Article

Using Twitter for Data Collection With Health-Care Consumers: A Scoping Review

Amy J. Zhang[1], Lauren Albrecht[2], and Shannon D. Scott[1]

Abstract

Background: Twitter is one of the most popular social media platforms. The growing use of Twitter by health-care consumers creates a novel venue to understand patient experiences. To understand the potential for this platform to be utilized in patient- and family-oriented health research, this study reviewed published literature on the use of Twitter in health research. **Methods:** In collaboration with the research team, a research librarian designed and implemented a search strategy in eight databases. Primary and secondary screenings were conducted using predetermined criteria by one reviewer. A second reviewer verified screening decisions in 10% of the studies. Evidence tables were created to synthesize across the following study elements: research design, data collection techniques, analytic approaches, and author's insights on Twitter as a data collection method. Descriptive narrative analysis was used to synthesize data. **Results:** The search strategy captured 618 articles; 233 were eliminated in primary screening and 366 articles were eliminated during secondary screening. Verification by the second reviewer resulted in very good agreement ($\kappa = .980$). Seventeen articles were included in the final data set. Synthesis across the studies demonstrated that Twitter is currently used to search and mine research data, while active recruitment strategies on Twitter are just beginning to emerge. **Conclusion:** The novelty of Twitter for study recruitment and data collection with health-care consumers presents advantages and challenges that differ from traditional methods of data collection.

Keywords

Twitter, social media, data collection, health-care consumers, patient-oriented research, family-oriented research, recruitment

What Is Already Known?

Twitter, a popular and growing social media platform, has been increasingly utilized in health contexts including public awareness of natural disasters, syndromic surveillance, assessing public misunderstanding surrounding antibiotic use, and as a cost-effective means of research recruitment. The use of these 140 character microblogs known as "tweets" and their rapid exchange of information are only a few characteristics of Twitter that may provide novel methods for data collection, which have yet to be fully explored.

What This Paper Adds?

This manuscript is the first review of published research on the use of Twitter to collect health research data with health consumers. From this study, we concluded that use of Twitter in health research, including participant recruitment, data collection methods, and type of data collected, is similar across different qualitative and quantitative study designs. Nevertheless, Twitter is ideal for qualitative research, as data are presented with unbiased text giving insight to underlying reasons and opinions of a phenomenon. In addition, Twitter deviates from traditional methods of data collection (i.e., in-person focus groups, interviews) in terms of how data are collected, how data are verified, ethical considerations, and accessibility of research. Thus, this manuscript provides new considerations for health research and researchers using Twitter and other social media platforms.

Background

As the use of social media proliferates, researchers have begun to explore these platforms as tools to efficiently disseminate information to vast and diverse audiences in a cost-effective manner (Robillard, Johnson, Hennessey, Beattie, & Illes,

[1] Faculty of Nursing, University of Alberta, Alberta, Canada
[2] Department of Pediatrics, University of Alberta, Alberta, Canada

Corresponding Author:
Shannon D. Scott, Edmonton Clinic Health Academy, Faculty of Nursing, 5-187, University of Alberta, 11405 87 NW Edmonton, Alberta, Canada T6G 1C9.
Email: shannon.scott@ualberta.ca

2013). Although social media is approximately 15 years old, these platforms have already changed approaches to health communication (Stump, Zilch, & Coustasse, 2012). In addition to information dissemination, social media has influenced the collection of health information (Wittmeier et al., 2014) and these platforms have been increasingly explored as mechanisms to increase engagement in health education and to conduct health research (Finfgeld-Connett, 2015). The strength of social media lies in its ability to develop relationships and communities around shared interests (i.e., rare disease and chronic illnesses) even if groups are geographically dispersed as well as being able to develop communities asynchronously (Wittmeier et al., 2014). There is increasing interest in understanding the full impact of social media platforms on health research (Finfgeld-Connett, 2015).

Twitter, one social media platform, is a form of microblogging with brief 140 character "tweets" that may consist of images, text, and links (O'Connor, Jackson, Goldsmith, & Skirton, 2014). Registered users follow other accounts to quickly exchange information and see updates; unregistered users may only read tweets (Stump et al., 2012). Established in 2006, Twitter reports 1.3 billion registered users, 313 million active registered users, and 500 million daily tweets (Finfgeld-Connett, 2015). Twitter is used by 23% of online adults, with 37% of users between 19 and 29 years old and 25% between the ages of 30 and 49 years (Finfgeld-Connett, 2015). The growth rate of Twitter has been considerable, with no evidence of slowing down (Stump et al., 2012).

Twitter has been increasingly utilized in health (Finfgeld-Connett, 2015), in contexts ranging from spreading awareness during natural disasters to tweeting real time in hospitals (Holt, 2011; Parker-Pope, 2009). "TwitterCare" includes using the platform for diabetes management, blood glucose monitoring, drug safety alerts, chronic condition self-management, diagnostic brainstorming, infant care tips, and post discharge follow-up care (Fisher & Clayton, 2012).

Twitter has also played various roles in health research initiatives (Robillard et al., 2013), including tracking flu epidemics by conducting syndromic surveillance (Eysenbach, 2009) and assessing public misunderstandings surrounding antibiotic use (O'Connor et al., 2014). As a tool for research recruitment, Twitter is a cost-effective platform with the potential to engage hard-to-reach populations (Clinical Digest, 2014). Although tweets are limited to 140 characters, they include useful data and metadata for researchers, such as the account holder's language, geolocation, their handles (account names), and the number and handles of followers as well (Finfgeld-Connett, 2015), thus maximizing useful data with minimal effort.

As one of the most popular social media platforms and a popular tool for health-care communication, Twitter presents a possible new venue to engage health-care consumers in health research. The ability of Twitter to connect individuals around shared interests and the unique qualities of this particular social media platform warrant the need for further exploration.

In particular, Twitter is ideal for qualitative research because it is one of the most well-established platforms used to discuss a variety of topics (Salzmann-Erikson, 2017). Twitter content is also already in text form, thus making it increasingly valuable to health researchers (Kilaru et al., 2016) seeking insight on underlying reasons and opinions of a phenomenon. The tweets, themselves, are naturally occurring contextual responses that can be collected with or without prompts from researchers and encompass cross-cultural and international views of the general public (Emma Hilton, 2017). To understand and design health interventions, it is imperative to broaden and examine diverse populations and audiences (Lyles, Lopez, Pasick, & Sarkar, 2013) and Twitter provides a well-established medium for this type of investigation. Hence, the purpose of this scoping review was to systematically identify and describe published health research evaluating the use of Twitter as a data collection method with health consumers.

