

# Making consent for electronic health and social care data research fit for purpose in the 21st century

Philip Anthony Heslop ,<sup>1,2</sup> Karen Davies,<sup>1</sup> Avan Sayer,<sup>1,2</sup> Miles Witham<sup>1,2</sup>

**To cite:** Heslop PA, Davies K, Sayer A, *et al.* Making consent for electronic health and social care data research fit for purpose in the 21st century. *BMJ Health Care Inform* 2020;**27**:e100128. doi:10.1136/bmjhci-2020-100128

Received 08 January 2020  
Revised 12 May 2020  
Accepted 18 May 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>AGE Research Group, Translational and Clinical Research Institute, Biomedical Research Building, Campus for Ageing & Vitality, Newcastle University, Newcastle University, United Kingdom

<sup>2</sup>NIHR Newcastle Biomedical Research Centre, Faculty of Medical Sciences, Newcastle University and Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom

## Correspondence to

Dr Philip Anthony Heslop;  
philip.heslop@ncl.ac.uk

## INTRODUCTION

There has been significant effort in linking different electronic datasets within health-care and between health and social care,<sup>1–6</sup> both for research purposes and for clinical care. These electronic data are referred to as ‘routinely collected’ as they are collected during usual clinical practice, in contrast to ‘research’ data which are collected in a bespoke manner within the context of a research project. Such datasets (here referred to as routinely collected electronic health and social care data) are already used extensively in epidemiological research, to assess the impact of healthcare interventions in real-world practice, and to provide outcome measures in clinical trials.<sup>7–9</sup>

In wider society, the actions of large personal data processors have brought public scrutiny on the use and misuse of personal data. Previous attempts at large-scale use of electronic health and social care data (eg, the national Electronic Health Record (EHR) in England) have been controversial in concept and problematic in execution, with consequent damage to public confidence in the ability of organisations to act as trustworthy stewards of personal data.<sup>10–13</sup> These negative experiences damage the relationship between individuals, their data and the research community.<sup>14</sup> We argue for moving towards a model of consent and use based on respect for digital personhood, and review some of the technical and governance solutions that could enable this transition.

## Differences in data generation and use

Routinely collected electronic health and social care data differ from data generated by traditional research studies. The primary expectation of research participants in the traditional research model is that their data are collected and used, with consent, for research purposes. In contrast, patients and service users may not know that routinely

collected electronic health and social care data may be used for research. In a traditional research study, there is opportunity to maintain regular contact and therefore foster collaborative, ongoing consent.<sup>15</sup> For researchers using routinely collected data scope for collaborative data use is limited. As routinely collected data are not limited to a specific research need or time period, the volume of information is much greater than would typically be collected in a traditional research study.

## Public attitudes

Most people in the UK are happy to consent to use of their routinely collected electronic health and social care data for ethically approved research by university and National Health Service (NHS) researchers,<sup>16</sup> reflecting the high levels of trust in health-care professionals and researchers.<sup>17 18</sup> Public acceptance is lower for commercial research data use,<sup>19 20</sup> despite the potential added value that commercial partners bring to improving healthcare.<sup>21</sup> This added value is at risk if the public are not convinced that their data will be used in an ethical way.<sup>21</sup> Lack of trust in commercial data providers explains global trends towards greater regulation—for instance, the General Data Protection Regulation (GDPR)<sup>22</sup> in Europe, California’s Consumer Privacy Act (California, 2018) and even companies such as Facebook now openly asking for more government control.<sup>23</sup>

## Digital personhood

The way that personal data are considered within ethical and legal frameworks has undergone an important shift over the last few decades. Personal data are often now considered part of a person in the same way that body parts and tissue samples are part of a person.<sup>24</sup> This concept of ‘digital personhood’ supposes that the data about a person, and the transactions on that data, are an

integral part of their persona, and that rights pertaining to personhood should be extended to incorporate these data.<sup>25</sup>

Ensuring that digital personhood is fully respected hinges on how consent is both conceptualised and operationalised. In a traditional research study, voluntary, uncoerced, transparently and honestly acquired consent is sought at the onset of the research. Consent requires capacity to understand, retain and weigh the relevant information. Close interaction with the research team affords opportunities to ensure that consent is fully informed, opt-in, provides opportunities for participants to change their mind, and in many cases allows participants to personalise the components of the study that they consent to. Each further interaction with the research team presents an opportunity to reaffirm, modify or withdraw consent. This ongoing, two-way process of consent maximises choice and autonomy for the participant.

