

Digital Orphans: Data closure and openness in patient-powered networks

This is the author-accepted manuscript of the paper published in Biosocieties.

Please cite as follows:

Tempini, N., Del Savio, L., in press. Digital Orphans: Data closure and openness in patient-powered networks. *BioSocieties*.

Niccolò Tempini (corresponding author)

University of Exeter

Byrne House

St German's Rd

Exeter (Devon) EX4 6TJ

United Kingdom

n.tempini@exeter.ac.uk

Lorenzo Del Savio

Christian-Albrechts-Universität zu Kiel, and King's College of London

lorenzo.delsavio@iem.uni-kiel.de

Word count: 11134 (9078 excluding references but including abstract and endnotes)

Niccolò Tempini is Research Fellow in Data Studies at Egenis, University of Exeter. He researches big data medical research and digital infrastructures, investigating the specific knowledge production economies, organization forms and data management innovations that these projects engender with an interest in their social and epistemic consequences. More information at www.tempini.info.

Lorenzo Del Savio earned his PhD in "Life Sciences: Foundations and Ethics" at the University of Milan. He did postdoctoral work on the "Social, ethical and regulatory aspects of citizen science" at the University of Kiel (Grant 01GP1311 of the German Federal Ministry of Education and Research).

Acknowledgements

We are indebted to the anonymous reviewers and the editor, who with their supportive and constructive comments helped us to better clarify and highlight the argument of the article. We would like to also thank friends and colleagues who have offered valuable comments and suggestions on early drafts of this paper. We would like to especially thank Barbara Prainsack, Sabina Leonelli, Alena Buyx, and David Teira. This research is funded by the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement number 335925, and the German Federal Ministry of Education and Research (grant number 01GP1311).

Digital Orphans: Data closure and openness in patient-powered networks

ABSTRACT

In this paper, we discuss an issue linked to data-sharing regimes in patient-powered, social-media-based networks, namely that much of the data that patient users share are not used to research scientific issues or the patient voice. This is not a trivial issue, as participation in these networks is linked to openness in data sharing, which would benefit fellow patients and contributes to the public good more generally. Patient-powered research networks are often framed as disrupting research agendas and the industry. However, when data that patients share are not accessible for research, their epistemic potential is denied. The problem is linked to the business models of the organisations managing these networks: models centred on controlling patient data tend to close networks with regard to data use. The constraint on research is at odds with the ideals of a sharing, open and supportive epistemic community that networks' own narratives evoke. This kind of failure can create peculiar scenarios, such as the emergence of the '*digital orphans*' of Internet research. By pointing out the issue of data use, this paper informs the discussion about the capacity of patient-powered networks to support research participation and the patient voice.

Keywords: patient-powered; orphan diseases; social media; data openness; patient participation; participatory research.

Social media and biomedical research

Over the past few years, we have witnessed how web-based digital technologies, social media and networks such as forums, blogs, wikis and sharing sites change the way in which social information is produced, shared and accessed (Gerlitz and Helmond, 2013; Tempini, 2015; van Dijck, 2013). These institutional and social changes contribute to the reconfiguring of the foundations of the modern diversification of expertise and the division of labour (Callon et al., 2009; Wynne, 1996; Del Savio, 2015). Especially in terms of the web, the existence of an increasing number of organisations that base their work processes and expert services on the contributions of a distributed user base seems to question the ways in which the relationship between public and professional experts has been structured in the production, retrieval, interpretation and use of personal data (Kallinikos and Tempini, 2014; Prainsack, 2014a). Hence, it has the potential to reconfigure the roles that citizens play as social interaction is translated through digital infrastructures.

One domain of application for such infrastructures is biomedical and health research, which some claim they are pervasively transforming (Nielsen, 2012; O'Connor, 2013; Swan, 2012; Lee and Cawley, 2009; Eysenbach, 2008). Ubiquitous devices for the collection of data at an individual level (Lucivero and Prainsack, 2015) and the emergence of online health-related services and data analytics are seen as paving the way for more precise healthcare and accelerating clinical discovery (Topol, 2012), as well as offering greater resources to promote citizen empowerment and autonomy (Prainsack, 2014b; Wyatt et al., 2013a). This convergence of the imaginaries of big data, precision medicine and participation is a significant trend in how contemporary innovation in health is portrayed, and one that has attracted the attention of policymakers and analysts (Richards et al., 2015; Collins and Varmus, 2015).

In this panorama, a set of *social-media-based* health research networks stands out, also called *patient-powered networks*; well-known examples are PatientsLikeMe and CureTogether. These enterprises claim to have the potential to accelerate medical research and, at the same time, broaden its field of participation by opening the research agenda to the priorities of patients, who are the ultimate recipients of scientific discoveries and innovations. Patients of chronic conditions have considerable medical expertise, including experiential knowledge and an in-depth acquaintance with scientific literature; it is not surprising, therefore, that online patient expert groups have first emerged among chronic patients (Fox et al., 2005). With the help of such networks, some patient groups have participated in, and sometimes initiated, systematic patient experience research. This has included off-label self-experimentation with drugs that eventually resulted in top-tier scientific journal publications (e.g. Wicks et al., 2011). According to arguments advanced by their leaders (Wicks et al., 2014; Wicks and Heywood, 2014), these networks continue, in a new organizational form, a trend of patient activism and participation in research that started to emerge well before the advent of social media (e.g. Epstein, 1996; Feenberg et al., 1996; Rabeharisoa and Callon, 2002). Patients are invited to engage in “data donation” for the benefit of others (Ruckenstein and Schüll, 2017). The rhetoric of participation purports a continuation of the trajectory of radical patient movements fighting for empowerment, especially in the case of ‘orphan diseases’, conditions (rare or disputed) that are unable to mobilise the attention and investments necessary to develop effective treatments. This argument seems to complement the observed trend of patient organisations shifting their attention from the care and management of illnesses to the active production of knowledge, a phenomenon that has been described as “evidence-based activism” (Rabeharisoa et al., 2013). Social observers are interested in understanding to what extent these networks can deliver on their promises (Prainsack, 2014a, 2017).

We start from the basic assumption that social-media-based, patient-powered research networks (health networks that “funnel” data towards research purposes) are organised around the generation, collection, interpretation and operationalisation of self-reported data. They adopt an innovative organisational model based on the systematic involvement of the ‘lay’ patient in the crucial tasks of observation and data collection (Kallinikos and Tempini, 2014). To understand the social and political consequences of these initiatives, then, one must ‘follow the data’. It is by tracing the data journeys (Leonelli, 2016; Büscher et al., 2014) that we can answer pressing questions. *Who gets to use the data? For what use, and what are they able to achieve? Who decides who gets access? What conditions shape the epistemic life of patient-generated data?* If we want to seriously consider the issue of patient-generated data as epistemic and political resources, we need to ask these kinds of questions. The conditions shaping what can be done with patient-generated data, in turn, shape how the patient voice is expressed (and therefore its socio-scientific value). Critical studies based on a critique of participation as labour (e.g. Cooper, 2012; Lupton, 2013) have advanced insightful questions of power, exploitation and value capture in the new media (Ruckenstein and Schüll, 2017), but what we want to pay attention to here is whether patient voice and contributions *as represented in data* are in fact used to create some other form of value through scientific research, or not. In other words, if patients “donate” data for a cause but there is no consequence, we think this is worth reflecting upon.

The article argues that the use of the data, for research or the expression of the patient voice, that participatory health social media networks routinely collect is often constrained – in ways that are at odds with the ideals of a sharing, open and supportive epistemic community that the

narratives of the network proponents evoke. We will argue that an important factor in conjuring up this contradiction is the difficulty of reconciling, in an inclusive data governance model, of other heterogeneous requirements such as patient protection and quality control and, most importantly, the principle of monopoly on data use that the business models of many Web 2.0 networks are centred on. The high expectations that the participatory visions of these networks elicit risk being frustrated, as these ventures can generate new divides. These divides resemble those observed in the health research economies that pre-dated the advent of the web, with the marginalization of underserved 'orphan disease' communities, yet they can be configured differently, along redrawn boundaries. Because of the dynamics that we observe to the economies of digital health research, this industry as well can generate its own orphan communities, which here we refer to as the *digital orphans*.¹ The paper is structured as follows. In the next section, we provide a brief review of certain sources in the broader discussion on health social media. We then touch upon examples of patient-powered networks that can be related to the concerns of this paper. We discuss our overview and reflect on some potentially promising scenarios in the concluding sections.