Method

A scoping review is a knowledge synthesis methodology that is intended to review literature addressing broad topics inclusive of many different study designs (Arksey & O'Malley, 2005). Scoping reviews are used to assess emerging and established fields by evaluating the landscape of a body of knowledge and identifying gaps in existing literature (Arksey & O'Malley, 2005; Colquhoun et al., 2014). The goal of a scoping review is to provide a descriptive overview, and often strict limitations are not placed on search terms and critical appraisal is not used with identified literature (Arksey & O'Malley, 2005; Pham et al., 2014). Scoping reviews are a relatively new methodology and are becoming increasingly popular (Arksey & O'Malley, 2005). This scoping review sought to provide an overview of previously published literature describing Twitter as a data collection method with health-care consumers and provide researchers with considerations when potentially using this data collection approach.

Search Strategy

The research team and a health research librarian developed and implemented search strategies in eight electronic databases OvidEpub, Ovid MEDLINE (R) Daily and Ovid MEDLINE(R), Ovid Embase, Ovid PsycINFO, Ovid EBM, EBSCO CINAHL, Elsevier Scopus, and Thomas Reuters Web of Science Core Collection, using language (English) and date restrictions (2010–2015; Appendix A). The decision to restrict studies to English was informed by recent systematic research evidence suggesting there is no empirical evidence of bias if papers written in languages other than English are excluded (Morrison et al., 2012). These date restrictions reflect the recent increase in popularity of Twitter within health care (Stump et al., 2012).

Study Inclusion Criteria

Studies were included if they met the criteria outlined in Table 1. Studies were not excluded based upon research design. Primary research studies were included if they used Twitter to collect research data with health consumers (Appendix C). Studies were

Table 1. Inclusion Criteria.

Features	Must Be/Include
Research type	Primary research of all study designs.
Participants	Health consumers were defined as recipients of health care and/or a participant in the health decision-making process (i.e., patient, parent, caregiver, and other family member)
Study focus	Include only Twitter, discuss methods used, and be related to health research
Key words	Search study parameters included key words such as social media, Twitter, health research, and data collection method
Language	English

excluded if they assessed other social media platforms (i.e., Facebook, Instagram) focused on research dissemination or included health-care providers. The large, heterogeneous population that is health-care consumers was selected to assess the state of the science. Even with this broad target population, only 17 studies were included in the final set; therefore, analysis of subpopulations was not warranted.

Study Selection

Primary screening of titles and abstracts was completed by one reviewer (A.Z.). Each article was rated as include, exclude, or unclear using a primary screening tool (Appendix C). Full-text articles were retrieved for secondary screening if they were classified as include or unclear in primary screening. Full articles were assessed using standard forms and predetermined inclusion criteria (Table 1). Articles included during secondary screening could only discuss Twitter, focus on only data collection, and target health-care consumers (see Appendix C). A second independent reviewer (X.W.) randomly screened 10% of all primary and secondary screened articles, using the secondary screening tool (Appendix C). All disagreements were resolved by discussion or third-party adjudication (L.A.).

Data Extraction

Data were extracted by one reviewer (A.Z.) for the following variables: research design, data collection techniques, analytic approaches, tools used to analyze and collect data, and the author's conclusions on Twitter as a health research method. These variables were selected to explore whether there is a predominant research design used, whether there are least common and most common methodologies being employed, and draw attention to any conclusions or insights authors noted after using this method. After analyzing these variables, we aim to provide an overview of how Twitter is being used to reach this population for research.

Data Analysis

Descriptive narrative analysis was used to identify potential patterns (e.g., similarities, anomalies) in terms of differences

of data collection techniques using Twitter. Evidence tables were built to facilitate synthesis across the studies. The descriptive analysis allowed us to (1) understand the current extent of how Twitter is used as a data collection method with health-care consumers, (2) evaluate the disadvantages and advantages of this method, and (3) compare and contrast our own experience using Twitter as a data collection method to help develop this body of literature.

Results

Seventeen studies met our inclusion criteria (Figure 1). The interrater reliability of screening decisions between the two reviewers was "very good" with a κ coefficient of .980, standard error of 0.020, and 95% confidence interval [0.940, 1.000]. ("GraphPadQuickCalcs: Quantify interrater agreement with kappa", 2016). Table B1 (Appendix B) summarizes the 17 studies including important study elements and is organized by alphabetical order.

Quantitative ($n = 2$), qualitative ($n = 7$), and mixed methods ($n = 8$) research designs were represented in Table B1 (Appendix B). A wide range of health topics and research questions were explored including health challenges such as pain, migraines, and cancer (Ahlwardt et al., 2014; Nascimento et al., 2014; Parsons, Breckons, & Durham, 2015), social discourse of conditions like perceptions of portrayal of seizures (McNeil, Brna, & Gordon, 2011), and cyberspace in comparison with real-world phenomena (Nagel et al., 2013).

Mixed-Methods Studies

Mixed-methods studies ($n = 8$) were published within the last 3 years, with the exception of two articles published in 2011 (McNeil et al., 2011) and (Marton, 2012). Articles were all published in Western countries including Australia ($n = 2$; Beykikhoshk, Arandjelović, Phung, Venkatesh, & Caelli, 2014, 2015), United Kingdom ($n = 3$; Greaves et al., 2014; Parsons et al., 2015; Ramagopalan, Wasiak, & Cox, 2014), Canada ($n = 2$; Marton, 2012; McNeil et al., 2011), and the United States (Ahlwardt et al., 2014). Mixed-method studies included health focuses on pain (toothache, backache, earache, and headache; Ahlwardt et al., 2014); multiple sclerosis (Ramagopalan et al., 2014); autism spectrum disorder (Beykikhoshk et al., 2014, 2015); epilepsy (McNeil et al., 2011); pain tracking (Parsons et al., 2015); cancer patient tweets from Canadian Cancer Society (Marton, 2012); tweets to the hospitals in the English National Health Service about quality of care (Greaves et al., 2014).

