

Resisting big data exploitations in public healthcare: Free riding or distributive justice?

Paraskevas Vezyridis, Stephen Timmons

Nottingham University Business School, UK

Abstract

We draw on findings from qualitative interviews with health data researchers, GPs and citizens who opted out from NHS England's care.data programme to explore controversies and negotiations around data sharing in the NHS. Drawing on theoretical perspectives from science and technology studies, we show that the new socio-technical, ethical and economic arrangements were resisted not only on the basis of individual autonomy and protection from exploitation, but also as a collective effort to protect NHS services and patient data. We argue that the resulting opt-outs were a call for more personal control over data uses. This was not because these citizens placed their personal interests above those of society. It was because they resisted proposed arrangements by networks of stakeholders, not seen as legitimate, to control flows and benefits of NHS patient data. Approaching informed consent this way helps us to explore resistance as a collective action for influencing the direction of such big data programmes towards the preservation of public access to healthcare as well as the distribution of ethical decision-making between independent, trustworthy institutions and individual citizens.

Keywords

NHS; care.data; Informed consent; Big data; General practice; Medical records

Introduction

Data-based research from electronic health records (EHRs) has been growing steadily in the UK for more than two decades (Vezyridis and Timmons 2016a). After substantial investments in clinical information systems across the NHS, academic and governmental research databases (e.g. CPRD, THIN, QResearch, ResearchOne, CALIBER) are exploiting this data for epidemiological, pharmaceutical and health services research.

The UK government announced (in 2013) its own programme called ‘care.data’. Led by NHS England and the Health and Social Care Information Centre (HSCIC, now NHS Digital), its aim was to collect, de-identify and link, in one central database, clinical and administrative datasets across the NHS and social services (Taylor 2014) for various uses, including research. However, due to a public outcry, which resulted in approximately 1.5m opt-outs, the programme never fully materialised and was eventually terminated in 2016 after the National Data Guardian’s (NDG) review (Lea 2016). Care.data’s failure has been attributed to poor communication, security risks and ambiguous information governance and data dissemination practices (HSCIC 2014, Vezyridis and Timmons 2017). Similar national programmes, for example in Denmark (Wadmann and Hoeyer 2018), are also under attack for their opaque dissemination practices in support of governmental and commercial contexts of exploitation.

With the exception of some patient populations, the public seems to be largely supportive of health data sharing for research (Aitken *et al.* 2016), but at the same time confused and sceptical when it comes to claims of public benefit, transparency, anonymisation and consent, especially for commercial purposes (Castell and Evans 2016). The public is largely unaware of how its data is being used (Hill *et al.* 2013) by corporations and state agencies, often without clear public benefits (Aitken *et al.* 2016).

Most of the literature has focused on public *or* healthcare professional attitudes to health data sharing, but not on analysing these two groups together, or involving the researchers who are using these datasets. More studies, focusing specifically on patients-citizens who have opted out, are needed (Rosenbloom *et al.* 2013) as well as studies on the public’s acceptance of ‘authorisation’ – institutional decision-making for data releases without informed consent – as a data governance

mechanism (Aitken *et al.* 2016). This research was designed to address these gaps, particularly for *national* data sharing programmes in healthcare by focusing on three groups involved in care.data: health data researchers, General Practitioners (GPs) and citizens who have opted out from, or have campaigned against the programme. We do this by applying an Actor-Network Theory (ANT) (Latour 1988) approach to ask; what did resistance to care.data elucidate on how data from electronic health records could be exploited as a public good? What could be the role of opt-outs from big data programmes in the context of national health systems?

ANT is a set of methodological tools to explore the social as a heterogeneous entanglement of ever-stabilising and proliferating associations between human and nonhuman entities (Latour 1988). ANT has been used to study ethical *controversies* in biomedical innovation and research (Williams-Jones and Graham 2003, Hayden 2007, Gardner and Webster 2017, Heeney 2017). Rather than studying moral frameworks from idealist viewpoints, it explores the practices of *doing* ethics within the material-semiotic networks, without making value judgements and by being attentive to the ontologies and normativities involved in mundane practices (Mol 2013). Previously we examined the scientific impetus and rationale for exploiting NHS patient datasets (Vezyridis and Timmons 2016a), the state's technolegal facilitation of decontextualised NHS patient data movements (Vezyridis and Timmons 2017) and the design affordances of the opt-out form (Vezyridis and Timmons 2016b). Here we focus our analysis on 'the dark side of the translation process and the disruption of the actor-network' (Galis and Lee 2014: 155) to understand how concerned citizens understand and practice ethics, politics and dissidence via the opt-out. We do not do this in isolation but *in relation* to the other protagonists (Latour 1988) of this hybrid controversy (GPs and health data scientists) who are directly involved in the collection and analysis of NHS patient data and are also affected by the opt-outs. By avoiding linear and deterministic macro-sociological examinations of adoption and acceptance of technoscientific innovations, which often ignore complex power relationships, the emphasis here is on the (in)visible sociomateriality of this assemblage in the context of deliberative politics (de la Bellacasa 2011, Welsh and Wynne 2013). There are varied reasons for patients-citizens to opt out from big data programmes and we do not attempt to present an exhaustive list. Rather, we are

concerned with collective forms of resistance and how the opt-outs are ‘enacted’ (Mol 2002) *here and now* as one effect of controversial entanglements of ethical, social, political, economic and scientific practices; away from simplistic debates over altruism vs. egoism and closer to questions of (re)distributions of power and benefits.

