

BMJ Open Scoping review protocol on the use of social media for health research purposes

Charline Bour,¹ Susanne Schmitz,² Adrian Ahne,^{3,4} Camille Perchoux,⁵ Coralie Dessenne,¹ Guy Fagherazzi ¹

To cite: Bour C, Schmitz S, Ahne A, *et al.* Scoping review protocol on the use of social media for health research purposes. *BMJ Open* 2021;**11**:e040671. doi:10.1136/bmjopen-2020-040671

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-040671>).

Received 20 May 2020

Revised 27 October 2020

Accepted 21 January 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Population Health, Luxembourg Institute of Health, Strassen, Luxembourg

²Department of Population Health, Competence Center for Methodology and Statistics, Luxembourg Institute of Health, Strassen, Luxembourg

³Center for Research in Epidemiology and Population Health (CESP), Inserm U1018, Villejuif, France

⁴Epiconcept, Paris, France

⁵Urban Development and Mobility, Luxembourg Institute of Socio-Economic Research (LISER), Esch-sur-Alzette, Luxembourg

Correspondence to

Dr Guy Fagherazzi; guy.fagherazzi@gustaveroussy.fr

ABSTRACT

Introduction More than one-third of the world population uses at least one form of social media. Since their advent in 2005, health-oriented research based on social media data has largely increased as discussions about health issues are broadly shared online and generate a large amount of health-related data. The objective of this scoping review is to provide an evidence map of the various uses of social media for health research purposes, their fields of applications and their analysis methods.

Methods and analysis This scoping review will follow the Arksey and O'Malley methodological framework (2005) as well as the Joanna Briggs Institute Reviewer's manual. Relevant publications will be first searched on the PubMed/MEDLINE database and then on Web of Science. We will focus on literature published between January 2005 and April 2020. All articles related to the use of social media or networks for health-oriented research purposes will be included. A first search will be conducted with some keywords in order to identify relevant articles. After identifying the research strategy, a two-part study selection process will be systematically applied by two reviewers. The first part consists of screening titles and abstracts found, thanks to the search strategy, to define the eligibility of each article. In the second part, the full texts will be screened and only relevant articles will be kept. Data will finally be extracted, collated and charted to summarise all the relevant methods, outcomes and key findings in the articles.

Ethics and dissemination This scoping review will provide an extensive overview of the use of social media for health research purposes. Opportunities as well as future ethical, methodological and technical challenges will also be discussed based on our findings to define a new research agenda. Results will be disseminated through a peer-reviewed publication.

INTRODUCTION

Social media (SM) are interactive 'mobile and web-based technologies' which allow discussion, creation and sharing of information between individuals, online communities and networks.¹ General platforms such as Facebook, Twitter and YouTube have emerged around 2004–2006 and many others since. SM are now increasingly used by a large proportion of the global population, estimated to 2.61 billion users worldwide in 2018.^{2,3} To date, the most popular SM platform

Strengths and limitations of this study

- This will be the first scoping review about the overall use of social media for health research purposes.
- The evolution of social media interest in health research, the different cases and fields of application of social media use for research purposes and the different methodologies for social media data analysis will be discussed in the review.
- Our scoping review will conform to the rigorous methodology manual by the Joanna Briggs Institute.
- The identification and synthesis of data will be limited to published articles found on the PubMed/MEDLINE and Web of Science databases and snowball references.
- Because the present scope is focused on health research, other major uses of social media by patients, associations, organisations and healthcare professionals will not be included as such, but only put in perspective in the discussion.

is Facebook with more than 2.41 billion monthly active users in 2019.^{2,4} In 2018, the average time spent by users daily on SM is about 142 min, while it was 90 min in 2012.⁵

Thus, the broad use of SM around the world offers numerous applications. SM users continuously generate large amounts of data that can, for instance, be studied in the political, business or even policy contexts.⁶ Most importantly, data generated by SM (1) are of high potential for medical research purposes,^{6–8} (2) can help healthcare professionals and scientists to keep being informed about the latest scientific discoveries or remotely follow medical conferences,^{9,10} and (3) can reshape the way patients interact with their peers and exchange health-related information and tips to manage their disease.^{11,12} For physicians, SM can improve their knowledge and abilities as well as their interactions with patients.¹³ It has also been shown that, somehow, people use SM to fulfil the need to belong to one or several social groups, reflecting our primary biological needs and survival instinct.¹⁴ People can interact with their friends,

family and audiences of potentially unlimited sizes. Hence, patients can easily interact with their peers on SM about their conditions, search for support or even try to sensibilise others with prevention and storytimes.^{15 16} Such digital space with no obvious hierarchy between users opens the door to new discourses as well as access and sharing of medical information about the patient's health, feelings and symptoms, that would have been impossible to collect in a face-to-face setting with a physician or research investigator.