Data collection methods using Twitter. Three studies used retrospective data collection to examine past tweets (Beykikhoshk et al., 2014, 2015; Ramagopalan et al., 2014). Six studies collected data prospectively by setting a future date to begin collecting tweets as data (Ahlwardt et al., 2014; Greaves et al., 2014; Marton, 2012; McNeil et al., 2011; Nakhasi et al., 2016; Parsons et al., 2015). Seven of the eight studies used various tools and programs to search for and collect data in addition to
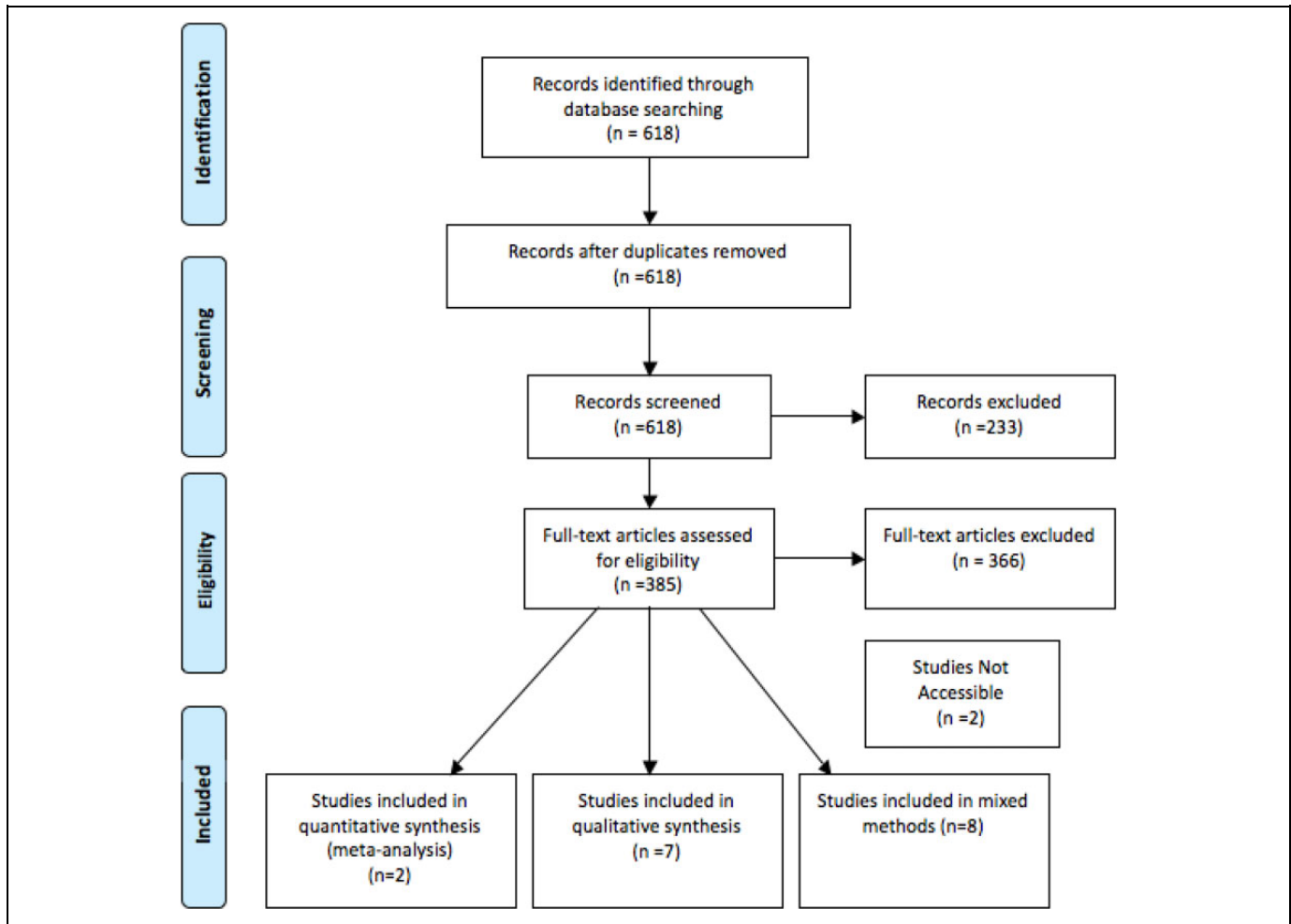
Figure 1. PRISMA diagram.

tweets including key words, phrases, or geolocations attached to the Twitter post or profile (Ahlwardt et al., 2014; Beyki-khoshk et al., 2014, 2015; Greaves et al., 2014; Marton, 2012; McNeil et al., 2011; Ramagopalan et al., 2014). One common study, included in both the eight mixed-method studies, and the total 17 studies in the final set, used active recruitment to identify participants on Twitter and collect data with predetermined individuals at a predetermined time, about a specific topic (Parsons et al., 2015). Alternatively, all other studies relied upon collecting data from tweets passively from any user without actively engaging them.

Qualitative Studies

Qualitative studies (*n* = 7) were published within the last 3 years, with the exception of one study in 2012 (Sugawara et al., 2012). One article was from Australia (Hewis, 2015), three from Japan (Sugawara et al., 2012; Tsuya, Sugawara, Tanaka, & Narimatsu, 2014), three from the United States (Nakhasi et al., 2016; Nascimento et al., 2014; Shive, Bhatt, Cantino, Kvedar, & Jethwani, 2013), one from New Zealand (Henzell, Knight, Morgaine, Antoun, & Farella, 2014). Qualitative studies examined health focuses including patient safety (Nakhasi et al., 2016);

infodemiology of self-reported migraine headache suffering (Nascimento et al., 2014); acne (Shive et al., 2013); orthodontics (Henzell et al., 2014; cancer (breast, leukemia, colon cancer, rectal cancer, colorectal, uterine, stomach, lung, and ovarian; Tsuya et al., 2014); magnetic resonance imaging patient's perspectives (Hewis, 2015); Twitter usage and its role in "wired" cancer patients (Sugawara et al., 2012).

