Electronic Commerce Research and Applications 27 (2018) 139–151

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Harnessing social media for health information management

Lina Zhou [a], Dongsong Zhang [c,a,*], Christopher C. Yang [b], Yu Wang [c]

[a] University of Maryland, Baltimore County, United States
[b] Drexel University, United States
[c] International Business School, Jinan University, China

ARTICLE INFO

Article history:
Received 13 December 2017
Received in revised form 14 December 2017
Accepted 14 December 2017
Available online 29 December 2017

Keywords:
Conceptual framework
Health information management
Data analytics
Social media

ABSTRACT

The remarkable upsurge of social media has dramatic impacts on health care research and practice. Social media are reshaping health information management in a variety of ways, ranging from providing cost-effective ways to improve clinician-patient communication and exchange health-related information and experience, to enabling the discovery of new medical knowledge and information. Despite some demonstrated initial success, social media use and analytics for improving health as a research field is still at its infancy. Information systems researchers can potentially play a key role in advancing the field. This study proposes a conceptual *framework for social media-based health information management* by drawing on multi-disciplinary research. With the guidance of the framework, this paper presents related research challenges, identifies important yet under-explored research issues, and discusses promising directions for future research.

1. Introduction

The remarkable upsurge of social media and online health communities has dramatic impacts on health care research and practice. We refer to *social media* as Internet-based tools or platforms that allow individuals and communities to gather and communicate with others and to generate, share, and distribute information, ideas, and experiences. According to a 2013 report of the Pew Internet & American Life Project, more than 70% of Internet users seek health information online. More and more people are going to social media websites to provide or seek knowledge about health, share personal experience with diseases, medical treatments, and medications, and communicate with healthcare professionals or other patients, etc. (Chretien and Kind, 2013; DeAndrea and Vendemia, 2016; Dizon et al., 2012; Fung et al., 2015; Kass-Hout and Alhinnawi, 2013), making social media a core element of "social health" (Andreu-Perez et al., 2015). In addition, pharmacovigilance, crowdfunding of healthcare services, and specialized, self-service healthcare tools also contribute to the increasing use of social media in healthcare.

Social media have potential to empower people to develop healthy life styles, make better and more informed medical decisions, and improve personal health management. We refer to *health information management* as the activities that people perform in order to acquire, organize, maintain, share, retrieve, and use health information items to complete healthcare tasks and fulfill their needs. Social media provide an unprecedented opportunity to advance health science and quality care by mobilizing broad social media users and enabling them to generate a vast amount of content. The type of support that social media can provide is not limited to informational, but nurturant and instrumental (Langford et al., 1997; Reblin and Uchino, 2008; Reid et al., 1989) as well. Nurturant support includes emotional and companionship support. Setoyama et al. (2011) found that both posters and lurkers in online Japanese breast cancer communities reported that they received moral support, emotional expression, advice, and insights/universality from peers in their online communities. Employing social media for health helps improve the efficiency of patient health management, as manifested by the reduction of financial expenditure and enhancement of health care in general. A patient has potential to differentiate various health care professionals that are available based on patient ratings and online reviews. Patients can join virtual communities, participate in research, receive financial support, set exercise goals, and track personal progress using social media (Ventola, 2014).

Although social media have been leveraged in a variety of ways aiming to improve health care, current research is still in the early stage. There are many technical, behavioral, and data management issues and challenges in social media use or analytics for health care that remain unsolved or not well solved; there is still

* Corresponding author.
E-mail addresses: zhoul@umbc.edu (L. Zhou), zhangd@umbc.edu (D. Zhang), ccy24@drexel.edu (C.C. Yang), twygs@jnu.edu.cn (Y. Wang).

https://doi.org/10.1016/j.elerap.2017.12.003
1567-4223/

insufficient understanding of economic and societal impact of social media usage for health care, and a lack of empirical evidence to demonstrate the effectiveness of social media in improving patients' health outcomes; and inconclusive, inconsistent, or even contradictory findings from different studies are not uncommon. As a result, it is essential and necessary to have a good understanding of the potential benefits of social media for health care and the challenges and research issues that need to be addressed in order to achieve and maximize those benefits.

The primary objective of this research is to create a structured and comprehensive view of the state-of-the-art research in this emerging field. Specifically, we propose a conceptual framework of *social media-based health information management* (SMHIM), and provide a road map for achieving its goals and realizing its impacts.

2. A Conceptual Framework for SMHIM

Aiming to facilitate a better understanding of SMHIM, we propose a conceptual framework, as shown in Fig. 1. The development of the framework draws upon epidemiology, sociology, economics, and public health research, and natural language processing, text mining, machine learning, social network analysis, and statistical modeling. In the framework, SMHIM is conceptualized as a set of *processes* in which *participants* who have concerns about specific health *issues* are engaged through social media *platforms* (4P) to pursue four objectives (4Cs). The 4Cs include improving *care quality and safety*, *communication efficiency*, *cost-effectiveness*, and *convenience of access*.

- **Care quality and safety:** social media are reshaping health *care* services with the potential to improve quality and safety of patient care. For instance, the broad reach of social media can overcome the limited generalizability of traditional clinical studies that typically draw conclusions based on a very specific population group. In addition, real-time social media data can be used to build public health surveillance systems to detect, track, and respond to infectious diseases.
- **Communication efficiency:** poor patient-physician *communication* has been a long-standing issue that causes many problems, such as low medication adherence of patients. Similarly, the lack of physician-physician *communication* also hinders the sharing of useful medical knowledge and best medical practices. Social media provide a ubiquitous online communication channel that can ease patient-physician, physician-physician, and patient-patient communications without time and location constraints.
- **Cost-effectiveness:** social media make the sharing and exchange of health information and knowledge (e.g., the appearance of certain symptoms from a disease) much easier and cheaper than ever before. Mining social media data enables researchers to identify disease cases, discover new knowledge (e.g., unknown *adverse drug reactions* (ADR)) or unmet needs of patients, and gain an instantaneous snapshot of the public's opinions and concerns about health-related issues.
- **Convenience of access:** social media can serve as a convenient and accessible venue for health intervention, patient education, patient health self-management, drug and health care service advertisement, etc. For example, social media have been used to educate patients and enable physically distributed patients to make connections with other patients who have the same disease to share their experience and provide emotional support.

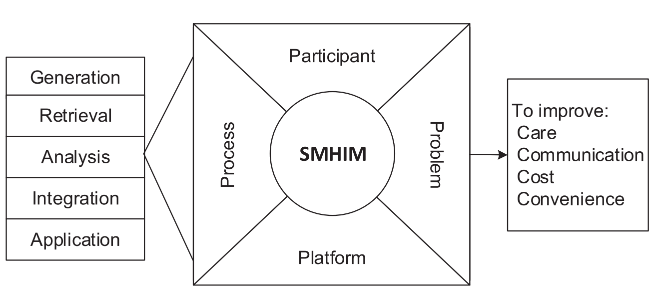

Fig. 1. A Conceptual Framework for SMHIM.

2.1. Participants

Social media offer a novel perspective to healthcare because they provide unique communication channels to patients, healthcare professionals, and general public. Patients often intend to achieve different goals from social media, including acquiring or sharing disease-specific knowledge or information, getting nurturant support from peer patients, sharing personal experience with healthcare providers, services, treatments, and drugs, engaging in social interactions, receiving online patient education and empowerment (e.g. network building, acceptance and belonging in a society, understanding and validation), etc. (Dosemagen and Aase, 2016). Social media also provide a low-cost means for patients to engage with health care professionals.

Health care organizations, such as hospitals, medical schools, professional health organizations, and other groups in the health care industry, have used social media to engage with patients or to reach out to potential consumers. Based on a survey study of 187 doctors in Australia (Brown et al., 2014), although 65.8% of the participants were hesitant to utilize social media to provide personal care for patients due to legal concerns, such as doctor-patient confidentiality, exposure to lawsuits, and the inability to regulate information about themselves, almost 2/3s of them routinely referred their patients to health care information available on social media; 69.7% reported discussing online resources with their patients. Similar results were also reported by U.S. doctors (Brown et al., 2014). Healthcare organizations have used social media to engage with patients, convey their missions and visions, conduct marketing campaigns, provide patient education, recruit workforce, and develop professional networking. It is estimated that 70% of health care organizations in U.S. use social media to develop better connections with patients (Househ, 2013). For example, Mayo Clinic uses social media to promote health, fight diseases, and improve healthcare via its social media network (socialmedia.mayoclinic.org). Some organizations are partnering with third-party social media sites such as PatientsLikeMe.com to communicate and collaborate with external stakeholders (Keckley and Hoffmann, 2010).

The increasing presence of pharmaceutical companies, biotechnology firms, and medical equipment manufacturers in social media is largely driven by the marketing needs – to increase the public awareness of their products and services. However, DeAndrea and Vendemia (2016) found that disclosing an affiliation with a drug company actually decreased consumers' trust in an organization that posted information about a drug, decreased trust in comments posted by other users about the drug, and made it more unlikely for people to recommend this drug to their family and friends. Some other researchers also argue that directly promoting health products and services in social media is linked with inappropriate medication use, overutilization, and increased spending on expensive drugs, and may endanger public health due to promotion of potentially dangerous products (Liang and Mackey, 2011). There is an essential need for more research and a better understanding of how the online presence of pharmaceutical companies and medical equipment manufacturers in online social media affects consumers' perceptions of their advertised health care products and services.

2.2. Problems

Social media have been used by patients to seek solutions to or feedback on a wide variety of health problems. Here are just some representative examples.