The argument, general in kind, that we present in this article is mainly built from publicly available information, and from a review of scholarly literature. Publicly available information includes Internet resources (e.g. corporate webpages, news outlets) and case studies of patient-powered networks published in peer-reviewed journals, including authors' own. The cases were selected as representative examples of patient-powered networks: the data that patients share will be consequentially used for some kind of application (esp. research) that the networks promise to pursue; and the only relationship occurring between the network-owner organization and the patient users is one based on data sharing. There is no financial transaction or sale of products and services between these two parties. Instead, the patient-powered networks we are interested in are aiming to be sustainable without requiring from the patients anything else than their data (and the effort it takes to share them). In these cases, we think, the contrast that we are highlighting is perhaps starker. Among the broader group of suitable cases, we selected cases for which information about their business model is available, and that cover different areas in the spectrum of initiatives: the first two networks are focused on consumer feedback services, the other two on data-based community building. We define them further in what follows.

Participation and openness in social media-based health networks

In response to a growing interest in patient-powered networks, scholars have often focused on the issues of patient protection and possible harms (due to unsupervised and risky experiments conducted through these networks and a lack of ethical monitoring) and on the methodological limitations of employing patient-generated data (e.g. Vayena et al., 2015; O'Connor, 2013; Vayena and Tasioulas, 2013a, 2013b; Janssens and Kraft, 2012; Lee and Cawley, 2009). When issues involving the control of data emerge, they are often in relation to possible misuse and not the potential conflicts of interests that can emerge in these initiatives. Web-based health data initiatives are frequently supported by arguments of empowerment, democratisation and citizen participation (Kelty and Panofsky, 2014). Their potential lies in empowering individuals by decreasing reliance on authoritative biomedical professionals in the generation, access, interpretation and clinical use of medical data (Feenberg et al., 1996; Kallinikos and Tempini, 2014; Prainsack, 2014b; Rabeharisoa et al., 2013; Salter et al., 2015; Wyatt et al., 2013a), as well as increasing patient autonomy, especially in combination with the emergence of direct-to-

consumer diagnostic services (Alison et al., 2012). However, it is important to pay attention to potential conflicts among the actors participating in these networks (i.e. patients from different communities, managers of these research networks, scientists, investors and the industry). In this respect, close cooperation with the industry can be a way of increasing innovation in research and services, through a faster feedback loop in healthcare. Data from patient-powered networks have been used for investigating comparative effectiveness and quality of life questions, but there have been also ‘virtual trials’ (Wicks et al. 2011). The transfer of data from patients to companies – data that can then be developed to deliver cutting-edge, better-targeted services and care back to patients – can be mutually advantageous. Innovation might especially benefit research on ‘orphan’ diseases – i.e. those diseases excluded from ‘standard’ research pipelines (Nielsen, 2012; Sablinski, 2014; von Hippel, 2005). The existence of orphan diseases is a notorious failure of market-based research structures, and hence the resolution of this cul-de-sac would be powerful evidence in favour of patient-powered research. Besides these beneficial research scenarios, an active involvement of patients in their own health management can help decrease the costs of healthcare (Richards et al., 2015; Robert et al., 2015). Social-media-based interaction can be valuable and, by and large, help integrate the uneven levels of care that patients have access to (Alison et al., 2012; Prainsack, 2014b).

The argument for a participatory turn enabled by digital technologies has not gone unchallenged, both in the health domain and elsewhere. Authors have questioned organisational arrangements, typical of the many crowdsourcing models, that assume free sourcing of labour. Lupton (2013) argues that the commodification of “patient opinion” is especially problematic when not accompanied by a full disclosure of financial interests to the network members. While she acknowledges that patients are now able to express their views and offer more effective feedback to care and research, she also highlights how new forms of exploitative valuation are emerging that sustain an extractive economy based on sharing and ‘prosumption’ (also Bonsu and Darmody, 2008; Thorpe, 2010; Thorpe and Gregory, 2010). Looking at the organisation of drug development research conducted through clinical trials and social-media-based networks, Cooper (2012) argues that the managers of these social-media-based networks – by retaining proprietary control on datasets and accordingly exploiting intellectual property regimes – are both betraying the rhetoric of openness that commonly supports web-based initiatives and exploiting participants as unpaid labourers.ⁱⁱ However, whether social interaction and data collection intermediated through social media technology qualifies as exploited labour is a debated issue (Ekbia and Nardi, 2015; Nardi, 2015; Rigi and Prey, 2015). The risk is to ignore how networks enable the public sphere (Jin and Feenberg, 2015), and how they are underscored by a wider set of values (Riso et al. 2017). Not all the value created by social media platforms from user data can be commodified and alienated from the users who contribute to their creation (Tempini, 2017; Heaton and Proulx, 2015). Social media might perhaps be understood as sourcing an organizational function outside of labour relation. Ekbia (2016), for instance, argues that labourers in social media are the waged staff.

Our argument is also interested in the political economy of these networks; however, we take a different perspective. We are interested in asking whether the effort in openness that is required of patients about their own lives (cfr. ‘shareveillance’ – see Birchall, 2017)ⁱⁱⁱ can be adequately matched by an effort by the network in opening data governance and in putting the data to use. What we argue is an interesting vantage point from which to consider this relationship: does the effort that patients put in for the ‘datafication’ of their own lives – for

purposes that, among others, include the public good – entitle them to a stake in the data's onward mobilisation and use in research? And, indeed, issues related to the right to self-determination already shape patients' stakes in respect to how the data they share are used (Aicardi et al., 2016). The issue we are drawing on here, in other words, is whether we should think of the failure to make use of systematic records of the patient voice as a problem; whether patients should allow for their data "donation" to be left to no consequence, or not.

In the world of social data and crowdsourced production, the assumption often seems to be that no data will go to waste and that all data and user contributions are productively used. Contrary to this, we do not think that the reuse of data happens seamlessly, and we argue that patient data are often not put to use as patients could hope. One well documented reason for this is that patient data are difficult to use, first because of the many potential sources of bias that can interfere in the representation of the medical situation at hand (issue of data quality), and second because of the intricacies involved in successfully resituating data in new contexts of use (issues of documentation and collaboration organization). This point has been put forward regarding data shared within specialist scientific communities – for instance, by Leonelli (2016; 2014) – but it is particularly sensitive in the case of distributed data collection of patient contributions via social media (Wyatt et al. 2013b): the standards and practices of patient recording can be very variable yet difficult to scrutinize, making it very labour-intensive to consolidate the data in aggregated that can be reliably used as evidence (see Tempini, 2015; Kallinikos and Tempini, 2014).

Here, however, we are suggesting a different issue. Patient-powered networks might have shortcomings that are related to the failure of market-based approaches to efficiently structure such networks. The institutional ecology (Star and Griesemer, 1989) that makes up and surrounds these social-media-based networks shapes the fate of the data. Different actors see data and their value differently. As Andrejevic (2015) observes, organisations managing social media are entirely dependent on data: control is key. A social media organisation operates within a broader socio-economic environment to which it is tied through multiple relations, the most influential of which are often those tied to funding streams. The strategy chosen to generate revenue from the use of data shapes the ways in which the patient voice is made valuable, and where the power is located shapes how the data is used. The issue of 'consequential sharing' raises important questions in the context of the sharing economy, where many patients donate data (and the time and effort involved in producing that data) so as to benefit a cause, while also – and not unimportantly – facing some risks (for instance, the consequences of disclosures). We argue that it can engender digital health's own kind of market failure, the *digital orphans* – orphan diseases of the digital research economy. The circumstances shaping the use of data that users generate might differentiate between patient groups in terms of their chances of participation and their contribution to research. Although all patients who partake in these networks are exposed to the same culture of "data donation" and "openness", we wonder if all data are created through equal conditions, as decisions about data use in research are often ultimately taken by social media managers who have to respond to multiple organisational imperatives and, hence, evaluations of sustainability, growth and revenue (Tempini, 2017). The sole and sheer degree of diversity across the patient base of some health networks suggests that it is extremely difficult to serve all participants equally, according to the ideals of openness, patient participation and empowerment. How is it possible to make sure that the data that patients contribute are put to good use and exploited for research? The

paper points out difficulties that business models centred on data closure, i.e. exercising a monopoly on the use of data, encounter in keeping pace with the expectations they raise. The issue of digital orphans is one that requires reflection on what kinds of partners come together for these networks, and how their worldviews shape the circulation and reuse of data for the public good.