Data collection methods using Twitter. Two studies used retrospective data collection (Hewis, 2015; Tsuya et al., 2014). Four studies collected data prospectively (Henzell et al., 2014; Nascimento et al., 2014; Shive et al., 2013; Sugawara et al., 2012). All seven qualitative studies also used various tools and programs to search for and collect data in addition to tweets including key words, phrases, or geolocations attached to the Twitter post or profile (Henzell et al., 2014; Hewis, 2015; Nakhasi et al., 2016; Nascimento et al., 2014; Shive et al., 2013; Sugawara et al., 2012; Tsuya et al., 2014).

Quantitative Studies

Quantitative studies (*n* = 2) were published within the last 3 years, and both were published in the United States (Love,

Himelboim, Holton, & Stewart, 2013; Nagel et al., 2013). One study examined tweets about vaccinations (Love et al., 2013). The second study examined interaction between cyberspace messages and real-world occurrences of influenza and pertussis (Nagel et al., 2013).

One study used retrospective data collection and one study used prospective data collection, respectively (Love et al., 2013; Nagel et al., 2013). Both quantitative studies used various tools and programs to search for and collect data in addition to tweets including key words, phrases, or geolocations attached to the Twitter post or profile (Love et al., 2013; Nagel et al., 2013).

Analytic approaches used for data collected via Twitter. The collected data were used for qualitative content and discourse analysis examining text, hashtags, and word frequency ($n = 7$). Quantitative statistical analysis was also used for linguistic analysis and applied to calculate sentiment, word frequency, and make comparisons (e.g., between tweets and disease; $n = 2$).

Discussion

This study highlights an emerging body of health research that is using Twitter, especially within the last 3 years. While qualitative, quantitative, and mixed-method study designs were represented in this review, our findings demonstrate that the way Twitter is used within health research, including participant recruitment, data collection methods, and type of data collected, is similar across study designs. For example, similar factors include the use of passive recruitment, data mining techniques, and collection of tweets, and user profiles display only one approach to collecting data. We will further explore these findings in the following paragraphs but we argue this highlights the novelty of Twitter as a data collection method. Many alternate approaches to using Twitter as a data collection method including approaches such as active recruitment, holding virtual focus groups and interviews, and collecting different data from tweets have yet to be utilized. Likewise, Twitter also deviates from traditional methods of data collection (i.e., in-person focus groups, interviews) particularly in the areas of participant recruitment, how data are collected, how data are verified, ethical considerations, and accessibility of research, which results in new considerations for health research and for researchers interested in using social media to enhance their work.

Participant Recruitment

Data collection using Twitter mostly focuses on passively collecting tweets from conversations that do not involve active participant recruitment (i.e., informed consent) into a study. Of the included articles, 16 studies conducted data mining, where new information that may previously be hidden is sought from large changing data sets, through Twitter searches (Barbier & Lui, 2011). Only one study actively recruited and consented participants and collected data through a preorganized, planned Twitter conversation (or chat; Parsons et al., 2015). This demonstrates that health research via Twitter does not typically involve active recruitment, and the typical recruitment technique involves tweeting to large audiences unaware that they are participating in research, termed covert research. Arguments for covert research focus on concerns of bias when participants are made aware of the researcher's professional identity, which affects the validity of observations and data (Parker & Crabtree, 2014). Opponents of covert research argue that it is a form of deception that eliminates participants' choice and that lack of transparency can lead to less than ideal research circumstances, which may cause stress and harm to participants and decrease credibility of future researchers (Parker & Crabtree, 2014). Active or overt participant recruitment on Twitter may involve contacting participants individually and conducting the informed consent process based on research ethics approval process, which takes valuable time and resources; however, it is argued that trustworthy data can only be created through establishing a trusting relationship between the researcher and participants (Pitts & Mille-Day, 2007).

Data Collection Methods

Sixteen studies of the included studies employed data mining techniques using Twitter's advanced search and stream functions and Twitter's application programming interface. These functions posed significant challenges that resulted in missing and/or excluding important data from their respective studies. For example, key words searched were not inclusive of all related word variations like alternate names and slang terms (Ramagopalan et al., 2014). Search strategies did not filter or properly eliminate nonhealth-related tweets, advertisements, and repetitive words, which required manual screening by research teams (Henzell et al., 2014).

Strategies to address data collection challenges and increase rigor were also identified. Searches were sometimes conducted by individuals with expertise in searching and extracting tweets (e.g., software developer) to increase speed and accuracy. Users' profiles were also examined for context (Hewis, 2015) to clarify tweet content and increase accuracy. Studies also used predetermined criteria to aid decision-making, when deciding to include or exclude tweets as part of the data (Ahltward et al., 2014; Sugawara et al., 2012). While these novel solutions overcame some challenges, there were difficulties scaling these solutions to larger data sets using Twitter's search and stream functions. Larger data sets that used search and stream functions used other strategies including having two reviewers and by hosting multiperson coding sessions to ensure interrater reliability (Ahlwardt et al., 2014; Nascimento et al., 2014).

Verifying Information

Using Twitter in health research, particularly in light of the passive and covert research methods typically used, means that researchers must trust that information provided in profiles and contained in tweets is accurate, since research wasn't the intended use of such information. For example, in some studies, key participant demographic data (i.e., age, gender, geography) could not be confirmed

(Nakhasi et al., 2016; Nascimento et al., 2014); therefore, researchers defined their population of interest by health concerns/interests (e.g., mining for key words like pain or cancer). Studies may also enforce rigor by contacting participants or reviewing their Twitter feed to verify inclusion or exclusion criteria (Hewis, 2015). However, this may be time-consuming and raise ethical concerns. Additional strategies like triangulation, thick description, member checks, and development of a coding system and interrater reliability can be used; however, Morse (2015) cautions that these may only be used with particular methods. An important limitation of collecting data covertly on Twitter is that data are difficult to verify and methods should be built into the research design to overcome this challenge.

Ethical Reconsiderations

Twitter is a public forum and all studies included in this review utilized public tweets; however, access to these data for research purposes is not well-defined in Twitter's user agreement (Hewis, 2015) and it is unclear whether Twitter users would actively consent and participate in the same research if they knew tweets were collected as research data. Because of this, increased transparency is warranted.

Accessibility

Although using Twitter in the context of health research poses several challenges, it also has the unique advantage of accessibility. Groups of individuals can quickly form conversations and discussions around shared interests and specific topics. Geographically dispersed participants can connect and jump in and out of a virtual space, which creates flexibility that is hard to achieve with in-person data collection. Twitter has 313,000,000 active users (Statista, 2017), which means that this method of data collection could reduce barriers to research participation based on geographical location of researchers and research resources. It can also maximize resources, including time, effort, and convenience.

