an online cancer survivor community using dynamic sentiment lexicon. In: Proceedings—2013 International Conference on Social Intelligence and Technology, SOCIETY 2013, 2013: 109–13. doi:10.1109/SOCIETY.2013.20.
44. Sokolova M, Matwin S, Schramm D. How joe and jane tweet about their health: mining for personal health information on Twitter. In: Proceedings of Recent Advances in Natural Language Processing, 2013: 626–32.
45. Beykikhoshk A, Arandjelovic O, Phung D, *et al*. Data-mining Twitter and the autism spectrum disorder: a pilot study. In: 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), 2014: 349–56. doi:10.1109/ASONAM.2014.6921609.
46. Bodnar T, Barclay VC, Ram N, *et al*. On the ground validation of online diagnosis with Twitter and medical records. In: Proceedings of the 23rd International Conference on World Wide Web, ACM, New York, NY, USA, 2014: 651–6. doi:10.1145/2567948.2579272.
47. Biyani P, Caragea C, Mitra P, *et al*. Identifying emotional and informational support in online health communities. In: Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, 2014: 827–36.
48. Chomutare T, Arsand E, Hartvigsen G. Mining symptoms of severe mood disorders in large internet communities. In: Proceedings of the IEEE Symposium on Computer-Based Medical Systems, 2015: 214–9. doi:10.1109/CBMS.2015.36.
49. De Choudhury M, De S, Choudhury MD, *et al*. Mental health discourse on reddit: self-disclosure, social support, and anonymity. In: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, 2014: 71–80.
50. Lin H, Jia J, Guo Q, *et al*. User-level psychological stress detection from social media using deep neural network. In: Proceedings of the ACM International Conference on Multimedia—MM '14, 2014: 507–16. doi:10.1145/2647868.2654945.
51. Nguyen T, Phung D, Dao B, *et al*. Affective and content analysis of online depression communities. *IEEE Trans Affective Comput* 2014; 5 (3): 217–26.
52. Opitz T, Aze J, Bringay S, *et al*. Breast cancer and quality of life: medical information extraction from health forums. *Stud Health Technol Inform* 2014; 205: 1070–4.
53. Paul MJ, Dredze M. Discovering health topics in social media using topic models. PLoS ONE 9(8): e103408. doi:10.1371/journal.pone.0103408.
54. Tuarob S, Tucker CS, Salathe M, *et al*. An ensemble heterogeneous classification methodology for discovering health-related knowledge in social media messages. *J Biomed Inform* 2014; 49: 255–68.
55. Wilson ML, Ali S, Valstar MF. Finding information about mental health in microblogging platforms. In: Proceedings of the 5th Information Interaction in Context Symposium on - IIiX '14, 2014: 8–17. doi:10.1145/2637002.2637006.
56. Adrover C, Bodnar T, Huang Z, *et al*. Identifying adverse effects of hiv drug treatment and associated sentiments using Twitter. *JMIR Public Health Surveill* 2015; 1 (2): e7.
57. Beykikhoshk A, Arandjelović O, Phung D, *et al*. Using Twitter to learn about the autism community. *Soc Netw Anal Min* 2015; 5: 22. doi:10.1007/s13278-015-0261-5.
58. Burnap P, Colombo W, Scourfield J. Machine classification and analysis of suicide-related communication on Twitter. In: Proceedings of the 26th ACM Conference on Hypertext and Social Media—HT '15, 2015: 75–84. doi:10.1145/2700171.2791023.
59. Davis MA, Anthony DL, Pauls SD. Seeking and receiving social support on Facebook for surgery. *Soc Sci Med* 2015; 131: 40–7.
60. Choudhury MD. Anorexia on Tumblr: a characterization study studies on anorexia. In: Proceedings of the 5th ACM Digital Health Conference DH'15, 2015: 43–50. doi:10.1145/2750511.2750515.
61. Guan L, Hao B, Cheng Q, *et al*. Identifying Chinese microblog users with high suicide probability using internet-based profile and linguistic features: classification model. *JMIR Ment Health* 2015; 2(2): e17. doi:10.2196/mental.4227.
62. Hu Q, Li A, Heng F, *et al*. Predicting depression of social media user on different observation windows. In: Proceedings of the 2015 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, WI-IAT 2015, 2016: 361–4. doi:10.1109/WI-IAT.2015.166.
63. Huang X, Li X, Zhang L, *et al*. Topic model for identifying suicidal ideation in Chinese microblog. In: 29th Pacific Asia Conference on Language, Information and Computation, 2015: 553–62.
64. Jimeno-Yepes A, MacKinlay A, Han B, *et al*. Identifying diseases, drugs, and symptoms in Twitter. *Stud Health Technol Inform* 2015; 216: 643–7. doi:10.3233/978-1-61499-564-7-643.
