The Rise of Citizen Science in Health and Biomedical Research

Andrea Wiggins & John Wilbanks

Published online: 24 Jul 2019.

Submit your article to this journal

Article views: 793

View Crossmark data

Citing articles: 19 View citing articles

OPEN ACCESS

Check for updates

Target Article

The Rise of Citizen Science in Health and Biomedical Research

Andrea Wiggins (ID)**,** University of Nebraska at Omaha
John Wilbanks (ID)**,** Sage Bionetworks

Citizen science models of public participation in scientific research represent a growing area of opportunity for health and biomedical research, as well as new impetus for more collaborative forms of engagement in large-scale research. However, this also surfaces a variety of ethical issues that both fall outside of and build upon the standard human subjects concerns in bioethics. This article provides background on citizen science, examples of current projects in the field, and discussion of established and emerging ethical issues for citizen science in health and biomedical research.

INTRODUCTION

"Citizen science," a range of participatory models for involving nonprofessionals as collaborators in scientific research, is receiving increased attention for its ongoing successes. Citizen science is well developed in fields like ecology and astronomy, where there is a long history of nonprofessionals making substantive contributions to the scholarly knowledge base. Western health and biomedical research have historically had a different relationship to nonprofessionals (as subjects or service recipients), however, with protective regulations and oversight further constraining interactions between parties. But making individual lifestyle changes for health management is an ancient practice, and the advent of digital communications networks makes it possible to broadcast these actions to the world, and to follow other people with similar interests. Increasingly, a "populist rhetoric" that calls for patient-centric practices and lay involvement is being embraced as a means of encouraging broader participation in the health system (Woolley et al. 2016).

The history of citizen science's emergence has been well documented (Miller-Rushing et al. 2012); the common definition describes a form of research collaboration involving the public in scientific research to address real-world problems (Bonney et al. 2009), which neither requires the research to be initiated by members of the public nor presupposes a "crowdsourcing" model of engagement. Notably, the word "citizen" in citizen science refers refer to voluntary participation by nonprofessional contributors. Originally expressed in hyphenated form as "citizen-science" to refer to a type of partnership between formal science and individual members of the public, the nonhyphenated phrase is now used in reference to a wide range of models of public engagement in research.

This article does not answer ethical questions related to citizen science in health and biomedical research, but exposes the complexity of issues that ethicists and practitioners need to consider and flags emergent concerns that are poorly managed, if at all. For example, many health data donation projects can be considered citizen science. Online communities such as PatientsLikeMe are often inhabited by patient-experts who share information, generate hypotheses based on common experiences, conduct N-of-1 experiments, and support health data sharing for research on various conditions (Wicks et al. 2010). Platforms such as 23andme and uBiome draw inspiration from venture-backed consumer data platforms, promising consumer-friendly services, network effects, and data return. These direct-to-consumer health information platforms appear similar to citizen science

Address correspondence to Andrea Wiggins, Information Systems and Quantitative Analysis, University of Nebraska at Omaha, 177C PKI, 6001 Dodge St, Omaha, NE 68182-0116, USA. E-mail: wiggins@unomaha.edu

but, as discussed later, are not always seen as such, and inspire concerns about surveillance and obscure terms of service, among others.

As citizen science evolves, the expectations of diverse stakeholders for public participation in scientific research generate new pressures on researchers, presenting both impetus and models for more collaborative forms of engagement. Some models have evolved completely outside the formal epistemic and moral systems of science, with little a priori conversation about salient topics such as peer review and the risks of amplifying problematic claims, or the paucity of informed consent about risks and benefits. Current citizen science models often shift assumed risk from the admittedly calcified systems of traditional health research to the individual, while also representing new opportunities where taking on uncertainty and risk may deliver new benefits:

- People with access to their own data streams can develop a unique familiarity that can support interpretive fluency.
- Developing relationships with engaged individuals can create new opportunities for rich longitudinal data streams, both informing their own choices and allowing larger cohorts to emerge from individual streams.
- Human cognitive processing is uniquely suited to, and in some cases provides superior performance in, such tasks as image analysis, graph interpretation, and puzzle solving.
- Funding in health and medicine typically follows a "long tail" distribution of investment, leaving a huge gap for citizen or community led research into underresearched spaces.
- Many members of the public have already demonstrated a strong interest in and willingness to contribute to meaningful scientific research.

Few research domains are as meaningful to the public as health and medicine, which should therefore be well positioned for citizen science engagement. Health and biomedical research encompass a vast range of potential inquiry, much of which is becoming newly accessible to nonprofessionals after a 40-year technology boom: From DNA sequencing to keto diets, the options for at-home inquiry have exploded. But these opportunities raise complex ethical questions that are rarely considered in disciplines such as ornithology or astronomy: Are citizen scientists doing "basic" research without risk to themselves or other humans; are they performing observations and interventions on themselves; are they sharing their personal or aggregate observations, interventions, and results with audiences around the world? The published work on ethics in citizen science, however, currently takes a fairly narrow view of relevant ethical implications, focusing on a limited set of issues specific to a given field or model of participation. Dominant themes in the current conversations include interrelated issues of recognition

(Riesch and Potter 2014), protections, and researcher responsibilities to participants (European Citizen Science Association 2015), which are just a fraction of the ethical issues that need to be engaged.

This article provides background on citizen science with examples of current initiatives and discussion of ethical considerations for citizen science participation. We highlight the citizen scientist's place outside the mainstream of scientific epistemology and modern bioethics, which assume the use of institutionalized systems such as peer review and ethics review. Additional concerns include the unique considerations of repurposing proprietary platform technologies for citizen science without full consideration of potential impact. The following sections discuss the related research practices that provide a foundation for citizen science in health and biomedical research, examples of current citizen science projects, and ethical concerns.

