Review

Research data management in health and biomedical citizen science: practices and prospects

Ann Borda, Kathleen Gray, and Yuqing Fu

Health and Biomedical Informatics Centre, Melbourne Medical School, The University of Melbourne, Melbourne, Australia

Corresponding Author: Ann Borda, Victorian Comprehensive Cancer Centre, Level 13, 305 Grattan St, Parkville, Victoria 3010, Australia; aborda@unimelb.edu.au

Received 1 March 2019; Revised 9 July 2019; Editorial Decision 12 September 2019; Accepted 30 September 2019

ABSTRACT

Background: Public engagement in health and biomedical research is being influenced by the paradigm of citizen science. However, conventional health and biomedical research relies on sophisticated research data management tools and methods. Considering these, what contribution can citizen science make in this field of research? How can it follow research protocols and produce reliable results?

Objective: The aim of this article is to analyze research data management practices in existing biomedical citizen science studies, so as to provide insights for members of the public and of the research community considering this approach to research.

Methods: A scoping review was conducted on this topic to determine data management characteristics of health and biomedical citizen science research. From this review and related web searching, we chose five online platforms and a specific research project associated with each, to understand their research data management approaches and enablers.

Results: Health and biomedical citizen science platforms and projects are diverse in terms of types of work with data and data management activities that in themselves may have scientific merit. However, consistent approaches in the use of research data management models or practices seem lacking, or at least are not prevalent in the review.

Conclusions: There is potential for important data collection and analysis activities to be opaque or irreproducible in health and biomedical citizen science initiatives without the implementation of a research data management model that is transparent and accessible to team members and to external audiences. This situation might be improved with participatory development of standards that can be applied to diverse projects and platforms, across the research data life cycle.

Key words: citizen science, crowdsourcing, participatory health, self-quantification, research data management

BACKGROUND

Citizen science

Public engagement in health and biomedical research is being influenced by the paradigm of “citizen science,” that is, active participation in research teams by members of the general public with no formal training in the field of research concerned. Different labels have been used to describe citizen science and related forms of public participation in scientific research, in which citizens may contribute actively to science with their intellectual effort, or their specialized knowledge, or their tools and resources. The online Oxford English Dictionary defines citizen science in the context of data activity: “The collection and analysis of data relating to the natural world by members of the general public, typically as part of a collaborative project with professional scientists.”

Citizen science dates back over a century in some fields of research - for example natural history, where the US Audubon
Society’s Christmas Bird Count began in 1900. The potential of citizens and experts working together to improve the development of science and policy gained prominence in such publications as Alan Irwin’s “Citizen Science: A Study of People, Expertise and Sustainable Development.” Citizen science activity has dramatically increased in the 21st century, influenced by societal and technological changes and participatory democracy. Critically it has enabled the large-scale collection and processing of scientific data and widespread dissemination of scientific knowledge and discoveries, notably in environmental sciences, ecology, and astronomy.

This is further supported by Internet access and connected communication devices which have been pivotal in making science more accessible to more people and which has, in turn, shaped motivational factors for participation. The concept and practice of citizen science has also become increasingly formalized through the establishment of citizen science organizations worldwide.

Citizen science projects can be differentiated according to the extent of responsibilities that citizens undertake for research activities, such as defining research questions, collecting and analyzing data, and interpreting and disseminating results. Other strategies classify citizen science projects according to their approach, such as investigative, online, or educational, or they can take further account of digital and distributed participation levels of “ownership.” Other new forms of citizen science participation are also arising such as “serious games” developed by professional scientists in which participants are engaged in problem-solving.

Data quality and management

Data management practices and tools in citizen science projects can take many forms but it is largely recognized that like any scientific project, public participation in research requires an understanding and implementation of appropriate data management and stewardship guidelines. The FAIR Data Principles: findable, accessible, interoperable, and reusable (www.go-fair.org/fair-principles) are among the guiding standards acknowledged as good data practice.

Ensuring data quality—for example, data of high quality, accuracy, consistency, and completeness to serve an intended purpose—remains a pervasive issue in data management. The challenge is primarily due to the fact that a lay public is generally considered to be untrained in scientific data management or research integrity, or may be prone to systematic errors which can impact data quality in processing and analysis tasks, for instance. Data quality can also be highly context dependent (i.e., “fitness for use”). The application of data quality methods in citizen science is most often dependent on the types of contributory work being carried out. These contributions can take the form of data collection (manual or automated) or data processing (e.g., classification, coding, and annotation). For example, data quality in data collection may be ensured by providing training or close supervision where feasible in a fieldwork context; and cross-checking for consistency with existing literature and/or with expert observations. One survey has examined mechanisms for data quality assurance and validation in the data collection process to inform a framework to apply to the design of citizen science activity, and taking into account the potential for error at both participant and protocol levels.

A high level of quality assurance is often associated with the use of “crowdsourcing”—a type of citizen science approach in which multiple people carry out the same work or task, such as contributing to peer review or replication of an analysis, for example, image identification. Such an approach is desirable across the sciences for validation, accuracy, and in reducing bias.

Whereas data quality mechanisms in the data collection and processing activities of citizen science have been fairly documented, the application of research data management practices to a whole project data lifecycle is much less in evidence. Published examples include the citizenscience.gov toolkit which contains general guidelines on data management for US federal agencies engaged in citizen science, and the citizen science initiative DataONE (www.dataone.org) guidelines which use a data lifecycle approach specific to that project’s goals and context. The largest data management study to date was conducted by the European Commission’s Joint Research Centre in 2015. Of 121 projects that responded, 84% of these represented areas of environmental research, and range from local neighborhood to multinational levels, showing the potential for citizen science to generate different scales of data. Over 60% of projects were reported to have data management plans in place, but there is a notable gap in understanding what these plans comprise. Data re-use licensing was a particular concern across the respondents. Some suggest that open data and open standards could better ensure reuse of project results.

As an overarching statement, it has been said that “the evidence that some citizen science programs produce high quality data of immediate use to science . . . does not translate into the conclusion that all citizen science programs can.” These are among the points which have led to the focus of the present paper.

Citizen science in health and biomedicine

Citizen science is advocated increasingly in health and biomedical research. The emerging connections between biotechnology and citizenship have added steadily to the bioscience research and policy discourse for two decades, for example, Rabinow’s coining of “biosociality” in 1996. Petryna’s account of biological citizens after Chernobyl and Rose and Novas’ “biological citizenship.” Reflecting the rise of consumer-oriented Internet platforms and connected tools for personal health data management, the idea of the so-called “participatory biocitizen” was proposed in 2012 as a way to realize personalized medicine by sharing life-logging and self-quantification data through social media channels. Such advocacy has translated slowly into peer-reviewed research outputs, not least hampered by differing definitional boundaries of what constitutes health and biomedical citizen science in the literature, and for the purposes of policy-making.