65. Kanouchi S, Komachi M, Okazaki N, *et al*. Who caught a cold?—identifying the subject of a symptom. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015: 1660–70.
66. Tamersoy A, De Choudhury M, Chau DH. Characterizing smoking and drinking abstinence from social media. In: Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT '15, 2015: 139–48.

67. Tuarob S, Tucker CS, Salathe M, *et al*. Modeling individual-level infection dynamics using social network information. In: Proceedings of the International Conference on Information and Knowledge Management, 2015: 1501–10. doi:10.1145/2806416.2806575.
68. Yang W, Mu L. GIS analysis of depression among Twitter users. *Appl Geogr* 2015; 60: 217–23.
69. Zhou X, Coiera E, Tsafnat G, *et al*. Using social connection information to improve opinion mining: identifying negative sentiment about HPV vaccines on Twitter. *Stud Health Technol Inform* 2015; 216: 761–5. doi:10.3233/978-1-61499-564-7-761.
70. Ben-Sasson A, Yom-Tov E. Online concerns of parents suspecting autism spectrum disorder in their child: content analysis of signs and automated prediction of risk. *J Med Internet Res* 2016; 18(11): e300. doi:10.2196/jmir.5439.
71. Braithwaite SR, Giraud-Carrier C, West J, *et al*. Validating machine learning algorithms for Twitter data against established measures of suicidality. *JMIR Ment Health* 2016; 3(2): e21. doi:10.2196/mental.4822.
72. Bui N, Yen J, Honavar V. Temporal causality analysis of sentiment change in a cancer survivor network. *IEEE Trans Comput Soc Syst* 2016; 3 (2): 75–87.
73. Chancellor S, Lin Z, Goodman EL, *et al*. Quantifying and predicting mental illness severity in online pro-eating disorder communities. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing—CSCW '16, 2016: 1169–82. doi:10.1145/2818048.2819973.
74. Daniulaityte R, Chen L, Lamy FR, *et al*. "When 'bad' is 'good'": identifying personal communication and sentiment in drug-related tweets. *JMIR Public Health Surveill* 2016; 2 (2): e162.
75. Dao B, Nguyen T, Venkatesh S, *et al*. Discovering latent affective dynamics among individuals in online mental health-related communities. In: Proceedings of the IEEE International Conference on Multimedia and Expo, 2016: 1–6. doi:10.1109/ICME.2016.7552947.
76. De Choudhury M, Kiciman E, Dredze M, *et al*. Discovering shifts to suicidal ideation from mental health content in social media. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI '16, 2016: 2098–110. doi:10.1145/2858036.2858207.
77. He L, Luo J. What makes a pro eating disorder hashtag': using hashtags to identify pro eating disorder tumblr posts and Twitter users. In: Proceedings of the 2016 IEEE International Conference on Big Data, Big Data 2016, 2016: 3977–9. doi:10.1109/BigData.2016.7841081.
78. Kavuluru R, Ramos-Morales M, Holaday T, *et al*. Classification of helpful comments on online suicide watch forums. In: Proceedings of the 7th ACM International Conference on Bioinformatics. Computational Biology, and Health Informatics—BCB '16, 2016: 32–40. doi:10.1145/2975167.2975170.
79. Krishnamurthy M, Mahmood K, Marcinek P. A hybrid statistical and semantic model for identification of mental health and behavioral disorders using social network analysis. In: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), IEEE, 2016: 1019–26. doi:10.1109/ASONAM.2016.7752366.
80. Lee H, McAuley JH, Hübscher M, *et al*. Tweeting back: predicting new cases of back pain with mass social media data. *J Am Med Inform Assoc* 2016; 23 (3): 644–8.
81. Marshall SA, Yang CC, Ping Q, *et al*. Symptom clusters in women with breast cancer: an analysis of data from social media and a research study. *Qual Life Res* 2016; 25 (3): 547–57.
82. Niederkrotenthaler T, Gould M, Sonneck G, *et al*. Predictors of psychological improvement on non-professional suicide message boards: content analysis. *Psychol Med* 2016; 46 (16): 3429–42.
83. Ping Q, Yang CC, Marshall SA, *et al*. Breast cancer symptom clusters derived from social media and research study data using improved k-medoid clustering. *IEEE Trans Comput Soc Syst* 2016; 3 (2): 63–74.
84. Rus HM, Cameron LD. Health communication in social media: message features predicting user engagement on diabetes-related Facebook pages. *Ann Behav Med* 2016; 50 (5): 678–89.
85. Saha B, Nguyen T, Phung D, *et al*. A framework for classifying online mental health-related communities with an interest in depression. *IEEE J Biomed Health Inform* 2016; 20 (4): 1008–15.