2.2.1. Chronic diseases

Management of chronic diseases (e.g., cancer, chronic pain, arthritis, diabetes, and depression) continues to be one of the main concerns in public health. According to the Centers for Disease Control and Prevention, nearly half of adults in the U.S. are living with at least one chronic health condition (www.cdc.gov/chronicdisease/index.htm). Patient behavior change plays an active role in health management. Social media create new opportunities to help patients with a chronic disease manage their health condition. The use of social media in managing chronic diseases can be self-guided, fostered through peer-to-peer interaction, or guided by an external source (e.g., a facilitator or health professional) (Seeman, 2008). The limited preliminary studies on the impact of social media on patients with chronic diseases showed some positive results (e.g., promoting self-care and fostering self-confidence). Benefits of social media use to psychosocial management have also been reported.

Serious mental illness (e.g., schizophrenia, bipolar disorder) is considered a chronic disease associated with debilitating symptoms of anxiety and depression. People suffering from mental illness are often reluctant to talk about their illness with others in the real world or go to health care providers when they need help due to societal stigma and prejudice about mental diseases and the rising health care cost. The anonymity nature of social media, however, encourages those people to share their thoughts and seek help or support online. Research has shown that people with mental health problems, especially those with depression, have high social media use (Davila et al., 2012; Rosen et al., 2013). They are increasingly turning to social media to share their illness experiences or to seek advice from others with similar health conditions (Naslund et al., 2016). Thus, social media provide a potential turning point for these people by helping them overcome barriers such as social isolation and reluctance to use formal health care services and get timely help.

The potential benefits of social media to mental health include detecting mental status of individuals, developing social relationships online without the stress that face-to-face interactions provoke, fostering connectedness, and boasting mutual support groups for people suffering from mental illness (Cottle, 2016). Online peer-to-peer connections may afford opportunities for individuals with serious mental illness to challenge stigma and increase engagement with online interventions (Naslund et al., 2016). In addition, analyzing people's behavior on social media websites may provide a genuine and proactive way to discover depression or even suicidal ideation of an individual, enabling timely intervention. For example, Yates et al. (2017) applied deep learning to identifying users with depression and estimating self-harm risks in online communities. Their classification model consisted of a shared architecture based on a *convolutional neural network (CNN)*, a merge layer, model-specific loss functions, and an output layer. It focused on learning representations of user posts and integrating representations into an overall representation of user activities. The evaluation of their proposed method showed that the CNN model achieved better classification performance, measured by recall and F1-score, than baselines including *support vector machine* (SVM) and multinomial naïve Bayes models trained with bag-of-words and other features.

Although there has been increasing research on social media analytics for mental health detection as illustrated above, the relationship between mental health and social media use may be intricate. Several studies have indicated that the prolonged use of social networking sites may cause depressive symptoms, anxiety, low self-esteem (Pantic, 2014), and addiction (Hormes et al., 2014). In addition, cyber-bullying on social media sites increases risk of depression and anxiety for children. Facebook use is also reported to be associated with the incidence of body image disturbance among adolescent girls (Cottle, 2016). To avoid social media-induced depression, users are advised to be aware of the risks of using those sites as a tool of comparison, and of the fact that most people are presenting a biased version of reality on social media.

2.2.2. Public health

One of the key benefits of using social media for health is public health surveillance (Moorhead et al., 2013), which has shown great potentials to influence health policy making. Traditionally, public health surveillance relied on surveys and aggregating primary data collected from healthcare providers and pharmacists. The lack of timely data and limited understanding of the emergence of global epidemics make monitoring and forecasting the spread of infectious diseases very difficult (Kotov, 2015). Public health surveillance can use social media as a data source to identify and target populations with illnesses of high prevalence and public health impact (e.g., influenza-like illness, infectious diseases), keep track of epidemic disease activities, detect health-related hot topics, identify adverse events related to medications, vaccines, and other drug uses (i.e., pharmacovigilance), and deliver social media based intervention. For example, CDC maintains an active presence on Twitter and Facebook to track tweets that might indicate a flu outbreak and to share updates about such incidents.

Social media may serve as valuable tools for improving public health professionals' ability to detect disease outbreaks in a more timely manner than traditional methods and shorten outbreak response time (Charles-Smith et al., 2015). Prediction of influenza outbreaks is critical for developing effective strategies for prevention, intervention, and countermeasures. While health reports are available on a weekly or monthly basis, tweets and online search query logs can be obtained almost instantly, making the latter a more up-to-date source for disease trend analysis or prediction of the spread or severity of influenza (Broniatowski et al., 2013; Santos and Matos, 2014; Xu et al., 2017). Google's Flu Trends consistently estimated the level of weekly influenza activities in each region of the U. S. with a 1-day reporting lag, which was about 1–2 weeks ahead of the reports from the CDC's U.S. Influenza Sentinel Provider Surveillance Network; the physical location of an online search query's origin, which can be identified from its associated IP address, can help health agencies and government identify where an outbreak is occurring and how fast it is spreading. According to a literature review of social media-based surveillance studies (Bernardo et al., 2013), 66% of reviewed studies reported that social media-based surveillance had performance comparable to traditional surveillance programs. Incorporating social media into surveillance systems along with standardized metrics to evaluate their performance is crucial to the advancement of these early warning systems (Velasco et al., 2014).

Social media have also been used to target adolescents who demonstrate health risk behaviors associated with tobacco use, substance abuse, and sexual activities (Charles-Smith et al., 2015). Social media applications (e.g., discussion boards and online chat rooms) are promising to online weight management interventions through facilitating education, user engagement, and peer support (Chang et al., 2013).

2.2.3. Rare diseases

Social media can be a lifeline for patients with rare diseases (Stone, 2015). Patients with rare diseases often have to travel long distances to receive specialized care, which is exhaustive and expensive. Not having a diagnosis takes a huge emotional toll. Those patients all have some levels of fear in meeting a new doctor and discussing their symptoms, and constantly worry the re-occurrence of their disease. Identifying healthcare solutions for close relatives who have rare diseases was actually the inspiration of founding popular health social media sites such as PatientsLikeMe.com and MedHelp.com. Through those social media sites, patients or caregivers are able to acquire useful information about rare diseases and develop an expanding network of friends. Sometimes, physicians also need help from a community to come up with an appropriate treatment plan. By taking advantage of social media platforms, genetics companies are able to recruit more people easier and faster for critical research on rare diseases in an effort to find an eventual solution.

2.3. Platforms

Social media platforms can be referred broadly to social media websites, infrastructure, and communication technology, etc. Numerous social media platforms are accessible to the general public, patients, and healthcare professionals. Participation or interaction is the key to the success of any social media platforms. The affordance of social media includes identity, flexibility, structure, narration, and adaptation (Merolli et al., 2013b). An analysis of social media content can reveal important dimensions of social media platforms, such as audience reach (e.g., followers and subscribers), page commenting policy, post source (e.g., consumer, pharmaceutical company), post format, and post interactivity (e.g., the number of "Likes") (Tyrawski and DeAndrea, 2015).

Prior studies on social media applications for health care categorized social media into 10 different types, including blogs (e.g., WordPress), microblogs (e.g., Twitter), social networking sites (e.g., Facebook), professional networking sites (e.g., LinkedIn), thematic networking sites (e.g., 23andMe), wikis (e.g., Wikipedia), mashups (e.g., HealthMap), collaborative filtering sites (e.g., Digg), media sharing sites (e.g., YouTube) (Grajales et al., 2014), and others (e.g., virtual reality and gaming environments) (Ventola, 2014). There are also other social networking communities (e.g., Sermo and Doximity) that support health care professionals to communicate with colleagues and allow point-of-care information crowd sourcing (Chretien and Kind, 2013). It remains under studied about whether specific types of social media may be more effective for addressing specific types of health problems. For instance, it might be the case that online medical forums and consumer reviews are the best social media sources of information for the detection of ADR, while microblog services such as Twitter are more useful to find population-level health signals (Paul and Dredze, 2011). On the other hand, Facebook, specialized chat rooms, websites, and Twitter have been explored for identifying health risk behavior associated with substance abuse and sexual activities (Charles-Smith et al., 2015). It is worthy of future research to investigate which types of social media content or social media platform are more effective and reliable sources for addressing a particular health care need or problem than other counterparts. The answers to this research problem can provide theoretical guidance to researchers when studying health related issues in a social media context.

2.4. Processes

A knowledge system generally consists of four sets of socially enacted knowledge processes: creation, storage/retrieval, transfer, and application (Alavi and Leidner, 2001; Holzner and Marx, 1979; Pentland, 1995). Similarly, the process of SMHIM is comprised of (1) generation, (2) retrieval, (3) extraction, (4) integration, and (5) application. It should be noted that analytical techniques can play a role across all of these processes. Advances in analytics of health-focused, natural language data allow for analyzing social media in order to understand why patients make certain decisions and gain insights into patients' health needs.

2.4.1. Generation

The core value of social media lies in user-generated content that reflects users' interests, emotions, and thoughts (McCord et al., 2014). Users adopting social media intend to share information about their health condition and experience to benefit themselves and others. Social media have quickly become the preferred method for communication and information sharing (Pillow et al., 2014). They provide an additional informal source of data that can be used to identify health information not reported to physicians or other health agencies and to reveal people's thoughts on health-related topics, especially those with a sensitive nature (Charles-Smith et al., 2015). Potential research questions may include, but not limited to, what are the factors driving people to go to social media for peer support? How do people communicate or interact with others, mostly strangers, on social media? What are the perceived quality, benefits, and risks of interaction on social media? How to protect patient privacy while getting needed, valuable insights from those interactions? and how to make social media platforms easier to use and more accessible?