### Informed, ongoing consent

How can the current processes for research using routinely collected health and social care data be adapted to ensure that digital personhood is fully respected? We propose mechanisms by which the consent relationship can be maximised while preserving the ability to conduct efficient and generalisable research.

### Reciprocal communication

For research using routinely collected electronic health and social care data, reciprocal communication can be challenging, as the individualised pathway to gain and reaffirm consent may not exist at scale in current EHR systems. Technology can help however—communication can be facilitated through ‘patient portal’ applications, allowing researchers to inform participants about research achievements, allow participants to provide feedback and allow participants to exercise their rights over the use of the data, such as the right to withdraw from a particular research use.<sup>26</sup> This would resemble the traditional researcher–participant relationship more closely than the current model.

As with traditional research, upholding the right not to participate in research risks losing participants, compromising the validity and generalisability of research. Nonetheless, this is a crucial part of developing and maintaining trust. Attempts to avoid developing these research relationships are more likely to lead to a loss of trust in the whole research process, with individuals rescinding or restricting access potentially in large numbers. Building these relationships with individuals may also stimulate new opportunities—particularly for individuals to put forward and then shape research questions of importance, facilitating the emergence of genuine codesign and coanalysis in research using routinely collected data.

### Understandable information

Access to transparent information using plain language is essential to support understanding about the potential

use of routinely collected electronic health and social care data in research. Information is currently delivered at population level, via advertising campaigns, information leaflets and information boxes on clinical appointments. Such methods have arguably failed to deliver individualised, tailored information. There is a growing expectation that individuals should be able to access their own health and social care data, and applications such as the patient portal<sup>1</sup> can help to deliver such information. Applications to manage chronic disease are another example of this approach (eg, the My Diabetes My Way application).<sup>27</sup>

### Managing consent

The design of routinely collected data consent systems would benefit greatly from public and patient involvement, either through extensive consultation or fully participatory codesign. The process of consent should incorporate mechanisms that support ease of access for individuals to their data. For consent to be meaningful, individuals need to have a sense of control over their data. They need to be able to request the correction or removal of inaccurate or inappropriately held data, rights currently embedded in GDPR legislation.<sup>22</sup> Mechanisms must enable individuals to withdraw or change consent. The increasing complexity of health and social care records means that it may be technically possible for individuals to express a preference as to what parts of their data can be accessed by different parties, as is the case with the patient portal application being deployed as part of the Great North Care Record.<sup>1</sup> Legislation also requires that preferences are processed in a timely manner, are shown to have been executed and that data no longer relevant to the original consented purpose are removed from linked datasets derived for research analysis (including backups). Supporting these processes carries costs, both in time and money, and these costs need to be acknowledged by researchers and funders.

### Regulatory frameworks

Governance and regulatory frameworks exist to protect the safety and rights of individuals, with most research being submitted, reviewed and approved on a project-by-project basis, an approach developed for traditional research involving small numbers of participants consenting face to face. Balancing regulation against the burden on research teams, funders and institutions is challenging.<sup>28 29</sup> Research involving routinely collected health and social care data typically involves multiple uses of the same datasets, often with access by multiple research groups. An opportunity exists to move away from traditional project-by-project approval to centralised or federated data warehouses, platforms and governance approvals. Governance and regulatory structures currently lag behind these technological advances in data management and research use. The focus of governance should shift to platform level rather than project level—an approach used successfully by the Dundee Health Informatics Centre,<sup>2</sup> in which a single umbrella ethics

approval covers a wide range of health informatics projects, with per-project approval devolved to a data access committee with external oversight. Other related examples are of large scale, reusable data repositories of anonymised medical data (eg, UK Biobank and the Scottish GO-SHARE—Survey of Health, Ageing and Retirement in Europe—project<sup>30 31</sup>). Developing and using such platforms also provide an opportunity for in-depth and sustained collaboration between research teams and the public in the design, governance and delivery of informatics research in a way that may not be possible with multiple standalone projects.