In 2010 in the USA, 80% of adults used the internet to search for health-related information and 11% of SM users posted comments, queries or information about health or medical content.¹⁷ It is possible to join virtual communities, to participate in research, to receive moral support and to track personal progress.⁹ Such actions generate data that can be used notably in health research. 'Health research' refers to all kinds of research performed to learn more about human health, prevent or treat disease, test ideas, improve treatments and answer questions.^{18 19} Among all subfields of health and medical research, epidemiology and public health are the two most important disciplines that can potentially benefit from the use of SM. 'Infodemiology' is an early 2000s term²⁰ which describes a new approach for public health based on Big Data monitoring.^{21 22} Public health, as the science of improving, protecting the health and the well-being of people and communities from a population-level perspective, can directly and easily benefit from accessing large datasets of health-related information on large samples.¹⁸ Researchers can recruit study participants on SM^{23 24} to collect data²⁵ and to disseminate research.²⁶ Moreover, tracking health-related, treatment-related and feelings-related posts or discussions on SM can develop new methods to improve healthcare.²⁷⁻³⁰ Not only have SM improved researchers' communication with individuals and peers, but it also has a high potential to improve their research (eg, collecting data, understanding public perceptions) and their impact.^{31 32} Still, using SM for research may raise ethical issues such as getting consent of online users, protecting users' privacy or preserving anonymity of study participants.^{33 34}

PROTOCOL DESIGN

This scoping review will follow the methodological framework introduced by Arksey and O'Malley in 2005³⁵ and the methodology manual published by the Joanna Briggs Institute for scoping reviews.³⁶ The present protocol and future corresponding scoping review are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Review guidelines.³⁷ Thus,

this review will follow five of these six stages: (1) identification of the research question; (2) identification of relevant studies; (3) selection of eligible studies; (4) charting the data; and (5) collating, summarising and reporting of the results. There is an optional stage 6 (consultation with stakeholders) in order to identify additional references about potential studies to include and to collect feedback about the findings uncovered by the review, but we will not include it because of time constraint.

Stage 1: identification of the research questions

Through consultation with the clinical research team, the overall research questions are:

1. How SM have modified or complemented traditional health research?
2. What are the different fields of application of this approach?
3. What are the different methodologies for SM data analysis?

Stage 2: identifying relevant studies

This review will use the PCC (Population, Concept, Context) framework suggested by the Joanna Briggs Institute.³⁶ We will base our search strategy on the PCC framework described on table 1.

For the scoping review, we do not have any restriction on the population of interest, we will take any relevant publications regardless of the age, the origin or the gender of the studied populations. The concept is the use of SM. We are looking for any potential benefits related to the use of SM, such as using the online available data or the features developed by SM. Lastly, both these elements have to be linked with health research.

The databases chosen for this review are PubMed/MEDLINE and Web of Science. An initial exploratory search strategy based on the PCC framework will be developed on PubMed to determine some relevant terms and articles. Database and other searches will combine terms from two themes: SM (eg, Twitter, Facebook) and health research (eg, medicine). The Medical Subject Headings (MeSH) terms will be screened, sorted by pertinence and frequency.

A second search strategy will be developed thanks to the most relevant MeSH terms. Some keywords will be searched both in the title, abstract and subject headings (eg, MeSH) on PubMed and as topics on Web of Science. Other terms such as "Humans" and "Clinical trial" might further be used as filters. We will focus on articles published in English between January 2005 and April 2020. The pilot search strategy is shown in online supplemental appendix A. Lastly, reference

Table 1 PCC framework of our scoping review

PCC element	Definition	Example
Population	All humans (no restrictions)	NA
Concept	Use of SM	Extracting Twitter data and metadata related to a specific keyword of interest
Context	Health research	Public health

NA, not applicable; SM, social media.

lists from the retrieved reviews on related topics will be used as an additional source for snowball searching for additional articles.

Stage 3: selection of eligible studies

All papers derived from the search process will be uploaded to EndNote in order to remove all duplicates. Then, a two-step screening will be performed. The first part consists of screening titles and abstracts, thanks to the research strategy, in order to define the eligibility of each article. Publications with title or abstract not meeting the eligibility criteria will be excluded. During the second part, the full texts having passed the first step will be screened and only relevant articles will be kept. The remaining ones will get full text screened. Screening will be conducted with CADIMA,³⁸ a free web tool to facilitate the conduct and the documentation of literature reviews.³⁹ Two reviewers will screen every article independently and consistency checks will be performed. In case of inconsistency, CADIMA will display the rating differences and prompt each reviewer to review the article a second time. In case of disagreement, both reviewers will discuss the relevance of the article to decide if it should be included or not.