Benefits of citizen science in health and biomedical research are said to be reciprocal. Professional scientists may gain increased research capacity (e.g., the Cochrane Crowd initiative) and more accurate findings from incorporating lay people’s local knowledge or patients’ expertise. For public participants, the benefits may be improved scientific literacy, more empowered communities, more engaged policy and decision-making, and more opportunity for knowledge co-production in areas some areas of biomedical science, such as heritability.

Research data management

Proponents of citizen science in health and biomedicine argue that “If millions of biocitizens were streaming data to the cloud, they would build the most powerful dataset for preventive and precision medicine the world has ever known.” However, there is a large leap from volumes of data to standards of evidence in health.
credibility of conventional health and biomedical research relies on rigorous research data management tools and methods, such as the Harvard Biomedical Data Management and Lifecycle Model (data-management.hms.harvard.edu) that comprises cyclical steps of data creation, analysis, distribution (eg, sharing), retention, storage and secondary use, for example. Indeed, the improvement of research data management is a major focus of biomedical informatics.49,50

This article therefore examines the requirements for data management, and the approaches taken to it, in health and biomedical citizen science, to identify key considerations for engaged and intending citizens and biomedical researchers, and to pave the way for informatics advances that can improve this kind of research.

METHODS

To explore the prevalence and extent of use of research data management in health and biomedical citizen science, we conducted two studies. First a scoping review was undertaken encompassing literature reviews, conceptual articles, book chapters, and reports. First, a manual search was conducted by the authors which generated 75 papers. All items were read and relevant terms for a database search were extracted. This resulted in the following terms: citizen science, crowdsourcing, crowdfunding, participatory health, personalized medicine, and self-quantification. We further refined this result following a recent taxonomy of health research crowdsourcing behaviors,53 using the search strings: [participat* OR crowd* OR "citizen" OR self*] AND [scienc* OR *omic* OR *quantif*] AND ["medic*" OR health* OR wellness OR "well-being" OR wellbeing].

Five online databases (PubMed, Scopus, ACM Digital Library, IEEE Xplore, and Web of Science) were searched. The search was restricted to English language publications between 2012 when the term “participatory bio citizen” appeared in the literature41 to the end of 2018. By screening Title and Abstract fields, we excluded papers that addressed only nonhuman health research, such as veterinary and ecological sciences, and where there was no evidence of human health related outcomes. We further excluded papers addressing only low-level public participation as defined in general classifications of public participation in scientific studies.2,5,11,19 for example, where citizens had no other involvement in a research project beyond completing a questionnaire or being interviewed, or beyond having their pre-existing online social media data mined and analyzed by researchers. We additionally excluded papers focusing on microwork for monetary incentives. We retrieved the full text of all accessible papers and manually searched reference lists for additional works that our search may have missed. A total of 146 papers comprised our final result set. References #40–#186 represent the literature review results, and are further tabled below.

We analyzed the content of these papers to identify key characteristics of the research projects described in them that may shape their research data management practices.

Second, through the scoping review and additional Internet searching, we identified five indicative online platforms for health and biomedical citizen science and we selected projects supported by each platform. Indicative platforms had specific mentions in the literature review papers and in associated internet searches of health and biomedical citizen science projects supported by these platforms. The online accessibility of the platforms and of their hosted projects in terms of web site information available to participants was another consideration. We used deductive thematic analysis of information from the web sites to describe the research data management facilities afforded by these platforms and projects. Finally, we synthesized the findings from these two studies to identify strengths, weaknesses, opportunities and threats of research data management in health and biomedical citizen science. Methods and findings from the platforms and projects review study are further discussed following the literature review.

RESULTS AND DISCUSSION

Literature review

The terms “citizen science” and related participatory health research were most commonly used in connection with public health research; crowdsourcing, with cancer and genomic research and/or rare disease identification; and self-quantification were typically associated with genomic and personalized medicine research. The most prevalent descriptors of health and biomedical citizen science were used to group the result sets as listed in Table 1. In some papers, several descriptors were in use to describe health and biomedical citizen science approaches.

The availability and use of suitable MeSH terms was a limiting factor on our review. The MeSH term “Community-Based Participatory Research” (scoped as a “collaborative process of research involving researchers and community representatives”) was associated with only a few relevant papers;17,77,133,143,150 Community-based participatory research (CBPR) is often used for environmental health investigations and/or as an approach in which researchers work closely with the local community in developing and implementing research likely to be of concern to members of that community.47,138

In regard to conceptual or methodological characteristics described, papers showed no strong pattern. For instance, papers reported on research reporting protocols,186 human-computer interactions,129,179 platform design,23,76,110,131 and social empowerment policies.63,146 Also reported were works in progress,51,148 simulated patient-led n = 1 trials,123 comparative effectiveness research,95 and feasibility studies.80,149 One conclusive clinical trial was reported117 and several papers reported on citizen science approaches to clinical trial development.45,114,116,117,126,132,148 The open public and open knowledge philosophy of some citizen science initiatives mean that their descriptions may be published outside the scientific literature, for example, cases described in.40,96,135,142

We found 16 literature review papers.44,49–51,70,71,96,103,111,115,121,129,132,145,157,159 Most of these review papers focused on crowdsourced health and biomedical research, including gamification, while others covered citizen science and participatory health, and self-quantification. Platforms cited in the review papers include Zooniverse,70,96 PatientsLikeMe,51,70,96,121,157,159 23andMe,70,121 and initiatives such as Cancer Research UK’s Trailblazer,96 Mark2Care,49 uBiome,70 DIYbio,70 and the serious game, Foldit.70,82,96,121 None of the literature had a primary focus on research data management, although aspects of data management, such as data collection and processing, are described in examples further in the sections below.

Participatory roles

We found that the participatory roles of health and biomedical citizen scientists could be categorized in three main ways.40–42,51,70,82,96,121,132

The first is their number and spread, that is, whether public engagement in a project is one sole person, or a community in a localized setting or a global effort. The literature we reviewed reflected at
Table 1. Descriptors of health and biomedical citizen science papers extracted from title/abstract

| Descriptor(s)              | Paper references (repeating numbers represent papers with more than one descriptor) |
|---------------------------|-----------------------------------------------------------------------------------|
| Citizen science           | 42, 43, 44, 46, 55, 62, 63, 64, 66, 70, 75, 77, 78, 80, 81, 84, 85, 88, 90, 93, 94, 95, 98, 102, 103, 107, 108, 110, 113, 114, 117, 118, 119, 120, 121, 131, 136, 138, 143, 145, 146, 147, 148, 150, 154, 156, 158, 167, 168, 169, 184, 186 |
| Crowdsourcing             | 40, 41, 44, 52, 53, 54, 62, 68, 72, 77, 81, 86, 90, 96, 99, 100, 103, 111, 115, 116, 119, 122, 123, 124, 125, 130, 134, 135, 137, 140, 142, 144, 146, 149, 157, 158, 159, 160, 161, 164, 166, 168, 172, 173, 174, 177 |
| Participatory health      | 40, 41, 45, 47, 48, 51, 61, 69, 76, 78, 84, 85, 87, 89, 93, 104, 105, 106, 107, 109, 110, 112, 114, 119, 128, 132, 138, 139, 141, 145, 148, 153, 162, 165, 177, 181, 182, 183, 185 |
| Self-quantification       | 40, 41, 56, 57, 58, 59, 60, 65, 67, 71, 73, 74, 79, 82, 83, 91, 92, 97, 101, 119, 126, 127, 129, 133, 151, 152, 153, 161, 163, 170, 171, 175, 176, 177, 178, 179, 180 |

*Articles and reports are referenced by number and listed by this number with full citation in the References.

least 10 studies of each size, but about half of the papers gave no indication of the number of citizen scientists intended or involved. Participant characteristics were variably or minimally reported, and logistics of participation were not reported consistently.