### CHALLENGES, SOLUTIONS AND EXAMPLES

Using routinely collected healthcare and social data for research poses challenges to the consent model used in traditional research studies—both of conceptualisation and of operationalisation. The current UK legal framework (the GDPR) allows use of such data for research without consent under ‘fair use’ provisions. This principle attempts to balance wider societal benefits against the rights contained in the concept of digital personhood. To limit the adverse impact of such use, these provisions stipulate that data must be acquired in a lawful manner, be used exclusively for the set purpose, and that the amount of data collected, and the length of time for which the data are kept, must not exceed what is necessary to achieve that purpose. What is lacking is the opportunity for individuals to give or withhold consent to some or all research uses of their data across an extended time period. Even where such provision does exist, the processes required to enable ongoing consent may be difficult and time consuming, as evidenced by the parallel example of operationalising the ‘right to be forgotten’ provisions within GDPR<sup>32</sup> highlighted by recent cases.<sup>33</sup>

An individualised consent model may not be suitable for all populations, for example, those who cannot give consent, vulnerable groups, those without digital access or skills or those who have died. For these reasons, the application of such a mechanism to a specific research project should be a matter for the ethics committee, allowing certain studies to use alternative mechanisms. Even in such cases, individualised mechanisms would still be valuable as an information or communication channel between researchers and participants. In addition, specific mechanisms to support those who are less digitally literate in exercising their consent choices need to be built into platforms and governance systems.

There is an inherent tension between individual rights and the public good; the concern with moving to a system of individualised consent for use of routinely collected data is that such a system would impair the ability of researchers to deliver generalisable research as a public good. It has recently been argued<sup>34</sup> that the current model is not only sufficient, but that the ethical locus of control for such studies is appropriately sited within authority or governance structures rather than at the level of an

individual contributing data—and that moving away from this current model is not practical. The practical issues can now be addressed using technological solutions, and the ethical arguments in favour of the current system are irrelevant if the public does not concur. A model giving choice to individuals needs to encompass changes of preference over time needs to include feedback on how data are used and to what benefit, thus building the ongoing relationship between researchers and participants.<sup>35</sup> Failure to do so could threaten the trust required for clinical care, and in the long term would render such data useless for research if a significant proportion of the population withheld their data. Although there are concerns that withdrawal of consent by individuals might jeopardise historical data or data already published, these issues are already managed within traditional research governance structures which give clear the limitations on how existing analyses will be amended. Initiatives such as the Wellcome Trust ‘Understanding Patient Data’<sup>36</sup> are essential to raising public awareness and engagement on how data are used without explicit consent for healthcare research and build widespread public buy-in; nonetheless such initiatives should not detract from developing models of explicit consent for such work.

Appropriate technical infrastructure is an integral part of any solution to enabling widespread use of routinely collected health and social care data. Trusted Research Environments (‘Safe Havens’ or ‘Walled Gardens’) are a popular model enabling research use of routinely collected data in a secure and controlled way.<sup>2-5</sup> They are usually implemented as a virtual environment, or for very sensitive data, using a terminal located in a controlled physical environment. They allow control of access, control what analytical tools can be used and importantly control what data can enter or leave the environment (ie, no identifiable data; aggregate results only)—including statistical disclosure control (ie, individuals cannot be identified retrospectively from aggregate outputs if there are rare conditions or small populations). These environments are now being incorporated into electronic healthcare records via projects such as Connected Health Cities.<sup>3</sup> Such environments also enable individuals to withdraw consent for use of their routinely collected healthcare and social data without a risk of identification by researchers.