Studies will be included if they describe the use of SM for health or medical research purposes. Articles will be excluded if they deal with the use of SM among patients, patient associations, organisations and healthcare professionals for their day-to-day practice. Studies about non-human subjects and grey literature will be excluded as well. Papers will be excluded if not one of the following: clinical study, journal article, letter or observational study. This exclusion criteria might change depending on the relevance of the studies.

Stage 4: charting the data

Still using CADIMA, two independent reviewers will conduct this process. First, relevant studies will be selected from all the remaining papers in order to develop agreement on what information should be extracted. We will focus on the different fields of application of SM use by health researchers as well as the developed tools to achieve data collection and analysis. Then, data extraction will be performed after defining critical appraisal criteria and results will be stored in a table. The data extraction table produced will include at least the following key elements:

1. Author(s).
2. Year of publication.
3. Origin/country of origin.
4. Aims/purpose.
5. Type of study.
6. Studied population(s) (eg, young adults).
7. Type of SM studied.
8. Methodology/methods.
9. Outcomes and details of these (eg, symptoms surveillance, medical concepts).
10. Key findings that relate to the scoping review questions (eg, tools used or developed, quality of SM use domains).

Stage 5: collating, summarising and reporting of the results

The purpose of this scoping review is to collect the findings and present an overview of the research rather than to evaluate the quality of the studies. As a result, our overall assessment of the strength of the evidence will be narrative instead of quantitative. The results of the previous stages will be synthesised to describe the progress of research, thanks to SM from 2005 to 2020, all the research fields where SM are helpful, and the methods to collect and analyse data. The PCC inclusion criteria will guide the map of the data. Thus, at least two tables will be carried out to introduce the data. The first table will be a bubble plot describing the number of research publications published per year on PubMed from 2005 to 2019 considering first, SM in their totality and then specific SM (eg, Twitter, Facebook). The second one will summarise the different approaches to collect SM data and the developed processes to investigate it. A descriptive summary will accompany the tabulated results and describe how the results apply to our scoping review questions. Results will then be classified into categories depending on the research field they link to.

Patient and public involvement

No patient involved.

ETHICS AND DISSEMINATION

Results of this scoping review will provide an overview of all the applications in health research of the use of SM. Thus, it will be informative for various stakeholders: researchers, data scientists, public health agencies and governments will easily capture the big picture of the field, the different SM uses and methodologies for health research, and have an extensive presentation of the benefits, usefulness and potential of SM. Ethical issues will also be outlined as they remain fundamental in health research. In terms of dissemination activities, the scoping review will be submitted for publication in a scientific journal. Overall, it will help future researchers to better shape their future projects using SM data or for other researchers to consider this source of information as a valuable option to answer their research question.

Since the scoping review methodology consists of reviewing and collecting data from publicly available materials, this study does not require ethics approval.

Twitter Adrian Ahne @Adrahne

Contributors Design of the protocol—CB, SS, CD and GF. Draft of the manuscript—GF and CB. Review and final approval of the manuscript—CB, SS, GF, AA and CP.

Funding This work is supported by the Luxembourg Institute of Health.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability

of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Guy Fagherazzi <http://orcid.org/0000-0001-5033-5966>