FOUNDATIONS OF HEALTH AND BIOMEDICAL CITIZEN SCIENCES

The label "citizen science" describes a number of phenomena across a variety of areas of research and practice; well-established variations include volunteer monitoring, community science, and participatory research, among others. It can also refer to more recently developed technology-supported projects where crowdsourcing techniques are applied to scientific data or analysis. In practice, most citizen science is based on collective or explicitly collaborative models of participation, rather than singular or "do-it-yourself" (DIY) forms of inquiry, the implications of which are further discussed in the following. These strategies for engaging the public as co-researchers are not actually new to science, and the novelty of modern citizen science often lies in the scope and scale of involvement by nonprofessionals. However, the diverse forms of citizen science in operation today clearly relate to existing methodologies such as participatory action research, community-based participatory research, and action science.

Participatory action research (PAR) methods are applied in a much broader category of critical research and differs from action research by involving subjects as co-researchers (Argyris and Schön 1989). PAR's goals include understanding and improving the world through change, and its primary affiliated disciplines are social sciences (Baum, MacDougall, and Smith 2006). The logistical and practical drawbacks of fully collaborative research involving the community at every step (Walter 1998) strongly resemble those in community-based citizen science. And participatory research is an imperfect fit for traditional ethical review (Vayena and Tasioulas 2013).

In public health research, community-based participatory research (CBPR) is best known for applications to address health inequities (Israel et al. 1998), often drawing on critical and constructivist ways of knowing for greater

compatibility with the social realities of participating community members. CBPR is approximately equivalent to community-based participatory action research (CB-PAR), and a primary goal of both approaches is bridging the science–practice gap through community engagement (Wallerstein and Duran 2006). Many of the same challenges observed in CBPR are endemic to citizen science as well (Viswanathan et al. 2004).

A handful of additional variations share many of the same philosophies and goals, where part of the definition of success often focuses on processes rather than outputs. For example, "action science" is a form of action research that emphasizes participants' theories-in-use (Argyris and Schön 1989) and is often incorporated into CBPR (Wallerstein and Duran 2006). Similar forms of community science have appeared under labels such as "popular epidemiology" and community-based health assessments (Den Broeder et al. 2016).

Twenty-first-century citizen science has developed its own brand, often with surprisingly little direct influence from these foundations due to disciplinary differences; nonetheless, many principles and practices are convergent with the preexisting methods. In the following sections, we discuss several dimensions of citizen science that provide a window into relevant ethical considerations. These dimensions complement an existing typology of citizen science in health and biomedical sciences (Fiske et al. 2019), which focuses on the types of contributions that participants make in the project.

TASK TYPES

The types of tasks that participants do can be roughly divided into data collection and data processing or analysis, noting that this is a dramatic simplification of the full range of activities that citizen scientists may actually engage in. Since the same terminology is often used to refer to different types of participation (Woolley et al. 2016), the examples and discussion here focus on the two most common categories of active engagement based on the role of the public in the scientific process, described as *contributory* and *co-created*, from a popular citizen science typology (Shirk et al. 2012). Contributory citizen science projects focus on engaging participants in a limited portion of the scientific research process, using a common citizen science model that "engages a dispersed network of volunteers to assist in professional research using methodologies that have been developed by or in collaboration with professional researchers" (Cooper et al. 2007, 2) in a top-down structure, with the explicit expectation that volunteers are involved primarily in data collection to address researchers' questions.

Co-created citizen science applies a very different strategy, with members of the public (usually framed as a community) engaged in most or all of the scientific inquiry process, with or without the involvement of a professional scientist. These projects often take on a bottom-up or grass-roots structure and a substantively different approach from conventional science. Typical goals can include empowerment and direct action focused on shared concerns (Shirk et al. 2012, Wandersman 2003).

Data Collection

Data collection projects dominate in environmentally focused citizen science but are inherently more challenging in health contexts, a concern that led the U.S. National Institutes of Health to host a workshop in 2015 on the ethical, legal, and social implications of citizen science (National Institutes of Health 2015). Ensuring that participants are not solely research subjects but can also contribute in an expanded role is a primary criterion for determining whether the term "citizen science" applies. Engaging the public in data collection can be accomplished through either a contributory or co-created model; contributory data collection projects typically benefit from the scale and distribution of volunteers, while co-created projects leverage participants' localized knowledge and specific interests. Currently, most data collection citizen science projects are observational studies of personal health data, microbiomes, and sensory pollution.

Ethically, data collection projects that are organized by professional scientists may encounter additional challenges due to social norms and research requirements around informed consent and privacy (Woolley et al. 2016). These forms of oversight do not engage the question of whether current paradigms of responsible conduct of research are actually appropriate for citizen science, however, and may limit the development of otherwise fruitful projects. Most citizen science focused on data collection that is led from an institutional context (typically academic or nonprofit) involves very limited forms of personal health data, or keeps the individual at arm's length, as with microbiome studies. Such projects may be relatively uncommon due in part to the constraints imposed by institutional ethics regimes. Community-driven citizen science is not necessarily subject to any such regulations, yet often exceeds the ethical standards imposed in institutional contexts. Many additional projects in ecology also focus on issues such as air and water quality that have direct impacts on, and are often motivated by concerns over, human health impacts (McKinley et al. 2017).

A range of existing projects provides illustrative examples. *100forParkinsons* involves self-tracking for 100 days via a smartphone app, while providing participants access to comparative data for the information they share. *Kinsey Reporter* elicits self-report of sexual behavior, with data available for offline analysis. *The American Gut Project* participants share microbiome samples and receive personalized analysis results; to

overcome the limits on participation posed by this model, a gamified puzzle-based smartphone app (*Colony B*) was developed as a companion project to support participant-driven data analysis and offer increased opportunities for in-depth engagement. These examples also provide a sense of the wide variety of inherent ethical issues: An identified citizen scientist sharing her microbial gut data poses radically different risks than the same citizen scientist sharing her sexual behavior data.

Data Processing

Data-processing projects move data from initial to analyzable states. These projects have benefitted substantially by borrowing from video games to support puzzle-based problem solving, a cognitively intensive task that is not otherwise well represented in citizen science. By nature, these projects apply a contributory model (due to the reliance on purpose-built tools and data sources) and focus on harnessing human attention at scale.