86. Sarker A, O'Connor K, Ginn R, *et al*. Social media mining for toxicovigilance: Automatic monitoring of prescription medication abuse from Twitter. *Drug Saf* 2016; 39 (3): 231–40.
87. Yang F-C, Lee AJT, Kuo S-C. Mining health social media with sentiment analysis. *J Med Syst* 2016; 40 (11): 236.
88. de Quincey E, Kyriacou T, Pantin T. #hayfever; a longitudinal study into hay fever related tweets in the UK. In: Proceedings of the 6th International Conference on Digital Health Conference—DH '16, 2016: 85–9. doi:10.1145/2896338.2896342.
89. Alimova I, Tutubalina E, Alferova J, *et al*. A Machine learning approach to classification of drug reviews in Russian. In: 2017 Ivannikov ISPRAS Open Conference (ISPRAS), 2017: 64–9. doi:10.1109/ISPRAS.2017.00018.
90. Alnashwan R, Sorensen H, O'Riordan A, *et al*. Multiclass sentiment classification of online health forums using both domain-independent and domain-specific features. In: Proceedings of the Fourth IEEE/ACM International Conference on Big Data Computing, Applications and Technologies - BDCAT '17, 2017: 75–83. doi:10.1145/3148055.3148058.
91. Benton A, Mitchell M, Hovy D. Multitask learning for mental health conditions with limited social media data. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, 2017: 152–62.
92. Birnbaum ML, Ernala SK, Rizvi AF, *et al*. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *J Med Internet Res* 2017; 19 (8): e289.
93. Cheng Q, Li TM, Kwok C-L, *et al*. Assessing suicide risk and emotional distress in Chinese social media: a text mining and machine learning study. *J Med Internet Res* 2017; 19(7): e243. doi:10.2196/jmir.7276.
94. Cohan A, Young S, Yates A, *et al*. Triaging content severity in online mental health forums. *J Assoc Inf Sci Technol* 2017; 68 (11): 2675–89.
95. Cronin RM, Fabbri D, Denny JC, *et al*. A comparison of rule-based and machine learning approaches for classifying patient portal messages. *Int J Med Inf* 2017; 105: 110–20.
96. De Choudhury M, Sharma SS, Logar T, *et al*. Gender and cross-cultural differences in social media disclosures of mental illness. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing—CSCW '17, 2017: 353–69. doi:10.1145/2998181.2998220.
97. Du J, Xu J, Song H, *et al*. Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. *J Biomed Semantics* 2017; 8: 9. doi:10.1186/s13326-017-0120-6. https://jbiomedsem.biomedcentral.com/articles/10.1186/s13326-017-0120-6
98. Huang X, Xing L, Brubaker JR, *et al*. Exploring timelines of confirmed suicide incidents through social media. In: Proceedings of 2017 IEEE International Conference on Healthcare Informatics, ICHI 2017; 2017: 470–7. doi:10.1109/ICHI.2017.47.
99. Lim S, Tucker CS, Kumara S. An unsupervised machine learning model for discovering latent infectious diseases using social media data. *J Biomed Inform* 2017; 66: 82–94.
100. Mariñelarena-Dondena L, Ferreti E, Maragoudakis M, *et al*. Predicting depression: a comparative study of machine learning approaches based on language usage. *Cuad Neuropsicol/Panam J Neuropsychol* 2017; 11 (3): 42–54.
101. Mowery D, Smith H, Cheney T, *et al*. Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. *J Med Internet Res* 2017; 19(2): e48. doi:10.2196/jmir.6895.
102. Nguyen T, O'Dea B, Larsen M, *et al*. Using linguistic and topic analysis to classify sub-groups of online depression communities. *Multimed Tools Appl* 2017; 6 (8): 10653–76.
103. Tapi Nzali MD, Bringay S, Lavergne C, *et al*. What patients can tell us: topic analysis for social media on breast cancer. *JMIR Med Inform* 2017; 5 (3): e23.

104. Oscar N, Fox PA, Croucher R, *et al*. Machine learning, sentiment analysis, and tweets: an examination of alzheimer's disease stigma on Twitter. *J Gerontol Ser B* 2017; 72 (5): 742–51.
105. Roccetti 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 (3): e51.
106. Salas-Zárate MDP, Medina-Moreira J, Lagos-Ortiz K, *et al*. Sentiment analysis on tweets about diabetes: an aspect-level approach. *Comput Math Methods Med* 2017; 2017: 5140631. doi:10.1155/2017/5140631.
107. Simms T, Ramstedt C, Rich M, *et al*. Detecting cognitive distortions through machine learning text analytics. In: Proceedings of the 2017 IEEE International Conference on Healthcare Informatics, ICHI 2017, 2017: 508–12. doi:10.1109/ICHI.2017.39.