Regimes of data sharing in patient-powered networks

Social-media-based, patient-powered networks are centred on both health data self-reporting and patient social networking. They collect large amounts of structured and unstructured data on people's symptoms, comorbidities, drug regimes, side effects, functional changes and other aspects of their lives. Patients can browse the network through a web of profiles, content and comments. These two aspects of health social networks – i.e. health data self-reporting and patient social networking – are inseparable sides of the same coin. The aim is to activate a feedback loop: social interaction is fostered by data, and data generation is fostered by social interaction (Gerlitz and Helmond, 2013). These initiatives innovate by easing the generation and retrieval of medical information, and change the ways in which information is collected, assembled, coded, collated and shared (Griffiths et al., 2012). These social media features shape the structure of incentives for participation in the network. Patients sharing data online are often motivated by the hope of contributing to finding a cure for their own condition or the condition of a family member, even if they may not benefit from it themselves. They might participate in medical research for altruism and other intrinsic motivations connected to the ideals of the scientific enterprise. In addition, they often hope that the data and their experience may prove insightful for other patients, thus furthering the contribution of their data beyond a simple research-only use. Finally, they might participate to offer and receive socialisation and support opportunities. Data generation and sharing are then supported by a diverse array of technological solutions.^{iv} Patients do not directly benefit from the economic value generated by their labour, while they face certain (opportunity) costs when they invest their time and energies in the collaboration. Also, any other benefit they might contribute to – such as scientific and care innovation – is projected in the future (Giannella, 2015). For their sustainability, they often pursue one or more of a number of monetisation options, including selling datasets or research services to healthcare corporations (pharmaceutical companies, insurance plans, etc.), advertising drugs and products, and recruiting for clinical trials.

Patient-powered networks occupy a niche in a broader constellation of initiatives regarding the collection of patient feedback, opinions and experience in the form of data and their delivery to healthcare and research service providers in the form of information services. Patients are well positioned to produce data about their experiences with regards to a number of health dimensions (e.g. symptoms, effects of drugs, co-morbidities, life course information), in addition to everyday coping and support. The niche is created by the possibility of feeding patient-generated data to research users. These networks have diverse combinations of different types of institutional nature, stated mission and legal status; and revenue strategy and data governance models. These conditions shape the alignment of different interests and values on a network (cfr. Riso et al., 2017), and set constraints on the creation of a research agenda and how data is reused. We briefly illustrate this using a few examples of patient-powered networks, starting with networks that are least orientated on scientific research.

Consumer feedback

Some networks that collect patient-generated data provide spaces for rating and commenting on the care and clinical services and products, including pharmaceuticals. The motivation for developing and using these networks seems to mirror that of other feedback networks in the sharing economy (e.g. Yelp! for rating local businesses and TripAdvisor for rating hospitality businesses). The model can be used to support healthcare providers in terms of the quality of their services, or express disappointment if the service was deemed inappropriate. For example, Care Opinion^v, a UK non-profit network funded by the Scottish government, allows patients to leave feedback about healthcare and social-care providers, including the UK National Health Service. Its initial funding was provided by the Scottish government, but the network charges health care services organisations fees for advanced services such as feedback management. The network aims at making patient experience “available as widely as possible”. On this basis, review data can be browsed and republished on the web for free.^{vi} Patient review data are automatically forwarded to care providers, health care regulators, patient watchdogs and patient organisations. If providers implement changes based on patient feedback, they can publish the information on the website (Munro 2017). The full access to datasets requires payment of a fee. Care Opinion also sells advanced feedback analysis and management services to health care services organisations.

Also sourcing patient data in a similar way are networks dedicated to pharmaceuticals, often supported by corporate entities. One leading case is MediGuard, a platform collecting patient feedback on pharmaceuticals that is owned by the biggest contract research organisation worldwide, Quintiles. Founder Hugo Stephenson argued that Mediguard was created because “people were asking the consumer what they thought of drug safety” (Barriaux 2007). Its 2,699,000 users learn about the service both online and from their medical providers, and join to provide ratings and reviews. Although MediGuard warns against interpreting the information made available as medical advice, the website staff consolidate and publish risk ratings and profiles about the pharmaceuticals that patients review. Users can then browse the risk information, along with reported drug alerts issued by regulators such as the US Food and Drug Administration. Patient data are used to profile patients, both to improve user experience, but most importantly to be able to market patients as audience for clinical trial enrolment campaigns (a core market interest of the parent company Quintiles). Thus data are shared “with our clinical research partners to assist them in better designing their websites and clinical trials”.^{vii} It was to seize this original market proposition and relationship to patients, that Quintiles purchased the technology startup.

Patient social networks

Another kind of patient-powered network is focused on community building through the development of a social space where patients can meet others, as well as interact and learn. Within this group, Griffiths et al’s (2015) review of research participation in health social networks singles out two patient-powered social networks geared towards research-focused utilisation of network data: CureTogether and PatientsLikeMe.^{viii}

CureTogether

CureTogether was set up to facilitate the exchange of information on symptoms, diagnoses and treatments between patients. The platform was originally conceived to be disease-specific and focused on three pain-related conditions only, but was soon enlarged upon the requests of users. Users of CureTogether describe their symptoms in disease-specific surveys, post on the forum, seek and interact with other similar users on the basis of common conditions or

symptoms, and browse the summary information curated by the staff of CureTogether. The summary information includes graphs of symptoms, treatments, side effects and causes for each condition. Such information is exclusively aggregated from data posted online by users. While CureTogether declares that it does not publish medical findings,^{ix} it claimed – until late 2012 – to produce “research findings” based on the statistical analyses of datasets. In one case, CureTogether claimed to have replicated a study on the effects of low-dose naltrexone on fibromyalgia patients.^x The platform’s orientation towards research is also suggested (in a witty manner) by its claim that the website would have been called “Patient-driven research” if it were not for the long URL.^{xi} The mission statement explicitly endorses this claim: “Partnering with universities, research organizations, and self-experimenters will help us to make new discoveries and connections faster, ask better questions about patient data, and validate that online methods for data gathering are effective for research purposes.”^{xii} The data are however a proprietary asset of the company: the exploitation in research is dependent on their financial evaluation. Armstrong et al. (2012) evaluated the reliability of CureTogether’s crowdsourced data focused on a specific condition – the skin disease, rosacea – and concluded that “crowdsourcing has the potential to yield high-volume data from diverse patient populations” but that chief limitations include “low number of responses, unspecified medication diagnoses and duration, and unknown adherence rates” (2012:15). They also emphasised the differences between outcome measures in crowdsourcing sites and traditional clinical trials that prevent accurate comparisons.^{xiii}

CureTogether and its four million phenotypic data points were eventually acquired by a personal genomics company, 23&Me, in July 2012 in an explicit effort to combine 23&Me’s large genomic databases with epidemiological and phenotypic data. Today, CureTogether has a user base of ca. 39,000 people collecting data on about 637 conditions.^{xiv} After the acquisition, CureTogether’s operations were integrated with 23&Me. Infographics based on the datasets have been published on 23&Me’s website^{xv} in a page reporting an overview of each pairwise significant association between CureTogether users’ medical conditions. While the website still runs, there have been very few updates of the blog, its research sections and the live feeds. The utilization of the data in research now depends on the alignment with 23&Me’s strategic interests. The business model of 23&Me includes the sale of genomic and ancestry personal testing, collaboration with the industry for medical R&D based on users’ genomic and self-reported data, and, eventually, earning royalties from patented discoveries (see Harris et al., 2016).