The puzzle-based projects in proteomics and neurology rely heavily on gamification and sophisticated bespoke systems to support complex problem-solving tasks such as protein folding and neural mapping. These projects are true innovators with novel designs that may be applicable more broadly. For example, *Foldit* players can work individually or in teams and can share their "recipes" or strategies for analyses, a contributing factor to gamers solving a long-standing primate HIV enzyme protein folding problem in a matter of days (Gilski et al. 2011). *Eyewire* stands out for exceptional application of gamification to engage players in mapping a retinal neuron in three dimensions (3D), and adaptation of its code base could hypothetically support similar 3D mapping tasks. *EteRNA* uses algorithmic evaluation of player contributions to rank submissions for lab synthesis, an approach that could be applied much more broadly in tools for data quality filtering. These projects appear to pose relatively little ethical "risk" in the traditional sense for individual contributors other than time invested, a common trait in projects toward the pure research end of the spectrum that keep distance from identifiable individuals.

Projects focused on extracting content from images through transcription, annotation, or classification tasks remain relatively uncommon in health research. Although health information privacy constraints are likely one cause of the slow uptake of crowdsourcing in biomedical research, recent research has also identified clear needs for expertise to interpret some medical imaging (e.g., mammograms; Baâzaoui, Barhoumi, and Zagrouba, 2017) and cases where machine classification can outperform humans (e.g., retinal scans; Mudie et al. 2017), unlike other fields where nonexpert human classification performs quite well (Swanson et al. 2016). These strategies are employed primarily in basic research studies where the volume of data requiring human annotation vastly outstrips the expert capacity or computational

sophistication to handle it. Crowdsourcing approaches are typically leveraged to digitize historic collections, classify not personally identifiable image content, and make data more accessible. *Mark2Cure* volunteers extract content from and add structured vocabulary to published biomedical research literature (Tsueng et al. 2016). Mark2Cure is notable in its foundational premise that members of the public are capable of reliably extracting information from highly technical research content; *Cochrane Crowd* employs a similar model. Other examples of classification-focused projects include *Cell Slider*, which identifies cancerous cells in slides of cell samples, and *Stall Catchers*, whose volunteers view video clips online to identify stalled blood flow in mouse neurons with the goal of advancing Alzheimer's research. Like the puzzle-based and problem-solving games, these classification and annotation projects raise few ethical concerns as long as materials are properly deidentified.

RESEARCH FOCUS

In addition to the distinctions between citizen scientist participation in data collection and data processing, the overarching nature of the research has substantive implications for ethics in health and biomedical citizen science. Research conducted with an observational focus typically raises fewer ethical flags than interventional research.

Observational Research

Observational studies, in which a citizen scientist observes a situation or organism and collects data about it, form the basis for most established citizen science and are often conducted in a contributory mode. Data collection can be distributed across a wide set of individuals, creating new opportunities for research at previously unheralded geographic, temporal, and demographic scales. The resulting data are usually then sent to a professional scientist or submitted to a database for automated analysis as part of the epistemic conversion to knowledge. Such projects often have a primary goal of producing science outputs, including traditional scholarly products and applied outcomes such as data-driven decisions. At the other end of the spectrum, co-created observational projects may operate similarly but focus on specific issues of local interest, or involve only community members without engaging a professional researcher.

In ethical terms, participation in observational research about the environment around us can pose a distinct set of risks, primarily in revealing new information about the world, such as the location of a rare species or archeological ruins, which can in turn be exploited. Secondarily, and often discussed in the citizen science practitioner community, are risks related to geolocated data associated with individual observers, for which a number of solutions have been explored and applied with good results (Bowser and Wiggins 2015).

Table 1. Classification of participation models by number of participants and source of inquiry.

	Professional driven	Public driven
Independent participation	Traditional science	N-of-1/DIY science
Collective participation	N-of-many-1's/contributory	N-of-we/co-created

When the focus is turned from the environment to individuals, the additional ethical considerations are reasonably well circumscribed for projects guided by institutional actors, but increasingly problematic for independent community-driven projects.

The known ethical issues related to sharing personal health data apply to many cases; however, the participants may not be the only "observers" in these projects. Thanks to the flexibility and popularity of existing social technologies, systems like Facebook and fitness tracker platforms are often repurposed by emergent community-based groups to address their data collection and management infrastructure needs. Individuals' observations can then reflect information back to aggregators in ways unexpected by participants. When using consumer-oriented systems for data collection rather than purpose-built (professional) scientific systems, the same data incur very different potential risks. While the trade-off is worthwhile for some individuals, for most people there isn't enough information presented in the right ways to make an informed decision about the benefits and risks of using these platforms for recording and sharing data: Monetization occurs in obscurity, many steps removed from the individual, and presented in impenetrable language, so potential consequences are rarely understood (Obar & Oeldorf-Hirsch 2018; Proferes 2017). This reflects the complete absence of informed consent in most consumer health data collection systems, typically intentionally marketed as "wellness" tools to avoid any such regulation or ethical constraint. Nonetheless, these tools offer a unique method for data collection and are rapidly driving development of citizen science projects to study health, regardless of whether they are organized top-down as in contributory models, or bottom-up in a co-created style.

Interventional Research

Interventional research, in which an intervention is made during the course of the study, is exceedingly rare in citizen science outside of health and biomedical sciences: Experimental designs (the closest approximation in other sciences) are difficult to design for effective execution by dozens to thousands of enthusiastic co-investigators. As a result, prior analysis of ethical issues in interventional citizen science is almost nonexistent. Many community-driven data collection projects in other disciplines mirror interventional models because they collect data both before and after a change, but the change is usually precipitated by an external force, such as hydraulic fracturing, rather than by the citizen scientists themselves.

By contrast, intervention is a cornerstone of health and biomedical research, and intervention signals increased risk. These projects always involve data collection tasks as a primary component of the scientific contribution and there are no examples of interventional research composed exclusively of data-processing tasks. The different nature of risk in an interventional study engenders substantial caution in contexts subject to ethical oversight, so community-driven projects may be more common, but also riskier. The nature of the ethical issues can vary substantially by the participation model, outlined in the next section.