The second is their status in the project as producers, instigators, and/or mobilisers (e.g., are they named as the chief investigators, is this a patient-driven investigation, or is this a school curriculum project, for instance). We found that most studies were instigated by professionally trained researchers; the roles of citizens and patients as co-producers were not fully described in most citizen science and crowdsourced research designs. Community scaled participatory research projects more often provided information on respective roles and responsibilities across research activities.

The third is the functions citizens carry out in the project, such as crowdfunding of resources, donation of data from self-tracking, self-reporting in participatory health, disease surveillance, and crowdsourcing of data analysis, among others. Crowdsourcing in particular was most commonly applied to the field of biomedical research and supporting analytical tasks, for example, image processing, sequence alignment, or molecular folding.

Crowdsourced data processing could further involve both lay people and those knowledgeable in a discipline, particularly where complex tasks (e.g., forms of annotation or relational tasks) or problem-solving are applied. Serious games and crowdsourced systematic literature review platforms were among these identifiable activities.

Health and biomedical citizen science further targeted health conditions where participants undertook self-reporting using mobile apps and wearable sensors. This included widespread conditions, such as cancer and mental illness, and specific conditions, such as Crohn’s disease and colitis. Additionally, research aimed to investigate a wide range of health conditions which can be informed by personal genomic data that people may generate through direct-to-consumer testing services.

Not all citizen science research used digital tools or online platforms, for example, Science café settings, but such technologies have widened the possibilities greatly.

Data management
Our analysis found three characteristics that were most likely to shape the research data management needs of such projects: research aims and objectives; participant roles and functions; research platforms and tools used.

Aims and objectives of public-health-oriented citizen science, for instance, included health service access, health literacy improvement, workplace health issues, environmental factors in human health, and identifying research priorities to optimize the design and delivery of patient-centered health services.

As outlined elsewhere in the paper, aspects of data management were closely associated with types of participatory roles and functions. Data collection and data processing were pervasive activities across the literature. Other specialized tasks included surveillance, monitoring, and problem-solving.

Among self-quantifying citizens, wearable devices were used commonly for personal data collection, for example, Fitbit and other fitness trackers, and sleep monitoring devices while other apps were designed for manual logging of personal health data, such as mood state or dietary intake and weight management.

In the present literature review, the citizen science platform Zooniverse (www.zooniverse.org) was the most cited generalizable platform supporting citizen science projects. Also prominent were platforms developed to support online communities of interest, for example, around personal genomics, the health social media platform PatientsLikeMe (www.patientslikeme.com), and the direct-to-consumer genomic service 23andMe (www.23andme.com). Table 2 provides a summary of literature review citations by platform and associated descriptors.

Other communities of practice were supported through wearable device manufacturers’ platforms and gaming platforms, to cite a few. Mainstream social media like Facebook, Twitter, and WhatsApp also were used for participant networking, data sharing, and task management.

Crowdsourcing and citizen science activities were most associated with quantitatively measurable tasks, whereas participatory health (especially CBPR) were primarily qualitative in design, and quantified self activity comprised mixed methods approaches to the objective biometric data and the subjective experience of the impact of these data.

Data quality and validation aspects of data collection and analysis were a focus in a number of papers, including articles reporting on the effect of training on accuracy, how various characteristics of participants affected their accuracy, and aspects of project design on
Table 2. Platforms cited in the literature review and associated descriptors

| Platform                  | Literature review citations | Associated descriptor(s) |
|---------------------------|----------------------------|--------------------------|
| Zooniverse www.zooniverse.org | 42, 43, 46, 94, 95, 119, 148, 149, 156, 158, 168, 184, 186 | Citizen science: 42, 43, 46, 94, 95, 119, 148, 156, 168, 184, 186  Crowdsourcing: 119, 149, 158, 168 |
| PatientsLikeMe www.patientslikeme.com | 40, 41, 42, 48, 52, 53, 56, 57, 58, 76, 94, 122, 133, 136, 140, 151, 158, 159, 160, 161, 175, 177 | Crowdsourcing: 40, 41, 52, 53, 122, 140, 158–159, 161, 177  Self-quantification: 56–58, 133, 151, 160, 161, 175, 177  Citizen science: 42, 94, 136,  Participatory health: 48 |
| 23andMe www.23andme.com | 40, 41, 42, 43, 53, 57, 59, 90, 94, 119, 133, 140, 159, 160, 161, 165, 166, 175 | Crowdsourcing: 40, 41, 53, 119, 140, 159, 160, 161, 166  Self-quantification: 57, 59, 133, 175  Citizen science: 42, 43, 90, 94, 119 |
| Scistarter https://scistarter.com/ | 108, 148, 159 | Citizen science:108, 148  Crowdsourcing: 159 |
| Citscibio https://citscibio.org/ | 123, 158 | Crowdsourcing: 123, 158 |

Data sharing of personal health and genomic data and associated ethical and legal challenges (eg, consent, re-use, and exploitation) were considered across a number of papers. These overarching challenges appear across a spectrum of projects and platforms, particularly those associated with self-reported data and specimen data contributions, such as 23andMe, PatientsLikeMe, and the American Gut Project. In the European Commission survey on data management, for example, only 10 projects (8%) out of those surveyed asked their participants to sign an explicit informed consent form. Equally there were considerations of the complexities of data governance and ownership, and possible approaches to these issues, such as the establishment of a participant-driven data commons.

Platforms and projects

Method

At the conclusion of the scoping review, we selected five online platforms and projects that would allow us to consider the range of responsibilities and resources for research data management that are found in health and biomedical citizen science at present. The chosen platforms were based on their frequency as identified examples in the literature review results; the prevalence of health and biomedical projects supported through these platforms; and the use of public participatory approaches in the research process. Of the selected platforms, Citscibio, SciStarter, and Zooniverse are not-for-profit and largely publicly funded. PatientsLikeMe and 23andMe are examples of for-profit crowdsourced patient platforms. We extracted information from the web pages of the platforms and projects (Supplementary Appendix 1 lists the source URLs), from related research institution websites, and from relevant publications, where available. This extraction task was undertaken in August 2018. We summarized and compared their research data management environments and approaches, following the Harvard Biomedical Data Management Lifecycle model comprising:

- Data creation
- Data analysis
- Data distribution
- Evaluation for data retention
- Organisation for long-term storage
- Archiving for research audit and for secondary use.