PatientsLikeMe

PatientsLikeMe is one of the most popular social-media-based health research network. The way patient contributions and data generation and consumption activities are organized has already been well documented (Tempini 2015, 2017; Kallinikos and Tempini, 2014). We will refer here to those resources but for essential introduction and some key remarks related to the concern of this paper with business models and data governance.^{xvi} Through the site, patients can connect to other patients, keep a log of their health history, and participate in medical research that the research team of the organisation conducts and publishes in peer-reviewed journals and other outlets. Data collection involves a number of health-relevant dimensions: conditions, treatments, symptoms, quality of life, hospitalisations, lab results, weight, and disease-specific patient-reported outcome measures (PROMs). This for-profit, ad-free network is sustained through funding from partnerships, mainly with the pharmaceutical industry,

centred on the sharing of pseudonymised data and the execution of custom research. Patient involvement in research is limited to the collection of data and the opportunity to collaborate in the development of interfaces and tracking tools. However, the organisation simply would not be able to run without the continuous collaboration with its patient members; and patients are not bound to participation by any kind of contract, are physically distributed all over the planet and are extremely diverse not only due to the kinds of patient experiences they are going through, but also due to the kinds of literacy they mobilise to make sense of it. After a long undertaking to restructure the network's architecture, in 2011 PatientsLikeMe became a unique social network that could connect all patients of all conditions together (Tempini, 2017).^{xvii} It was not a painless transformation: due to the incompatibility of the new data architecture with data that patients had previously collected about their co-morbidities, the implementation necessitated the removal of some of the old data from the website, to be 'parked' for good (Tempini, 2014) – an example of the kinds of uneasy data governance trade-offs that patient-powered networks needs to face in their day-to-day management.

Importantly, PatientsLikeMe's new architecture allowed patients of any given condition to join the network *before*, instead of *after*, any research activity on a given condition had been conceived. Rather than launching a community as a consequence of starting a research project, it made it possible to conceive of research *because* determinate patient groups aggregated on the network. However, the onward sharing and use of data in research would need to wait for a new project – funded by a partnership – that allowed R&D efforts to begin. While the number of hosted conditions exploded with the new architecture from ca. 25 to more than 1,400 in the first year alone, as did the number of patient members, their data were still scarcely used in any research activity unless a funded project had been started that allowed data to be generated and worked with according to the standards and processes that medical research requires. The research staff's deeper experience in the study of neurodegenerative diseases meant that, even years later, a great amount of the studies generated from PatientsLikeMe data focused only on diseases such as Amyotrophic Lateral Sclerosis, Multiple Sclerosis and Parkinson's Disease. While the company made available its web tools to patient communities of all kinds of diseases, and the it was able to develop its research capabilities further, patient 'data donations' became more uncertain, overall, to be used in research. The network has grown further, launching the Open Research Exchange, a system that allows more systematic involvement of patients in the design and development and use of Patient-reported Outcome Measures (PROMs), with the aim to accelerate the development process and to design tools that are closer to the patient experience and thus more meaningful for end users. The pre-existing data governance arrangement subsists; the organisation controls access to patient data because its business model is centred around its proprietary exploitation.

Digital economy, digital orphans?

The aim of this brief overview is to illustrate how the use in research of data from health social media is constrained by business models and the institutional ecology that they support and are embedded in. On the one hand, patient-powered networks are seen as complementary to the vision of a sharing and supportive society, echoing the ideals of a commons-based community and bottom-up research and empowerment. On the other hand, they are developed on a for-profit basis for faster development and market share. Their key challenge is to integrate the high expectations evoked by a participatory vision with the business imperatives of for-profit environments. As Prainsack argues (2017:120), the issue in this kind of convergence is not the

convergence of for-profit and not-for-profit elements *per se*: the question is rather who benefits, and who is in charge. Such tensions do not undermine the viability of these networks and their potential for scientific innovation. In this respect, we wonder if, in the future, there could be more innovation in terms of organisational forms that more directly address the issue of data governance and the democratization of the reuse of data that we have been trying to delineate. While the for-profit model can be advantageous for accelerating the growth of these networks, we suspect that tight coupling between control over data (power over the research agenda) and control over the infrastructure of participation (power over technology and community development) is the core of the issue at stake here. Decisions regarding the research uses of patient data are assigned to the new gatekeepers, i.e. network developers, who need to justify action against organisational imperatives that are related to business partnerships with clients, the research interests of participating researchers and their respective areas of expertise, and the perception of the opportunity for an impactful contribution (cfr. Riso et al., 2017).

It must be noted that the problem might be compounded by other reasons to favour some form of control over access to patient data, relating to issues of data quality, and of documentation and organization of collaborations. Patient data can fluctuate widely in quality and reliability, as patients can operate with very different combinations of epistemic standards, background knowledge, and motivations. It is labour-intensive to make available these data to different actors in a form that can be reliably used as evidence. Wyatt and colleagues (2013b) also highlight the careful juggling act of data managers as they need to foster trust in the scientific enterprise from both patients, investors and the scientific community. A great amount of curation and interpretation is necessary to make the vast amounts of data valuable for certain uses, and trusted by their audiences. The overall population sample has its own peculiar biases (in the health-related Internet, it is often white, female, educated, etc., but can fluctuate) (boyd and Crawford, 2012), and the deep familiarity that specialists can develop with data is key to making the data reusable (Fleming et al., 2017). The expertise of specialists is crucial to document the data, including the application of standardized metadata. Successful and repeatable reuse of data depends on them to know the data in-depth and intermediate with those other users that access the data for the occasional research project; who can help and guarantee the quality of research, controlling for biases as well as blatant misuse. If access is uncontrolled, misuse that could harm both patients and the legitimacy of patient self-reported research, and which could originate from many actors, including both internal patient members and external members of the public and organisations.

The development of this human resource is an expensive undertaking and yet, somewhat ironically, it might not suffice to control for the most radical risks that derive from self-report, self-select Internet data, that is, intentional misinformation. A controversy opened when the BBC reported how staff of a healthcare trust in the UK submitted most of the patient reviews about the trust itself that were available on Care Opinion.^{xviii} CureTogether's model of allowing users to submit anonymous data on pharmaceuticals and other treatment options seems also vulnerable to targeted manipulation. But as nearly any social media service on the Internet, none of the four examples we reviewed here require any 'hard' proof of identity. As the general public has reckoned with in numerous ways over the last few years, of concern is not the lone act but also the possibility of organized, focused actions of misinformation. The most incendiary debate is over Russia's manipulation of the public sphere over Facebook. But the fakes industry is an underwater sector of the web: fake product reviews, fake followers, etc. have been well

documented. In regard to security and patient protection as well, it is not fully clear that centralized control of the data is necessary or advisable.^{xix} Openness can elicit broad and diverse scrutiny and reporting.

Despite these obligatory considerations, our argument is not directly concerned with the validity of the model (at least not directly) of knowledge production through patient-powered networks, but it does reflect instead on its magnitude. We wonder if contemporary examples of social-media-based participatory research really afford hope for the kind of disruptive change that can shift the orientations of an industry. PatientsLikeMe has produced a considerable amount of research, but it is the only example, in our set, to systematically achieve this. Still, what is of concern is not so much the character of the research that is actually done than the research that is *not* done – the missed opportunities that cannot be explored due to the existing regimes of data sharing. Not all data are created equal. Scientifically valuable venues of research might be turned down in cases where no industrial interest supports them. In the absence of such interest, the opportunity to make use of patient data remains quite small. The problem that scientific research data too easily ‘go to waste’ has been well documented (Leonelli, 2013).

The imperatives of for-profit business models based on the monetisation of data – as well as driving research strategies and monopolist decisions regarding the use of the data collected through these patient-powered networks – define constraints that shape the collaboration’s infrastructure and outcomes. The cost in trying to mediate a marriage of patient and industry interests is to never (disruptively) exceed, in the use of the data, what the wider institutional ecology of commercial and research partners is willing to engage in. Contrary to the hypothesis of these networks functioning as the centre of evidence-based activism (Rabeharisoa et al., 2013) with patients in control of the research agenda for data reuse, there is a risk here that much of the patient input will never be used further than for fuelling the web-based computational infrastructures that support online community interaction. While pre-emptive data collection might be to some level ‘use-agnostic’, business models are not (Aaltonen and Tempini, 2014). Much of data sharing or “donation” could have no consequence in research.