PARTICIPATION MODELS

The three modes of participation discussed next can intersect with both observational and interventional research, but create the more substantial ethical concerns for health and biomedical citizen science due the potential consequences for participants engaging in interventional studies. These three categories reflect a combination of the source of inquiry, whether driven by a professional researcher or the public, and whether participants' experience is collective versus independent, as shown in Table 1.

N-of-1

Perhaps the most mature citizen science in health is known as "self tracking" or "N-of-1" research. Notably, this model is almost entirely absent in other scientific disciplines, where the nearest recognizable phenomenon is "DIY science" that focuses on satisfying individual curiosities (e.g., Russell 2014). It therefore fits into neither the contributory nor the co-created categories, but is philosophically in keeping with co-created projects. N-of-1 and quantified self (QS) or quantified relationship (QR) studies develop a line of inquiry to satisfy personal interests or needs, based primarily on self-tracking.

Self-tracking is what it sounds like: using a variety of means to record one's daily actions, and observing the outcomes of different actions or interventions over time. Health self-tracking is an ancient practice, and nearly 70% of Americans are believed to track at least one health indicator over time (Fox and Duggan 2013). QS studies often rely on proprietary or pay-for-play sensors and systems (Swan 2013), making them more individualist than a typical citizen science project, and particularly when they reflect on intimate activities, subject to privacy concerns (Danaher, Nyholm and Earp 2018).

Although most N-of-1 studies today are reliant on proprietary platforms that implicitly aggregate individuals' data, the citizen scientists designing these inquiries and interventions are not doing so in concert but rather as solo investigators, and it is easy to imagine "analog" versions of their research that are less reliant on technology platforms.

Since citizen science is primarily associated with collective models of participation, N-of-1 studies are less likely to be recognized as citizen science until becoming visible through coordination or sharing of results. For example, the Soylent Diet self-experimentation community shares results of N-of-1 experiments and could be considered citizen science (Dolejšová and Kera 2017). However, N-of-1 tracking for personal health made a massive incursion into the lives of many individuals with the advent of ubiquitous, cheap computing and sensing systems marketed for consumers, which raises ethical concerns about surveillance as well as potential harms from interventions. For example, fitness and calorie tracking apps have been associated with disordered eating related to self-imposed interventions based on N-of-1 data (Simpson and Mazzeo 2017), yet the same tools also support positive health interventions and outcomes for others, and the ethical responsibility for these impacts is assigned to the individuals making interventions rather than to those whose tools support the interventions.

Individuals who run N-of-1 observational and interventional studies may focus their inquiry on diets, environments, sleep, medicines, bathroom habits, and more, but rarely equate their inquiry with the citizen science phenomenon as such. They increasingly repurpose consumer-grade devices such as Fitbits, smartphones, and smart watches to track steps, workouts, and movement through space and time. These research data are increasingly uploaded to platforms such as Strava to realize their epistemic value through graphs that render patterns of change over time visible, statistical analyses and comparisons against other platform users in aggregate, and more. Public outrage over data reuse and exposure (e.g., 23andme partnership with GlaxoSmithKline; Ducharme 2018), however, confirms prior research showing that many users fail to realize the lack of protections on their data and potential uses that they agree to as part of the trade, which are demonstrably far more extensive and intrusive than discussed here (Peppet 2014) and raise specific ethical concerns for health care (Denecke et al. 2015). As these examples suggest, while N-of-1 studies focus on one individual, they represent the seeds from which a co-created collective effort could sprout (Mukhija 2010), or around which larger studies can be designed.

N-of-We

The usual complement to N-of-1 inquiry is the community-driven "N-of-we" study, which shares a co-created participation model with many localized and issue-oriented citizen science projects in other disciplines. These projects may begin in an observational mode and move to interventional, or focus primarily on interventional research.

Moving from N-of-1 studies to generalizable knowledge is a core goal for many citizen science advocates, and is increasingly achievable without the support of professional researchers as social media platforms allow people to find one another and self-aggregate into larger groups. Such emergent collaborations are often repurposed by citizen scientists and health advocates to form "N-of-we" communities where observational results and tips can be exchanged in a supportive environment; they are labeled "public-driven" in Table 1 because the major impetus for the development of these groups comes from its members, although professional health researchers have also explored the use of social networking sites for research data collection (Alshaikh et al. 2014). The dominant platform for self-organized N-of-we observational projects is Facebook (Lupton 2015), although Twitter chats also form large ad hoc communities over time (#BCSM, #HCLDR). Health data companies such as PatientsLikeMe and 23andme represent platforms specifically designed to facilitate N-of-we formation, offering groups and forums for various diseases, ancestry, and more, which are often notable for their collective spirit.

When they emerge from a community interest group without professional scientists' involvement, these studies may lack the research design rigor that is often necessary to tease out conflating factors, an ethical issue in itself: Did the Soylent itself drive weight loss, or was it exercise? How long did the benefits last, and how would one be able to predict whether the same effects would apply to another person? What are the risks of publishing a short-term, N-of-1 outcome to an unfiltered feed for amplification within an enthusiast community? Conversely, groups like Citizen Science Belleville demonstrate the ability of citizen science communities to adhere to higher research standards than many professional researchers, leveraging peer-reviewed registered research reports for replication studies to demonstrate rigor (Santos-Lang 2018).

Like N-of-1 investigations, N-of-we communities relying on repurposed social technologies are also exposed to reuse of their data in ways they often didn't imagine. Among these platforms, PatientsLikeMe works the hardest to inform their users about data reuse, but recent publicity over Facebook breast cancer group data breaches (Fazzini and Farr 2018) and 23andme's pharmaceutical development deals (Ducharme 2018) represent only the initial wave of outrage over unexpected data uses from participants who technically consented but were not truly informed. Social technology companies by nature are constantly pressured to monetize contributed data, and to profile users with more and more accuracy for their other customers, corporate data users

(Zuboff 2015). Thus N-of-we research is constantly at risk of inappropriate or unwanted data reuse unless using its own tools or facilitated by platforms with significantly different governance and business models, such as a data cooperative, for example, Savvy (Anzilotti 2018), or a nonprofit such as Sage Bionetworks. Co-created citizen science in other disciplines often thrives without the use of social technologies, but this is often because the focus of its research is geographically localized enough to make in-person meetings feasible, or small enough in scale that using spreadsheets for data collection and aggregation is reasonable. These conditions are less likely to be true of health and biomedical citizen science projects formed around shared interests.