Finally, Twitter is an emerging venue of qualitative data for researchers. Sixteen of the 17 articles screened contained qualitative elements (i.e., qualitative mixed-methods studies). The dominant use of Twitter for qualitative research demonstrates the potential of Twitter to provide researchers access and valuable insights about health-care consumers in 140-character tweets. Participants are able to overcome geographical and physical barriers, and researchers can obtain a comprehensive and representative group of participants for their studies. This alone makes Twitter a valuable resource for qualitative health researchers; however, caution must be taken in the design and implementation of health research studies to ensure transparency, trustworthy data, and authentic health-care consumer engagement.

Conclusion

Utilizing Twitter as a recruitment and data collection tool in health research remains largely unexplored. Currently, Twitter is predominantly used for passive and covert data collection, but there is

potential to gather data through more active and overt methods with some safeguards in place. Given Twitter's unique accessibility advantages, future research should explore these active and overt participant recruitment and data collection methods to determine best practices. Further methodological refinement is required to fully realize the potential of Twitter in health research.

Appendix A

Search Strategies

Search Strategies: Effectiveness of twitter as data collection method among health-care consumers. The following databases were searched on May 24, 2016:

- Ovid Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946to Present> - 272 items retrieved
- Ovid Embase<1974 to 2016 May 23> - 240 items retrieved
- Ovid PsycINFO<2002 to May Week 3 2016>154 items retrieved
- Ovid EBM Reviews—Cochrane Central Register of Controlled Trials April 2016—11 items retrieved
- EBSCO CINAHL Plus with Full-text, 1937-Current – 133 items retrieved
- Elsevier Scopus, 1960-Current—311 items retrieved
- Thomson Reuters Web of Science Core Collection, 1900-Current—344 items retrieved

The total number of items retrieved from the database searches was 1455; 618 items remained after the duplicates were removed.

Epub Ahead of Print, In-Process and Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

Search Strategy:

1. (twitter* or tweet*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
2. Patient Satisfaction/ or Consumer Behavior/ or Consumer Health Information/ or Consumer Advocacy/ or exp ConsumerParticipation/
3. exp Patients/
4. (patient* or client* or caregiver* or carer* or consumer* or ((healthcare or care or health) adj3 recipient*) or ((research or study or studies or trial*) adj3 (participant* or recruit*))).mp.
5. 2 or 3 or 4
6. 1 and 5
7. limit 6 to english language
8. tweetable abstract.ab.

9. 7 not 8
10. social media/ or (social media or twitter).mp.
11. 8 and 10
12. 9 or 11
13. limit 12 to yr="2001 -Current"
14. limit 13 to news
15. 13 not 14

Embase 1974 to 2016 May 23

Search Strategy:

1. (twitter* or tweet*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]
2. (patient* or client* or caregiver* or carer* or consumer* or ((healthcare or care or health) adj3 recipient*) or ((research or study or studies or trial*) adj3 (participant* or recruit*))).mp.
3. 1 and 2
4. limit 3 to (conference abstract or conference paper or conference proceeding or "conference review")
5. 3 not 4
6. limit 5 to english language
7. tweetable abstract.ab.
8. 6 not 7
9. social media/ or (social media or twitter).mp.
10. 7 and 9
11. 8 or 10

PsycINFO 2002 to May Week 3 2016

Search Strategy:

1. (twitter* or tweet*).mp.
2. client participation/ or exp patients/
3. (patient* or client* or caregiver* or carer* or consumer* or ((healthcare or care or health) adj3 recipient*) or ((research or study or studies or trial*) adj3 (participant* or recruit*))).mp.
4. 2 or 3
5. 1 and 4
6. limit 5 to english language
7. limit 6 to ("0100 journal" or "0400 dissertation abstract")

EBM Reviews - Cochrane Central Register of Controlled Trials April 2016

Search Strategy:

1. (twitter* or tweet*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary

concept word, rare disease supplementary concept word, unique identifier]
2. Patient Satisfaction/ or Consumer Behavior/ or Consumer Health Information/ or Consumer Advocacy/ or exp ConsumerParticipation/
3. exp Patients/
4. (patient* or client* or caregiver* or carer* or consumer* or ((healthcare or care or health) adj3 recipient*) or ((research or study or studies or trial*) adj3 (participant* or recruit*))).mp.
5. 2 or 3 or 4
6. 1 and 5
7. limit 6 to english language
8. tweetable abstract.ab.
9. 7 not 8
10. social media/ or (social media or twitter).mp.
11. 8 and 10
12. 9 or 11
13. limit 12 to yr="2006 -Current"

EBSCO CINAHL Plus with Full-text

((MH "Consumer Health Information") OR (MH "Consumer Advocacy") OR (MH "Consumers") OR (MH "Consumer Participation") or patient* or client* or caregiver* or carer* or consumer* or (healthcare or care or health) n3 recipient* or (research or study or studies or trial*) n3 (participant* or recruit*)) AND (twitter* or tweet*)

Limit to English Language, Academic Journals, Dissertations and CEU publication. Brief Items excluded. No records published before 2009 were retrieved so no date limit applied.

Scopus

(TITLE-ABS-KEY(twitter* or tweet*) AND TITLE-ABS-KEY(patient* or caregiver* or carer* or ((healthcare or care or health) w/3 (recipient* or consumer*)) or (((consumer* or client*) w/3 (satisf* or participat* or advoca*)) and health*)) or (((research or study or studies or trial) w/3 (participant* or recruit*)) and health*)) AND NOT ABS("tweetable abstract") AND (LIMIT-TO(PUBYEAR,2016) OR LIMIT-TO(PUBYEAR,2015) OR LIMIT-TO(PUBYEAR,2014) OR LIMIT-TO(PUBYEAR,2013) OR LIMIT-TO(PUBYEAR,2012) OR LIMIT-TO(PUB-YEAR,2011) OR LIMIT-TO(PUBYEAR,2010) OR LIMIT-TO(PUBYEAR,2009) OR LIMIT-TO(PUBYEAR,2008)) AND (LIMIT-TO(LANGUAGE,"English")) AND (LIMIT-TO(SRCTYPE,"j"))

Web of Science Core Collection

twitter* or tweet*
AND
patient* or caregiver* or carer* or ((consumer* or client*) near/3 (satisf* or participat* or advoca*) and health*) or ((research or study or trial) near/3 (participant* or recruit*) and health*)

Appendix B

Table B1. Analysis and Overview of Full-Text Studies.