Our review suggests that not all web-based participatory research activities have an equal chance if the data shared by patients from certain condition communities are more intensively used than other data. The scenario is one where the disruptive, innovative web economy can produce its own kind of market failures. Even in the highly inclusive environment of patient-powered research networks, a patient community can become an ‘orphan’. In the context of the web networks, it seems that the problem of orphan diseases might persist in a new form – notably, to be eventually shifted between different disease communities. On the one hand, web networks are potentially helping some ‘traditional’ orphan patient communities to increase research activity around their causes (for instance, ALS is the most striking example of successful mobilisation through PatientsLikeMe). On the other hand, the new convergence of actors in these digital environments makes what one might call *digital orphans* of those diseases (which might not have been orphans before the Internet) that, relative to patient-powered research, now do not enjoy the convergence needed to make the digital scientific enterprise work.

In the observed examples, we can think of at least four main factors that make such a convergence possible. *First*, the level of “engagement” among the patient population can vary enormously (size, data-productive levels of activity). *Second*, the expertise of staff matters, as

staff specialise in certain medical phenomena more than others, and are as a result differently able to seize opportunities. *Third*, the character and epistemological status of the patient experience and how trustworthy its measurements are considered to be affects how much self-reported patient observations are trusted beyond the patient group itself. *Fourth*, and drawing on the previous three points, the eventual presence of an industrial interest is needed to materialise a research project by providing funding. These networks, despite being continually improved and built upon to include more participants, seem to have to necessarily differentiate between patients so as to initiate, participate and contribute to research. This is no small consequence and suggests that, even on the web, participation by patients is a necessary, but not sufficient, condition to foster research regarding a disease. It makes one challenge the idea that patients (or citizens) are ‘in charge’.

It must be noted that probably the most acclaimed research achievement of patient-powered research networks – a PatientsLikeMe-authored study of the drug lithium-carbonate published in *Nature Biotechnology* (Wicks et al., 2011) – was indeed made possible only through the involvement of research scientists on a long labour-intensive project (still shorter than a traditional trial), although patients had arguably initiated the research and made the opportunity possible. Patients are undeniably important actors in such research, but published research from web networks, in most instances, is not completed or co-authored by patients ‘in the wild’. While we agree that web-based networks are making it easier for “patient-led research” and other disruptive initiatives to happen (e.g. Vayena et al., 2015; Vayena and Tasioulas, 2013a, 2013b), we believe that the kind of innovation that these networks are driving is better captured by the PatientsLikeMe staff’s own use of the language of “patient-powered” research or networks. It tracks the organisational reality more closely, also conceding to the potential exclusion of patients from key roles. In our view, the term “patient-led” instead risks concealing any divergence between users’ and network managers’ interests, and makes us forget about the network development work that happens at the backstage of the web (see Ekbia, 2016; Tempini, 2015).

Towards open patient-powered research

If we consider web-based participatory networks as infrastructures for distributed research, then the question of control regarding access and the use of the data is key. It must be noted that the constraints put on data sharing and reuse in these networks are linked to a more general problem that besets any form of production and the sharing of common goods, of which data is a paradigmatic example (Eschenfelder and Shankar, 2016).^{xx} There is a tension between the cost to maintain, develop and run platforms, and the fact that data could often be made non-rivalrous in use. Also, a completely distributed data governance framework, where every data donor is in charge of the micro-management of the use of their data by others, arguably offloads on the individual risks and trade-offs that are very difficult to evaluate individually (Aicardi et al., 2016; Evans, 2017).

Maintaining that making patient data available to just about anyone would be much worse (just the privacy concerns are enough issues to consider), we wonder whether it would be more beneficial if networks separated the management and development of the web infrastructure from decisions about the use of data for research purposes. If this were organisationally sustainable, it could be the first step for patient-powered research networks to more fully espouse openness at both layers: to match the openness with which patients share details about

their lives with an openness about the research avenues and data reuse opportunities that the network would be able to support. This could allow patient-powered networks to better adopt the ideal of patients engaging in “evidence-based activism” and contributing *from within* to the “collective inquiry associating patient/activists and specialists/professionals in the conjoint fabrics of scientific statements and political claims” (Rabeharisoa et al., 2013:2). Forms of federated, hybrid partnerships where trusted patient-activist groups are delegated powers over data governance might be able to improve oversight and multiply research projects while protecting the patients’ interests (Vayena and Blasimme, 2017). We are very aware that the representativeness of patient groups is often questioned. Patient-activist NGOs have been criticised for biased representation. Still, this might be an improvement over the data governance models that are common to the networks we have surveyed as examples. To address this conundrum at least in part, some novel arrangements are being experimented with to distribute management over data use for data producers (i.e. patients). Patients are then endowed with a larger latitude of control over what is done with their data.

One promising example of emerging initiatives is “data cooperatives” (Hafen et al., 2014; Riso et al., 2017). Data cooperatives are organisations, owned collectively, that are delegated the day-to-day management of the data of their constituents, including data management tasks and information governance decisions. They are fiduciaries of their members’ data. Data cooperatives can be set up at varying levels of locality, and can refer to different kinds of constituency criteria. The data cooperative infrastructure is developed with the aim of allowing patients to upload their health records, share them with others, and make them available for research for those projects they find worthy, and thus tries to strike a balance between granular control and day-to-day delegation to representatives. Their rationale is to simultaneously tackle two interconnected issues of data reuse – namely, inefficient data sharing and a lack of distributed decision-making: “Inaccessibility of data is not only a cause of inefficiency and increased costs in the healthcare system, it also prevents the citizens to control their data and thus engage more actively in their own health” (Hafen et al., 2014:82). The model aims to be self-sustainable, creating revenue by charging access fees to research users while keeping the control over data governance with the patients, who are likely to be more interested in multiplying and diversifying the research uses of their data. Interestingly, the chosen policy of sustaining development through cost recovery fees is an *ad hoc* solution (it is not prerogative of cooperative organisations) and may be experimented with by other forms of organisations as well.

Other current medical research governance experiments are aimed at transferring control of data reuse to patients without overly tinkering with data ownership policies. The US-based Patient-Centered Outcomes Research Institute (PCORI), a non-profit and non-governmental organisation created as part of the Patient Protection and Affordable Care Act, involves patients as co-investigators in comparative effectiveness research (Robbins et al., 2016). According to online documentation, proposals can be submitted by patients, caregivers, clinicians and other healthcare stakeholders, and are subsequently reviewed to identify and prioritise significant topics. Funding applications are reviewed by professionals and lay citizens. Any stakeholder can provide commentaries on reports, proposed policies and other initiatives, and join the organisation’s advisory panel.^{xxi} An evaluation of 50 pilot projects has been published, showing that non-professional stakeholders were mostly engaged in topic solicitation, question development, study design and data collection (Forsythe et al., 2016), a result that suggests that

patients can actively contribute to increase the opportunities of patient data reuse. While PCORI does not govern medical datasets, it has overlapping aims with the Open Research Exchange initiative at PatientsLikeMe. However, the PCORI funding arrangement seemingly makes it unconstrained by profitability requirements. While PCORI is a private organisation, not a federal agency, it does receive public funding to fulfil its explicit mandate of collecting, evaluating and expanding citizens' inputs in medical research. The future outcomes of PCORI and similar initiatives are a second source of potentially valuable insights to address the tension between the aspiration of patient control on research, and its implementation.

Conclusions

The emphasis of our paper on the conditions in which patient data can be “donated” by patients and yet left unutilised helps to highlight an issue that seems to characterise web-based infrastructures, where a certain kind of data-intensive business model is combined with a participatory organisational arrangement. How more open governance structures like the ones we briefly touched upon will fare remains to be seen, but they provide alternative models to compare and discuss. Of course, we are wary about the enormous uncertainties that come with pioneering these models. Development is costly, ventures can be difficult to scale, and the venture can suddenly need to respond to emerging problems that are specific to distributed patient engagement and need to be addressed contingently through labour-intensive strategies (Tempini, 2015). We do not intend to downplay the costs and risks of running a patient-powered research enterprise through a different business model. However, we welcome more experimentation to explore how the development of such networks could be sustained through alternative, more consequential models. The implications of online and digital intermediation are deep and subtle (Bowker, 2013) and, as such, technology development needs to allow for rethinking and re-evaluation (Wajcman, 2015).