Power and ethics for data movements

Science and Technology Studies (STS) has produced an extensive set of analytical tools to demonstrate a situated, relational and performative understanding of power dynamics and their ethics (Galis and Lee 2014) in the making of technological projects and scientific knowledge. For example, Callon’s (1984) notion of ‘translation’ has been widely used to analyse sociotechnical processes of bringing together and forming stable material-semiotic networks of associated human and non-human actors responsible for the launch and maintenance of a technological innovation. STS emphasises that the fate of technoscientific projects is often unpredictable, and always contingent on power struggles weakening or strengthening network ties. These power struggles align interests and values and, thus negotiate and delimit, against many other possibilities (Williams-Jones and Graham 2003), ‘the identity of actors, the possibility of interaction and the margins of manoeuvre’ (Callon 1984: 203). In his review of approaches to power, Law (1991) argues that stable sociotechnical networks are the relational outcome of power balances (‘power storage’) where the actors involved have either accepted the new power dynamics or are able to exercise their own ‘power/discretion’. In this line of ANT research, Latour (2004b) considers translation as the (political) process of resolving the ‘matters of concern’ amongst enrolled actors so as to render the resulting technoscientific assemblage a hard-to-refute ‘matter of fact’, while for de la Bellacasa (2011) it is about actors being able to ‘care’ for the *things* they are not only materially and ethico-politically but are also *affectively* concerned with (Latour 2004b).

Thus, a programme like care.data is theorised (in this paper) as a political performance full of tensions around possibilities of data exploitations and ‘different ways of doing the good’ (Mol 2002: 176). The paper critically questions what (and how) actors attempt to become the ‘obligatory passage points’ (OPPs) (Callon 1984) for the ‘decontextualization, dissociation and detachment’ (Callon 1998: 19) of

NHS patient datasets to increase circulations (Leonelli 2016) within and beyond the network of public healthcare. The paper also highlights how ontologically and normatively multiple (Mol 2002; 2013) datasets and data subjects come into being, as they are framed (Callon 1998) and scripted (Akrich 1992) by scientific, governmental or corporate ‘centres of calculation’ (Latour 1988). As such, care.data provided the opportunity to examine not only the infrastructures and calculations attempted between actors (Callon 1998) for big data exploitations in public healthcare but also how (public) values are prioritised, stabilised or obscured (Gitelman 2013) as they are hybridised in programmes that blur the borders ‘between science and politics, culture and technology, morals and economy’ (Venturini 2010: 265).

For this, STS scholars’ understanding of publics and their values regarding power struggles between actors, as they set boundaries and facilitate intersections, (Dussauge *et al.* 2015, Montgomery 2017) are of importance to this paper. Rather than being self-evident or universal, values are theorised here as things made in practice (Leonelli 2016, Dussauge *et al.* 2015) which are meaningful only within the networks in which they are enacted (West and Davis 2011). While for Marres (2005) publics are ‘ungraspable’ and abstract entities, they are nevertheless formed by actors who gather together in ‘democratic collectives’ (Latour 2004a) to evoke pressure and negotiate their participation in boundary-making practices (Jasanoff and Kim 2009). In light of this, it could be argued that a public that has opted out from care.data did not necessarily pre-exist, but it was formed to respond (as Welsh and Wynne (2013) would argue) to any colonisations of meanings, matters of concern and values by particular (institutional) ‘spokespersons’ who sought to represent this public (Callon 1984). We thus highlight the inscriptions, power relationships and normativities that the opt-out mediated between the actors associated with this controversial big data programme.

Methods

The paper draws on the findings of a qualitative study carried out in 2016. Data were mainly collected through semi-structured interviews with three groups: primary EHR data researchers (n=11), GPs (n=7) and citizens (n=9) who had opted out from or campaigned against care.data. Participants (10 female and 17 male) were relatively older and well-educated. Researchers interviewed were working

Opting out of care.data

at the academic research databases mentioned previously, in roles including statistics, epidemiology, data architecture, administration, and research analysis. For the GPs, areas of expertise included clinical commissioning, data protection, ethics and academic health data-based research, while for citizens and campaigners their expertise included: information governance, business analysis, (non-) academic research, social work and public healthcare. Findings are also informed by observations at the organisations studied and from attending team meetings, conferences, public consultations and courses on conducting research with data from GP records.

Participant recruitment was completed via: (for citizens) social media (#caredata) and newsletters to a patient advocacy group (Healthwatch); (for researchers) targeted emails; (for GPs) targeted emails and bulletins. All the digitally audio recorded and anonymised interviews were conducted by the first author in 2016 via face-to-face (n=16), phone (n=10) and Skype (n=1) and lasted between 20 and 113 min (average duration 55 min). Following a constructivist grounded theory approach (Bryant and Charmaz 2007), data was analysed inductively, using QSR NVivo 11, to identify differences, relationships and patterns among participants' narratives, and construct the conceptual understanding of their important meanings faithfully, inclusively and contextually. The study obtained ethical approval from Nottingham University Business School REC.

Findings

Resisting exclusion from public good management

The study found support for the *principle* of re-using large conglomerations of personal health data for wider societal benefit. However, the potential value of these datasets translated by care.data as 'matters of fact' (Latour 2004b) for the improvement of health (and wealth) included a prescriptive policy normativity of exclusion of patients-citizens and GPs from co-shaping the exploitation of this public good (Jasanoff and Kim 2009, Welsh and Wynne 2013). The attempt by the programme's leaders to black-box the aggregation and assetisation of NHS patient datasets via an 'all or nothing' approach to informed consent (see also Vezyridis and Timmons 2017) and their limited engagement with public deliberation destabilised the hybrid (Latour 2004b). The 'illegal and completely cavalier way' (Citizen 11) care.data was to deal with *their* personal health data, and with the mandatory

extractions for GPs raised individual and societal concerns (Latour 2004b) around the construction of values and entanglements of NHS patient datasets within and beyond public healthcare.