Study	Country	Study Design	Health Focus	Timing of Data Collection	Data Collection Approaches	Analytic Approach and Types of Data Collection
Beykikhoshk et al., 2014	Australia	/	Autism spectrum disorder	X	Tools: Twitter application programming interface (API) "search" and "stream" functions. Key words: "autism," "adhd," "autismawareness," "Asperger," and "aspie"	Generic quantitative linguistic analysis. Content analysis of Tweet content – Word frequency – Hashtag analysis – Part of speech analysis
Beykikhoshk et al., 2015	Australia	/	Autism spectrum disorder		Tools: Twitter API (application programming interface) "search" and "stream" functions. Key words: "autism," "adhd," "Asperger," and "aspie."	
Hewis, 2015	Australia	*	MRI patient's perspectives	X	Tools: Advanced Twitter search functions. Key words: "MRI" or "magnetic resonance imaging"	In-depth qualitative content analysis – Manual coding. Thematic analysis – Emerged themes identified from coded tweets
Love et al., 2013	United States	*	Tweets about vaccinations	X	Tools: Network analysis tool Node XL(social Media research Foundation, Belmont, CA) to collect tweets. Key words: Words related to vaccination: "vaccine," "vaccination," and "immunization"	Content analysis – coded for key variables
Ramagopalan et al., 2014	UK	/	Multiple sclerosis	X	Tools: Topsy (houses all tweets since 2006). Key words: Searched names of MS treatments e.g., "Aubagio" "Rebif" and "Tysabri"	Mixed Methods. Sentiment and word frequency analysis – Tweets grouped into monthly time periods using twitter R package and Jeffrey Breen's sentiment analysis code – Word frequency analysis was completed using TagCrowd. Statistical analysis – Using lists of positive and negative words sentiment score for any tweet was calculated – Sentiment score = number of positive words − number of negative words
Tsuya et al., 2014	Japan	*	Cancer (breast, leukemia, colon cancer, rectal cancer, colorectal, uterine, stomach, lung, and ovarian)	X	Tools: Twitter API and data mining method Searched every publicly available Twitter profile in Japan for types of cancer. e.g., Used application programming interface and data mining method	Qualitative content analysis of tweets Tweets were broken down into morphemes ("words") using Japanese language morpheme analysis software ChaSen (from the Nara Institute of Science and Technology, Japan). Words are in their original forms. Noun groups were extracted. Generation of co-occurrence networks(relations between key words that appear together in each tweet)

Table B1. (continued)

Study	Country	Study Design	Health Focus	Timing of Data Collection	Data Collection Approaches	Analytic Approach and Types of Data Collection
Ahlwardt et al., 2014	United States	/	Pain (toothache, backache, earache, and headache)	o	Tools: Twitter search APIs Searched anatomical words and pain words	Qualitative content analyses and quantitative analysis – Word frequency calculator – Codebook development – Used descriptive statistics
Greaves et al., 2014	UK	/	Tweets to the hospitals in the English National Health Service about quality of care	o	Tools: Twitter API "streaming" (API) Collected tweets aimed at NHS hospitals by using "mentions" for example @username. Searched all NHS hospital organization names on Twitter	Qualitative directed content analysis – Coded thematically – Sentiment analysis – Codebook development
Henzell et al., 2014	New Zealand	*	Orthodontics	o	Tools: Twitter's advanced-search function. Used to limit posts to New Zealand only. Key words: "Braces," "orthodontist," and "orthodontics."	Qualitative discourse analysis – Meanings – Position – Underlying assumptions found in text
Marton, 2012	Canada	/	Cancer patient tweets from Canadian Cancer Society	o	Tools: Visited the Twitter feed of the Canadian Cancer Society	Qualitative Statistical analysis of several attributes: – Frequency of popular words – Key Twitter attributes – Calculated frequency of Twitter feed attributes for each day of the month
McNeil et al., 2011	Canada	/	Epilepsy	o	Tools: Twitter publicly available website Key words: "seizure," "seizures," "seize," "seizing," and "seizuring"	Prospective qualitative content analysis –Using seizure-related Twitter status updates.
Nagel et al., 2013	United States	–	Influenza and pertussis (explore interaction between cyberspace messages and real world occurrences of)	o	Tools: geo-search-enabled Twitter tools created to operate with Twitter search and streaming application programming interfaces (APIs). Key words: "flu," "influenza," "pertussis," "whooping cough"	Quantitative "calculated correlation coefficients between tweets or subgroups of tweets and disease"
Nakhasi et al., 2016	United States	*	Patient safety	o	Tools: Twitter search application program interface Key words: search terms based on model key phrases to capture most representative expressions describing health providers, consulted phrases with a patient safety expert	Descriptive qualitative analysis – Source of tweet – Source of error – Type of error – Emotional response – Reporting percentages used to summarize data

(continued)

9

Table B1. (continued)

Study	Country	Study Design	Health Focus	Timing of Data Collection	Data Collection Approaches	Analytic Approach and Types of Data Collection
Nascimento et al., 2014	United States	*	Infodemiology of self-reported migraine headache suffering	O	Tools: Twitter public search engine. Key words: "migraine"	Qualitative Coding categories: – Migraine headache – Commercial – Metaphor – Not related – Betweet – Third person's – Discussion – Blanks – Inconclusive – Analyzed time and date using GMT
Parsons et al., 2015	United Kingdom	/	Pain tracking	O	Participants were recruited through electronic advertising. Participants were asked to tweet at three set local times (8 a.m., 12 p.m., and 10 p.m.). Data extracted manually	Qualitative. Data for analysis were accessing the study's Twitter account archive, and copied to protected Microsoft Excel document
Shive et al., 2013	United States	*	Acne	O	Tools: Twitter Streaming Application Programming Interface (API). Keywords: "pimple, pimples, zit, zits, and acne."	Qualitative – Defined and analyzed high impact tweets – Word frequency comparisons – Content of tweets compared to acne patient information website published by the American Academy of Dermatology (AAD)
Sugawara et al., 2012	Japan	*	Twitter usage and its role in "wired" cancer patients.	O	Tools: "16 (one-six) Profile Search β, mentionmapp, and Whotwi Key words: "breast cancer," "leukemia," "colon cancer," "rectal cancer," "colorectal cancer," "cancers of the uterus," "malignant lymphoma," "brain tumor," "stomach cancer," "lung cancer," "thyroid cancer," "ovary cancer," "kidney cancer," "prostate cancer," "esophagus cancer," "bladder cancer," "liver cancer," "oral cancer," "pharyngeal cancer," "gallbladder cancer," "cholangiocarcinoma," "laryngeal cancer," "skin cancer," and "multiple myeloma."	Qualitative Reviewed all Twitter user accounts with names of cancer described in user's Twitter profile. – Reviewed relationships between users – Reviewed user-generated Twitter content from 51 "power accounts"