REFERENCES

- Kietzmann JH, Hermkens K, McCarthy IP, *et al*. Social media? get serious! understanding the functional building blocks of social media. *Bus Horiz* 2011;54:241–51.
- McFadden C. A Chronological History of Social Media. In: interestengineering.com [Internet]. Available: <https://interestingengineering.com/a-chronological-history-of-social-media> [Accessed 16 Oct 2018].
- Ortiz-Ospina E. The rise of social media. In: ourworldindata.org [Internet]. Available: <https://ourworldindata.org/rise-of-social-media> [Accessed 18 Sep 2019].
- Clement J. Social media - Statistics & Facts. In: statista.com [Internet]. Available: <https://www-statista-com.proxy.bnl.lu/topics/1164/social-networks/?pds=74202092424141745339986600588747> [Accessed 4 Sep 2019].
- Clement J. Daily time spent on social networking by internet users worldwide from 2012 to 2019. In: statista.com [Internet]. Available: <https://www-statista-com.proxy.bnl.lu/statistics/433871/daily-social-media-usage-worldwide/?pds=84202085923141846464863258420388> [Accessed 26 Feb 2020].
- Conway M, O'Connor D, Media S. Social media, big data, and mental health: current advances and ethical implications. *Curr Opin Psychol* 2016;9:77–82.
- Thomas VL, Chavez M, Browne EN, *et al*. Instagram as a tool for study engagement and community building among adolescents: a social media pilot study. *Digit Health* 2020;6:205520762090454.
- Wetsman N. How Twitter is changing medical research. *Nat Med* 2020;26:11–13.
- Ventola CL. Social media and health care professionals: benefits, risks, and best practices. *P T* 2014;39:491–520.
- Collins K, Shiffman D, Rock J. How are scientists using social media in the workplace? *PLoS One* 2016;11:e0162680.
- Wilson S, Mogan S, Kaur K. Understanding the role of Facebook to support women with endometriosis: a Malaysian perspective. *Int J Nurs Pract* 2020;26:e12833.
- Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. *Health Info Libr J* 2017;34:268–83.
- Alanzi T, Al-Yami S. Physicians' attitude towards the use of social media for professional purposes in Saudi Arabia. *Int J Telemed Appl* 2019;2019:1–6.
- Baumeister RF, Leary MR. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol Bull* 1995;117:497–529.
- PubMed - NCBI. Connected motherhood: social support for Moms and Moms-to-Be on Facebook. Available: <https://www.ncbi.nlm.nih.gov/pubmed/25665177> [Accessed 8 Apr 2020].
- Shelby A, Ernst K. Story and science: how providers and parents can utilize storytelling to combat anti-vaccine misinformation. *Hum Vaccin Immunother* 2013;9:1795–801.
- Susannah F. The Social Life of Health Information, 2011. In: <https://www.pewresearch.org/internet/2011/05/12/the-social-life-of-health-information-2011/> [Accessed 12 May 2011].
- Harvard Countway Library. Participating in Health Research Studies. In: guides.library.harvard.edu/healthresearch [Internet]. 6 Mar 2020. Available: <https://guides.library.harvard.edu/healthresearch> [Accessed 6 Mar 2020].
- Institute of Medicine (US) Committee on Health Research and the Privacy of Health Information: The HIPAA Privacy Rule. Beyond the HIPAA privacy rule: enhancing privacy, improving health through research 2010.
- Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009;11:e11.
- Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. *J Med Internet Res* 2017;19:e228.
- Mavragani A. Infodemiology and Infoveillance: Scoping review. *J Med Internet Res* 2020;22:e16206.
- Sanchez C, Grzenda A, Varias A, *et al*. Social media recruitment for mental health research: a systematic review. *Compr Psychiatry* 2020;103:152197.
- Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. *J Med Internet Res* 2012;14:e28.
- Alshaikh F, Ramzan F, Rawaf S, *et al*. Social network sites as a mode to collect health data: a systematic review. *J Med Internet Res* 2014;16:e171.
- Breland JY, Quintiliani LM, Schneider KL, *et al*. Social media as a tool to increase the impact of public health research. *Am J Public Health* 2017;107:107.
- Park A, Conway M. Tracking health related discussions on Reddit for public health applications. *AMIA Annu Symp Proc* 2017;2017:1362–71.
- Tricco AC, Zarin W, Lillie E, *et al*. Utility of social media and crowd-intelligence data for pharmacovigilance: a scoping review. *BMC Med Inform Decis Mak* 2018;18:38.
- Adrover C, Bodnar T, Huang N, *et al*. Identifying adverse effects of HIV drug treatment and associated Sentiments using Twitter. *JMIR Public Health Surveill* 2015;1:e7.
- Sewalk KC, Tuli G, Hswen Y, *et al*. Using Twitter to examine web-based patient experience Sentiments in the United States: longitudinal study. *J Med Internet Res* 2018;20:e10043.
- Chelaru SV, Orellana-Rodriguez C, Altingovde IS. Can social features help learning to RANK YouTube Videos? *WISE* 2012:552–66.
- Thackeray R, Neiger BL, Smith AK, *et al*. Adoption and use of social media among public health departments. *BMC Public Health* 2012;12:242.
- Denecke K, Bamidis P, Bond C, *et al*. Ethical issues of social media usage in healthcare. *Yearb Med Inform* 2015;10:137.
- Hunter RF, Gough A, O'Kane N, *et al*. Ethical issues in social media research for public health. *Am J Public Health* 2018;108:343–8.
- Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8:19–32.
- Peters MDJ, Godfrey C, Mclnerney P. JBI Reviewer's manual. In: [wiki.joannabriggs.org](https://www.joannabriggs.org/display/MANUAL/Chapter+11%3A+Scoping+reviews) [Internet], 2019. Available: <https://www.joannabriggs.org/display/MANUAL/Chapter+11%3A+Scoping+reviews>
- Tricco AC, Lillie E, Zarin W, *et al*. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;169:467–73.
- CADIMA. CADIMA vers. 2.2.1. Available: <https://www.cadima.info/> [Accessed 13 May 2020].
- Kohl C, McIntosh EJ, Unger S. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. *Environ Evid* 2018.