..at some point I might opt back in, it's not that I'm against the principle, I just don't think [the programme's leaders] have got it right yet. (Citizen 1)

.. most people are altruistic, put up a good suggestion and you'll get cooperation, but if you come in all heavy-handed and say, we are going to, because we can, you're going to get resistance... what [programme's leaders] said was, we're going to use it for economic regeneration and we're not going to ask you first. (GP 1)

With its lack of tiered consent for patients-citizens, and without any meaningful choice about potential *uses* and *users* of NHS patient data (see also Sterckx *et al.* 2016), the opt-out attempted to inscribe in this network an unacceptable amplification of power asymmetries. By transferring rights of (obscured) exploitations to certain institutions, this 'interessement' device (Latour 1988) of (inadequate) personal control failed to enrol them, while making visible the lack of 'power/discretion' for the local network to participate actively in decision-making. Data transmissions were irrevocably destabilised (Law 1991), as some GPs and members of the public were seeking reassurances that these datasets would facilitate what they considered socially responsible research for public benefit, rather than appropriating and restricting access to these datasets for personal benefit (cf. Sterckx *et al.* 2016).

I think everybody wants the ability to say, I feel good about donating it because I know what's happening to it ... But if that organisation then says, 'you don't have a say in how we spend your money', then you're not going to give them the money. (GP 2)

I would like to understand better who's going to benefit from it. So, if it's in the public interest ..., I would always say yes ... I wouldn't necessarily be happy that any information I gave would be able to be used by any researcher; for example, if a company was then going to buy my data and restrict the use of that information ... because they were then going to be able to make a big profit from it ... I would want to understand what would that mean for that drug company if they profit by using the information in Britain, but they can't get that same

Opting out of care.data

information in developing countries ... are you then encouraging drug companies to only research ... first world countries' problems? (Citizen 2)

In effect, the opt-out was translated as an ethical, political and affective rendering and used as an active commitment in mediating responsible data exploitations (de la Bellacasa 2011) by restoring power imbalances (Latour 1988). The programme could then be reconfigured to address these citizens' matters of concern and care, rather than just protecting themselves at the expense of the society at large (cf. Sterckx *et al.* 2016). The attempt to place data exploitations in a transactional narrative where data extractions to the state, in exchange for NHS services, bifurcated research and care. This prioritised data collections and exploitations at the expense of the local relationships that produce the data (Montgomery 2017). The use of the NHS was understood as an absolute entitlement of citizenship, independent of any other obligation, and there was a material and affective attachment to the NHS (de la Bellacasa 2011). The attempted enactment of a new reality (and normativity) (Mol 2002) in the use of NHS services as an indistinguishable entanglement of healthcare provision, participation in data-based research and wealth creation, also shifted responsibility for maintenance from institutions to supportive individuals and publics (see also Marris 2015). This was perceived as 'a really radical change to the social contract that we have in this country' (Citizen 3).

... it is another sneaky manoeuvre that makes me lose confidence in the ethics of this programme ... We have the right to privacy and confidentiality which means we have the right to exercise it without being shamed for it. And I think a lot of the people who have opted out would be inclined to opt in if their concerns about confidentiality and ethics were addressed. That's not freeloading, that's just being sensible. (Citizen 4)

... but there are all sorts of more objectionable ways that people get free rides at the moment. We don't deny transplants to people that don't carry a donor card, and I would argue that that's a much worse kind of free ride. (Researcher 1)

Enforcing transparency, accountability and trustworthiness

The care.data controversy also opened up the black-box of data extractions and releases. Matters of concern around (information) governance, agreements with obscure actors, assetisations, accountability and oversight were exposed and challenged, revealing complex and dynamic relationships that refused reductions. Socio-material practices of data exploitations, political and economic asymmetries between data producers and users, prioritisations and stabilisations of ethical and economic values were all rendered visible and scrutinised for accuracy and conflicts of interest. Publicity campaigns, particularly via ‘junk mail’ methods, were deemed unsatisfactory (cf. Hays and Daker-White 2015). Citizens perceived as distorted NHS England’s claims of completely adequate data security and that no unintended harms would be caused by the proposed centralisation of all personal health data into one database. Confusing vocabularies and terminologies around data(bases) in the public debate excluded concerned citizens further from problematising and defining (Galis and Lee 2014) what is public and what is private, or good and bad uses. For policy-makers, care.data was only an issue of public acceptability and a question of which communication strategy about benefits could help the public overcome its ‘privacy paranoia’ and ‘free riding’, so as to *manage* the controversy away from more inclusive NHS patient data exploitations for public benefit (cf Marris 2015).

What everybody’s frightened of is that if people ever were asked to opt in, a lot of them wouldn’t. (Researcher 1)

[Programme’s leaders] have used those terms interchangeably as if there’s no difference like opt out and various terms like, pseudonymised, anonymous, these terms have, in my view, purposefully been misused and used interchangeably to try to prevent debate and to try to paint anyone who raises criticisms as a conspiracy theorist. Because they’re using terms that are meaningless. (Citizen 3)

The report by the Institute of Actuaries (Banthorpe *et al.* 2013) on accurate health insurance pricing, based on the analysis of hospital data provided by HSCIC, and the review of all NHS patient data

Opting out of care.data

releases that followed (HSCIC 2014) had been instrumental in raising public anxiety about data disseminations.

These publications constituted an ‘ethical moment’ (Heeney 2017) that brought to the fore not only questions about the normative framework and institutional practices of data releases but also decisions to opt out. They made visible not only the incongruity between the programme’s stated aims and actual practices of the institution responsible for these datasets, but also *unknown* relationships between various ‘centres of calculation’ (Latour 1988) and the NHS *already* operating in the background. Concerned about unknown multiplications of their medical histories and exploitation by different interrelated actor-networks (Mol 2002) beyond public healthcare, these members of the public lost their trust in the programme before any actual harms materialised:

And care.data is not necessarily egregiously different to anything else, but it’s the straw that broke the camel’s back ... The fire storm around care.data was only partially about care.data. A lot of it was what do you mean they are already selling hospital data? The reason everybody was so outraged about care.data was when the NHS said this is what we’ve been doing for 25 years, people said, what do you mean, this is what we’ve been doing for 25 years? (Citizen 5)

..when it came to care.data [Patient Participation Group members] were angry because they felt a betrayal of trust, ... that their records were being taken and they were being told that it was for scientific purposes when actually the government were selling it on. They were fearful that insurance companies could actually access their individual record, without permission.