The Harvard Biomedical Data Management Lifecycle model was selected for its use in health and biomedical research data management contexts, and applicability across different types of data activities that were captured in the literature review.

Results and discussion

Platforms in health and biomedical citizen science have been in operation for over a decade. For example, both PatientsLike Me and 23andMe were launched in 2006, and Zooniverse was launched in 2007, SciStarter in 2014, and CitScibio in 2016. Analysis of these platforms enabled us to illustrate the current span and reach of health and biomedical citizen science, as shown in Table 3. The authors felt it particularly critical to include commercial for-profit as well as not-for-profit platforms. Biomedical research in general spans both kinds of organizations as well as government-owned organizations, and not without controversies related to its orientation. PatientsLike Me and 23andMe have attracted both popular and critical attention for their scientific activities, and we wanted to examine them through the lens of responsible research data management.

The operating models of platforms vary—for instance, sourcing opportunities to collect and analyze data, or pooling personal health data, or providing access to health data from other sources, or sharing data with peers, or on-selling data for third-party research and development—and this influences the way they support collaboration among members of the public and professional researchers. Some run formal research programs where they are data custodians and where people agree to donate their own data for relevant studies, while others have a less direct relationship to research projects.
This research project is supported by the Ludwig Boltzmann Gesellschaft (Society) of Austria. A crowdsourcing program invites patients to share their experiences with healthcare experts, to help the latter to set research priorities. The first “Tell-us!” study, focusing on mental health issues, launched in 2015 and finished in 2016. The second study, underway in 2018, brings together traumatologists and people who have experiences of injury.189 Participants also needed to complete regular questionnaires and submit their medical history and physician contacts details to the research team.

### Zooniverse
https://www.zooniverse.org/
This was established and is hosted by a group of professional scientists from Oxford University (UK), the Adler Planetarium in Chicago (USA), and records the contributions of citizen scientists.76 People to find projects relevant to their location, age and available devices, and records the contributions of citizen scientists.76

### Scistarter
https://scistarter.com/
This is a multidiscipline citizen science platform with over 2700 available citizen science projects and events begun in 2014. Scistarter has collaborations with the US National Science Foundation, Arizona State University’s Center for Engagement and Training in Science and Society, NASA, Girl Scouts of America, and others. Volunteers cannot directly access and contribute on the platform but are directed to websites of each separate project. Scistarter lists projects that offer citizen science opportunities, educates the general public about citizen science, helps people to find projects relevant to their location, age and available devices, and records the contributions of citizen scientists.76

### PatientsLikeMe
https://www.patientslikeme.com/
This online community for patients to exchange information is a private enterprise founded by lay people (Jamie Heywood, Ben Heywood, and Jeff Cole) in 2004. It launched its first online community in 2006 initially to enable Amyotrophic Lateral Sclerosis (ALS) patients to pool information. The philosophy of PatientsLikeMe is that the more information is shared by patients, the more possible it is to conduct high-impact research. In 2011, the platform expanded to all patients and all health conditions; in 2018, over half a million patients have reported on thousands of conditions, treatments and symptoms.134

### Lupus
https://www.23andme.com/lupus/
23andMe has collaborated with pharmaceutical company Pfizer and the Lupus Research Institute to study lupus and DNA. This study aimed to recruit 5000 American eligible lupus patients, and the goal was achieved in 17 months.171 Participants could either be an existing client of 23andMe, or other lupus patients who would be provided with a free 23andMe genetic test. Participants also needed to complete regular questionnaires and submit their medical history and physician contacts details to the research team.

### Mark2Cure
https://mark2cure.org/
This project was designed by Dr Ginger Tsung at the Su Lab at The Scripps Research Institute in California. Anyone who can read English can help research, annotate, and extract crucial information from the biomedical literature of the online biomedical database PubMed. Currently, Mark2Cure studies focus on rare disease conditions, that is, N-glycanase 1 deficiency. Volunteers do a tutorial to understand the interface and the tasks, then mark up the literature on how this disease relates with other conditions or treatments. Volunteers contributed over one million annotations by 2018.98

### Amyotrophic lateral sclerosis (ALS)
https://www.patientslikeme.com/conditions/als
12,000 ALS patients are registered on this platform to share their experience. Some reported experimenting with lithium carbonate to slow the progression of ALS. Data from 596 patients in all (149 in the treatment group, and 447 in the control group) subsequently underwent two analyses by trained researchers: an intent-to-treat analysis of 149 patients who reported taking lithium for at least 2 months (but may have discontinued taking the drug or died within 12 months), and an analysis of the subset of 78 patients who stayed on lithium for a full 12 months or died within that period.183

### “Tell-us!”
https://scistarter.com/project/19871-Tell-us-about-injuries
This research project is supported by the Ludwig Boltzmann Gesellschaft (Society) of Austria. A crowdsourcing program invites patients to share their experiences with healthcare experts, to help the latter to set research priorities. The first “Tell-us!” study, focusing on mental health issues, launched in 2015 and finished in 2016. The second study, underway in 2018, brings together traumatologists and people who have experiences of injury.189

### Bash the bug
https://www.zooniverse.org/projects/mrnibio/bash-the-bug
Dr Philip W Fowler, in the Modernising Medical Microbiology Group at Oxford University, is the lead researcher. To date over