Centralising the technology and governance frameworks requires strategic investment. Health services have already recognised this need; NHS England has commenced the Global Digital Exemplar initiative,<sup>37</sup> and examples such as the Great North Care Record are creating a regional single point of access for health data.<sup>1</sup> While primarily enabling more joined-up health and social care, these platforms are also well suited for use by researchers. Research organisations such as the National Institute for Health Research Biomedical Research Centres and the Health Data Research UK Digital Innovation Hubs are well placed to oversee and enable this work. Embedding these research environments within clinical

records also provide the necessary universal coverage for population-level research in a way that separate opt-in schemes (eg, GO-SHARE)<sup>31</sup> may not achieve.

## CONCLUSION

The research value of routinely collected electronic health and social care data depends on maintaining high levels of public trust in how researchers use such data. Crucial to maintaining trust is ensuring that digital personhood is respected: robust, ongoing consent processes must be developed to facilitate this. Challenges in how such consent processes are operationalised can be overcome by adopting appropriate organisational, technical and governance structures. Changing public expectations around data use and digital personhood mandate that researchers, funders and health and social organisations implement solutions to support ongoing, flexible and accessible consent processes for routinely collected data research. The alternative—a loss of trust and a consequent loss of this valuable research ability—would harm all of us who use health and social care services.

**Twitter** Philip Anthony Heslop @philipheslop

**Acknowledgements** All the authors acknowledge support from the NIHR Newcastle Biomedical Research Centre.

**Contributors** PAH and MW cowrote the draft of the paper; all authors critically revised the content. All authors contributed significantly to the manuscript.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