(GP 3)

Although the researchers interviewed considered information governance procedures in the NHS as strict, to the extent that they restricted ‘worthwhile’ research, they nonetheless believed that it was important for them to be strict so that data did not get into the ‘wrong hands or isn’t used for the wrong purposes’ (Researcher 2). They thought that it was better to be over-cautious, especially when the omnishambles of care.data had already had a negative impact on their work. Others understood information governance not as a barrier but as a kind of “ethical data hygiene”: necessary preparatory

steps that researchers should take to make sure that their use of data is socially safe and ethically sound, ‘like washing your hands before having dinner’ (Researcher 6).

... 5 years ago if you wanted to access any of these data sources was much ... easier, the process now is lengthy ... maybe because of care.data, but it is something that is supposed to be like this so ... you cannot say that it made it bad, it changed it ... (Researcher 5)

..some organisations pulled out of [the database], took their consent back and it stopped us really being able to go out there and promote it ... But it was just so stupid the way they did it, I mean they really did treat people as though they just had the right to do all these things. (Researcher 3)

Protecting public healthcare from predatory market forces

The care.data controversy and the resulting resistance to the proposed entanglement of science, technology, politics, economics and ethics for NHS patient data assetisation moved beyond the ethics of individual privacy and confidentiality. It called for an examination of how the proposed arrangements for data exploitations would continue to support other actor-networks entangled with this one, such as access to (public) healthcare and also responsible scientific research. Here, public alienation was fuelled not by profit-making *per se* (see also Aitken *et al.* 2016) but by a lack of transparency and engagement in deliberations about the ‘ethos of investment’ and innovation (Muniesa 2017, Marris 2015) from these datasets. They felt uneasy with the idea that these public datasets could be used by private actors to reap the economic benefits directly, without giving back to the country that had made them available in the first place. Interviewees thought that the concept of care.data was good, notwithstanding issues of communication and meaningful choice. Citizens and some GPs accounted for their criticism in terms of an ‘execution [that was] much too commercially focused and it [seemed] to be much more for the benefit of the private sector than the benefit of the health for people in the UK’ (Citizen 4).

Opting out of care.data

... and the metaphor of course is the immortal life of Henrietta Lacks. So the idea that someone could take something made out of British healthcare records and become fabulously wealthy is something that grates with the average GP and patient. (GP 4)

This is not to say that private involvement in the NHS was entirely unacceptable to participants. It was that making the selling of data ‘an industry [was] the wrong approach’ (Citizen 6). They refused to accept the attempted stabilisation of this network through a silencing of their objections as well as private actors obtaining more power and knowledge of the NHS than could possibly be justified (Callon 1984). They feared that the NHS is threatened by the introduction of a private insurance model that will create additional problems of access to healthcare for many. Care.data was, therefore, seen not as a big data scheme that will ultimately benefit people, but one that will also facilitate the demise of the public character of the NHS via the provision to private companies of ‘information at population wide level ... to be able to have enough business intelligence .. to understand how the system works’ (Citizen 2). Most citizens and some frontline GPs were worried that this private involvement was more likely to cause, rather than solve, problems that may be detrimental to the sustainability of the NHS (see also Montgomery 2017).

Because I don't think there's any trust. I think you can say, giving it to these big pharmaceuticals, giving it to [private healthcare provider], increase the economy, but people say, well I don't want [private healthcare provider] to have this information, I don't want them to be any more better off and in a position to cause detriment to the NHS. (GP 2)

What are big pharma using these data for? Well, they're using them ... for marketing, which is ... objectionable. So, if they've got a new, expensive antidepressant that's not being prescribed to certain people, they could find that out from the GP data, then adjust their marketing to try and cost the NHS more money. (Researcher 1)

Such *covert* profit-making practices based on these data are antithetical to the founding principles of the NHS. Care.data was seen as a new database which would have brought together different social, technical and regulatory actors to form new heterogeneous networks to expand the scope of questionable secondary uses of data, and exert new powers through specific orderings, as one GP

suggested. For most participants, if any kind of exploitation was to take place, it should not be at the expense of the individual, whether that was advertising private healthcare, stopping welfare benefits or excluding patients from care.

I wonder if new databases make new rules and new rules means the ability to do things that other databases might not do, like export data to other institutions, countries, private companies, etc. which is something that existing databases might not do. So the conspiracy theorist in me wonders whether this is something that's economically driven. (GP 4)

I don't believe that it would benefit people with chronic long-term health conditions in any way. I think the insurance companies will cherry pick healthy people to provide coverage to and people like me will be left with no healthcare or bottom of the barrel scraps of healthcare, like in America. (Citizen 4)

Citizens expected that their responsibility to share their data would be matched by the responsibility of data users to share the NHS's foundational values (e.g. universal care). These private actors were not understood as residing *outside* society. They are members of the same assemblage and are expected to continue performing the meanings and values associated with the NHS. If not, then the opt out is mobilised as an OPP for safeguarding the NHS and protecting its assets from unacceptable political and market forces (Latour 1988).

Reclaiming responsibility from institutional decision-makers

Care.data became a question of what kind of normativities should accompany the ontologies of patients-citizens, datasets, regulations and institutional decision-makers. The controversy was the outcome of two different 'ontonorms' in tension (Mol 2013). On the one hand, these members of society enacted the patient-citizen as a societal *active* subject in an accountable and mutually beneficial data sharing relationship with institutional decision-makers and *specific* data actors; on *agreed* purposes of use as well as benefits distribution. On the other hand, the patient-citizen was scripted (Akrich 1992), via the opt-out, as an individual *passive* object of data extractions by representative (unaccountable) institutional decision-makers. In *anticipation* of maximum return on

Opting out of care.data

investment and potential wider societal benefits as well as risks of abuse, data exploitations had to be left fluid and open to *unknown* purposes.