2.4.2. Retrieval

One of the major objectives of using social media is to search, share, and distribute health related information. Today, people are shifting from passive receivers of health care and medical treatments to active seekers and learners of health knowledge that can enable them to make more informed health-related decisions. Experience-based information empowered by social media's ability to aggregate individuals' experiences or opinions allows people to find or follow peers who experience similar medical conditions or concerns (Song et al., 2016). Many people have developed habitual consumption of social media content related to their condition and integrated it into their daily social media use.

Social media content varies in health focus. For instance, medical forums are generally more health related than general social media platforms like Twitter. When analyzing social media data, it is common that researchers first build a filter to retrieve health related posts before proceeding with any further analysis (Paul and Dredze, 2011). If the focus is to detect ADR of a specific drug, researchers would first filter out all the tweets that do not mention the target drug (Bian et al., 2012).

Providing feedback is another aspect when users respond to the user-generated content. Feedback comes in the forms of users' opinions about the quality or relevance of a user's content, answers to a user's question, or a follow up, etc. Common examples are "like/not like," "thumbs up/down," dig it, star ratings, social commentary, tagging, flagging and badging, and answers. A user can leverage such feedback to adjust his/her current and future behavior to achieve the desired result. Users can potentially co-create collective and experiential knowledge through interactions.

2.4.3. Integration

Integration concerns linking, sorting, and archiving information collected from various social media platforms. Integration involves defining knowledge types or developing a knowledge taxonomy. Existing medical ontologies and knowledge resources can facilitate this process. For example, Lu et al. (2013) used MetaMap to map biomedical text to concepts in the UMLS Meta thesaurus to obtain medical terminologies automatically. Similarly, if one could map the data retrieved from social media platforms to an ontology, then the alignment of each platform to the ontology would pave the way for integrating data from different platforms. Nevertheless, these ontologies are insufficient to address at least two challenges in using social media data. One is heterogeneity in health-related terminologies used by health care professionals and those used by the lay public - patients usually use lay or slang terms (e.g., can't sleep), rather than scientific medical terms (e.g., "insomnia"), when describing their symptoms and treatment experience in text on social media. As a result, social media analytics programs may have difficulty in interpreting and understanding expressions used in social media by the lay public, which can potentially lead to undesirable misunderstanding and consequences. Filling such a terminology gap is one common challenge facing textual social media data analysis. One possible approach to solving the above language mismatch problem is to use term co-occurrence to identify the relationships between lay terms and medical terms. However, the dimension of such a co-occurrence matrix will increase as the size of the vocabulary increases, which does not scale well. Therefore, instead of capturing global statistics from data, word embeddings (e.g., Mikolov et al., 2013) can be more suitable to suggest similar words and help extend original medical lexicons. Another viable, and already attempted, solution is to develop a dictionary of lay or slang terms used in social media, such as Consumer Health Vocabulary (CHV), that links UMLS standard medical terms to patients' colloquial language (Kuhn et al., 2010). The challenge to the use of CHV lies in its costly maintenance – lay terms keep changing and evolving, and consequently CHV needs to be constantly updated. The set of colloquial phrases can also be created through manually crafted patterns (e.g., Yates et al., 2013), manual extraction (e.g., Leaman et al., 2010), or both (e.g., Benton et al., 2011).

Another need for integration concerns mobile health, which uses mobile and wearable devices and mobile apps along with social media as a channel for users to monitor personal health, and for health practitioners to track patients remotely (Santoro et al., 2015). There have been suggestions to integrate social media and mobile health in health promotion and healthcare programs (Burke-Garcia and Scally, 2014). The prevalence and pervasiveness of smartphones make them a powerful platform for patient monitoring and individualized health care delivered at the patient's convenience. What is more appealing to combine social media data and mobile health applications is that researchers and physicians can possibly acquire offline behavior and other personal characteristics of individuals through their interactive activities on mobile devices and align them with their online behavior in social media to gain a more comprehensive picture about the health of individuals. We believe that this is a promising research area that offers great potentials in continuous patient monitoring and intervention, especially for patients with mental disorders.

2.4.4. Analysis

Due to the overwhelming and ever-increasing volume of social media content, it is practically impossible and cost-prohibitive to manually wade through all the content in order to find hidden yet valuable information or knowledge. The automated analysis of social media content may provide a cost-effective way to discover patterns, correlation, trends, preferences, and other useful information that can help patients and healthcare professionals make better decisions. The discovery of such knowledge requires natural language processing, information extraction (e.g., adverse drug effect detection), machine learning (e.g., clustering, classification, and prediction), and other techniques (e.g., sentiment analysis and link analysis).

Text is the most popular format of social media data. Natural language processing (NLP) techniques are often used to preprocess textual social media data (e.g., Bian et al., 2012; Leaman et al., 2010). These preprocessing techniques include sentence splitting, tokenization, stop word removal, part-of-speech tagging, and syntactic parsing. However, analyzing user-generated textual social media content poses significant challenges for analytical tasks using general-purposed NLP tools (e.g., NLTK toolkit; MALLET). Those tools, traditionally developed based on formal written text, do not adapt well to informal language commonly used in social media (Nguyen et al., 2017). Spelling errors, abbreviations, and other informal language use such as phrase construction irregularities and layman terms make it even more challenging to process health social media data.

Information extraction has been widely used to extract medical entities such as drug names, symptoms, and disease names from textual social media content (e.g., blogs). The approaches to medical entity extraction can be classified into lexicon- and knowledge-based methods, machine learning methods, and hybrid methods. Most studies (e.g., Nguyen et al., 2017) used lexicons and even ontologies constructed using existing resources, such as UMLS (Unified Medical Language System), MedDRA, SIDER, MeSH, RxNorm, NCI Thesaurus, and CHV (e.g., Benton et al., 2011; Yates et al., 2013).

Building on the results of medical entity extraction, relation extraction (e.g., a drug was taken to treat a symptom/disease or a drug causes a certain symptom as ADRs) requires the detection and classification of semantic relationship mentions among medical entities. Liu and Chen (2015) categoried biomedical relation extraction techniques into four types, including co-occurrence analysis, rule-based approaches, statistical learning approach, and hybrid approach (i.e., a combination of rules and statistical learning). For extraction of ADE relations, most prior studies adopted co-occurrence analysis approaches (e.g., Benton et al., 2011). Co-occurrence analysis identifies relations among biomedical entities based on their probability of co-occurrence in the same text message. The method assumes that if two entities are frequently and closely mentioned in the same text, there would be an underlying biological relationship between them (Mao et al., 2013). Normally, only lexical information is considered in co-occurrence analysis due to their simplicity and flexibility (Liu and Chen, 2015). Some researchers (e.g., Nikfarjam and Gonzalez, 2011) have applied association rule mining algorithms to identify ADRs based on frequent co-occurrence of drug and adverse effect terms.

In rule-based approaches, researchers identify relations based on rules that incorporate syntactic or semantic information. Semantic indicators of biomedical relations consist of certain trigger words developed by experts. A pair of entities that satisfies a certain predefined template will be considered to have a relation.

Machine learning approaches have been widely used for discovering knowledge from health social media data (Zhou et al., 2017). Those approaches generally build a classification model by learning from a corpus of annotated documents tagged by human experts. Statistical learning can be categorized into feature-based methods (e.g., bags of words) and kernel-based methods (e.g., tree kernel, shortest dependency path) (Liu and Chen, 2015). In addition, ensemble learning techniques and conditional random fields have also been employed for medical information extraction. For example, Tuarob et al. (Tuarob et al., 2014) proposed an ensemble learning method that combined multiple classifiers (e.g., Naïve Bayes, SVM, and Random Forest) to discover health-related information from social media. Each classifier was trained to learn a semantically different aspect of the data. They used five heterogeneous feature types, including N-gram features, dictionary based compound features, topic distribution features, sentiment features, and combined features, as model input. In view of its great promise in other applications, word embedding has been used to classify medical entities in recent years (e.g., Nguyen et al., 2017; Nikfarjam et al., 2015; Xie et al., 2017). For example, ADRMine (Nikfarjam et al., 2015) used conditional random fields to mine ADR from online user posts. It generated word embeddings from unlabelled social media data using a deep learning technique, then used them to cluster words for modeling word semantic similarities. Nguyen et al. (2017) generated word vectors that captured relationships among terms to derive an extended lexicon of ADR terms by identifying the closest word vectors to the ADR in the original lexicon. A recent study employed deep learning (i.e., bidirectional long short-term memory (Bi-LSTM) recurrent neural network) to extract e-cigarette adverse events from social media (Xie et al., 2017). The Bi-LSTM model achieved the best performance compared to three baseline models, with a precision of 94.10%, a recall of 91.80%, and an *F*-measure of 92.94%; Xu et al. (2017) deployed four models, including deep learning with feedforward neural networks, a generalized linear model, LASSO (least absolute shrinkage and selection operator), and autoregressive integrated moving average, to forecast new influenza-like illness cases in general outpatient clinics 1 or 2 weeks in advance. Their evaluation showed that deep learning delivered the most competitive predictive performance among the four individual models.

Sentiment analysis is a branch of natural language processing that involves automatically ascribing sentiments (i.e., emotions) of an individual to portions of his written text that express opinions. It has been commonly used in social media analytics and opinion mining to understand public attitudes toward public health issues and policies (Paul et al., 2016). For example, sentiment analysis has been used to understand public attitudes toward a new school meal policy (Kang et al., 2017) and drug abuse (Thompson et al., 2015). Yang et al. (2016) proposed a three-phase framework for analyzing user-generated content in a health community to find public's sentiments toward certain public health issues. Sentiment analysis was the core of the framework. The first phase was to extract medical terms, including conditions, symptoms, treatments, effectiveness, and side effects, to form a virtual document for each question asked in the community. Next, a modified Latent Dirichlet Allocation topic modeling approach was used to cluster virtual documents with similar medical term distributions into a conditional topic. Finally, the clustered topics were analyzed to identify their sentiment polarities. The findings showed that most conditional topics with strong sentiments provided valuable insights.