N-of-Many-1's

In contrast to primarily public-driven forms of citizen science, an emergent model integrates traditional observational research, such as a longitudinal cohort study, with the contributory citizen science methods popular in other disciplines. This model does not offer a community-driven full-process research experience like the N-of-we studies, but instead standardizes methods and aggregates individual experiences so that the N-of-1 data can aggregate up into more epistemologically conventional research analyses (Nafus 2017), typically carried out by professional scientists.

When the projects follow a more typical contributory citizen science participation model where research advantages for traditional science are fairly clear, the ethical question of "who benefits?" is forefronted and rewards to contributors are explicitly considered. Some of these projects offer personalized services like those in the N-of-1 space (customized or individualized data displays are most common), others provide access to raw data for export, and still others provide no explicit rewards to participants beyond the intrinsic satisfaction of contributing to scientific research. The AllofUs Research Program and the UK Biobank Initiative are perhaps the largest of these, but other initiatives include Google's Project Baseline, Geisinger's MyCode, and more. Some of these projects arise from efforts to capture potential scientific value of aggregating across N-of-1's, with platforms assisting by nudging participants to add data in standard formats for aggregation, and providing rewarding data outputs. This is a specific goal of research platform technologies such as Apple's ResearchKit, Android's ResearchStack, and is an implicit goal in many digitally enabled "traditional" research studies.

Most projects discussed in the preceding as examples of observational data collection projects resemble this model. In contrast to the American Gut Project, uBiome is a seemingly similar microbiome research venture marketed as a citizen science project, essentially providing a direct-to-consumer health service much like 23andme.

Pay-to-play services like uBiome and 23andme are rarely considered citizen science, potentially because they are marketed primarily for self-discovery rather than for collective development of scientific knowledge. The price point for these services may also substantially bias access and participation, which is often incompatible with core values of inclusion that are central to most citizen science. When the price point is low enough to enable independent citizen science, the sponsoring service may have surveillance capitalism imperative (Zuboff 2015), much like the social technology platforms, to use obscure terms of service and aggregate individuals' detailed and potentially revealing data into a larger, sellable database. Deeply personal data disappear into the market for resale when these companies go out of business, often with no notice to the individuals, creating new and often unknowable risks over time. In these cases, the question of "who benefits" is only partially answered for participants, who are rarely exposed to adequate information to make a truly informed decision about the risk-benefit tradeoffs of participation.

Such N-of-many-1's data collection projects are close to being considered "not citizen science" because recruitment and enrollment of participants are often conducted in ways that do not clearly disclose details around data access and participant benefits prior to registration, itself a common and prominent ethical problem in citizen science more broadly. The issue is often one of marketing versus reality. If volunteers are limited to donating personal data, the project essentially follows a standard human subjects research study model with uncompensated, self-selecting participants, which both is ethically fraught and threatens research quality. In reality, many (but not all) actually do provide access to data and information with potential to support participants' own independent inquiry, but offer little actual support for participation beyond data donation. Better supporting citizen science inquiry requires systemic intervention to overcome the new digital divide of data, and not just technocratic approaches such as better interfaces and visualizations for data exploration and analytical tools. For example, teaching data science skills in high schools and community colleges, and facilitating their use in communities, would better support participants in pursuing their own investigations while also supporting scientists.

The emergence of smartphone app-based research platforms also offers increasing opportunities for both N-of-1 and N-of-many-1's research by aggregating data through more centralized study providers (Rothstein, Wilbanks, and Brothers 2015), which can include traditional research institutions and groups more clearly aligned with citizen science, like patient advocacy organizations. The N-of-many-1's contributory platforms also exhibit a wider range of governance models than those often seen in N-of-we studies. AllofUs, Baseline, and MyCode each build on a very traditional bioethics basis, with independent review board oversight and in-depth

informed consent processes more akin to traditional research study enrollment than the cursory terms of service used by platform data companies.

But the day-to-day experience of an N-of-many-1's research participant—whether using platforms or apps—is typically an N-of-1 experience, since the traditional goal of protecting participant privacy in formal studies means most providers are not allowed to facilitate meaningful connections between participants, which is a common motivation for engagement in citizen science. By extension, these projects are rarely able to provide avenues for individual inquiry and participation beyond the role of a subject, precisely because of the traditional institutional ethics paradigms in which they operate, making them questionable as "citizen science" since they still closely resemble traditional subject studies.

Mixed Models

Compared to the foregoing examples, the Denver Museum of Nature & Science essentially stands alone with its Genetics of Taste Lab project (Garneau et al. 2014), which enlisted members of the public in multiple stages of a large-scale ongoing biomedical study under standard institutional oversight. Museum visitors provided crowdsourced health data, while trained citizen science volunteers enrolled and consented participating museum visitors; processed, prepared, and sequenced the DNA samples; and worked with research staff on analysis. The project features top-down inquiry and research design with in-depth engagement of volunteers throughout the remaining steps of the inquiry process. For most members of the public, this project is an N-of-many-1's experience, but for the volunteers who assist in data collection and study execution, it resembles a traditional lab study where they take on the role of research assistants. The ethical considerations in this type of basic research project are not much different from those of a mirror case where the work is carried out exclusively by professional researchers. As other mixed models of citizen science engagement emerge, they are likely to carry a dual burden for ethics, inheriting the ethical considerations for each component participation model that they incorporate.