However, care.data, as a sociotechnical actor-network in the wider governmental network, was not criticised in isolation, but *in relation to* other, past and present, translations (Callon 1984) that lacked wider public and clinical support: national technology programme failures (e.g. NPfIT, SCR) (cf. Hays and Daker-White 2015, Greenhalgh *et al.* 2010) or controversial government data sharing schemes (e.g. Home Office, Google DeepMind) (see also Sexton *et al.* 2017). This entanglement not only destabilised this network but, more importantly, called for the re-problematisation of NHS patient data exploitations as an issue of institutional trustworthiness and sociotechnical competence, necessitating a public to rise to the occasion (Marres 2005), take *responsibility* for its data and reclaim a more mediating, and OPP, role (Callon 1984) in decision-making about data releases. For example, some participants, especially those working in the NHS or in the information governance sector, expressed their disappointment for the fact that ‘privacy by design’ (Citizen 7) was missing from NHS patient data systems. Tracking responsibility, or exercising control was becoming increasingly difficult in public services with too many databases where the burden is always placed on citizens ‘to have to know them all, know where the fair processing notice is, know how to opt out’ (Citizen 1).

It wasn't necessarily care.data exclusively ... the rest of government is no better than the NHS was. We've seen that the rest of government is just as screwed up with data as HSCIC was two years ago. The problem is that people cared about the NHS more and the NHS was first. It was not worse. (Citizen 5)

I don't trust anybody with it, because I'm a Data Protection Officer, and I know what mistakes are made. The fact that I can no longer trust the NHS, really disappoints me, ... and I would probably trust someone like [grocery retailer] to look after my data ... because to them, it's money, they can't afford data breaches, because of their reputational damage, and the costs that they might incur. I think public sector people have been too lackadaisical, and have not realised the value of data. (Citizen 1)

Most citizens were not confident that they could trust institutions such as HSCIC or the Information Commissioner's Office (ICO), to act as their 'spokespersons' (Callon 1984). For example, the current NDG was highly regarded by participants due to her experience in the field of information governance in healthcare as well as her 'code of ethics' (Citizen 4), but contrary to the recent NDG review (Lea 2016) most citizens felt that some institutions' proclaimed independence and trustworthiness had been compromised too many times in the past by government. Therefore, they were not 'complacent in giving up individual rights to an institution' (Citizen 2).

... HSCIC, they have no credibility from my perspective because they're not independent ... I think I would place my trust in [Caldicott review body] because they have raised these issues and Dame Fiona Caldicott has had a critical independent stance on care.data ... But ideally, I would like to see individual patients being able to take control of their care record. (Citizen 3)

I don't think the [ICO], for example, appears to instil a great deal of fear into corporations. I think they know that they are an over-stretched organisation ... I just think they're not the best. They may have some of the sort of legal powers behind them, but I don't think they make much use of them, and I think what powers they have are fairly ineffectual. (GP 5)

Unlike other programmes of data collection for scientific analyses, the stability and durability of this network required *constant local* work and maintenance from all actors involved for continuous extraction of patient data. The expectation was of an ongoing relationship between patients, the NHS and institutional decision-makers. In effect, consent did not exist a priori but it was understood as an *ongoing relational achievement* of stable socio-material relationships (Mol 2002; Leonelli 2016) across and beyond the NHS. Citizens were aware of the technical and financial challenges in setting up and maintaining these relationships. They also understood that not everyone might be sufficiently 'educated' or 'articulate' to make informed decisions. However, they considered 'a one size that fits all' (Citizen 1) model of control and consent unsatisfactory. They would like to see a model where the individual has the opportunity to make a choice, while for those who do not wish to take up control other arrangements, such as 'the GP to have control of that data' (Citizen 3), can be put in place. Within this more collaborative governance framework, institutional experts could provide them with

their informed *opinions* about potential uses and users of their data, but it should be ‘the people with the final say over where their information goes’ (Citizen 9).

Conclusion

Care.data was a big data programme that attempted to disrupt and re-configure, unsuccessfully, the current epistemological, political, economic and social practices of big data-based research in biomedicine (Leonelli 2016). By mobilising normative assumptions about citizenship and economics (Woods 2016), it brought new expectations and promises (Brown 2007) of innovative research for capitalising NHS patient data. However, just as with other national (EHR) centralised implementations, it took ‘on a civic character’ (O’Doherty *et al.* 2011: 368) and followed an unpredictable path through a complex and dynamic storm of social, technical, ethical and legal challenges (Greenhalgh *et al.* 2010) translated as decisions about which material-semiotic reality should be made durable and visible, and which not (Strathern 2000).

We used an ANT approach to illuminate care.data’s failure to consider the fact that any values of such programmes for the public good are not inherent and, therefore, accepted as given by the socio-material networks they attempt to re-configure. They are always enacted in practice by specific stabilised networks against a backdrop of many other possibilities (Mol 2002). In developing one of the biggest healthcare databases in the world, embedded norms and values need to be translated as differences can be expected among ambivalent actors (Singleton and Michael 1993) who are provoked to react (Callon 1984). From an STS perspective, care.data can be described as a typical example of an imposed ‘sociotechnical imaginary’ (Jasanoff and Kim 2009) that for sceptical voices was more of an institutional reassurance of self-acquired ethical legitimacy. It failed to engage with the full sociomateriality of NHS patient datasets and, failed to consider (in public deliberation and policy-making) all matters of concern, entanglements of values (Dussauge *et al.* 2015), affective attachments from, and invisible work by, those who care for the production of the data (de la Bellacasa 2011). Therefore, it could not develop *collectively* a wider *common* vision and practice for data exploitations by debating the extent to which (and by whom) the marketisation of datasets and the public good will coincide (Gardner and Webster 2017).