2.4.5. Application

In general, we can classify existing research on social media for health into two broad categories: one is the use of social media as a platform for health information communication and knowledge sharing, and the other is social media analytics for discovering medical knowledge or building predictive models, as shown in Table 1.

Research focusing on the use of social media includes clinician-patient and patient-patient communication; health information search, sharing, and dissemination; chronic disease management; marketing; professional training; patient education, and so on. For example, social media have a great potential to help increase the awareness of the public about common health issues. Personal experiences related to a certain health condition shared by users online can educate or remind other people about symptoms that they should not ignore before it is too late. Social media can also be a great tool in support of long-term health goals such as weight loss, and management of chronic conditions such as diabetes. For example, they allow users to track each other's accomplishments on social media and compete with one another.

Applications of social media data analytics include public health surveillance, discovery of health-related information or knowledge, discovery of adverse drug events, disease trend prediction, disease intervention, and so on. For example, ADR can result in significant consequences, even death of patients (Wester et al., 2008). Unfortunately, not all ADR of a drug can be discovered in lab tests and small clinical trials with limited test durations and types of patients prior to entering the market. Current practice of post-market drug surveillance uses two different approaches: passive and active (Sarker et al., 2015). Active methods seek to automatically generate safety reports based on different data sources. They complement passive approaches that rely on individual reports of potential ADR from sources such as health professionals, pharmaceutical companies, and medical literature. Active surveillance has received increasing attention in recent years to overcome some inherent limitations of passive methods such as underreporting (Hazell and Shakir, 2006) and lack of timeliness in detection (Kotov, 2015). In support of active discovery, social media are emerging as a promising data source for ADR detection in addition to patient health records and medical and pharmacy claim databases (Sarker et al., 2015). Combining them makes it possible to provide a more complete set of patient reported ADR and to discover unknown ADR earlier (Wu et al., 2013). A large portion of ADRs are attributed to *drug-drug interactions* (DDI), which are often caused by their shared action mechanisms and metabolic pathways. Unknown DDI accounts for up to 30% of unexpected ADRs. Research has explored the detection of DDI from social media (Correia et al., 2016). For example, Yang and Yang (2013) applied association rule mining (co-occurrence of a pair of drugs and consequential ADR in the posts of an online health community as an association) to detect DDIs from social media.

Table 1
A summary of social media use for healthcare application.

General Categories	Research Focus	Sample Studies
The use of social media for health information communication and knowledge sharing	Patient-patient and patient-clinician communication	Naslund et al. (2016), Setoyama et al. (2011), van Rensburg et al. (2016)
	Health information search, sharing, and dissemination	Chretien and Kind (2013), Fergie et al. (2016), Shaw and Johnson (2011), Song et al. (2016), Ventola (2014)
	Chronic disease management	Merolli et al. (2013c), Patel et al. (2015), Stellefson et al. (2013)
	Marketing, professional training, and patient education	DeAndrea and Vendemia (2016), Househ (2013), Stellefson et al. (2014)
Social media analytics for knowledge discovery and predictive modeling	Public health surveillance	Davila et al. (2012), Jashinsky et al. (2014), Rosen et al. (2013)
		Fung et al. (2015), Kang et al. (2017), Paul et al. (2016), Thompson et al. (2015)
	Discovery of health related information or knowledge	Lu et al. (2013), Tuarob et al. (2014)
	Discovering adverse drug events	Correia et al. (2016), Nguyen et al. (2017), Paul et al. (2016), Wu et al. (2013), Yang et al. (2015)
	Disease trend prediction	Broniatowski et al. (2013), McGough et al. (2017), Nagar et al. (2014), Santos and Matos (2014), Xu et al. (2017)
	Disease intervention	Robinson et al. (2015), Tanner et al. (2016), Valimaki et al. (2016)

3. Challenges, open research issues, and promising directions

With guidance from the proposed 4P framework, this section discusses challenges, open research issues, and possible directions in SMHIM. Table 2 provides an overview of those research issues.

3.1. eHealth literacy

One major issue with SMHIM is users' eHealth literacy. According to (Norman and Skinner, 2006), eHealth literacy consists of health literacy (Sørensen et al., 2012), computer literacy, information literacy, media literacy, traditional literacy and numeracy, and science literacy, which influence the competence of users in leveraging online health information and tools (e.g., apps). A study of self-diagnosis apps concludes that self-diagnosis through social media tools such as apps is not a good practice, and users are not as competent as trained medical professionals (Lupton and Jutel, 2015). Thus, in addition to improving users' health literacy, how to adapt social media content to the level of user's eHealth literacy is an important yet under studied issue.

3.2. Cultural differences

There exist intercultural differences in health information on social media. Prior research has observed significant cultural differences in the level of trust in three experience-based information sources, including blogs, online support groups, and social networking sites (Song et al., 2016). More specifically, individuals with a holistic culture tend to have higher levels of trust in all these three sources than those with an analytic culture. In contrast, trust in expertise-based information sources, including online professional health sites, is universal, showing no cultural differences. Culture also contributes significantly to differences in searching health information on behalf of family members and in the goals of information searching. These findings point to the needs of understanding the culture of social media health consumers, and more importantly, the impact of culture on their behavior on social media.

3.3. Incorporating media sharing sites

Research on health social media has mainly focused on text-based content so far, which is partly attributed to the availability of natural language processing toolkits and related knowledge resources. In contrast, studies that focus on the use of or analytics of media sharing sites (e.g., Youtube.com or flicker.com) or multimedia social media content in the context of health care are sparse. One of the very few examples is Merolli et al. (2016)'s study, which used pain management videos on Youtube.com, along with a large chronic pain support community on Facebook and various chronic pain blogs, to study social media use for pain management; Yang and Luo (2017) used both images and text data to analyze behavior patterns and detect drug dealer accounts on Instagram. As the techniques for analyzing multimedia content advance (e.g., He et al., 2016), future studies should tap into media sharing sites to discover health-related information.

Table 2
Some challenges and open research issues.

4P		Research Issues
Participant		Health literacy
		Cultural differences
Platform		Media sharing sites
Problem		Chronic diseases
Process	Generation	Data quality
	Retrieval	Data privacy and confidentiality
	Analysis	Data preparedness
		Causal assessment
	Integration	Multi-modal multi-source multi-role data fusion
	Application	Patient education
		Disease intervention
		Health impact assessment
		Personalized health-management

3.4. Chronic diseases

While the future of social media in support of chronic disease management is optimistic, there is limited concrete evidence indicating whether and how social media use significantly improves patient outcomes (Merolli et al., 2013c). Not many studies have empirically evaluated the impact of social media on the patients with chronic diseases. A few preliminary studies reported positive impact on health status, such as improvement of psychosocial health (e.g., improved psychosocial management, sense of hope, connectedness, and relief from social media use). For example, Merolli et al. (2016) investigated perceptions of patients with chronic pain towards using social media for pain management while waiting for clinic access. Participants highlighted social media's affordance of "exploration", "connection", "narration", and "adaptation", without any negative comments on social media use or any adverse events. However, a few other studies reported neutral or negative impact of social media caused by inaccurate information or biases in social media content (Patel et al., 2015). Further, there is little evidence of benefits of social media for physical condition management to date (Merolli et al., 2013b). Those inconsistent or even conflicting results imply that social media use may not address all types of health care needs. Therefore, studying effectiveness and measurable benefits of social media under different scenarios and with different health care needs can provide valuable insights for future research and practice.

There exist very few rigorous theoretical frameworks that provide guidance on how to gather evidence of improved health outcomes from social media use in chronic disease management (Merolli et al., 2013a). A review study (Stellefson et al., 2013) found that the most popular theoretical framework used in prior studies was Social Cognitive Theory, followed by Social Ecological Theory. In addition, the research methodologies used to examine the impact of social media on chronic disease patients vary. Among the 19 reviewed studies in (Merolli et al., 2013c), seven examined the discourse or content of communication on specific social media platforms via thematic content analysis; three were randomized controlled trials; and one was a randomized longitudinal design. Further research of high methodological quality is required to investigate the affordance of social media and how social media can best serve patients with chronic diseases and physicians. Evidence-based practice with support from social media data may be considered.

3.5. Data quality

Arguably the biggest concern with SMHIM is poor quality of data (Moorhead et al., 2013; Ventola, 2014) and uncertainty about the reliability of data collected from social media (Antheunis et al., 2013). User-generated content in social media is often unreferenced, inconsistent, biased, incomplete, or even inaccurate due to lack of medical knowledge, health status, etc. (Nguyen et al., 2017; Wang and Jiang, 2017). In comparison to well-structured and documented electronic health records and health professionals' reports, there are no requirements regarding writing and structuring descriptions of pharmacovigilance-related events on social media (Lardon et al., 2015). For instance, the main reason of false negatives of drug name extraction from social media lies in misspelled drug names and the abbreviations and acronyms used (Segura-Bedmar et al., 2015). This data quality problem may not only concern or even mislead patients who look for health related information from social media, but also make health care professionals quickly lose faith in social media and become more apt to reject important, counter-intuitive implications that may emerge from social media analytics. Moreover, misinformation may be circulated rapidly given the nature of social media (Oyeyemi et al., 2014). Thus, the information exchanged via social media needs to be checked and possibly rectified for quality and trustworthiness.