BEYOND BELMONT: ETHICAL PRINCIPLES IN CITIZEN SCIENCE

Modern professional health and biomedical research are governed by a patchwork quilt of laws and regulations intended primarily to protect research participants from potential abuses by scientists and institutions. The Belmont Report (NCPHSBBR 1978) was written more than 40 years ago, predating the modern social technologies that are fundamentally shifting participatory research, and it still dominates the bioethics landscape. The report created guidelines for research involving human subjects based on three essential principles: respect for persons, beneficence, and justice. Its findings influence U.S. federal laws such as the Common Rule and provide an overall framework for ethics review by independent boards. Commonly, research scientists and institutions take the principles to require informed consent for participants, an honest and transparent assessment of risks and benefits by researchers, and an equitable process to select and enroll participants. In addition to these standard categories, Fiske et al. (2019) also raise the issues of social or scientific value and scientific validity as additional concerns, noting that these are typically addressed in top-down researcher-led projects and are more likely to be problematic in community-driven projects. Vayena et al. further reinforce the need for a new "social contract" for participant-led research (Vayena et al. 2016).

In citizen science the underlying principles are the same: The open and voluntary nature of participation forefronts autonomy, but does not relieve researchers of the responsibility of informing participants about the risks and benefits of participation. However, there are certain ethical elements introduced inside these concepts when citizens themselves are driving the research or take on core contributor roles beyond serving as subjects. For example, all three of these principles support expectations for collegial relationships, rather than clinician–patient or researcher–subject interactions, which incur responsibilities for clearly communicating research progress and results, along with faithfully completing the planned work. While developing and maintaining collegial relations with a large pool of contributors carries its own challenges, the benefits can be substantial: For example, patient communities sometimes identify more interesting or relevant questions than researchers (as do citizen science contributors), advancing the pace of research and availability of solutions (Chiauzzi et al. 2015; Wicks and Little 2013). At the same time, community-led projects may engage their members effectively while overlooking some of the customary disclosures expected of professionals.

These principles further extend into still-evolving practices around attribution and credit for work performed, which are typically represented as professional obligations between scientists but now extend to citizen science partners as well. Current conversations around data ownership, access, and acknowledgment do not necessarily reflect the concerns of participants, many of whom are satisfied to entrust the project team with stewardship of the data (Ganzevoort et al. 2017), but participants' perspectives on the distribution and use of data they share are far from uniform and far from neutral. Multiple strategies for allocating credit are being tested in citizen science with no clear consensus on a "best" model for attribution, but the topic cannot be ignored without consequences, not least because participant-generated content can trigger intellectual property

considerations. The resulting questions around data ownership and apportioning credit further complicate the landscape (Riesch and Potter 2014), although much of the inherent issue is about control and access rather than ownership (Evans 2016).

These issues of access and benefit sharing also represent an expansion of traditional concepts of beneficence: The citizen science practitioner community as a whole is deeply concerned with ensuring inclusivity, which cannot be adequately entertained without considering the costs of participation and direct benefits to participants. As a result, project designs typically emphasize minimal barriers to participation, low-cost equipment, and increasing focus on issues of vital importance to underrepresented communities whose members do not have the wealth and leisure to participate in science as a hobby. These expansions on the otherwise limited interpretations of institutional ethics bodies have previously been noted as critiques of the Belmont Report by advocates of community-based participatory research (Shore 2006). Likewise, new considerations for research with online data sources are emerging, raising such questions as whether social media platforms' terms of service can be considered a proxy for informed consent (Vitak, Shilton, and Ashktorab 2016).

Citizen health science also introduces a variety of new and emergent issues that can and should be part of the ethics conversation as part of the "ethics gap" between traditional research and citizen science (Rasmussen 2016). As noted earlier, wearables manufacturers, platform providers, and data aggregators regularly leverage faulty models of consent that would never pass independent ethics review, and may deliberately position themselves to avoid such review. The impacts of weak, obscure consent models are amplified when they are ignored not just by the average consumer but also by citizen scientists running more and more powerful N-of-1 studies, aggregated into N-of-we or N-of-many-1's research. The more that a consumer tool is reused this way, the greater is the potential for unexpected or harmful consequences from sale and reuse of data, as it is nearly impossible to fully predict future data uses.

Another emergent issue not captured by the Belmont framework is a side effect to repurposing social media and wellness platforms to aggregate into N-of-we: Citizen scientists can publish their observations and results directly to millions without a filter. And while the peer review process of scientific publishing combined with the slow pace and high cost of clinical study has limited broader participation in science, it also often blocked the spread of misinformation. From vaccine denialism to Goop's pseudoscience, there is a real health risk from the spread of faulty information emerging from bad science, regardless of where it originates. And as predictable as this potential risk may be, there are no standards of ethical accountability for platform providers whose infrastructures enable independently organized community-based health studies, and no clear solutions for preventing the diffusion of problematic interpretations of health information.

In addition, we highlight the false dichotomy of N-of-1 versus N-of-we models for public engagement in health and biomedical citizen sciences: Involvement of the public in solo and in group modes in projects where participants contribute at all stages of the research are not the only options. In fact, they are among the least common models in other sciences. The N-of-many-1's represents a complementary participation model more congruent with citizen science successes in other disciplines. In these projects, participants take part in a carefully designed study with institutional oversight or extrainstitutional ethics review, but also have opportunity to engage with the research in a capacity beyond that of simple subjects. The landscape of such projects is increasingly well developed in other disciplines, where reporting on oneself is not the focus of data collection, but woefully limited in health and biomedical sciences due in part to the mismatch of current institutional oversight paradigms to truly participatory inquiry. Most of the examples in this article that resemble the N-of-many-1's model actually offer very little opportunity for participation beyond the role of a research subject: Either the citizen science label is potentially misleadingly co-opted or the adaptations from a traditional research model to accommodate participants as co-researchers are underdeveloped. Fully harnessing the power of the public in health and biomedical research will require developing a different model of ethical oversight that recognizes the autonomy of participants as a given, continues to ensure that they are well informed on an ongoing basis as the research evolves, and also addresses the problematic assumptions baked into an ethics framework that assumes docile subjects instead of engaged co-researchers.