References

- Aaltonen, A., Tempini, N., 2014. Everything counts in large amounts: a critical realist case study on data-based production. *Journal of Information Technology* 29, 97–110.
<https://doi.org/10.1057/jit.2013.29>
- Aicardi, C., Del Savio, L., Dove, E.S., Lucivero, F., Tempini, N., Prainsack, B., 2016. Emerging ethical issues regarding digital health data. On the World Medical Association Draft Declaration on Ethical Considerations Regarding Health Databases and Biobanks. *CMJ* 57, 207–213. <https://doi.org/10.3325/cmj.2016.57.207>
- Alison, H., Brand, A., Holgate, S.T., Kristiansen, L.V., Lehrach, H., Palotie, A., Prainsack, B., 2012. The future of technologies for personalised medicine. *New Biotechnol.* 29, 625–33.
doi:10.1016/j.nbt.2012.03.009
- Andrejevic, M., 2015. Personal Data: Blind Spot of the “Affective Law of Value”? *The Information Society* 31, 5–12. <https://doi.org/10.1080/01972243.2015.977625>
- Armstrong, A.W., Harskamp, C.T., Cheeney, S., Schupp, C.W., 2012. Crowdsourcing for research data collection in rosacea. *Dermatol Online J.* 18(3):15
- Barriaux, M. Traffic-light medicine risk website to launch. *The Guardian*, Tue 2 Oct 2007.
- Benkler, Y., 2014. “Between Spanish Huertas and the Open Road: a Tale of Two Commons?”. In Frischmann, B.M., Madison, M.J., Strandburg, K.J., *Governing knowledge commons*. Oxford University Press, Oxford, New York, pp. 69–98
- Birchall, C., 2017. *Shareveillance*. University of Minnesota Press, Minneapolis.
- Bonsu, S.K., Darmody, A., 2008. Co-creating second life: Market-Consumer cooperation in contemporary economy. *Journal of Macromarketing*, 28(4), 355–368
- Bowker, G.C., 2013. Data Flakes: An Afterword to “Raw Data” Is an Oxymoron, in: Gitelman, L. (Ed.), “Raw Data” Is an Oxymoron. MIT Press, Cambridge, MA, pp. 167–172

- boyd, d. m., Crawford, K., 2012. Critical Questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15, 662–679.
- Boyle, J., 2007. "Mertonianism Unbound?: Imagining Free, Decentralized Access to Most Cultural and Scientific Material". In: Hess, C., Ostrom, E., (eds.) 2007. *Understanding Knowledge As A Commons*. MIT Press, Cambridge, pp. 137–151
- Brives, C., 2013. Identifying ontologies in a clinical trial. *Social Studies of Science* 43, 397–416. <https://doi.org/10.1177/0306312712472406>
- Büscher, M., Perng, S., Wood, L., 2014. How to Follow the Information? A Study of Informational Mobilities in Crises. *Sociologica* 1, 1–37. doi:10.2383/77044
- Callon, M., Lascoumes, P., Barthe, Y., 2009. *Acting in an uncertain world: an essay on technical democracy*. MIT Press, Cambridge, Mass
- Collins, F.S., Varmus, H., 2015. A new initiative on Precision Medicine. *N Engl J Med* 2015; 372:793–795. doi: 10.1056/NEJMp1500523
- Cooper, M., 2012. The Pharmacology of Distributed Experiment - User-generated Drug Innovation. *Body Soc.* 18, 18–43. doi:10.1177/1357034X12446380
- Del Savio, L., 2015. Crowdmind and the nature of expert teams. Available at: <https://citizenbiomedicine.wordpress.com/2016/01/25/crowdmind-and-the-nature-of-expert-teams/> [Accessed 30/06/2016]
- Ebeling, M.F.E., 2016. *Healthcare and Big Data*. Palgrave Macmillan US, New York.
- Ekbja, H.R., 2016. Digital inclusion and social exclusion: The political economy of value in a networked world. *The Information Society* 32, 165–175. <https://doi.org/10.1080/01972243.2016.1153009>
- Ekbja, H., Nardi, B., 2015. The Political Economy of Computing: The Elephant in the HCI Room. *Interactions* 22, 46–49. <https://doi.org/10.1145/2832117>
- Epstein, S., 1996. *Impure Science: AIDS, Activism, and the Politics of Knowledge*. University of California Press, London
- Eschenfelder, K.R., Shankar, K., 2016. Designing Sustainable Data Archives: Comparing Sustainability Frameworks. Presented at the iConference 2016, iSchools, Philadelphia, PA, USA. doi:10.9776/16243
- Evans, B., 2017. Barbarians at the Gate: Consumer-Driven Health Data Commons and the Transformation of Citizen Science. *American Journal of Law and Medicine* 42, 651–685.
- Eysenbach, G., 2008. Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness. *Journal of Medical Internet Research* 10(3): e22
- Feenberg, A.L., Licht, J.M., Kane, K.P., Moran, K., Smith, R.A., 1996. The online patient meeting. *J. Neurol. Sci., Amyotrophic Lateral Sclerosis/Motor Neurone Disease* 139, Supplement, 129–131. doi:10.1016/0022-510X(96)00093-7
- Fleming, L.E., Tempini, N., Gordon-Brown, H., Nichols, G., Sarra, C., Vineis, P., Leonardi, G., Golding, B., Haines, A., Kessel, A., Murray, V., Depledge, M.H., Leonelli, S., 2017. Big Data in Environment and Human Health. *Oxford Encyclopedia of Environment and Human Health*.
- Fox, N.J., Ward, K.J., O'Rourke, A.J., 2005. The 'expert patient': empowerment or medical dominance? The case of weight loss, pharmaceutical drugs and the Internet. *Soc Sci Med.* 60(6):1299–309
- Forsythe, L.P., Ellis, L.E., Edmundson, L., Sabharwal, R., Rein, A., Konopka, K., 2016. Patient and Stakeholder Engagement in the PCORI Pilot Projects: Description and Lessons Learned. *J Gen Intern Med.* 2016;31(1):13–21
- Frischmann, B.M., Madison, M.J., Strandburg, K.J., 2014. *Governing knowledge commons*. Oxford University Press, Oxford. 2014
- Gerlitz, C., Helmond, A., 2013. The Like economy: Social buttons and the data-intensive web. *New Media Soc.* 15, 1348–1365. doi:10.1177/1461444812472322
- Giannella, E., 2015. Morality and the Idea of Progress in Silicon Valley. *Berkeley J. Sociol.* 59
- Griffiths F., Cave, J., Boardman, F., Ren, J., Pawlikowska, T., Ball, R., 2012. Social networks--the future for health care delivery. *Soc Sci Med.* 75(12):2233–41