---

### Table 3. Overview of selected health and biomedical citizen science platforms and associated projects

| Platform       | Project                                           |
|----------------|---------------------------------------------------|
| 23andMe        | Lupus                                             |
|                | https://www.23andme.com/en-int/                   |
|                | This direct-to-consumer genetic testing company was founded in 2006. Now it has more than 3 million customers and over 5 million genetic results stored in its databases. Users of this service buy a test kit, then place their tissue sample in a collector and send it back to the lab. Several weeks later, they receive their genetic results online through their 23andMe account portal. Customers have the option to receive health analysis results and/or ancestry results. 23andMe has partnerships with biomedical R&D institutions to study the relationships between DNA and some health conditions. In 2012, 23andMe acquired the crowdsourced patient advocacy startup “CureTogether.” Over 80% of its clients give explicit consent for 23andMe to use their data in this way. It now has ongoing projects in Parkinson’s disease, inflammatory bowel disease, lupus, fertility and major depressive disorders.173 |
| CitSciBio      | Mark2Cure                                         |
|                | https://citscibio.org/                           |
|                | CitSciBio—The Biomedical Citizen Science Hub is sponsored by the Division of Cancer Biology and the Division of Cancer Control and Population Sciences at the National Cancer Institute, at the National Institutes of Health, USA. Launched in 2016, it is an open platform for the general public to find and add relevant resources, projects, and events. According to the website, volunteer coordinators have started over 700 projects that have contributed nearly 1 million measurements for analysis to answer local, regional and/or global questions. Citizen science programs provided by Citscibio are devolved from the platform itself; the platform provides only brief descriptions and web links. Citscibio also collaborates with the citizen science data management working group of the DataONE (Data Observation Networked Earth) program (https://www.dataone.org/) to facilitate data sharing and stewardship.31 |
| PatientsLikeMe | Amyotrophic lateral sclerosis (ALS)              |
|                | https://www.patientslikeme.com/                  |
|                | This online community for patients to exchange information is a private enterprise founded by lay people (Jamie Heywood, Ben Heywood, and Jeff Cole) in 2004. It launched its first online community in 2006 initially to enable Amyotrophic Lateral Sclerosis (ALS) patients to pool information. The philosophy of PatientsLikeMe is that the more information is shared by patients, the more possible it is to conduct high-impact research. In 2011, the platform expanded to all patients and all health conditions; in 2018, over half a million patients have reported on thousands of conditions, treatments and symptoms.134 |
| Scistarter      | “Tell-us!”                                        |
|                | https://scistarter.com/                          |
|                | This is a multidiscipline citizen science platform with over 2700 available citizen science projects and events begun in 2014. Scistarter has collaborations with the US National Science Foundation, Arizona State University’s Center for Engagement and Training in Science and Society, NASA, Girl Scouts of America, and others. Volunteers cannot directly access and contribute on the platform but are directed to websites of each separate project. Scistarter lists projects that offer citizen science opportunities, educates the general public about citizen science, helps people to find projects relevant to their location, age and available devices, and records the contributions of citizen scientists.76 |
| Zooniverse     | Bash the bug                                      |
|                | https://www.zooniverse.org/                      |
|                | This was established and is hosted by a group of professional scientists from Oxford University (UK), the Adler Planetarium in Chicago (USA), and records the contributions of citizen scientists.76 |
and monitor citizens’ work with research data in less detail. Almost all platforms state that they collect standard website user data, such as site visit and navigation information of visitors, and some use cookies to record the time and duration of each visit, and so on. At this meta-level, the platforms are purposely built to support basic data management needs for web analytics by their host organization.

Information about the research data management aspects of these platforms and their representative projects is summarized following Table 3, in terms of data creation, data analysis, data distribution, evaluation, long-term storage, and reuse. A list of web pages consulted for each platform and project is listed in Supplementary Appendix 1.

### Data creation

Data may be generated by participants, in the form of personal biometric measurements self-reported or tested independently (eg, 23andMe), or they may be narrative observations and perceptions about their health (eg, TellUs!), or they may be data about a family member or another person they care for (eg, in rare childhood conditions). Datasets may also be compiled by professional researchers for citizen scientists to act upon (as in Mark2Cure), or they may be openly available for a research team to work with (as with increasing access to governments’ aggregated health datasets). Research data creation in health and biomedical citizen science is diverse, and data quality management processes that would ensure rigor at this stage in the data lifecycle must vary accordingly. Although specific data formats have been proposed for some categories of data, data quality assurance is not clearly addressed on the websites, and some are published externally, such as the biocuration aspects of Mark2Cure. Patient advocacy stories by donor participants were identifiable in several of the examples, such as 23andMe. 23andMe, and others.

### Data quality management processes

Data quality management processes that would ensure rigor at this stage in the data lifecycle must vary accordingly. Although specific data formats have been proposed for some categories of data, data quality assurance is not clearly addressed on the websites, and some are published externally, such as the biocuration aspects of Mark2Cure. Patient advocacy stories by donor participants were identifiable in several of the examples, such as 23andMe. 23andMe, and others.

### Data analysis

In citizen science projects where nonprofessionals undertake data analysis, this work may be done as a form of recreation (as in gamified projects), or in conjunction with health literacy education and training (eg, CitSciBio), or by expert patients or high-performing athletes (eg, within health self-quantification groups). A citizen scientist’s role may be limited to very basic annotation or contextuali-
to the project specificity of data protocols in each case. Zooniverse has published a set of success criteria for projects, which includes data related processes,\textsuperscript{191} for example. Some projects involve research data of a different nature than personal health data (eg, Mark 2 Cure). In 23andMe, personal and registration information is stored separately from any genetic information to reduce the likelihood that a consumer could be identified; if a consumer opts in to the 23andMe research program and completes a survey, their genetic information is de-identified and stored with their survey response data in a separate research environment. Zooniverse, too, makes clear that it maintains separate data storage of personal information, users’ contributions, and project data on various servers. 23andMe uses software, hardware and physical security measures to protect the computers where consumer data are stored, and robust authentication methods for access to its systems. 23andMe’s privacy measures include avoiding data re-identifiability, limiting personal data accessibility, and detecting and vulnerabilities in real time. PatientsLikeMe servers are located in the United States and a range of data security measures is described,\textsuperscript{191} however, no further information is provided about methods of data storage.

23andMe states that it does not sell, lease or rent individual-level information to any third party without explicit consent, but that it does share de-identified aggregated data with third parties. It is noted in this respect that there is controversy in the nontransparent way in which 23andMe, for example, has applied for patents based on participant-supplied genomic and phenotypic data.\textsuperscript{43} PatientsLikeMe is the only platform that explains how it profits commercially from data supplied by the community. “What happens to research data in citizen science projects over the long term?” Is a question very few people seem to ask, and few would be able to answer, based on what information other platforms provide to the public. None of the five platforms or associated projects fully explain or account for the entire research data lifecycle in a way suggestive of data management practices that might more confidently assure the quality of the research across the data lifecycle. This gap in available information further suggests that not all data processes are presently transparent or open to the public or potential participants.

LIMITATIONS

First, this scoping review was not intended to provide a comprehensive or systematic review of the topic—but rather a representative one. As such, our review of the literature was not designed to generate effect sizes or aggregate metrics. Nevertheless, given the context of identifying and exploring research data management practices across the breadth of citizen science, participatory health, crowdsourcing, and self-quantification, the review methodology that we followed was deemed an appropriate lens. This has inevitably taken into account that there are differing definitional boundaries of what constitutes health and biomedical citizen science in the literature.\textsuperscript{5,21,40,43,51,92} As stated in Fiske et al.,\textsuperscript{70} it is challenging to develop a meaningful typology of citizen science initiatives in biomedicine.

Our analysis of data management processes of the selected platforms was also constrained due to limited access to mechanisms by which we could request more information than was publicly available on platform web sites.

We further recognize that our analysis of research data management practices is only one avenue to provide insights into the opportunities and challenges of citizen science particularly in relation to the for-profit platforms, PatientsLikeMe and 23andMe, both of which are undergoing increasing scrutiny by bioethicists, for example.\textsuperscript{70,192,193}

Other limitations can be directly associated with the challenges of undertaking such a multidimensional topic. The authors acknowledge that there is much on this topic which requires a fuller investigation, and not least what the implications are for each of the separate and diverse approaches of citizen science, participatory health, crowdsourcing, and self-quantification. The educative aspects in relation to the instigators of a health and biomedical citizen science project and the participants themselves in data management practices are a further consideration which was not scoped in the present study.