CONCLUSION

This article identifies several dimensions of citizen science that point to a variety of ethical considerations, many of which are currently overlooked, to open an overdue conversation around ethical engagement of the public in science. These issues are ripe for discussion as there are few clear solutions in sight despite the growth of citizen science in health and biomedical research. We also highlight an imperative for pluralistic approaches, since current ethics regimes are often mismatched to the fundamentals of citizen science. Pluralistic ethics implies a need for not only multiple modes of ethical engagement, but also effective strategies for determining which ones apply to a given situation, which is no small challenge given the growing diversity of citizen science models. We also encourage researchers and ethicists to acknowledge that the Pandora's box of citizen science

has already been flung open wide, and ignoring the consequences would be denying the realities of a changing research landscape. Despite the challenges posed by citizen science, its payoffs can be substantial, and hold potential to guide the evolution of research ethics in an ultimately rewarding direction.

CONFLICTS OF INTEREST

FUNDING

ORCID

Andrea Wiggins (iD) http://orcid.org/0000-0003-4082-4138
John Wilbanks (iD) http://orcid.org/0000-0002-4510-0385

REFERENCES

Alshaikh, F., F. Ramzan, S. Rawaf, and A. Majeed. 2014. Social network sites as a mode to collect health data: a systematic review. *Journal of Medical Internet Research* 16(7): e171.

Anzilotti, E. 2018, July 27. This co-op lets patients monetize their own health data. *Fast Company*. https://www.fastcompany.com/90207550/this-co-op-lets-patients-monetize-their-own-health-data

Argyris, C., and D. A. Schön. 1989. Participatory action research and action science compared: a commentary. *American Behavioral Scientist* 32(5): 612–623. doi: 10.1177/0002764289032005008.

Baâzaoui, A., W. Barhoumi, and E. Zagrouba. 2017. Towards semantic visual features for malignancy description within medical images. In 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), 397–402. IEEE.

Baum, F., C. MacDougall, and D. Smith. 2006. Participatory action research. *Journal of Epidemiology and Community Health* 60(10): 854–857. doi: 10.1136/jech.2004.028662.

Bonney, R., C. B. Cooper, J. Dickinson, et al. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* 59(11): 977–984. doi: 10.1525/bio.2009.59.11.9.

Bowser, A., and A. Wiggins. 2015. Privacy in participatory research: Advancing policy to support human computation. *Human Computation* 2(1): 19–44.

Chiauzzi, E., C. Rodarte, and P. DasMahapatra. 2015. Patient-centered activity monitoring in the self-management of chronic health conditions. *BMC Medicine* 13(1): 77.

Cooper, C., J. Dickinson, T. Phillips, and R. Bonney. 2007. Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society* 12(2): 11.

Danaher, J., S. Nyholm, and B. D. Earp. 2018. The quantified relationship. *The American Journal of Bioethics* 19(2): 3–19.

Den Broeder, L., J. Devilee, H. V. Oers, A. Jantine Schuit, and A. Wagemakers. 2016. Citizen science for public health. *Health Promotion International* 33(3): 505–514.

Denecke, K., Bamidis, P. C., Bond E., et al. 2015. Ethical issues of social media usage in healthcare. *Yearbook of Medical Informatics* 24(01): 137–147. doi: 10.15265/IY-2015-001.

Dolejšová, M., and D. Kera. 2017. Soylent diet self-experimentation: Design challenges in extreme citizen science projects. In Proceedings of the 2017 ACM International Conference on Computer Supported Cooperative Work and Social Computing, 2112–2123. ACM.

Ducharme, J. 2018. July 26. A major drug company now has access to 23andMe's genetic data. Should you be concerned? *Time*. http://time.com/5349896/23andme-glaxo-smith-kline/.

European Citizen Science Association. 2015. Ten principles of citizen science. European Citizen Science Association: Berlin, Germany. https://ecsa.citizen-science.net/sites/default/files/ecsa_ten_principles_of_citizen_science.pdf.

Evans, B. J. 2016. Barbarians at the gate: consumer-driven health data commons and the transformation of citizen science. *American Journal of Law and Medicine* 42(4): 651–685. doi: 10.1177/0098858817700245.

Fazzini, K., and C. Farr. 2018, July 12. Facebook recently closed a loophole that allowed third parties to discover the names of people in private, 'closed' Facebook groups. *CNBC*. https://www.cnbc.com/2018/07/11/facebook-private-groups-breast-cancer-privacy-loophole.html

Fiske, A., L. Del Savio, B. Prainsack, and A. Buyx. 2019. Conceptual and ethical considerations for citizen science in biomedicine. In N. Heyen, S. Dickels, and A. Brüninghause eds., *Personal health science*, 195–217. Wiesbaden, Germany: Springer VS.

Fox, S., and M. Duggan. 2013. Tracking for health. *Pew Research Center's Internet & American Life Project*. http://www.pewinternet.org/2013/01/28/tracking-for-health/

Ganzevoort, W., R. J. van den Born, W. Halffman, and S. Turnhout. 2017. Sharing biodiversity data: citizen scientists' concerns and motivations. *Biodiversity and Conservation* 26(12): 2821–2837. doi: 10.1007/s10531-017-1391-z.

Garneau, N. L., T. M. Nuessle, M. M. Sloan, S. A. Santorico, B. C. Coughlin, and J. E. Hayes. 2014. Crowdsourcing taste research: genetic and phenotypic predictors of bitter taste perception as a model. *Frontiers in Integrative Neuroscience* 8: 33. doi: 10.3389/fnint.2014.00033.

Gilski, M., M. Kazmierczyk, S. Krzywda, et al. 2011. High-resolution structure of a retroviral protease folded as a monomer. *Acta Crystallographica Section D Biological Crystallography* 67(11): 907–914. doi: 10.1107/S0907444911035943.

Israel, B. A., A. J. Schulz, E. A. Parker, and A. B. Becker. 1998. Review of community-based research: assessing partnership approaches to improve public health. *Annual Review of Public Health* 19(1): 173–202. doi: 10.1146/annurev.publhealth.19.1.173.