The validation of large, noisy social media data poses enormous challenges (Kass-Hout and Alhinnawi, 2013). Several methods have been proposed to solve these challenges. Assessing outliers based on unusual effects or information is a common approach. Kim et al. (2016) proposed a framework and a reporting standard that researchers can use to assess the quality of social media data used in a study. The framework consisted of three major steps in collecting social media data, including development, application, and validation of search filters. In addition, quality assessment criteria including precision and recall may not always be practical because of the difficulty in acquiring the gold standard in the first place. Thus, there is a need for more formal methods or guidelines for assessing social media data quality. HONcode (Health on the Net Foundation, 2017) guidelines for posting trustworthy health information on the Internet may be extended to social media content. Other recommendations for monitoring and enhancing the quality and reliability of health-related communication via social media include the need to determine the relative effectiveness of different types of social media for communication using randomized control trials and to explore potential mechanisms (Moorhead et al., 2013).

The quality of social media data is also affected by the unavailability of information concerning patients' characteristics (i.e., age, gender, and medical history), drugs (e.g., dosages and date of treatment initiation), or medication adherence (Lardon et al., 2015). As a result, it is difficult to examine the relationships between patient characteristics and their behavior on social media, and to understand the impact of social media on patient health. This limitation is related to the privacy and confidentiality issues, which will be discussed next.

3.6. Data privacy and confidentiality

Social media use or analytics for health involves a number of privacy, trust, legal, and ethical issues, which is even more so when social media data are integrated with data collected from mobile and wearable devices and sensors, or even mobile cloud. Despite numerous advantages of social media applications for healthcare, the privacy and confidentiality of users have to be preserved (Denecke et al., 2015). Due to privacy concerns, many public-health related social media studies have been conducted using publicly accessible data (Fung et al., 2015). Nevertheless, information about patients' characteristics in social media is still lacking for the same reason. As a result, the generalizability of the findings of prior studies remains a question.

Even though social media data are publicly available, users may still have expectations of privacy (Liu et al., 2011). Potential breaches of patient privacy and confidentiality and resulting legal issues are among the key challenges and risks to patients and health care professionals (Moorhead et al., 2013; Ventola, 2014). One of the key concerns hinges on the extent to which social media data should be treated as public versus private data. The distinction between public and private data becomes further complicated given that machine learning algorithms can infer private attributes from public data, even if they are not explicitly stated in the latter (Horvitz and Mulligan, 2015). Seeking balance between sharing public health information and protecting social media users' privacy remains an ethical challenge for health agencies.

Concerns regarding the use of social media by health care professionals frequently center on the potential for negative repercussions resulting from the breach of patient confidentiality under federal HIPAA and state privacy laws. The widespread use of social media has introduced new legal complexities. Many health care institutions and professional organizations have issued various guidelines to prevent these risks (Ventola, 2014). However, depending on what and how social media are used in support of health (e.g., peer mental support vs. depression detection and intervention), different types of social media data of individual users can be involved and analyzed. Therefore, guidelines for privacy protection may vary for different applications.

3.7. Data preparedness

To enable effective and efficient extraction of information for SMHIM, we often need to classify data. As required by all machine-learning based classification tasks, researchers have to use labelled data for training. Unfortunately, there are very few annotated corpora available because of the high corpora construction cost and privacy concerns. Moreover, preparing labelled datasets in the context of social media is non-trivial. Messages related to a specific health topic or event in social media are highly distributed, while unrelated messages are topically diverse. In an earlier study, researchers found only six studies that made their annotations publicly available (Sarker et al., 2015). The very limited sharing of labelled datasets makes the comparison among approaches and findings of different research studies difficult, if not practically impossible.

One way to alleviate the above problem is to augment training data through a partially supervised learning algorithm (i.e., semi-supervised learning), which constructs classifiers based on mostly unlabeled data and a small number of labelled positive examples (Chapelle et al., 2010). The unlabeled examples are mixed with both positive and negative examples. It requires good labeling heuristics for identifying both positive and negative examples from unlabeled datasets to train a classifier. By leveraging Latent Dirichlet Allocation modeling and partially supervised classification approaches, Yang et al. (2015) trained various classifiers by varying the number of positive examples and the number of topics. Top-ranked posts from each classifier were pooled and the resulting set of posts was reviewed by a domain expert to evaluate the classifiers. Their proposed approach performed significantly better than the alternative approaches using supervised learning methods and three general-purpose partially supervised learning methods. Segura-Bedmar et al. (2015) proposed a system based on distant supervision that does not need annotated data. The learning process was supervised by a database. Distant supervision hypothesizes that if two entities occur in the same sentence, then they may involve in a relation. Other advanced machine learning techniques such as active learning and transfer learning can also be explored to alleviate the problem of scarcity of labelled social media data.

Another concern with SMHIM is imbalanced data. Health problems are generally considered as rare events. Accordingly, social media posts that pertain to specific health problems are much fewer than those posts that do not, which create unbalanced data. As a result, accuracy of classification models trained with such imbalanced data would only reflect underlying data distribution but may not tell the whole story, likely leading to the poor performance of a model for the minority class. Various strategies have been proposed in the traditional machine learning literature, including oversampling and under sampling, adopting additional performance metrics such as F1-score and ROC scores, selecting classification algorithms that are insensitive to data imbalance issues, and adjusting the cost model of a classification technique by biasing it toward the minority class.

3.8. Causal assessment

One major step after identifying associations among different medical entities is to establish causality, which will ultimately help guide clinical practice. For instance, for a pair of drug and sign/symptom terms appearing in the same sentence of a social media post, the sign/symptom could be a side effect of the drug, or an indication of the drug treatment, depending on their semantic relationships. General causality assessment in the medical domain is normally guided by Bradford-Hill criteria (Hill, 1965). The criteria have been adapted for ADR causality, which contain the following aspects (Anderson and Borlak, 2011): strength of association, consistency of association, dose/response relationship, temporal relationship, coherence and specificity, and plausibility. The causality assessment is often done manually because domain expertise plays an important role in this process. The Causal Association Rule Discovery method has recently been introduced to identify causal relationships between drug combinations and adverse events extracted from adverse event reports in a spontaneous reporting system (Cai et al., 2017). Nevertheless, such attempts have yet to be made for social media data. Additionally, it also requires incorporating multiple information sources (e.g., international databases) for establishing a relationship because information about some contributing factors may be missing in the data itself (Karimi et al., 2015).

3.9. Multi-modal, multi-source, and multi-role data fusion

Social media exhibit all characteristics of big data: high volume, velocity, and variety, and low veracity. The objective of SMHIM is to create value for health consumers. Bian et al. (2012) analyzed 2 billion tweets to identify potential users of a drug and possible adverse effects. Chee et al. (2011) exploited 27,290 public Health and Wellness Yahoo! groups with a total of 12,519,807 messages, spanning 7 years, to identify dangerous drugs.

As far as variety is concerned, most social media platforms allow users to post multimedia content. There are also various types of social media that may provide complementary information. Thus, it is promising to fuse multimodal multi-source social media data to gain a more comprehensive understanding of health-related issues. McGough et al. (2017) combined information from Zika-related Google search queries, Twitter microblogs, and a HealthMap digital surveillance system with historical Zika suspected case counts to track and predict weekly Zika cases during the 2015 $\sim$ 2016 Latin American outbreak. They showed that models that combined Google (and Twitter data when available) with autoregressive information showed the best predictive accuracy for 1-week ahead predictions. Nguyen et al. (2017) focused on identifying ADRs of the top ten most frequently prescribed psychiatric drugs by collecting and filtering a corpora of documents collected from Twitter, Reddit, and LiveJournal. However, technical challenges associated with social media data fusion across different modalities, sources, and platforms, such as heterogeneity of data, need to be addressed. The linked data framework (Bizer et al., 2009) that links social media data from different sources through typed links represented in the Resource Description Framework as triples (i.e., subject, predicate, object) is one possible direction, but there is an obvious need for exploring more effective approaches.

Sources of health information for integration are not limited to social media. With the rapid growth of ubiquitous and mobile computing technologies, many mobile and wearable devices have been used to capture continuous human physical, physiological, and even functional behavior. Advances in wireless sensor networks have enabled the monitoring of daily activities and behaviors of individuals. Continuous monitoring helps health professionals better understand patients' behavioral patterns and provide them with more reliable data for intervention. To date, there has been little research that combines social media data and m-health applications, which provides a lot of future research opportunities.

Integrating data from multiple sources or mapping social media expressions from one role (e.g., health professionals) to another (e.g., patients) has significant research and practical implications. Vocabulary difference between health professionals and consumers remains a major challenge in health information searching and sharing in social media.

3.10. Patient education

Social media have been advocated as an inexpensive means for patient education; however, very few studies have examined the impact of social media based patient education, especially on health outcomes. Stellefson et al. (2014) analyzed 223 education videos on YouTube for chronic obstructive pulmonary disease patients. Two independent coders evaluated each video to determine topics covered, media source(s) of posted videos, information quality, and viewer exposure/engagement metrics. The findings suggest that existing education video content on the disease vary significantly in quality. There is a strong need of studying not only how to deliver more effective and helpful patient education through social media, but also how to assess its impact on patient health outcomes.

Health professionals play a key role in patient education. Physicians have begun to develop an interest in interacting with patients online for the purpose of educating and monitoring patients and for encouraging behavioral changes and drug adherence, etc. (Househ, 2013). Yet, some studies have shown that considerable resistance to using social media to interact with patients still exists due to ethical issues or other concerns (Dizon et al., 2012). Some researchers even recommend physicians not to discuss clinical topics with any patients via Twitter's direct messaging mode due to security concerns and non-existence of a real patient-physician relationship (Peters et al., 2015). To mitigate risks when interacting with patients over social media, Grajales et al. (2014) suggest four principles for healthcare professionals to follow: maintain professionalism at all times; be authentic, have fun, and do not be afraid; ask for help; and focus, grab attention, engage, and take actions. Additional research is needed with regard to how to engage physicians and patients in social media based patient education.