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

## ORCID iD

Philip Anthony Heslop <http://orcid.org/0000-0001-9607-102X>

## REFERENCES

- NHS. Great North Care Record [Internet], 2019. Available: <https://www.greatnorthcarerecord.org.uk/> [Accessed 21 Jun 2019].
- Dundee University. Dundee Health Informatics Centre [Internet], 2019. Available: <https://www.dundee.ac.uk/hic/> [Accessed 21 Jun 2019].
- CHC. Connected Health Cities [Internet], 2019. Available: <https://www.connectedhealthcities.org/> [Accessed cited 2019 Jun 21].
- Farr Institute. Farr Institute annual report 2016/2017. *Annu Rep* 2017;52.
- Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
- Health data research UK, 2020. Available: <https://www.hdruk.ac.uk/> [Accessed 2 May 2019].
- Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37–43.
- Clarke GM, Conti S, Wolters AT, et al. Evaluating the impact of healthcare interventions using routine data. *BMJ* 2019;365:12239.
- Walker H, Patton A, Bayne G, et al. Reduction in post-operative acute kidney injury following a change in antibiotic prophylaxis policy for orthopaedic surgery: an observational study. *J Antimicrob Chemother* 2016;71:2598–605.
- BBC. NHS IT system one of “worst fiascos ever”, say MPs [Internet], 2013. Available: <https://www.bbc.co.uk/news/uk-politics-24130684> [Accessed 21 Jun 2019].
- Greenhalgh T, Morris L, Wyatt JC, et al. Introducing a nationally shared electronic patient record: case study comparison of Scotland, England, Wales and Northern Ireland. *Int J Med Inform* 2013;82:e125–38.
- Hendy J, Reeves BC, Fulop N, et al. Information in practice information technology (NPfIT): a qualitative study. *BMJ* 2005;331–6.
- Peltu M, Eason K, Clegg C. How a Sociotechnical approach can help NPfIT deliver better NHS patient care. *Br Comput Soc* 2008;1–28.
- van Staa T-P, Goldacre B, Buchan I, et al. Big health data: the need to earn public trust. *BMJ* 2016;354:i3636–4.
- Davies K, Collerton JC, Jagger C, et al. Engaging the oldest old in research: lessons from the Newcastle 85+ study. *BMC Geriatr* 2010;10:64.
- Riordan F, Papoutsis C, Reed JE, et al. Patient and public attitudes towards informed consent models and levels of awareness of electronic health records in the UK. *Int J Med Inform* 2015;84:237–47.
- Ipsos M. Ipsos mori Veracity index 2018, 2018November):1–12. Available from. Available: [https://www.ipsos.com/sites/default/files/ct/news/documents/2018-11/veracity\\_index\\_2018\\_-\\_topline\\_v1\\_public.pdf](https://www.ipsos.com/sites/default/files/ct/news/documents/2018-11/veracity_index_2018_-_topline_v1_public.pdf)
- Aitken M, de St. Jorre J, Pagliari C, et al. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Med Ethics* 2016;17:1–24.
- YouGov, Lexington. NHS MPs Poll 11th - 23rd March [Internet]. YouGov, 2019. Available: [https://d25d2506sfb94s.cloudfront.net/cumulus\\_uploads/document/ewt22o1jag/LexComms\\_tables\\_MPs\\_Mar19\\_forsite.pdf](https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/ewt22o1jag/LexComms_tables_MPs_Mar19_forsite.pdf)
- YouGov, Lexington. NHS Fieldwork: 11th-12th March [Internet]. YouGov, 2019. Available: [https://d25d2506sfb94s.cloudfront.net/cumulus\\_uploads/document/rfiqqee3vv/Results\\_for\\_Lexington\\_Comms\\_\(NHS\)\\_133\\_12.3.2019.pdf](https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/rfiqqee3vv/Results_for_Lexington_Comms_(NHS)_133_12.3.2019.pdf)
- The Wellcome Trust. *The One-Way Mirror: Public attitudes to commercial access to health data*, 2016: 1–161.
- McCall B. What does the GDPR mean for the medical community? *Lancet* 2018;391:1249–50.
- The Economist. Mark Zuckerberg says he wants more regulation for Facebook [Internet]. The Economist, 2019. Available: <https://www.economist.com/business/2019/04/06/mark-zuckerberg-says-he-wants-more-regulation-for-facebook>
- Price D. The human tissue act 2004. *Mod Law Rev* 2005;68:798–821.
- Clarke R. The digital persona and its application to data surveillance. *The Information Society* 1994;10:77–92.
- Miller T, Bell L. Consenting to what? issues of access, gate-keeping and informed consent. *Ethics Qual Res* 2002;53–69.
- NHS Scotland. My Diabetes My Way [Internet], 2019. Available: <https://www.mydiabetesmyway.scot.nhs.uk/>
- McMahon AD, Conway DI, Macdonald TM, et al. The unintended consequences of clinical trials regulations. *PLoS Med* 2009;3:e1000131.
- Vose JM, Levit LA, Hurley P, et al. Addressing administrative and regulatory burden in cancer clinical trials: summary of a stakeholder survey and workshop hosted by the American Society of clinical oncology and the association of American cancer Institutes. *J Clin Oncol* 2016;34:3796–802.
- Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
- Börsch-Supan A, Hank K, Jürges H. A new comprehensive and international view on ageing: introducing the 'Survey of Health, Ageing and Retirement in Europe'. *Eur J Ageing* 2005;2:245–53.
- General Data Protection Regulation (GDPR). Right to erasure ('right to be forgotten') [Internet], 2019. Available: <https://gdpr-info.eu/art-17-gdpr/>
- Court of Justice. Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González. [Internet], 2014. Available: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:62012CJ0131>
- Sheehan M, Thompson R, Fistein J, et al. Authority and the future of consent in population-level biomedical research. *Public Health Ethics* 2019;12:1–12.

- 35 Salvo MJ. Ethics of engagement: User-Centered design and Rhetorical methodology. *Technical Communication Quarterly* 2001;10:273–90.
- 36 Wellcome Trust Understanding Patient Data [Internet], 2020. Available: <https://understandingpatientdata.org.uk/> [Accessed 2 May 2020].
- 37 Kelsey T, Cavendish W. Personalised health and care 2020: using data and technology to transform outcomes for patients and citizens. A framework for action. *NHS Engl Natl Inf Board* 2014:1–66.