Data research charities and institutions, private companies and government departments failed to eliminate the stark contrast between a ‘regime of hope’ (Moreira and Palladino 2005, Brown 2007) for speculative investments in scientific discoveries in the future and a ‘regime of truth’ where, for some of our participants, the risks of data exploitations, knowledge monetisation and the privatisation of NHS services were already evident. By focusing too much on raising awareness of the scientific value of these datasets, which none of our participants rejected in principle, they failed to reach consensus about other important ‘matters of concern’ (Latour 2004b). A convincing discourse and practice, for example, in support of the public character of NHS services or how these datasets would be protected from profitable ‘closed-data and closed-algorithm business models in health’ (Wilbanks and Topol 2016: 347) was missing.

The programme aimed at securing the bare minimum of trust while maximising potential returns on investment. It thus quickly dismissed privacy and respect for individual autonomy as individualistic rights opposing wider prosperity, rather than seeing them as principles of social trust *and* public engagement (Taylor 2014). The rigidity of the programme’s design and the ‘blurring of the distinction between ownership of technical innovation and ownership of information’ (Munns and Basu 2013: 130-131) propagated the ‘myth’ of a paranoid public of free riders. At the same time, shifts in alliances between the NHS, patients-citizens and *unknown* future users (Zuboff 2015) of these datasets were taking place beyond the context of public healthcare. In that respect, the care.data controversy was also about ‘about power, both power over data and power over the outputs data can produce’ (Taylor 2016: 11). It was about choice with regards to the kinds of logics and subsequent entanglements of exploitation in which concerned citizens, GPs and even health data researchers were expected to enrol (Callon 1984) without any ‘spaces of contestation’ (Barry 2002: 270).

The programme instilled in parts of the public a sense that it was a politically and economically driven project, rather than a clinical-scientific one that would benefit patients-citizens (Greenhalgh *et al.* 2010). The opt-out was the materialisation of critique, protection and distrust towards institutional representatives and designers of such programmes (Greenhalgh *et al.* 2010) that extended ‘beyond individual modes of opting out to collective forms of conscientious objection’ (Benjamin 2016: 968).

Opting out of care.data

It was a material and affective commitment by dissenters to mediate between prescriptive and exclusionary science and policy normativities (Jasanoff and Kim 2009, Welsh and Wynne 2013) and affirm that these datasets are utilised for wider public benefit, with minimal social and economic harm for patients-citizens. As such, it was an ethical, social, political and economic *collective* activity against a “conscripted marketisation” of NHS patient data. Personal responsibility and control were mobilised as a way for the public to maintain its OPP role (Callon 1984) for secondary data uses as well as for adequate representation of the public’s interests and values (Gottweis *et al.* 2011) when it comes to the distribution of public benefits (Benjamin 2016). In effect, it moved the discussion from the ethics of public altruism to questions around loyalty to NHS and, the trustworthiness and transparency of people and practices behind such programmes to represent public interests and values (Zuboff 2015).

As Winickoff (2016: 54) argues ‘benefit sharing attempts to stitch a distributive norm at the seam of the market and gift economies’. It calls for both distributive justice *and power* (O’Doherty *et al.* 2011), away from individualistic understandings of autonomy and protection (Dove *et al.* 2017). In this way, we can begin exploring ways of including in governance structures (of national big data programmes) ‘the collectivity as sovereign ethical subject’ (Hayden 2007: 744). We can then develop, for example, a ‘partnership governance’ (Winickoff 2016) that goes beyond occasional public consultations, to provide them with opportunities to co-configure public meanings, risks and values of these datasets (Jasanoff and Kim 2009), co-formulate governance policies (cf. Dove *et al.* 2012) and exert a share in decision-making on how best to distribute the benefits they have co-produced (Winickoff 2016). These would be based on mutually agreed boundaries of data exploitations and standards of responsive accountability (Welsh and Wynne 2013).

Based on our findings (and the fate of care.data) we suggest that choice should be supported to act as the necessary ‘mediator’ (Latour 1988) that could bring this material-semiotic network together. As de la Bellacasa (2011) would assert, establishing and maintaining the technoscientific assemblage of a national healthcare database should not be about deprecating concerns, distorting agendas (Galis and Lee 2014) or excluding from translations those who disagree with it, based on binary (moral and

epistemological) understandings of caring (or not) to ‘save lives’. If treated as a collective action in the context of deliberative politics (Welsh and Wynne 2013), choice could *act* as, for example, a barometer of ‘social acceptability’ (Floridi and Taddeo 2016) for big data programmes in healthcare – not only the ethical right of the public but also a distributed form of *enforced* public awareness, accountability, oversight and social responsibility.

In a political-economic environment where a clear determination of ‘public interest’ is lacking, growing commercial interests are increasingly structured around proprietorial control of such data assets (Taylor 2016) and the knowledge produced for the financialisation of biomedicine (Birch 2017). Data collections are increasing in size and scope across and beyond healthcare, while public confidence in privacy protection is reducing (Phillips *et al.* 2017). Simultaneously, shared decision making between clinicians and patients is promoted to move from a paternalistic to a more patient-centred model of communication (Spatz *et al.* 2016). Removing a meaningful and collectively agreed opt-out confines data users’ relationship with the public solely to communication of benefits and assurances of sufficient data protection practices (cf Lea 2016). This may result in another form of (data) paternalism and a ‘governance by elites’ (Woods 2016) model of NHS patient data use where data users and policy-makers set the usage agenda *and* the conditions of public research participation, increasing the ‘agency gap’ (Winickoff 2016). Consequently, public distrust and resistance to national big data programmes in healthcare grows.