3.11. Disease intervention

Recent studies have explored the use of social media for patient intervention. For instance, people with serious mental disorders do spend time online for the purposes of disclosure, information gathering, or gaming (Valimaki et al., 2016). For those patients, online interventions may have the potential to disseminate care and support patients' participation in group interactions (Rotondi et al., 2010). However, there is a lack of coherent knowledge and understanding of the effects of social media use in patient treatment (Robinson et al., 2015).

According to the findings of a systematic review and meta-analysis of the effects of social media interventions for supporting mental health and well-being among people with schizophrenia (Valimaki et al., 2016), only two studies evaluated social media intervention with random trials. In both studies, the participants in the intervention group used peer-support forums, while the participants in the control group underwent usual treatment or were on the waiting list. At 3 months, participants with schizophrenia in the intervention group reported lower perceived stress levels and showed a higher level of perceived social support. Nevertheless, those who reported more positive experiences with the peer support group also reported higher levels of psychological distress. Another major challenge in studying social media based intervention is the difficulty in isolating its effect from the effect of other treatments, because the former is unlikely to be the sole intervention that a patient receives. As a result, future research on developing rigorous research methodologies and metrics for evaluating the impact of social media based intervention is critically needed.

It is worth mentioning that early detection of mental disorders can also be done through applying many text mining and natural language processing techniques to user generated content. There has been increasing research on analyzing social media for detection of suicidal ideation. For example, Jashinsky et al. (2014) explored bloggers' suicidal ideation based on Twitter data. The study detected strong correlations between the results of a suicide prediction model built with Twitter data and the actual suicide rates of individual states reported by the Centers for Disease Control and Prevention in the U.S. Accordingly, preventive treatments can be recommended to the people with suicidal ideation detected in a timely manner before tragedies may happen.

3.12. Health impact assessment

Although social media have shown some promise in improving positive psychotic symptoms, hospital admissions, socialization, social connectedness, depression, and medication adherence of people, the issues of heterogeneity, poor study quality, and the early stage of previous research preclude any definite conclusions. The vast majority of existing studies on SMHIM either did not include a formal evaluation, or primarily focused on user perceptions and adoption intentions. There is a lack of empirical studies that have examined the effectiveness of social media use in improving patient health with objective health related measures.

Obviously, conducting such studies in an open environment like social media websites is challenging in multiple aspects, including patient recruitment and engagement, difficult control of confounding factors, and increased non-usage attrition. Providing incentives for participation is a common strategy, but it is not an ideal solution because it raises questions about system sustainability. In the long run, users will only use social media if they feel that social media can really benefit them. Although it is suggested that the majority of patients would prefer interacting with health care providers electronically, the extent and manner in which patients may wish to engage is less understood. The findings of a qualitative study (van Rensburg et al., 2016) showed that psychiatric patients were acceptive of the idea of communicating with mental health providers over social media, and highlighted how social media could provide an easy and less anxiety-provoking mode of communication and allow for constant access to a mental health provider. However, the study also identified many potential problems and risks of communication through social media, such as potential anxiety if a provider does not respond immediately and a sense that online interactions are not very rich. The perceived problems indicate the need for further work and protocol development in order for social media to be a feasible and effective venue for communication between health care providers and adolescents with psychiatric illness, or among patients themselves.

Several studies have indicated that the prolonged use of social networking sites, such as Facebook, may be related to signs and symptoms of depression, anxiety, and low self-esteem (Pantic, 2014). One of the reasons why time spent on social networking sites may be associated with depressive symptoms is the fact that social media communication may lead to the altered (and often wrong) impression of physical and personality traits of other users. The impact of social media on health may be extended to family members or caregivers. Future research should take a holistic view when examining the impact of social media on health.

3.13. Personalized health management

Although it has been suggested that personalized health-management and intervention tailored to the needs of individual patients would lead to positive improvements in instances of pain and activity limitation (Schubart et al., 2011), research along this line in the context of social media has been relatively scarce.

There are six ways that health information can be personal, which are illustrated with examples in Table 3. We highlight three key characteristics of personal health information in social media. First, it covers a broad range of information that is owned by, about, directed toward, sent/posted by, experienced by, and relevant to an individual patient. For instance, social media content generated by other users is a major source of personal health information, particularly for health information seeking. This encompasses seeking out information about a disorder, the applicable treatment for a disorder, and related medications and dosages. Many online health communities allow users to ask questions to seek responses from other users (e.g., patients or physicians) in a community. Second, personal health information scales from limited 'personal' information to 'big' social media data. Management and analysis of big data represent both a main challenge to the current healthcare practice and a strategic opportunity to innovate information on which medical decisions or even health policy-making processes are based. Third, the management of personal health information using social media is a continuous process. It is because social media run on a 24/7 basis and user-generated content can reach out to other social media users instantaneously. In contrast, patients only pay periodic visits to clinical offices, hospitals, and emergency departments. Social media provide a venue for sharing and seeking of information about what happens in-betweens such as post-treatment experience with adverse effects and pre-surgery consultation.

The concept of WeCare (Tanner et al., 2016) has been proposed as a tailored intervention that uses a community-based participatory research approach to improve care linkage and retention and health outcomes of youth with HIV. Its design uses theory-based messages tailored to each participant's place on the HIV care continuum to communicate with individual participants via a combination of Facebook messenger, text messaging, and app-based instant messages. If a participant expresses discomfort, WeCare will help increase self-efficacy by identifying and then reducing perceived barriers. However, WeCare remains at a conceptual design with no actual development and evaluation of a prototype system. Future research may consider incorporating user feedback

Table 3
Dimensions and illustrations of personal health information.

Dimensions	Example with social media
Owned by patient	Social media identifier and profile
About patient	Description of medical history
Directed toward patient	Responses to a patient's post in an online health community
Sent/posted by patient	Peer emotional support in social media
Experienced by patient	Posts about a patient's treatment and outcomes
Relevant to patient	A blog on chronic kidney disease to a patient in the disease condition

mechanisms and triggers that provide automated, personalized messages or interventions to motivate and inspire users to conduct healthier and self-management behaviors. There also needs more research to investigate which types of personalized patient self-management support can encourage user engagement, and what is the impact of such personalized support on health outcomes and patient satisfaction.

4. Conclusion

The role of social media in the health care sector is far reaching, and many questions in terms of data analytics, governance, ethics, professionalism, privacy, confidentiality, and information quality remain unanswered. SMHIM can be manifested in customized patient education based on skill level, cultures, and health literacy levels of individual patients; social media-based personalized health recommendation systems; social media based intervention; and managing the health of people with chronic conditions on a day-to-day basis, so on and so forth. Future research on SMHIM will be required to understand the synergies between social media and health promotion.

Acknowledgment

This research was partially supported by the National Science Foundation under Grant No. SES-152768 and CNS-1704800, by the National Institutes of Health under Grant No. 2R42CA168107-02, and by the National Natural Science Foundation of China under Grant No. #71772075 and #71672074. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the above funding agencies.

References

Alavi, M., Leidner, D.E., 2001. Review: Knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Q. 25, 107–136.

Anderson, N., Borlak, J., 2011. Correlation versus causation? Pharmacovigilance of the analgesic flupirtine exemplifies the need for refined spontaneous ADR Reporting. PLOS One 6, e25221.

Andreu-Perez, J., Poon, C.C.Y., Merrifield, R.D., Wong, S.T.C., Yang, G.Z., 2015. Big data for health. IEEE J. Biomed. Health. Inf. 19, 1193–1208.

Antheunis, M.L., Tates, K., Nieboer, T.E., 2013. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ. Couns. 92, 426–431.

Benton, A., Ungar, L., Hill, S., Hennessy, S., Mao, J., Chung, A., Leonard, C.E., Holmes, J.H., 2011. Identifying potential adverse effects using the web: a new approach to medical hypothesis generation. J. Biomed. Inf. 44, 989–996.

Bernardo, T.M., Rajic, A., Young, I., Robiadek, K., Pham, M.T., Funk, J.A., 2013. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. J. Med. Internet Res. 15, e147.

Bian, J., Topaloglu, U., Yu, F., 2012. Towards large-scale twitter mining for drug-related adverse events. In: Proc. Conference on Information & Knowledge Management, 25.

Bizer, C., Heath, T., Berners-Lee, T., 2009. Linked data: the story so far. Int. J. Semantic Web Inf. Syst. 5, 1–22.

Broniatowski, D.A., Paul, M.J., Dredze, M., 2013. National and local influenza surveillance through twitter: an analysis of the 2012–2013 influenza epidemic. PLOS One 8, e83672.

Brown, J., Ryan, C., Harris, A., 2014. How doctors view and use social media: a national survey. J. Med. Internet Res. 16, e267.

Burke-Garcia, A., Scally, G., 2014. Trending now: future directions in digital media for the public health sector. J. Public Health (Oxford) 36, 527–534.

Cai, R., Liu, M., Hu, Y., Melton, B.L., Matheny, M.E., Xu, H., Duan, L., Waitman, L.R., 2017. Identification of adverse drug-drug interactions through causal association rule discovery from spontaneous adverse event reports. Artif. Intell. Med. 76, 7–15.

Chang, T., Chopra, V., Zhang, C., Woolford, S.J., 2013. The role of social media in online weight management: systematic review. J. Med. Internet Res. 15, e262.

Chapelle, O., Schlkopf, B., Zien, A., 2010. Semi-Supervised Learning. The MIT Press.