108. Smith RJ, Crutchley P, Schwartz HA, *et al*. Variations in Facebook posting patterns across validated patient health conditions: a prospective cohort study. *J Med Internet Res* 2017; 19(1): e7. doi:10.2196/jmir.6486.
109. Stanovsky G, Gruhl D, Mendes PN. Recognizing mentions of adverse drug reaction in social media using knowledge-infused recurrent models. In: Proceedings of Conference of the European Chapter of the Association for Computational Linguistics, 2017: 1142–51. doi:10.18653/v1/E17-1014.
110. Stewart SA, Abidi SSR. Leveraging medical taxonomies to improve knowledge management within online communities of practice: the knowledge maps system. *Comput Methods Programs Biomed* 2017; 143: 121–7.
111. Strapparava C, Mihalcea R. A computational analysis of the language of drug addiction. In: Proceedings of the 5th Conference of the European Chapter of the Association for Computational Linguistics; 2017: 136–42.
112. Sulieman L, Gilmore D, French C, *et al*. Classifying patient portal messages using Convolutional Neural Networks. *J Biomed Inf* 2017; 74: 59–70.
113. Vedula N, Parthasarathy S. Emotional and linguistic cues of depression from social media. In: Proceedings of the 2017 International Conference on Digital Health—DH '17, 2017: 127–36. doi:10.1145/3079452.3079465.
114. Wang Y, Tang J, Li J, *et al*. Understanding and discovering deliberate self-harm content in social media. In: Proceedings of the 26th International Conference on World Wide Web - WWW '17, 2017: 93–102. doi:10.1145/3038912.3052555.
115. Wang T, Brede M, Ianni A, *et al*. Detecting and characterizing eating-disorder communities on social media. In: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining—WSDM '17, 2017: 91–100. doi:10.1145/3018661.3018706.
116. Workewych AM, Ciuffetelli Muzzi M, Jing R, *et al*. Twitter and traumatic brain injury: a content and sentiment analysis of tweets pertaining to sport-related brain injury. *SAGE Open Med* 2017; 5: 1–11. doi:10.1177/2050312117720057.
117. Yazdavar AH, Al-Olimat HS, Ebrahimi M, *et al*. Semi-supervised approach to monitoring clinical depressive symptoms in social media. In: Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017—ASONAM '17, 2017: 1191–8. doi:10.1145/3110025.3123028.
118. Zhang S, Kang T, Qiu L, *et al*. Cataloguing treatments discussed and used in online autism communities. In: Proceedings of the 26th International Conference on World Wide Web—WWW '17, 2017: 123–31. doi:10.1145/3038912.3052661.
119. Zhang S, Grave E, Sklar E, *et al*. Longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks. *J Biomed Inform* 2017; 69: 1–9.
120. Zhu L, Yang H, Yan Z. Extracting temporal information from online health communities. In: Proceedings of the 2nd International Conference on Crowd Science and Engineering—ICCSE'17, 2017: 50–55. doi:10.1145/3126973.3126975.
121. Abdellaoui R, Foulquie P, Texier N, *et al*. Detection of cases of noncompliance to drug treatment in patient forum posts: topic model approach. *J Med Internet Res* 2018; 20: e85. doi:10.2196/jmir.9222.
122. Bryan CJ, Butner JE, Sinclair S, *et al*. Predictors of emerging suicide death among military personnel on social media networks. *Suicide Life Threat Behav* 2018; 48 (4): 413–30.
123. Karisani P, Agichtein E. Did you really just have a heart attack?: towards robust detection of personal health mentions in social media. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18, 2018: 137–46.
124. Yadav S, Ekbal A, Saha S, *et al*. Medical sentiment analysis using social media: towards building a patient assisted system. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018), 2018: 2790–7.
125. Nguyen T, Dao B, Phung D, *et al*. Online social capital: mood, topical and psycholinguistic analysis. In: Proceedings of the Seventh International AAAI Conference on Web and Social Media (ICWSM 2013), 2013: 449–56.
126. Roccetti M, Casari A, Marfia G. Inside chronic autoimmune disease communities: A social networks perspective to Crohn's patient behavior and medical information. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015, 2015: 1089–96. doi:10.1145/2808797.2808813.
127. Mitra T, Counts S, Pennebaker JWJ. Understanding anti-vaccination attitudes in social media. In: Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016), 2016: 269–78.
128. Zhang S, Bantum E, Owen J, *et al*. Does sustained participation in an online health community affect sentiment? *AMIA Annu Symp Proc* 2014; 2014: 1970–9.
129. 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: 195–8.
130. McIver DJ, Hawkins JB, Chunara R, *et al*. Characterizing sleep issues using Twitter. *J Med Internet Res* 2015; 17 (6): e140.
131. McKee R. Ethical issues in using social media for health and health care research. *Health Policy* 2013; 110 (2–3): 298–301.