Note. Study design: * = qualitative, - = quantitative, / = mixed methods. When data were collected X = retrospective, O = prospective.

Appendix C

Primary Screening Checklist

Guiding Screening Question: Does the study evaluate the use of Twitter as a method of data collection with health-care consumers in health research?

Yes = include
No = exclude

To answer "Yes" to the above guiding question, the study must meet the following parameters:

1. Is the following article about Twitter?.

Yes = move to next question
No = exclude study

2. Is the following article about health research?.

Yes = move to next question
No = exclude study

3. Study Written in English.

Yes = include study
No = exclude study

Secondary screening

Step 1—Verify Criteria: 10% screen of studies ($n = 618$); organized database by title and selected the top 618 articles.
Step 2—Determine if criteria adjustments are required based on 10% screen
Step 3—Revise secondary screening criteria as needed
Step 4—Screen remaining articles

Guiding Screening Question: Does the study include the use of Twitter as a method of data collection with health care consumers in health research?

Yes = include
No = exclude

To answer "Yes" to the above guiding question, the study must meet the following parameters:

1. Is only Twitter discussed? Other forms of social media (Facebook, Instagram), and blogs should not be mentioned in the study and should be excluded if mentioned. The focus is only Twitter.
   Yes = move to the next question
   No = exclude study
2. Is the use of Twitter as a data collection method discussed? The focus of data collection evaluation is "was Twitter used to collect information" not as dissemination or education.

Yes = move to next question
No = exclude study
3. Is the target audience for data collection a recipient of healthcare and/or a participant in the health decision-making process for/with a health-care recipient? Data collection must occur with health care consumer, not a health-care provider.
   Yes = move to next question
   No = exclude study
4. Do the research findings focus and discuss the use of Twitter as a data collection method sufficiently to add insight to this body of literature?

Yes = include study
No = exclude study

References

Ahlwardt, K., Heaivilin, N., Gibbs, J., Page, J., Gerbert, B., & Tsoh, J. Y. (2014). Tweeting about pain. *Journal of the American Dental Association (JADA)*, *145*, 737–743. doi:10.14219/jada.2014.30

Arksey, H., & O'malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*, 19–32. doi:10.1080/1364557032000119616

Barbier, G., & Liu, H. (2011). Data mining in social media. *Social network data analytics* (pp. 327–352). United States: Springer.

Beykikhoshk, A., Arandjelović, O., Phung, D., Venkatesh, S., & Caelli, T. (2014, August). Data-mining Twitter and the autism spectrum disorder: A pilot study. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, *2014* (pp. 349–356). doi:10.1109/ASONAM.2014.6921609

Beykikhoshk, A., Arandjelović, O., Phung, D., Venkatesh, S., & Caelli, T. (2015). Using Twitter to learn about the autism community. *Social Network Analysis and Mining*, *5*, 1–17. doi:10.1007/s13278-015-0261-5

Clinical Digest. (2014). Nurses find Twitter a useful recruitment tool for research in hard-to-reach groups. *Journal of Advanced Nursing*, *28*, 16–17. doi:10.1111/jan.12222

Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., & Moher, D. (2014). Scoping reviews: Time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology*, *67*, 1291–1294. doi:10.1016/j.jclinepi.2014.03.013

Emma Hilton, C. (2017). Unveiling self-harm behaviour: What can social media site Twitter tell us about self-harm? A qualitative exploration. *Journal of Clinical Nursing*, *26*, 1690–1704. doi:10.1111/jocn.13575

Eysenbach, G. (2009). Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *Journal of Medical Internet Research*, *11*. doi:10.2196/jmir.115

Finfgeld-Connett, D. (2015). Twitter and health science research. *Western Journal of Nursing Research*, *37*, 1269–1283 15p. doi: 10.1177/0193945914565056

Fisher, J., & Clayton, M. (2012). Who gives a tweet: Assessing patients' interest in the use of social media for health care. *Worldviews on Evidence-Based Nursing*, *9*, 100–108 9p. doi:10.1111/j.1741-6787.2012.00243.x

GraphPadQuickCalcs: Quantify interrater agreement with kappa. Retrieved from: https://www.graphpad.com/quickcalcs/kappa1.cfm. (Accessed 23 November, 2016).

Greaves, F., Laverty, A. A., Ramirez Cano, D., Moilanen, K., Pulman, S., Darzi, A., & Millett, C. (2014). Tweets about hospital quality: A mixed methods study. *BMJ Quality & Safety*, *23*, 838–846. doi:10.1136/bmjqs-2014-002875

Henzell, R., Knight, M., Morgaine, K. C., Antoun, J., & Farella, M. (2014). A qualitative analysis of orthodontic-related posts on Twitter. *The Angle Orthodontist*, *84*, 203–207. doi:10.2319/051013-355.1

Hewis, J. (2015). Do MRI patients tweet? Thematic analysis of patient tweets about their MRI experience. *Journal of Medical Imaging & Radiation Sciences*, *46*, 396–402. doi:10.1016/j.jmir.2015.08.003

Holt, C. (2011). Emerging technologies: Web 2.0. *Health Information Management Journal*, *40*, 33–35.