Charles-Smith, L.E., Reynolds, T.L., Cameron, M.A., Conway, M., Lau, E.H.Y., Olsen, J.M., Pavlin, J.A., Shigematsu, M., Streichert, L.C., Suda, K.J., Corley, C.D., 2015. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. PLOS One 10, e0139701.

Chee, B.W., Berlin, R., Schatz, B., 2011. Predicting adverse drug events from personal health messages. AMIA Annu. Symp. Proc. 2011, 217–226.

Chretien, K.C., Kind, T., 2013. Social media and clinical care: ethical, professional, and social implications. Circulation 127, 1413–1421.

Correia, R.B., Li, L., Rocha, L.M., 2016. Monitoring potential drug interactions and reactions via network analysis of instagram user timelines. In Proc. Pacific Symp. Biocomputing 21, 492–503.

Cottle, J., 2016. Facebook and Mental Health: Is Social Media Hurting or Helping? MentalHelp.net.

Davila, J., Hershenberg, R., Feinstein, B.A., Gorman, K., Bhatia, V., Starr, L.R., 2012. Frequency and quality of social networking among young adults: associations with depressive symptoms, rumination, and corumination. Psychol. Popular Media Culture 1, 72–86.

DeAndrea, D.C., Vendemia, M.A., 2016. How affiliation disclosure and control over user-generated comments affects consumer health knowledge and behavior: a randomized controlled experiment of pharmaceutical direct-to-consumer advertising on social media. J. Med. Internet Res. 18, e189.

Denecke, K., Bamidis, P., Bond, C., Gabarron, E., Househ, M., Lau, A.Y.S., Mayer, M.A., Merolli, M., Hansen, M., 2015. Ethical issues of social media usage in healthcare. Yearb. Med. Inform. 10, 137–147.

Dizon, D.S., Graham, D., Thompson, M.A., Johnson, L.J., Johnston, C., Fisch, M.J., Miller, R., 2012. Practical guidance: the use of social media in oncology practice. J. Oncol. Pract. 8, e114–124.

Dosemagen, S., Aase, L., 2016. How Social Media is Shaking Up Public Health and Healthcare. The Blog, Huffington Post, January, p. 27.

Fergie, G., Hilton, S., Hunt, K., 2016. Young adults' experiences of seeking online information about diabetes and mental health in the age of social media. Health Expect. 19, 1324–1335.

Fung, I.C.-H., Tse, Z.T.H., Fu, K.-W., 2015. The use of social media in public health surveillance. West. Pacific Surveillance Response J. 6, 3–6.

Grajales 3rd, F.J., Sheps, S., Ho, K., Novak-Lauscher, H., Eysenbach, G., 2014. Social media: a review and tutorial of applications in medicine and health care. J. Med. Internet Res. 16, e13.

Hazell, L., Shakir, S.A., 2006. Under-reporting of adverse drug reactions: a systematic review. Drug Saf. 29, 385–396.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society Press, Washington, DC, pp. 770–778.

Health on the Net Foundation, 2017. The HON code of conduct for medical and health web sites (HONcode), December 2017. Available at: http://www.healthonnet.org/HONcode/Conduct.html.

Hill, A.B., 1965. The environment and disease: association or causation? Proc. R. Soc. Med. 58, 295–300.

Holzner, B., Marx, J., 1979. The Knowledge Application: The Knowledge System in Society. Allyn-Bacon, Boston.

Hormes, J.M., Kearns, B., Timko, C.A., 2014. Craving Facebook? Behavioral addiction to online social networking and its association with emotion regulation deficits. Addiction 109, 2079–2088.

Horvitz, E., Mulligan, D., 2015. Data, privacy, and the greater good. Science 349, 253.

Househ, M., 2013. The use of social media in healthcare: organizational, clinical, and patient perspectives. Stud. Health Technol. Inf. 183, 244–248.

Jashinsky, J., Burton, S.H., Hanson, C.L., West, J., Giraud-Carrier, C., Barnes, M.D., Argyle, T., 2014. Tracking suicide risk factors through Twitter in the U.S. crisis 35, 51–59.

Kang, Y., Wang, Y., Zhang, D., Zhou, L., 2017. The public's opinions on a new school meals policy for childhood obesity prevention in the U.S.: a social media analytics approach. Int. J. Med. Inf. 103, 83–88.

Karimi, S., Wang, C., Metke-Jimenez, A., Gaire, R., Paris, C., 2015. Text and data mining techniques in adverse drug reaction detection. ACM Comput. Surv. 47, 1–39.

Kass-Hout, T.A., Alhinnawi, H., 2013. Social media in public health. Br. Med. Bull. 108, 5–24.

Keckley, P., Hoffmann, M., 2010. Social networks in health care: communication, collaboration and insights. Deloitte Center for Health Solutions.

Kim, Y., Huang, J., Emery, S., 2016. Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. J. Med. Internet Res. 18, e41.

Kotov, A., 2015. Social media analytics for healthcare. In: Reddy, C.K., Aggarwal, C.C. (Eds.), Healthcare Data Analytics. first ed. Chapman and Hall/CRC, New York, pp. 309–340.

Kuhn, M., Campillos, M., Letunic, I., Jensen, L.J., Bork, P., 2010. A side effect resource to capture phenotypic effects of drugs. Mol. Syst. Biol. 6, 343.

Langford, C.P.H., Bowsher, J., Maloney, J.P., Lillis, P.P., 1997. Social support: a conceptual analysis. J. Adv. Nurs. 25, 95–100.

Lardon, J., Abdellaoui, R., Bellet, F., Asfari, H., Souvignet, J., Texier, N., Jaulent, M.C., Beyens, M.N., Burgun, A., Bousquet, C., 2015. Adverse drug reaction identification and extraction in social media: a scoping review. J. Med. Internet Res. 17, e171.

Leaman, R., Wojtulewicz, L., Sullivan, R., Skariah, A., Yang, J., Gonzalez, G., 2010. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In: Proceedings of the 2010 Workshop on Biomedical Natural Language Processing, Association for Computational Linguistics, pp. 117–125.

Liang, B.A., Mackey, T.K., 2011. Prevalence and global health implications of social media in direct-to-consumer drug advertising. J. Med. Internet Res. 13, e64.