- Griffiths F, Dobermann T, Cave JA, Thorogood M, Johnson S, Salamatian K, 2015. The Impact of Online Social Networks on Health and Health Systems: A Scoping Review and Case Studies. *Policy Internet*. 7(4):473-96
- Hafen E., Kossmann, D., Brand, A., 2014. Health data cooperatives - citizen empowerment. *Methods Inf Med*. 53(2):82-6
- Harris, A., Kelly, S., Wyatt, S., 2016. *CyberGenetics*. Routledge, London.
- Heaton, L., Proulx, S., 2015. Paradoxical Empowerment: Immaterial Labor Translated in a Web of Affective Connections. *The Information Society* 31, 28–35.
<https://doi.org/10.1080/01972243.2015.977629>
- Janssens, A.C, Kraft, P., 2012. Research conducted using data obtained through online communities: ethical implications of methodological limitations. *PLoS Med*. 9(10):e1001328
- Jin, D.Y., Feenberg, A., 2015. Commodity and Community in Social Networking: Marx and the Monetization of User-Generated Content. *The Information Society* 31, 52–60.
<https://doi.org/10.1080/01972243.2015.977635>
- Kallinikos, J., Tempini, N., 2014. Patient Data as Medical Facts: Social Media Practices as a Foundation for Medical Knowledge Creation. *Inf. Syst. Res.* 25, 817–833.
doi:10.1287/isre.2014.0544
- Kelty, C., Panofsky, A., 2014. Disentangling Public Participation in Science and Biomedicine. *Genome Medicine* 6:8. doi:10.1186/gm525
- Lee, S.S., Cawley, L., 2009. Research 2.0: Social Networking and Direct-To-Consumer (DTC) Genomics. *Am J Bioeth.* 9(6): 35-44
- Leonelli, S., 2016. *Data-Centric Biology: A Philosophical Study*. University of Chicago Press, Chicago, IL.
- Leonelli, S., 2014. What difference does quantity make? On the epistemology of Big Data in biology. *Big Data Soc.* 1, 2053951714534395. doi:10.1177/2053951714534395
- Leonelli, S., 2013. Global data for local science: Assessing the scale of data infrastructures in biological and biomedical research. *BioSocieties* 8, 449–465.
doi:10.1057/biosoc.2013.23
- Leonelli, S., Spichtinger, D., Prainsack, B., 2015. Sticks and carrots: encouraging open science at its source. *Geo Geogr. Environ.* n/a–n/a. doi:10.1002/geo2.2
- Lucivero, F., Prainsack, B., 2015. The lifestyleisation of healthcare? “Consumer genomics” and mobile health as technologies for healthy lifestyle. *Appl. Transl. Genomics*.
doi:10.1016/j.atg.2015.02.001
- Lupton, D., 2013. The commodification of Care Opinion: the digital patient experience economy in the age of big data. *Sociol Health Illn* 36(6): 856-869
- Milne, R., 2018. From People with Dementia to People with Data: Participation and Value in Alzheimer’s Disease Research. *BioSocieties* 1–17. <https://doi.org/10.1057/s41292-017-0112-x>
- Munro, J., 2017. Online opinions changing care. *Br J Nurs* 26(13): 722.
- Nardi, B., 2015. *Inequality and Limits*. Presented at the 1st Workshop on Computing Within Limits, June 15-16, Irvine, CA.
- Nielsen, M., 2012. *Reinventing Discovery: The New Era of Networked Science*. Princeton University Press
- O’Connor, D., 2013. The Apomediated World: Regulating Research When Social Media Has Changed Research. *J. Law. Med. Ethics* 41, 470–483. doi:10.1111/jlme.12056
- Ostrom, E., 1990. *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press, Cambridge
- Prainsack, B., 2014a. Understanding participation: the ‘citizen science’ of genetics. In: Prainsack, B., Schicktanz, S., Werner-Felmayer, G., editors. *Genetics as social practice*. Ashgate, Farnham, UK. 147–164
- Prainsack, B., 2014b. The powers of participatory medicine. *PloS Biol.* 12(4): e1001837
- Prainsack, B., 2017. *Personalized Medicine: Empowered Patients in the 21st Century?* New York University Press, New York.

- Rabeharisoa, V., Callon, M., 2002. The involvement of patients' associations in research. *Int. Soc. Sci. J.* 54, 57–63. doi:10.1111/1468-2451.00359
- Rabeharisoa, V., Moreira, T., Akrich, M., 2013. Evidence-based activism: Patients' organisations, users' and activist's groups in knowledge society (Working Paper No. 033), CSI Working Papers Series. Centre de Sociologie de l'Innovation, Mines ParisTech, Paris, France
- Richards, M., Anderson, R., Hinde, S., Kaye, J., Lucassen, A., Matthews, P., Parker, M., Shotter, M., Watts, G., Wallace, S., Wise, J., 2015. The collection, linking and use of data in biomedical research and health care: ethical issues. Nuffield Council on Bioethics, London
- Rigi, J., Prey, R., 2015. Value, Rent, and the Political Economy of Social Media. *The Information Society.* 31(5): 392-406. doi:10.1080/01972243.2015.1069769.
- Riso, B., Tupasela, A., Vears, D.F., Felzmann, H., Cockbain, J., Loi, M., Kongsholm, N.C.H., Zullo, S., Rakic, V., 2017. Ethical sharing of health data in online platforms – which values should be considered? *Life sciences, societies and policy* 13: 12.
- Robbins, M., Tufte, J., Hsu, C., 2016. Learning to "Swim" with the Experts: Experiences of Two Patient Co-Investigators for a Project Funded by the Patient-Centered Outcomes Research Institute. *Perm J.* 20(2):85-8
- Robert, G., Cornwell, J., Locock, L., Purushotham, A., Sturmey, G., Gager, M., 2015. Patients and staff as codesigners of healthcare services. *BMJ* 350, g7714. doi:10.1136/bmj.g7714
- Ruckenstein, M., Schüll, N.D., 2017. The Datafication of Health. *Annual Review of Anthropology* 46, 261–278. <https://doi.org/10.1146/annurev-anthro-102116-041244>
- Sablinski, T., 2014. Opening Up Clinical Study Design to the Long Tail. *Sci. Transl. Med.* 6, 256ed19–256ed19. doi:10.1126/scitranslmed.3009116
- Salter, B., Zhou, Y., Datta, S., 2015. Hegemony in the marketplace of biomedical innovation: Consumer demand and stem cell science. *Soc. Sci. Med.* 131, 156–163. doi:10.1016/j.socscimed.2015.03.015
- Star, S.L., Griesemer, J.R., 1989. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19: 387-420
- Swan M., 2012. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J Med Internet Res.* 14(2):e46. doi: 10.2196/jmir.1988
- Tempini, N., 2017. Till data do us part: Understanding data-based value creation in data-intensive infrastructures. *Information and Organization* 27, 191–210. <https://doi.org/10.1016/j.infoandorg.2017.08.001>
- Tempini, N., 2015. Governing PatientsLikeMe: information production and research through an open, distributed and data-based social media network. *Inf. Soc.* 31, 193–211. doi:10.1080/01972243.2015.998108
- Tempini, N., 2014. Governing Social Media: Organising information production and sociality through open, distributed and data-based systems (Doctoral Dissertation). London School of Economics and Political Science, London
- Tempini, N., 2013. "The Paradox of Context Flexibility: Balancing User Engagement and Semantic Context in Distributed Data Collection." In *29th EGOS Colloquium - European Group of Organizational Studies*. HEC, Montreal
- Thorpe, C., 2010. Participation as Post-Fordist politics: Demos, New Labour, and science policy. *Minerva* 48(4): 389-411
- Thorpe, C., & Gregory, J., 2010. Producing the Post-Fordist public: the political economy of public engagement with science. *Science as Culture* 19(3): 273-301
- Topol, E., 2012. *The Creative Destruction of Medicine*. Basic Books, New York
- van Dijck, J., 2013. *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press, New York, NY
- Vayena, E., Blasimme, A., 2017. Biomedical Big Data: New Models of Control Over Access, Use and Governance. *Bioethical Inquiry* 14, 501–513. <https://doi.org/10.1007/s11673-017-9809-6>

- Vayena, E., Brownsword, R., Edwards, S.J., Greshake, B., Kahn, J.P., Ladher, N., Montgomery, J., O'Connor, D., O'Neill, O., Richards, M.P., Rid, A., Sheehan, M., Wicks, P., Tasioulas, J., 2015. Research led by participants: a new social contract for a new kind of research. *J. Med. Ethics* medethics-2015-102663. doi:10.1136/medethics-2015-102663
- Vayena, E., Tasioulas, J., 2013a. The ethics of participant-led biomedical research. *Nat. Biotechnol.* 31, 786–787. doi:10.1038/nbt.2692
- Vayena, E., Tasioulas, J., 2013b. Adapting Standards: Ethical Oversight of Participant-Led Health Research. *PLoS Med.* 10, e1001402. doi:10.1371/journal.pmed.1001402
- von Hippel, E., 2005. *Democratizing Innovation: The Evolving Phenomenon of User Innovation*. The MIT Press, Cambridge, MA
- Wajcman, J., 2015. Who's to blame for the digital time deficit? *Aeon Mag.* Accessed May 26th, 2015 <http://aeon.co/magazine/technology/whos-to-blame-for-the-digital-time-deficit/>
- Wicks, P., Heywood, J.A., 2014. Data Donation Could Power the Learning Health Care System, Including Special Access Programs. *Am. J. Bioeth.* 14, 27–29. doi:10.1080/15265161.2014.957627
- Wicks, P., Vaughan, T.E., Massagli, M.P., Heywood, J., 2011. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat. Biotechnol.* 29, 411–414. doi:10.1038/nbt.1837
- Wicks, P., Vaughan, T., Heywood, J., 2014. Subjects no more: what happens when trial participants realize they hold the power? *BMJ* 348, g368–g368. doi:10.1136/bmj.g368
- Wyatt, S., Bier, J., Harris, A., van Heur, B., 2013a. Participatory Knowledge Production 2.0: Critical Views and Experiences. *Information, Communication & Society* 16, 153–159. doi:10.1080/1369118X.2012.746382
- Wyatt, S., Harris, A., Adams, S., Kelly, S.E., 2013b. Illness Online: Self-reported Data and Questions of Trust in Medical and Social Research. *Theory Cult. Soc.* 30, 131–150. doi:10.1177/0263276413485900
- Wynne, B., 1996. May the Sheep Safely Graze? A Reflexive View of the Expert-Lay Knowledge Divide, in: Lash, S., Szerszynski, B., Wynne, B. (Eds.), *Risk, Environment and Modernity. Towards a New Ecology*. Sage, London, pp. 44–83.