Acknowledgments

The authors would like to thank the participants, the anonymous reviewers and the Editor. This study was funded by the European Commission (H2020-MSCA-IF-2014-659478). ST was partially supported by the NIHR East Midlands CLAHRC. The views expressed are those of the authors and not necessarily those of the EC or the NIHR.

References

Aitken, M., de St. Jorre, J., Pagliari, C., Jepson, R., et al. (2016) Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies, *BMC Medical Ethics*, **17**, 1.

- Akrich, M. (1992) The de-scription of technical artifacts. In Bijker, W. E. and Law, J. (eds) *Shaping Technology/Building Society: Studies in Sociotechnical Change*. Cambridge, MA: MIT Press. pp. 205–224.
- Banthorpe, P., Cleverley, P., Fairall, C., Groyer, A., et al. (2013) *Extending the Critical Path: A report from the Critical Illness Definitions and Geographical Variations Working Party*. Institute and Faculty of Actuaries. Available at <https://web.archive.org/web/20190525190548/https://sias.org.uk/media/1161/extending-the-critical-path.pdf>
- Barry, A. (2002) The anti-political economy, *Economy and Society*, **31**, 2, 268–84.
- Benjamin, R. (2016) Informed Refusal: Toward a Justice-based Bioethics, *Science, Technology & Human Values*, **41**, 6, 967–90.
- Birch, K. (2017) Rethinking Value in the Bio-economy: Finance, Assetization, and the Management of Value, *Science, Technology, & Human Values*, **42**, 3, 460–90.
- Brown, N. (2007) Shifting Tenses: Reconnecting Regimes of Truth and Hope, *Configurations*, **13**, 3, 331–55.
- Bryant, A. and Charmaz, K. (eds.) (2007) *The SAGE handbook of grounded theory*. Los Angeles, CA: Sage Publications.
- Callon, M. (1984) Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay, *The Sociological Review*, **32**, 196–233.
- Callon, M. (ed.) (1998) *The laws of the markets*. Sociological review monograph series. Oxford; Malden, MA: Blackwell Publishers/Sociological Review.
- Castell, S. and Evans, H. (2016) *The One-Way Mirror: Public attitudes to commercial access to health data. Report prepared for the Wellcome Trust*. London: Ipsos MORI Social Research Institute. Available at <https://web.archive.org/web/20181028033617/https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf>
- de la Bellacasa, M.P. (2011) Matters of care in technoscience: Assembling neglected things, *Social Studies of Science*, **41**, 1, 85-106.
- Dove, E.S., Joly, Y. and Knoppers, B.M. (2012) Power to the people: a wiki-governance model for biobanks, *Genome Biology*, **13**, 5, 158.
- Dove, E.S., Kelly, S.E., Lucivero, F., Machirori, M., et al. (2017) Beyond individualism: Is there a place for relational autonomy in clinical practice and research?, *Clinical Ethics*, p.1477750917704156.
- Dussauge, I., Helgesson, C.-F., Lee, F. and Woolgar, S. (2015) On the omnipresence, diversity, and elusiveness of values in the life sciences and medicine. In Dussauge, I., Helgesson, C.-F. and Lee, F. (eds) *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press. pp. 1–28.
- Floridi, L. and Taddeo, M. (2016) What is data ethics?, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **374**, 2083, 20160360.
- Galis, V. and Lee, F. (2014) A sociology of treason: The construction of weakness, *Science, Technology, & Human Values*, **39**, 1, 154-79.

- Gardner, J. and Webster, A. (2017) Accelerating Innovation in the Creation of Biovalue: The Cell and Gene Therapy Catapult, *Science, Technology, & Human Values*, **42**, 5, 925-46.
- Gitelman L. (2013) *“Raw Data” is an Oxymoron*, London: MIT University Press.
- Gottweis, H., Gaskell, G. and Starkbaum, J. (2011) Connecting the public with biobank research: reciprocity matters, *Nature Reviews Genetics*, **12**, 11, 738-9.
- Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., et al. (2010) *The devil’s in the detail: final report of the independent evaluation of the Summary Care Record and Health Space programmes. 2010*. London: University College London.
- Hayden, C. (2007) Taking as Giving: Bioscience, Exchange, and the Politics of Benefit-sharing, *Social Studies of Science*, **37**, 5, 729-58.
- Hays, R. and Daker-White, G. (2015) The care.data consensus? A qualitative analysis of opinions expressed on Twitter, *BMC Public Health*, **15**, 1.
- Heeney, C. (2017) An ‘Ethical Moment’ in Data Sharing, *Science, Technology & Human Values*, **42**, 1, 3–28.
- Hill, E.M., Turner, E.L., Martin, R.M. and Donovan, J.L. (2013) ‘Let’s get the best quality research we can’: public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study, *BMC medical research methodology*, **13**, 72.
- HSCIC (2014) *Data Release Review*. Health and Social Care Information Centre. Available at https://web.archive.org/web/20190525191002/https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/367791/HSCIC_Data_Release_Review_PwC_Final_Report.pdf
- Jasanoff, S. and Kim, SH. (2009) Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea, *Minerva*, **47**, 2, 119-46.
- Latour, B. (1988) *Science in action: how to follow scientists and engineers through society*. Cambridge, Mass: Harvard University Press.
- Latour, B. (2004a) *Politics of nature: how to bring the sciences into democracy*. Cambridge, Mass: Harvard University Press.
- Latour, B. (2004b) Why has critique run out of steam? From matters of fact to matters of concern, *Critical Inquiry*, **30**, 225-48.
- Law, J. (1991) Power, discretion and strategy. In Law, J. (ed.) *A Sociology of Monsters: Essays on Power, Technology and Domination*. London: Routledge. pp. 165–91.
- Lea, W. (2016) *Review of Data Security, Consent and Opt-Outs*. National Data Guardian for Health and Care. Available at https://web.archive.org/web/20181125092716/https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF
- Leonelli, S. (2016) *Data-centric biology: a philosophical study*. Chicago & London: The University of Chicago Press.
- Marres, N. (2005) Issues Spark a Public into Being. In Latour, B. and Weibel, P. (eds) *Making Things Public: Atmospheres of Democracy*. Cambridge & London: MIT Press, pp. 208-17.
- Marris, C. (2015) The construction of imaginaries of the public as a threat to synthetic biology, *Science as Culture*, **24**, 1, 83-98.