Kilaru, A. S., Meisel, Z. F., Paciotti, B., Ha, Y. P., Smith, R. J., Ranard, B. L., & Merchant, R. M. (2016). What do patients say about emergency departments in online reviews? A qualitative study. *BMJ Quality & Safety*, *25*, 14–24. doi:10.1136/bmjqs-2015-004035

Love, B., Himelboim, I., Holton, A., & Stewart, K. (2013). Twitter as a source of vaccination information: Content drivers and what they are saying. *American Journal of Infection Control*, *41*, 568–570. doi:10.1016/j.ajic.2012.10.016

Lyles, C. R., López, A., Pasick, R., & Sarkar, U. (2013). "5 mins of uncomfyness is better than dealing with cancer 4 a lifetime": An exploratory qualitative analysis of cervical and breast cancer screening dialogue on Twitter. *Journal of Cancer Education*, *28*, 127–133. doi:10.1007/s13187-012-0432-2

Marton, C. (2012). Consumer health 2.0 in Canada: Tweeting about cancer. *Journal of the Canadian Health Libraries Association (JCHLA)*, *33*, 98–102.

McNeil, K., Brna, P. M., & Gordon, K. E. (2011). Epilepsy in the Twitter era: A need to re-tweet the way we think a about seizures. *Epilepsy & Behavior*, *23*, 127–130. doi:10.1016/j.yebeh.2011.10.020

Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., & Rabb, D. (2012). The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care*, *28*, 138–144. doi:10.1017/S0266462312000086

Morse, J. M. (2015). Critical analysis of strategies for determining rigor in qualitative inquiry. *Qualitative Health Research*, *25*, 1212–1222. doi:10.1177/1049732315588501

Nagel, A. C., Tsou, M., Spitzberg, B. H., An, L., Gawron, J. M., Gupta, D. K., . . . Sawyer, M. H. (2013). The complex relationship of realspace events and messages in cyberspace: Case study of influenza and pertussis using tweets. *Journal of Medical Internet Research*, *15*, e237. doi:10.2196/jmir.2705

Nakhasi, A., Bell, S. G., Passarella, R. J., Paul, M. J., Dredze, M., & Pronovost, P. J. (2016). The potential of Twitter as a data source for patient safety. *Journal of Patient Safety*. doi:10.1097/PTS.0000000000000253

Nascimento, T. D., DosSantos, M. F., Danciu, T., DeBoer, M., van Holsbeeck, H., Lucas, S. R., & DaSilva, A. F. (2014). Real-time sharing and expression of migraine headache suffering on Twitter: A cross-sectional infodemiology study. *Journal of Medical Internet Research*, *16*, e96. doi:10.2196/jmir.3265

O'Connor, A., Jackson, L., Goldsmith, L., & Skirton, H. (2014). Can I get a retweetplease? Health research recruitment and the Twittersphere. *Journal of Advanced Nursing*, *70*, 599–609. doi:10.1111/jan.12222

Parker, J., & Ashencaen Crabtree, S. (2014). Covert research and adult protection and safeguarding: An ethical dilemma? *The Journal of Adult Protection*, *16*, 29–40. doi:10.1108/JAP-07-2013-0029

Parker-Pope, T. (2009, July 8). Tweeting from the operating room. *The New York Times*. Retrieved from: https://well.blogs.nytimes.com/2009/07/09/tweeting-from-the-operating-room/.

Parsons, C. F., Breckons, M., & Durham, J. (2015). Twitter: A viable medium for daily pain diaries in chronic orofacial pain? *British Dental Journal*, *219*, 75–78. doi:10.1038/sj.bdj.2015.558

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, *5*, 371–385. doi:10.1002/jrsm.1123

Pitts, M., & Miller-Day, M. (2007). Upward turning points and positive rapport-development across time in researcher-participant relationships. *Qualitative Research*, *7*, 177–201. doi:10.1177/1468794107071409

Ramagopalan, S., Wasiak, R., & Cox, A. P. (2014). Using Twitter to investigate opinions about multiple sclerosis treatments: A descriptive, exploratory study. *F1000Research*, *3*. doi:10.12688/f1000research.5263.1

Robillard, J., Johnson, T., Hennessey, C., Beattie, B., & Iles, J. (2013). Aging 2.0: Health information about dementia on Twitter. *PLoS One*, *8*, e69861. doi:10.1371/journal.pone.0069861.g001

Salzmann-Erikson, M. (2017). Virtual communication about psychiatric intensive care units: Social actor representatives claim space on Twitter. *International Journal of Mental Health Nursing*, *26*, 366–374. doi:10.1111/inm.12253

Shive, M., Bhatt, M., Cantino, A., Kvedar, J., & Jethwani, K. (2013). Perspectives on acne: What Twitter can teach health-care providers. *JAMA Dermatology*, 149, 621–622. doi:10.1001/jamadermatol.2013.248

Statista. (2017). Graph illustration of number of monthly active Twitter users worldwide from 1st quarter 2010 to 2nd quarter 2016 (in millions), November 8, 2016. Twitter: number of monthly active users 2010–2017. Retrieved from: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/.

Stewart, L. A., Clarke, M., Rovers, M., Riley, R. D., Simmonds, M., Stewart, G., & Tierney, J. F. (2015). Preferred reporting items for systematic review and meta-analyses of individual participant data: The PRISMA-IPD Statement. *JAMA: Journal of the American Medical Association*, 313, 1657–1665. doi:10.1001/jama.2015.3656

Stump, T., Zilch, S., & Coustasse, A. (2012). The emergence and potential impact of medicine 2.0 in the health-care industry. *Hospital Topics*, 90, 33–38. doi:10.1080/00185868.2012.679909

Sugawara, Y., Narimatsu, H., Hozawa, A., Shao, L., Otani, K., & Fukao, A. (2012). Cancer patients on Twitter: A novel patient community on social media. *BMC Research Notes*, 5, 699. doi:10.1186/1756-0500-5-699

Tsuya, A., Sugawara, Y., Tanaka, A., & Narimatsu, H. (2014). Do cancer patients tweet? Examining the twitter use of cancer patients in Japan. *Journal of Medical Internet Research*, 16, e137. doi:10.2196/jmir.3298

Wittmeier, K., Holland, C., Hobbs-Murison, K., Crawford, E., Beauchamp, C., Milne, B., & Keijzer, R. (2014). Analysis of a parent-initiated social media campaign for Hirschsprung's disease. *Journal of Medical Internet Research*, 16, e288. Retrieved from http://doi.org/10.2196/jmir.3200