Lupton, D. 2015. Lively data, social fitness and biovalue: The intersections of health self-tracking and social media. In *The sage handbook on social media*, eds. Jean Burgess, Alice Marwick, and Thomas Poell. London, UK and New York, NY: Sage.

McKinley, D. C., A. J. Miller-Rushing, H. L. Ballard, et al. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 208: 15–28. doi: 10.1016/j.biocon.2016.05.015.

Miller-Rushing, A., R. Primack, and R. Bonney. 2012. The history of public participation in ecological research. *Frontiers in Ecology and the Environment* 10(6): 285–90. doi: 10.1890/110278.

Mudie, L. I., X. Wang, D. S. Friedman, and C. J. Brady. 2017. Crowdsourcing and automated retinal image analysis for diabetic Retinopathy. *Current Diabetes Reports* 17(11): 106. doi: 10.1007/s11892-017-0940-x.

Mukhija, V. 2010. N of one plus some: an alternative strategy for conducting single case research. *Journal of Planning Education and Research* 29(4): 416–426. doi: 10.1177/0739456X10362770.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (NCPHSBBR) 1978. *The belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Bethesda, MD.

National Institutes of Health 2015. Trans-NIH workshop to explore the ethical, legal and social implications (ELSI) of citizen science. https://www.genome.gov/27559982/elsi-transnih-workshop-to-explore-the-ethical-legal-and-social-implications-of-citizen-science/ (accessed March 9, 2018)

Nafus, D. 2017. N of many ones: Creating space for alternative data flows. Lecture presented December 4, 2017 at Algorithmic Fairness & Opacity Lecture Series in University of California Berkeley School of Information. Berkeley, CA.

Obar, J. A., and A. Oeldorf-Hirsch. 2018. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*: 1–20. doi: 10.1080/1369118X.2018.1486870.

Peppet, S. R. 2014. Regulating the internet of things: first steps toward managing discrimination, privacy, security and consent. *Texas Law Review* 93: 85.

Proferes, N. 2017. Information flow solipsism in an exploratory study of beliefs about twitter. *Social Media + Society* 3(1): 2056305117698493. doi: 10.1177/2056305117698493.

Rasmussen, L. M. 2016. Filling the 'Ethics Gap' in citizen science research: A workshop report. https://www.niehs.nih.gov/research/supported/translational/peph/webinars/ethics/rasmussen_508.pdf

Riesch, H., and C. Potter. 2014. Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Understanding of Science* (Bristol, England) 23(1): 107–120. doi: 10.1177/0963662513497324.

Rothstein, M. A., J. T. Wilbanks, and K. B. Brothers. 2015. Citizen science on your smartphone: An ELSI research agenda: Currents in contemporary bioethics. *The Journal of Law, Medicine & Ethics* 43(4): 897–903. https://doi.org/10.1111/jlme.12327

Russell, S. A. 2014. *Diary of a citizen scientist*. Corvallis, OR: Oregon State University Press.

Santos-Lang, C. 2018, August 20. Citizen science registered report passes peer-review. *Medium.com*. https://medium.com/citizen-science-belleville/citizen-science-registered-report-passes-peer-review-fde350b265bc

Shirk, J. L., H. L. Ballard, C. C. Wilderman, et al. 2012. Public participation in scientific research: A framework for deliberate design. *Ecology and Society* 17(2). doi: 10.5751/ES-04705-170229.

Shore, N. 2006. Re-conceptualizing the belmont report: A community-based participatory research perspective. *Journal of Community Practice* 14(4): 5–26. doi: 10.1300/J125v14n04_02.

Simpson, C. C., and S. E. Mazzeo. 2017. Calorie counting and fitness tracking technology: Associations with eating disorder symptomatology. *Eating Behaviors* 26: 89–92. doi: 10.1016/j.eatbeh.2017.02.002.

Swan, M. 2013. The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data* 1(2): 85–99. doi: 10.1089/big.2012.0002.

Swanson, A., M. Kosmala, C. Lintott, and C. Packer. 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology: The Journal of the Society for Conservation Biology* 30(3): 520–531. doi: 10.1111/cobi.12695.

Tsueng, G., S. Nanis, J. Fouquier, B. Good, and A. Su. 2016. Citizen science for mining the biomedical literature. *Citizen Science: Theory and Practice* 1(2). doi: 10.5334/cstp.56.

Vayena, E., R. Brownsword, S. J. Edwards, et al. 2016. Research led by participants: A new social contract for a new kind of research. *Journal of Medical Ethics* 42(4): 216–219.

Vayena, E., and J. Tasioulas. 2013. Adapting standards: Ethical oversight of participant-led health research. *PLoS Medicine* 10(3): e1001402.

Viswanathan, M., A. Ammerman, E. Eng, et al. 2004. Community-based participatory research: Assessing the evidence: Summary. https://www.ncbi.nlm.nih.gov/books/NBK11852/

Vitak, J., K. Shilton, and Z. Ashktorab. 2016. Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, 941–953. ACM.

Wallerstein, N. B., and B. Duran. 2006. Using community-based participatory research to address health disparities. *Health Promotion Practice* 7(3): 312–23. doi: 10.1177/1524839906289376.

Walter, M. 1998. Participatory action research. *Reason* 71: 1–8.

Wandersman, A. 2003. Community science: Bridging the gap between science and practice with Community-Centered models. *American Journal of Community Psychology* 31(3–4): 227–242. doi: 10.1023/A:1023954503247.

Wicks, P., and M. Little. 2013. The virtuous circle of the quantified self: A human computational approach to improved health outcomes. In P. Michelucci, ed., *Handbook of human computation*, 105–129. New York, NY: Springer.

Wicks, P., M. Massagli, J. Frost, et al. 2010. Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Internet Research* 12(2): e19.

Woolley, J. P., M. L. McGowan, H. J. Teare, et al. 2016. Citizen science or scientific citizenship? disentangling the uses of public engagement rhetoric in national research initiatives. *BMC Medical Ethics* 17(1): 33.

Zuboff, S. 2015. Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology* 30(1): 75–89.