- Mol, A. (2002) *The body multiple: ontology in medical practice*. Science and cultural theory. Durham: Duke University Press.
- Mol, A. (2013) Mind your plate! The ontonorms of Dutch dieting, *Social Studies of Science*, **43**, 3, 379-96.
- Montgomery, C.M. (2017) Clinical Trials and the Drive to Material Standardisation, *Science & Technology Studies*, **30**, 4, 30-44.
- Moreira, T. and Palladino, P. (2005) Between Truth and Hope: On Parkinson's Disease, Neurotransplantation and the Production of the 'Self', *History of the Human Sciences*, **18**, 3, 55-82.
- Muniesa, F. (2017) On the political vernaculars of value creation, *Science as Culture*, **26**, 4, 445-54.
- Munns, C. and Basu, S. (2013) The NHS Information Revolution: 'Choice of Control' to 'Choice' and 'Control', *International Review of Law, Computers & Technology*, **27**, 1-2, 124-60.
- O'Doherty, K.C., Burgess, M.M., Edwards, K., Gallagher, R.P., et al. (2011) From consent to institutions: Designing adaptive governance for genomic biobanks, *Social Science & Medicine*, **73**, 3, 367-74.
- Phillips, M., Dove, E.S. and Knoppers, B.M. (2017) Criminal Prohibition of Wrongful Re-identification: Legal Solution or Minefield for Big Data?, *Journal of Bioethical Inquiry*, **14**, 4, 527-39.
- Rosenbloom, S.T., Madison, J.L., Brothers, K.B., Bowton, E.A., et al. (2013) Ethical and practical challenges to studying patients who opt out of large-scale biorepository research, *Journal of the American Medical Informatics Association*, **20**, e2, e221-5.
- Sexton, A., Shepherd, E., Duke-Williams, O. and Eveleigh, A. (2017) A balance of trust in the use of government administrative data, *Archival Science*, **17**, 4, 305-30.
- Singleton, V. and Michael, M. (1993) Actor-networks and ambivalence: general practitioners in the UK cervical screening programme, *Social Studies of Science*, **23**, 2, 227-64.
- Spatz, E.S., Krumholz, H.M. and Moulton, B.W. (2016) The New Era of Informed Consent: Getting to a Reasonable-Patient Standard Through Shared Decision Making, *JAMA*, **315**, 19, 2063.
- Sterckx, S., Rakic, V., Cockbain, J. and Borry, P. (2016) 'You hoped we would sleep walk into accepting the collection of our data': controversies surrounding the UK care.data scheme and their wider relevance for biomedical research, *Medicine, Health Care and Philosophy*, **19**, 2, 177-90.
- Strathern, M. (2000) The tyranny of transparency, *British Educational Research Journal*, **2**, 3, 309-21.
- Taylor, L. (2016) The ethics of big data as a public good: which public? Whose good?, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **374**, 2083, 20160126.
- Taylor, M. (2014) Information Governance as a Force for Good? Lessons to be Learnt from Care.data, *SCRIPTed*, **11**, 1.
- Venturini, T. (2010) Diving in magma: How to explore controversies with actor-network theory, *Public Understanding of Science*, **19**, 3, 258-73.
- Vezyridis, P. and Timmons, S. (2016a) Evolution of primary care databases in UK: a scientometric analysis of research output, *BMJ Open*, **6**, 10, e012785.

- Vezyridis, P. and Timmons, S. (2016b) Dissenting from care. data: an analysis of opt-out forms, *Journal of Medical Ethics*, **42**, 792-796.
- Vezyridis, P. and Timmons, S. (2017) Understanding the care. data conundrum: New information flows for economic growth, *Big Data & Society*, **4**, 1, 2053951716688490.
- Wadmann, S. and Hoeyer, K. (2018) Dangers of the digital fit: Rethinking seamlessness and social sustainability in data-intensive healthcare, *Big Data & Society*, **5**, 1, 2053951717752964.
- Welsh, I. and Wynne, B. (2013) Science, Scientism and Imaginaries of Publics in the UK: Passive Objects, Incipient Threats, *Science as Culture*, **22**, 4, 540-66.
- West, K. and Davis, P. (2011) What is the public value of government action? Towards a (new) pragmatic approach to values questions in public endeavours, *Public Administration*, **89**, 2, 226-41.
- Wilbanks, J.T. and Topol, E.J. (2016) Stop the privatization of health data, *Nature*, **535**, 7612, 345-8.
- Williams-Jones, B. and Graham, J.E. (2003) Actor-Network Theory: a tool to support ethical analysis of commercial genetic testing, *New Genetics and Society*, **22**, 3, 271-96.
- Winickoff, D.E. (2016) From Benefit Sharing to Power Sharing: Partnership Governance in Population Genomics Research. In Kaye, J. and Stranger, M. (eds.) *Principles and Practice in Biobank Governance*. London & New York: Routledge. pp. 53-66.
- Woods, S. (2016) Big Data Governance: Solidarity and the Patient Voice. In Mittelstadt, B.D. and Floridi, L. (eds.) *The Ethics of Biomedical Big Data*. Cham: Springer International Publishing. pp. 221-38.
- Zuboff, S. (2015) Big other: surveillance capitalism and the prospects of an information civilization, *Journal of Information Technology*, **30**, 1, 75-89.