CONCLUSION

Health and biomedical citizen science platforms reflect diverse approaches to data management among both participants and researchers, as well as specific values associated with types of data activities, such as data sharing and re-use beyond the lifetime of a project. The lack of standard definitional boundaries of what constitutes health and biomedical citizen science contributes to this diversity of approach. These overall relationships are critical to understand in terms of what specific initiatives seek to achieve, whom they benefit, and what typologies or frameworks might be specifically aligned to participant roles and responsibilities in the data management lifecycle processes of health and biomedical citizen science. This must include the capacity to ensure data are “fit for purpose” and appropriate levels of data governance are supported. Future work requires further considered approaches to a comparative research data management focus which could be additionally supported through a public interest consensus about minimum standards that can be applied throughout the research data lifecycle. Without which many important data collections, analysis and subsequent reuse may be opaque, conflicted or untraceable.

On a practitioner level, the situation could be more widely improved with a concerted program of participatory development, producing clearer standards and reusable resources for health and biomedical citizen science research data management. Such a project might take a contributory, collaborative or co-created citizen science approach. Without a directed initiative in this context, health and biomedical citizen science is unlikely to derive optimal shared value, benefits, or sustained use of the enabling platforms, resources, and collective knowledge of its stakeholders.

SUPPLEMENTARY MATERIAL

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

AUTHOR CONTRIBUTIONS

AB and KG made substantial contributions to the study design, conception and drafting of the manuscript. YF implemented the methods and carried out the qualitative analysis. AB made critical revisions for final approval of the version to be published.

CONFLICT OF INTEREST STATEMENT

None declared.
REFERENCES
1. Bonney R. Citizen science: a lab tradition. *Living Bird* 1996; 15 (4): 7–15.
2. Bonney R, Ballard H, Jordan R, et al. *Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education*. A CAISE Inquiry Group Report. Washington, DC: Center for Advancement of Informal Science Education; 2009.
3. Science Europe. Briefing paper on citizen science: D/2018/13.324/2. 2018. https://www.scienceeuropa.org/wp-content/uploads/2018/07/SE_BriefingPaper_CitizenScience.pdf. Accessed 12 February, 2019.
4. Wiggins A, Crowston K. Surveying the citizen science landscape. First Monday 2014; 20 (1): doi.org/10.5210/fm.v20i1.5520; https://firstmonday.org/ojs/index.php/fm/article/view/5520/4194 Accessed October 1, 2019.
5. Kullenberg C, Kasperson RW, et al. What is citizen science? – A scientometric meta-analysis. *PLoS One* 2016; 11 (1): e0147152.
6. Oxford English Dictionary. Definition of citizen science in US English. 'citizen science'. https://www.lexico.com/en/definition/citizen_science. Accessed 2 October, 2019.
7. Silvertown J. A new dawn for citizen science. *Trends Ecol Evol* 2009; 24 (9): 467–71.
8. Irwin A. *Citizen Science: A Study of People, Expertise and Sustainable Development*. London: Routledge; 1995.
9. Bonney R, Cooper CB, Dickinson J. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *Bioscience* 2009; 59 (11): 977–84.
10. Masters K, Oh EY, Cox J, et al. Science learning via participation in online citizen science. *J Sci Commun* 15 (3): A07.
11. Haklay M. Citizen science and volunteered geographic information: overview and typology of participation. In: Sui D, Elwood S, and Goodchild M, eds. Crowdsourcing Geographic Knowledge. Dordrecht: Springer; 2013: 105–122.
12. Christian C, Lintott C, Smith A, Fortson L, Barmford S. Citizen science: contributions to astronomy research. organizations, people and strategies; 2012: arXiv: 1202.2577. https://arxiv.org/abs/1202.2577 Accessed February 12, 2019.
13. Reed J, Raddick MJ, Lardner A, Carney K. An exploratory factor analysis of motivations for participating in Zooniverse, a collection of virtual citizen science projects. In: Sprague RH, ed. Proceedings of the 46th Annual Hawaii International Conference on System Sciences. New York: IEEE; 2013: 610–619.
14. Wright DR, Underhill LG, Keene M, Knight AT. Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. *Soc Nat Resour* 2015; 28 (9): 1013–29.
15. Bonney R, Phillips TB, Ballard HL, Enck JW. Can citizen science enhance the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. *Public Underst Sci* 2016; 25 (1): 2–16.
16. Haklay M. *Citizen Science and Policy: A European Perspective*. Washington, DC: Woodrow Wilson International Center for Scholars; 2015.
17. Storksdeek M. Associations for citizen science: regional knowledge, global collaboration. *Citiz Sci Theory Pract* 2016; 1 (2): 1–10.
18. Shirk JL, Ballard HL, Wilderman CC, et al. Public participation in scientific research: a framework for deliberate design. *Ecol Soc* 2012; 17 (2): 29.
19. Wiggins A, Crowston K. From conservation to crowdsourcing: A typology of citizen science. In: proceedings of the Annual Hawaii International Conference on System Sciences (HICSS), Kauai, Hawaii, USA; 2011: 1–10.
20. Baker B. Frontiers of citizen science. *Bioscience* 2016; 66 (11): 921–7.
21. Cooper S, Khatib F, Treuille A, et al. Predicting protein structures with a multiplayer online game. *Nature* 2010; 466 (7307): 756–60.
22. Simperl E, Reeves N, Phethean C, Lynes T, Tinati R. Is virtual citizen science a game? *Trans Soc Comput* 2018; 1 (2): 1.
23. Wiggins A, Bonney R, LeBuhn G, Parrish JK, Welschn JF. A science products inventory for citizen-science planning and evaluation. *BioScience* 2018; 68 (6): 436–44.
24. Williams J, Chapman C, Leibovici DG, et al. Maximising the impact and reuse of citizen science data. In: Hecker S, Haklay M, Bowser A, et al., eds. *Citizen Science: Innovation in Open Science, Society and Policy*. London: UCL Press; 2018: 321–336.
25. Riesch H, Potter C. Citizen science as seen by scientists: methodological, epistemological and ethical dimensions. *Public Underst Sci* 2014; 23 (1): 107–20.
26. Kelling S, Fink D, La Sorte F, Johnston A, et al. Taking a ‘Big Data’ approach to data quality in a citizen science project. *Ambio* 2015; 44 (S4): 601–11.
27. Kosmala M, Wiggins A, Swanson A, Simmons B. Assessing data quality in citizen science. *Front Ecol Environ* 2016; 14 (10): 551–60.
28. Wiggins A, Newman G, Stevenson R, Crowston K. Mechanisms for data quality and validation in citizen science. In proceedings 7th IEEE International Conference on e-Science Workshops, eScienceW 2011, Stockholm, Sweden; 2011: 14–19.
29. Antelio M, Esteves M, Schneider D, Souza JM. Qualitocracy: a data quality collaborative framework applied to citizen science. In: proceedings of the International Conference on Systems, Man and Cybernetics (IEEE SMC 2012), Seoul, South Korea; 2012: 931–936.
30. Boulou MN, Resch B, Crowley DN, et al. Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples. *Int J Health Geogr* 2011; 10 (1): 67.
31. Watson D, Floridi L. Crowdsourced science: sociotechnical epistemology in the e-research paradigm. *Synthese* 2018; 195 (2): 741–64.
32. US Department of General Services Administration. Citizen science. https://www.citizenscience.gov/toolkit/. Accessed February 12, 2019.
33. Wiggins A, Bonney R, Graham E, et al. *Data Management Guide for Public Participation in Scientific Research*. Albuquerque, NM: DATA-ONE; 2013.
34. Schade S, Tsinaraki C. *Survey Report: Data Management in Citizen Science Projects*: EUR 27920 EN. Luxembourg (Luxembourg): Publications Office of the European Union; 2016.
35. Parrish JK, Burgess H, Weltzin JF, Fortson L, Wiggins A, Simmons B. Exposing the citizen science in citizen science: fitness to purpose and intentional design. *Integr Comp Biol* 2018; 58 (1): 150–60.
36. Caron-Flinterman JF, Broeze JE, Bunders JF. Patient partnership in decision-making on biomedical research: changing the network. *Sci Technol Human Values* 2007; 32 (3): 339–68.
37. Rabinow P. Artificiality and enlightenment: from sociobiology to biosociality In: Rabinow P, ed. *Essays on the Anthropology of Reason*. Princeton, NJ: Princeton University Press; 1996: 91–111.
38. Petryna A. *Life Exposed: Biological Citizens after Chernobyl*. Princeton, NJ: Princeton University Press; 2002.
39. Rose N, Novas C. Biological citizenship. In: Ong A, Collier, S, eds. Global Assemblages: Technology, Politics and Ethics as Anthropological Problems. Malden, MA: Blackwell; 2005: 439–63.
40. Pauwels E, Denton S. *The Rise of the Bio-Citizen*. 1st ed. Washington, DC: The Wilson Centre; 2018.
41. Swan M. Health 2050: The realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. *J Pers Med* 2012; 2 (3): 93–118.
42. Strasser B, Baudry J, Mahr D, Sanchez G, Tancoigne E. Citizen Science”? Rethinking science and public participation. *Sci Technol* 2018; 52–76.
43. Woolley JP, McGowan ML, Teare HJA, et al. Citizen science or scientific citizenship? Disentangling the uses of public engagement rhetoric in national research initiatives. *BMC Med Ethics* 2016; 17 (1): 33.
44. Strang L, Simmons RK. *Crowdsourcing-for-Semantic Reviews*. Cambridge: The Healthcare Improvement Studies Institute; 2018.
45. Bach M, Jordan S, Hartung S, et al. Participatory epidemiology: the contribution of participatory research to epidemiology. *Emerg Themes Epi demiol* 2017; 14 (1) :2.
46. Broeder Devilee DL, Van Oers J, Schuit H, Wagemakers AJ. Citizen science for public health. *Health Promot Int* 2018; 33 (3): 505–14.
47. Balaza CI, Morello-Frosch R. The three R’s: how community-based participatory research strengthens the rigor, relevance, and reach of science. *Environ Justice* 2013; 6 (1): 9–16.
89. Rosas LG, Salvo D, Winter SJ, et al. Harnessing technology and citizen science to support neighborhoods that promote active living in Mexico. J Urban Health 2016; 93 (6): 953–973.