Liu, X., Chen, H., 2015. A research framework for pharmacovigilance in health social media: identification and evaluation of patient adverse drug event reports. J. Biomed. Inf. 58, 268–279.
Liu, Y., Gummadi, K.P., Krishnamurthy, B., Mislove, A., 2011. Analyzing Facebook privacy settings: user expectations vs. reality. In: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference. ACM Press, Berlin, Germany, pp. 61–70.
Lu, Y., Zhang, P., Liu, J., Li, J., Deng, S., 2013. Health-related hot topic detection in online communities using text clustering. PLOS One 8, e56221.
Lupton, D., Jutel, A., 2015. It's like having a physician in your pocket!'A critical analysis of self-diagnosis smartphone apps. Soc. Sci. Med. 133, 128–135.
Mao, J.J., Chung, A., Benton, A., Hill, S., Ungar, L., Leonard, C.E., Hennessy, S., Holmes, J.H., 2013. Online discussion of drug side effects and discontinuation among breast cancer survivors. Pharmacoepidemiol. Drug Saf. 22, 256–262.
McCord, B., Rodebaugh, T.L., Levinson, C.A., 2014. Facebook: social uses and anxiety. Comput. Hum. Behav. 34, 23–27.
McGough, S.F., Brownstein, J.S., Hawkins, J.B., Santillana, M., 2017. Forecasting Zika incidence in the 2016 Latin America outbreak combining traditional disease surveillance with search, social media, and news report data. PLOS Negl. Trop. Dis. 11, e0005295.
Merolli, M., Gray, K., Martin-Sanchez, F., 2013a. Developing a framework to generate evidence of health outcomes from social media use in chronic disease management. Medicine 2.0 2, e3.
Merolli, M., Gray, K., Martin-Sanchez, F., 2013b. Health outcomes and related effects of using social media in chronic disease management: a literature review and analysis of affordances. J. Biomed. Inform. 46, 957–969.
Merolli, M., Gray, K., Martin-Sanchez, F., 2016. Patient participation in chronic pain management through social media: a clinical study. Stud. Health Technol. Inf. 225, 577–581.
Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. arXiv:1301.3781 [cs.CL].
Moorhead, A.S., Hazlett, E.D., Harrison, L., Carroll, K.J., Irwin, A., Hoving, C., 2013. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 15, e85.
Nagar, R., Yuan, Q., Freifeld, C.C., Santillana, M., Nojima, A., Chunara, R., Brownstein, J.S., 2014. A case study of the New York City 2012–2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. J. Med. Internet Res. 16, e236.
Naslund, J.A., Aschbrenner, K.A., Marsch, L.A., Bartels, S.J., 2016. The future of mental health care: peer-to-peer support and social media. Epidemiol. Psychiatr. Sci. 25, 113–122.
Nguyen, T., Larsen, M.E., O'Dea, B., Phung, D., Venkatesh, S., Christensen, H., 2017. Estimation of the prevalence of adverse drug reactions from social media. Int. J. Med. Inf. 102, 130–137.
Nikfarjam, A., Gonzalez, G.H., 2011. Pattern mining for extraction of mentions of Adverse Drug Reactions from user comments. Proc. Ann. Symp. AMIA, 1019–1026.
Nikfarjam, A., Sarker, A., O'Connor, K., Ginn, R., Gonzalez, G., 2015. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. J. Am. Med. Inform. Assoc. 22, 671–681.
Norman, C.D., Skinner, H.A., 2006. eHealth literacy: essential skills for consumer health in a networked world. J. Med. Internet Res. 8, e9.
Oyeyemi, S.O., Gabarron, E., Wynn, R., 2014. Ebola, Twitter, and misinformation: a dangerous combination? BMJ British Medical Journal 349.
Pantic, I., 2014. Online social networking and mental health. Cyberpsychol. Behav. Soc. Network. 17, 652–657.
Patel, R., Chang, T., Greysen, S.R., Chopra, V., 2015. Social media use in chronic disease: a systematic review and novel taxonomy. Am. J. Med. 128, 1335–1350.
Paul, M.J., Dredze, M., 2011. You are what you tweet: analyzing twitter for public health. In: Adamic, L.A., Baeza-Yates, R.A., Counts, S. (Eds.), ICWSM. AAAI Press, Palo Alto, CA.
Paul, M.J., Sarker, A., Brownstein, J.S., Nikfarjam, A., Scotch, M., Smith, K.L., Gonzalez, G., 2016. Social media mining for public health monitoring and surveillance. In: Proceedings of the Pacific Symposium on Biocomputing. World Scientific Publishing, Singapore, pp. 468–479.
Pentland, B.T., 1995. Information systems and organizational learning: the social epistemology of organizational knowledge systems. Accounting Manage. Inf. Technol. 5, 1–21.
Peters, M.E., Uible, E., Chisolm, M.S., 2015. A Twitter education: why psychiatrists should tweet. Curr. Psychiatry Rep. 17, 94.
Pillow, M.T., Hopson, L., Bond, M., Cabrera, D., Patterson, L., Pearson, D., Sule, H., Ankel, F., Fernandez-Frackelton, M., Hall, R.V., Kegg, J.A., Norris, D., Takenaka, K., 2014. Council of residency directors social media task forcesocial media guidelines and best practices: recommendations from the council of residency directors social media task force. West. J. Emergency Med. 15, 26–30.
Reblin, M., Uchino, B.N., 2008. Social and emotional support and its implication for health. Curr. Opin. Psychiatry 21, 201–205.
Reid, M., Landesman, S., Treder, R., Jaccard, J., 1989. My family and friends: six- to twelve-year-old children's perceptions of social support. Child Dev. 60, 896–910.
Robinson, J., Rodrigues, M., Fisher, S., Bailey, E., Herrman, H., 2015. Social media and suicide prevention: findings from a stakeholder survey. Shanghai Arch. Psychiatry 27, 27–35.
Rosen, L.D., Whaling, K., Rab, S., Carrier, L.M., Cheever, N.A., 2013. Is Facebook creating "iDisorders"? The link between clinical symptoms of psychiatric disorders and technology use, attitudes and anxiety. Comput. Hum. Behav. 29, 1243–1254.
Rotondi, A.J., Anderson, C.M., Haas, G.L., Eack, S.M., Spring, M.B., Ganguli, R., Newhill, C., Rosenstock, J., 2010. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr. Serv. 61, 1099–1105.
Santoro, E., Castelnuovo, G., Zoppis, I., Mauri, G., Sicurello, F., 2015. Social media and mobile applications in chronic disease prevention and management. Front. Psychol. 6.
Santos, J.C., Matos, S., 2014. Analysing Twitter and web queries for flu trend prediction. Theor. Biol. Med. Modell. 11. S6 S6.
Sarker, A., Ginn, R., Nikfarjam, A., O'Connor, K., Smith, K., Jayaraman, S., Upadhaya, T., Gonzalez, G., 2015. Utilizing social media data for pharmacovigilance: a review. J. Biomed. Inform. 54, 202–212.
Schubart, J.R., Stuckey, H.L., Ganeshamoorthy, A., Sciamanna, C.N., 2011. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput. Inf. Nurs. 29, 81–92.
Seeman, N., 2008. Web 2.0 and chronic illness: new horizons new opportunities. Healthcare Q. 11, 104–108.
Segura-Bedmar, I., Martínez, P., Revert, R., Moreno-Schneider, J., 2015. Exploring Spanish health social media for detecting drug effects. BMC Med. Inf. Decis. Making 15. S6 S6.
Setoyama, Y., Yamazaki, Y., Namayama, K., 2011. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. J. Med. Internet Res. 13, e122.
Shaw, R.J., Johnson, C.M., 2011. Health information seeking and social media use on the Internet among people with diabetes. Online J. Public Health Inf. 3.
Song, H., Omori, K., Kim, J., Tenzek, K.E., Morey Hawkins, J., Lin, W.Y., Kim, Y.C., Jung, J.Y., 2016. Trusting social media as a source of health information: online surveys comparing the United States, Korea, and Hong Kong. J Med Internet Res 18, e25.
Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., Brand, H., 2012. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 12, 80.
Stellefson, M., Chaney, B., Barry, E.A., Chavarria, E., Tennant, B., Walsh-Childers, K., Sriram, P.S., Zagora, J., 2013. Web 2.0 chronic disease self-management for older adults: a systematic review. J. Med. Internet Res. 15, e35.
Stellefson, M., Chaney, B., Ochipa, K., Chaney, D., Haider, Z., Hanik, B., Chavarria, E., Bernhardt, J.M., 2014. YouTube as a source of chronic obstructive pulmonary disease patient education: a social media content analysis. Chron. Respir. Dis. 11, 61–71.
Stone, J., 2015. Social media is a lifeline for patients with rare diseases, Forbes, May 2.
Tanner, A.E., Mann, L., Song, E., Alonzo, J., Schafer, K., Arellano, E., Garcia, J.M., Rhodes, S.D., 2016. weCARE: a social media-based intervention designed to increase HIV care linkage, retention, and health outcomes for racially and ethnically diverse young MSM. AIDS Educ. Prev. 28, 216–230.
Thompson, L., Rivara, F.P., Whitehill, J.M., 2015. Prevalence of marijuana-related traffic on Twitter, 2012–2013: a content analysis. Cyberpsychol. Behav. Soc. Network. 18, 311–319.
Tuarob, S., Tucker, C.S., Salathe, M., Ram, N., 2014. An ensemble heterogeneous classification methodology for discovering health-related knowledge in social media messages. J. Biomed. Inform. 49, 255–268.
Tyrawski, J., DeAndrea, D.C., 2015. Pharmaceutical companies and their drugs on social media: a content analysis of drug information on popular social media sites. J. Med. Internet Res. 17, e130.
Valimaki, M., Athanasopoulou, C., Lahti, M., Adams, C.E., 2016. Effectiveness of social media interventions for people with schizophrenia: a systematic review and meta-analysis. J. Med. Internet Res. 18, e92.
van Rensburg, S.H., Klingensmith, K., McLaughlin, P., Qayyum, Z., van Schalkwyk, G. I., 2016. Patient-provider communication over social media: perspectives of adolescents with psychiatric illness. Health Expect. 19, 112–120.
Velasco, E., Agheneza, T., Denecke, K., Kirchner, G., Eckmanns, T.I.M., 2014. Social media and Internet-based data in global systems for public health surveillance: a systematic review. Milbank Q. 92, 7–33.
Ventola, C.L., 2014. Social media and health care professionals: benefits, risks, and best practices. Pharm. Ther. 39, 491–520.
Wang, H., Jiang, Y., 2017. Quality of user-reported data in health virtual community. In: Proceedings of the 23rd Americas Conference on Information Systems, Boston, MA.
Wester, K., Jonsson, A.K., Spigset, O., Druid, H., Hagg, S., 2008. Incidence of fatal adverse drug reactions: a population based study. Br. J. Clin. Pharmacol. 65, 573–579.
Wu, H., Fang, H., Stanhope, S.J., 2013. Exploiting online discussions to discover unrecognized drug side effects. Methods Inf. Med. 52, 152–159.
Xie, J., Liu, X., Dajun Zeng, D., 2017. Mining e-cigarette adverse events in social media using Bi-LSTM recurrent neural network with word embedding representation. J. Am. Med. Inform. Assoc.
Xu, Q., Gel, Y.R., Ramirez Ramirez, L.L., Nezafati, K., Zhang, Q., Tsui, K.-L., 2017. Forecasting influenza in Hong Kong with Google search queries and statistical model fusion. PLOS One 12, e0176690.
Yang, F.C., Lee, A.J., Kuo, S.C., 2016. Mining health social media with sentiment analysis. J. Med. Syst. 40, 236.

Yang, H., Yang, C.C., 2013. Harnessing social media for drug-drug interactions detection. In: Proc. 2013 IEEE International Conference on Healthcare Informatics. IEEE Computer Society Press, Washington, DC, pp. 22–29.

Yang, M., Kiang, M., Shang, W., 2015. Filtering big data from social media: building an early warning system for adverse drug reactions. J. Biomed. Inform. 54, 230–240.

Yang, X., Luo, J., 2017. Tracking illicit drug dealing and abuse on instagram using multimodal analysis. ACM Trans. Intell. Syst. Technol. 8, 1–15.

Yates, A., Cohan, A., Goharian, N., 2017. Depression and self-harm risk assessment in online forums. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, pp. 2958–2968.

Yates, A., Goharian, N., Serdyukov, P., Segalovich, I., Braslavski, P., Kuznetsov, S.O., Kamps, J., Rüger, S., Agichtein, E., Yilmaz, E., 2013. ADRTrace: Detecting Expected and Unexpected Adverse Drug Reactions from User Reviews on Social Media Sites. In: Serdyukov, P., Segalovich, I., Braslavski, P., Kuznetsov, S. O., Kamps, J., Rüger, S., Agichtein, E., Yilmaz, E. (Eds.), Advances in Information Retrieval: Proceedings of the 35th European Conference on IR Research. Springer Verlag, Berlin, Germany.

Zhou, L., Pan, S., Wang, J., Vasilakos, A.V., 2017. Machine learning on big data: opportunities and challenges. Neurocomputing 237, 350–361.