ⁱ The orphan disease concept captures a situation where – in the context of market-led, capital-intensive pharmaceutical research – certain diseases and their communities are unable to align with industrial interests and resources to the sufficient extent required to make research for a cure happen. The rarity of a disease (implying a small market) is, for instance, one among several factors that could contribute to making the profitability of eventual breakthroughs uncertain.

ⁱⁱ A further line of critique is linked to arguments about the labour of research participants that takes place “outside” the Internet and yet “outside” the clinic at the same time. The labour of “enacting the body” in particular ways (Brives, 2013), and engage in certain sets of social interaction (Milne, 2018), to comply with the regimes of scientific research, must be acknowledged.

ⁱⁱⁱ Birchall (2017) provocatively observes how sharing “has to be understood today not as a conscious and conscientious act but as a key component of contemporary data subjectivity” (17); and is “not something we do after possessing data but is the basis on which having any relation with that data can be possible at all” (23).

^{iv} The ways in which a patient-powered network would typically employ solutions to stimulate further and active data production include: features for filtering the user base and displaying other users with similar health profiles, with whom the user is then encouraged to interact by tailored system interfaces; and features for encouraging data integration across the network, as patients log and report on their experiences (i.e. their symptoms and their severity), by cutting across search queries through synonym analysis and displaying example inputs from other patients. These are two examples among many others that make up a web of links connecting people to other people and medical phenomena, constructing a space of “possible” and “encouraged” social interaction that endlessly refers to something or someone in a chain that never ends (Gerlitz and Helmond, 2013; Kallinikos and Tempini, 2014; Tempini, 2013; van Dijck, 2013).

^v Formerly known as Patient Opinion.

-
- vi Accessed 27th June 2015, <https://www.patientopinion.ie/info/faq>.
- vii Accessed 27th June 2015, <https://www.mediguard.org/help/what-is-iguard/privacy>.
- viii Griffiths et al. (2015) wrote a systematic review of health social networks that explicitly interrogates whether and which online social networks aim to produce expert research, defined as “the use of information derived from activity within the particular network by professionals for research purposes” (2015:478). They found that 5 among 23 social networks declared that their objectives include the enablement of novel research: Psych Central, Mumsnet, Health Talk Online, PatientsLikeMe and CureTogether. Among these, the last three have internal research experts; only one network (PatientsLikeMe) is “focused on the collection, collation, and correction of information derived from the network itself”, while “almost all other identified networks emphasized the dissemination of existing information” (Griffiths et al., 2015:483). HealthTalk Online is a website that shares patient experiences, but only after these experiences are sought for, selected and coded by researchers who conduct interviews with the patients, with no data collection based on social networking. Within HealthTalk and Psych Central, research focuses on the development of a reliable health information portal, rather than on medical research. Mumsnet’s focus is instead broader than health, with surveys conducted on non-medical products and blog hosting.
- ix Accessed 29th June 2016, <http://curetogether.com/terms.php>.
- x Accessed 29th June 2016, <http://curetogether.com/blog/2011/10/05/stanford-fibromyalgia-study-on-ldn-replicated-at-curetogether/>.
- xi Accessed 29th June 2016, <http://curetogether.com/blog/ongoing-studies/>.
- xii Accessed 29th June 2016, <http://curetogether.com/blog/ongoing-studies/>.
- xiii Accessed 29th June 2016, <http://curetogether.com/blog/ongoing-studies/>.
- xiv No data about active users are released by the organisation on the website. Accessed 29th June 2016, https://www.23andme.com/en-gb/curetogether_connections/.
- xv Accessed 29th June 2016, https://www.23andme.com/en-gb/curetogether_connections/.
- xvi PatientsLikeMe’s user base includes more than 600,000 patient members affected by more than 2,800 conditions. Initially conceptualised as a website for amyotrophic lateral sclerosis (ALS), the founders quickly recognised the potential for the research model to be applied to more conditions. Expanding first into other neurodegenerative conditions (Parkinson’s disease and multiple sclerosis), by 2010, the network hosted approximately 25 distinct, ‘siloe’d’ patient communities. These included epilepsy, HIV and fibromyalgia, but also what we may call, for lack of a better word, ‘groups defined by patient experience’, such as the transplant community (for tracking the post-operative health of transplant recipients) and a mood disorders community (grouping together patients who suffer from bipolar disorder, depression or schizophrenia, among others).
- xvii PatientsLikeMe opened the system to patient-generated definitions, or ‘folksonomy’, for conditions, treatments and symptoms – i.e. medical entities as captured in patients’ own terms and circumstances. Expert staff reviewed any new definitions before allowing new data categories to be shared on the website and be adopted by other patients, whom could then aggregate their own data into these categories. Merging all communities into one had become a research and business development priority, as researchers wanted to better capture co-morbidity phenomena. Running siloe’d patient communities, each centred on a single condition, had long been a technical compromise with implications that needed to be tackled. The old model of community building was slow and labour intensive, and made the infrastructure development process too dependent on funding from clients for specific research projects. The staff had to conduct time-consuming research on the literature of each condition and review existing tools for patient self-reporting to then develop an appropriate metaphor for the patient experience in the set of self-reporting tools that could be selected as the new template for the patient profile. In the new architecture, a condition could be created in a few clicks by staff dedicated to the review of patient-generated conditions. A new condition would be attributed a set of basic tools that could be applied to any patient experience: symptoms, treatments and weight trackers, and quality of life and condition history questionnaires. For more in-depth discussion of this evolution, see Tempini (2017).
- xviii Trust spokesman argued that reviews were submitted on behalf of patients, acknowledging lack of transparency. Accessed 2nd Feb 2018, <http://www.bbc.co.uk/news/health-26229041>.
- xix The work of Mary Ebeling has also, chillingly, shown how health data trusted to an organization can be legally sold, unstoppably circulate on the Internet, and harm (Ebeling, 2016). The stakes of sharing have never been higher.
- xx Advocates of open data and open science argue for experimenting with a more distributed control of data sharing and use (Boyle, 2007) in order to fully reap the social benefits. The openness of a resource is related to the capacity of actors to access and use it, and, as such, can be limited by the

existence of barriers to its deployment. Openness can also be assessed by the degree of control a person or a group exercises on the resource: the more control is concentrated, the less the resource is open (Frischmann et al., 2014). In the case of data, one element necessary to achieve openness is law (i.e. intellectual property, contracts) – although this necessary step is often not sufficient, as openness is actually made possible only when the specific situations of practice dispose of the appropriate levels of skill, organisational resources, and methodological, ethical and technical commitments (Leonelli et al., 2015). The paradigmatic example of an openness regime is the public domain (Benkler, 2014), but other scholars have focused on governance models that combine the central control typical of public action with individual property rights regimes that underpin market organisation (Frischmann et al., 2014). These approaches, broadly inspired by pioneering work on common pool resources (Ostrom, 1990; Hess and Ostrom, 2007), have been recently proposed as a framework for governing personal health data (Evans, 2016).

xxi Accessed 29th June 2016, <http://www.pcori.org/research-results/how-we-select-research-topics/generation-and-prioritization-topics-funding-4>.