90. Roskams J, Popovic Z. Power to the people: addressing big data challenges in neuroscience by creating a new cadre of citizen neuroscientists. Neuron 2016; 92 (3): 658–664.

91. Rothstein MA, Wilbanks JT, Brothers KB. Citizen science on your smartphone: an ELSI research agenda: currents in contemporary bioethics. J Law Med Ethics 2015; 43 (4): 897–903.

92. Rowbotham S, McKinnon M, Leach J, Lambert R, Hawe P. Does citizen science have the capacity to transform population health science? Crit Public Health 2019; 29 (1): 118–128.

93. Senabre E, Ferran-Ferrer N, Perello J. Participatory design of citizen science experiments. Commun Media Educ J 2018; 26 (34): 29–38.

94. Shuttleworth S, Frampton S. Constructing scientific communities: citizen science. Lancet 2015; 385 (9987): 2568.

95. Smittenaar P, Walker AK, McGill S, et al. Crowdsourcing in human pathology. J Pathol Inform 2018; 120 (1): 27–33.

96. Lichten C, Ioppolo R, D’Angelo C, Simmons R, Jones M. Citizen Science: crowdsourcing for Research. 1st ed. Cambridge: The Healthcare Improvement Studies Institute; 2018: 1–17.

97. Tsueng G, Kumar A, Nanis M, Su AI. Aligning needs: integrating citizen science. JOM 2017; 25 (1): 45–60.

98. Tseng G, Nanis SM, Fouquier J, Good BM, Su AI. Citizen science for mining the biomedical literature. CiteSeer 2016; 2 (1): 14.

99. Van Vliet K, Moore C. Citizen science initiatives: engaging the public and demystifying science. J Microbiol Biol Educ 2016; 17 (1): 13–16.

100. Land-Zaandstra AM, Devilee JLA, Snik F, Buurmeijer F, van den Broek JM. Citizen science on a smartphone: participants’ motivations and learning. Public Underst Sci 2016; 25 (1): 45–60.

101. Cox J, Oh EY, Simmons B, et al. Defining and measuring success in online citizen science: a case study of Zooniverse projects. Comput Sci Eng 2015; 17 (4): 28–41.

102. Afshinnekoo E, Ahsanuddin S, Mason CE. Globalizing and crowdsourcing biomedical research. Br Med Bull 2016; 120 (1): 27–33.

103. Alayza R, Tavakkol S, Tavakkol E, et al. A review on the applications of crowdsourcing in human pathology. J Pathol Inform 2018; 9 (2): 2.

104. Bruggemann J, Lander GC, Su AI. Exploring applications of crowdsourcing to cryo-EM. J Struct Biol 2018; 203 (1): 37–45.

105. Carter RR, DiFeo A, Bogie K, Zhang GQ, Sun J. Crowdsourcing awareness: exploration of the ovarian cancer knowledge gap through Amazon Mechanical Turk. PLoS One 2014; 9 (1): e85508.

106. dos Reis FJC, Lynn S, Ali HR, et al. Crowdsourcing the general public for large-scale molecular pathology studies in cancer. Ebiomedi cine 2015; 2 (7): 681–9.

107. Garneau NL, Nuesse TM, Tucker RM, Yao M, Santorisco SA, Mattes RD. Taste responses to linoleic acid: a crowdsourced population study. Chem Senses 2017; 42 (9): 769–73.

108. Good BM, Su AI. Crowdsourcing for bioinformatics. Bioinformatics 2013; 29 (16): 2025–33.

109. Good BM, Loguercio S, Griffith O, Nanis M, Wu C, Su AI. The cure and benefit of a crowdsourcing game for gene selection in breast cancer survival prediction. JMR Serious Games 2014; 2 (2): 67.

110. Hossain SM, Hnai T, Saleheen N, et al. mcerebrum: a mobile sensing software platform for development and validation of digital biomarkers and interventions. In: proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, Vol. 7. SenSys’17, Delft, Netherlands; 2017: 1–14.

111. Johnson D, Deterding S, Kuhn KA, et al. Gamification for health and wellbeing: a systematic review of the literature. Internet Interventions 2016; 6 (6): 89–106.

112. Kahn MG, Brown JS, Chun AT, et al. Transparent reporting of data quality in distributed data networks. EGE15s 2015; 3 (1): 7.

113. Lambertson KF, Damiani SA, Mighl T, Shelton R, Terry SF. Participant-driven matchmaking in the genomic era. Hum Mutat 2015; 36 (10): 965–73.

114. Lawson J, Robinson-Vyas RJ, McQuillan JP, et al. Crowdsourcing for translational research: analysis of biomarker expression using cancer microarrays. Br J Cancer 2017; 116 (2): 237–245.

115. Lee YJ, Aida JA, Donovan HS. The application of crowdsourcing approaches to cancer research: a systematic review. Cancer Med 2017; 6 (11): 2595–2605.

116. Leiter A, Sablini T, Diefenbach M, et al. Use of crowdsourcing for cancer clinical trial development. J Natl Cancer Inst 2014; 106: 1–3.

117. Morris RR, Schueler SM, Picard RW. Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. J Med Internet Res 2013; 17 (3): e72.

118. Parry C, Beckjord E, Moser RP, Vieux SN, Padgett LS, Hesse BW. It takes a (virtual) village: crowdsourcing measurement consensus to advance survivorship care planning. Behav Med Prat Policy Res 2015; 5 (1): 53–59.

119. Parvanta C, Roth Y, Keller H. Crowdsourcing 101: a few basics to make you the leader of the pack. Health Promot Pract 2013; 14: 1673–1675.

120. Pianta M, Makrai E, Verspoor KM, Cohn T, Downie L. Crowdsourcing critical appraisal of research evidence (CrowdCARE) was found to be a valid approach to assessing clinical research quality. J Clin Epidemiol 2018; 104: 8–14.

121. Ranard BL, Ha YP, Miesel ZF, et al. Crowdsourcing—harnessing the masses to advance health and medicine, a systematic review. J Gen Intern Med 2014; 29 (1): 187–203.

122. Renwick MJ, Mossialos E. Crowdfunding our health: economic risks and benefits. Soc Sci Med 2017; 191: 48–56.

123. Riggare S, Unruh KT, Sturr J, et al. Patient-driven N-of-1 in Parkinson’s disease. Methods Inf Med 2017; 56 (6): e123–8.

124. Sauermann H, Franzoni C. Crowd science user contribution patterns and their implications. Proc Natl Acad Sci USA 2015; 112 (3): 679–684.

125. Snyder J, Mathers A, Crooks VA. Crowdsourced health research studies: an important emerging phenomenon. J Health Serv Res Policy 2016; 21: 27–30.

126. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res 2012; 14 (2): e46.

127. Swan M. DIYgenomics: crowdsourced health research studies: personal wellness and preventive medicine through collective intelligence. AAAI Spring Symposium - AAAI Technical Report SS-12-05; 2012: 54–59.

128. Swan M. The quantified self: fundamental disruption in big data science and biological discovery. Big Data 2013; 1 (2): 85–99.

129. Thomas J, Noel-Storr A, Marshall I, et al. Living systematic reviews: a combined human and machine effort. J Clin Epidemiol 2017; 91: 31–37.

130. Torr-Brown S. Crowdsourced science and genetic insights for personalized medicine: beginning to deliver the goods. Per Med 2012; 9 (6): 569–572.

131. Vera J, Santos G. Can gamification contribute to computer modeling-driven biomedical research? Front Physiol 2018; 9: 908.

132. Wicks P, Thorley EM, Simacek K, Curran C, Emmas C. Scaling PatientsLikeMe via a “generalized platform” for members with chronic illness: web-based matchmaking in the genomic era. J Gen Intern Med 2014; 29 (1): 187–203.
180. Vandenberghe B, Geerts D. Sleep monitoring tools at home and in the hospital: bridging quantified self and clinical sleep research. In: Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth ‘15, Istanbul, Turkey; 2015: 153–160.

181. Vegter MW. Towards precision medicine; a new biomedical cosmology. Med Health Care and Philos 2018; 21 (4): 443–56.

182. Wicks P, Little M. The virtuous circle of the quantified self: a human computational approach to improved health outcomes. In: Michelucci P, ed. Handbook of Human Computation. New York, NY: Springer; 2013: 105–29.

183. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. Nat Biotechnol 2011; 29 (5): 411.

184. Williams AD. Harnessing the quantified self movement for optimal mental health and wellbeing. In: paper presented at the Proceedings of the first Workshop on Lifelogging Tools and Applications LTA ’16, Amsterdam, The Netherlands; 2016: 37.

185. Zhang J, Dibia V, Sodnomov A, Lowry PB. Understanding the disclosure of private healthcare information within online quantified self 2.0 platforms. In: proceedings of the Pacific Asia Conference on Information Systems, PACIS 2015, Singapore; 2015: 140.

186. Lopez-Campos GH, Almalki M, Martin-Sanchez F. Proposal for a standardised reporting guideline to annotate health-related self-quantification experiments. Stud Health Technol Inform 2014; 202: 79–82.

187. European Parliament. General data protection regulation. Off J Eur Union 2016; 59: 1–88.

188. Harvard University. Harvard biomedical data management data lifecycle model; 2018. https://datamanagement.hms.harvard.edu/hms-data-lifecycle Accessed February 12, 2019.

189. Beck S, Brasseur TM, Poetz MK, Sauermann H. What’s the problem? How crowdsourcing contributes to identifying scientific research questions. Acad Manag Proc 2019; 2019: 1.

190. Simera I, Moher D, Hoey J, Schulz KF, Altman DG. The EQUATOR Network and reporting guidelines: helping to achieve high standards in reporting health research studies. Maturitas 2009; 63 (1): 4–6.

191. PatientsLikeMe. What safeguards does PatientsLikeMe have in place to secure data? 2015. https://support.patientslikeme.com/hc/en-us/articles/201186634-What-safeguards-does-PatientsLikeMe-have-in-place-to-secure-data Accessed February 12, 2019.

192. Hummel P, Braun M, Dabrock P. Data donations as exercises of sovereignty. In: Krutzinna J, Floridi L, eds. The Ethics of Medical Data Donation. Philosophical Studies Series. Cham: Springer; 2019 (137): 23–54.

193. Tempini N, Del Savio L. Digital orphans; data closure and openness in patient-powered networks. BioSocieties 2019; 14 (2): 205–27.
