



Oxford

helen.salisbury@phc.ox.ac.uk Follow

Helen on Twitter: @HelenRSalisbury

Cite this as: *BMJ* 2021;373:n1325<http://dx.doi.org/10.1136/bmj.n1325>

Published: 25 May 2021

PRIMARY COLOUR

Helen Salisbury: Should patients worry about their data?

Helen Salisbury *GP*

One of the better things to come out of the pandemic has been a flourishing of medical research, and much of this depends on access to patient data, often derived from general practice electronic records. Some of this research has been enabled by the 2002 Control of Patient Information Regulations, which allow for emergency processing of data—regardless of patient identifiability—to control the spread of an infectious disease.¹

If data are the new oil then the NHS is a very rich field, especially the well coded, digitised GP records that go back decades. But the current position on extracting patient data is a bit of a mess: a patchwork of agreements between different research bodies and individual practices (or groups of practices). For the past three years NHS Digital has worked with the Royal College of General Practitioners (RCGP), the BMA, and others to develop a more coherent structure and process. This new method of data collection, the General Practice Database for Planning and Research (GPDPR), will enable these activities to continue more broadly beyond the pandemic.²

There's now a deadline, 1 July 2021, after which very large volumes of data will be extracted from GP records. Patients who wish to opt out must do so by 23 June. Not everyone is happy that this extraction is going ahead before agreement with all relevant bodies: significantly, the RCGP and the BMA have not endorsed this process.

NHS Digital will be working with a number of industry partners. Data taken from a patient record are de-identified, but if a company with huge reach and access to other data banks wanted to, the data wouldn't be hard to re-identify. If a patient's social media posts reveal where she lives and her children's dates of birth, matching that to her patient record wouldn't be difficult given adequate computing power. This would, of course, be illegal under UK law.

If it comes down to balancing the risk of breaching medical confidentiality against the benefits of research, it's quite possible to argue in favour of the latter. However, it's not a simply binary choice, as there are other ways to do this work without extracting whole medical records to a data store. Techniques developed by Ben Goldacre and his team at OpenSafely allow researchers to ask specific questions and get answers—for example, about the prevalence of a certain condition in people of different ethnic origins or the effect of a drug on renal function—without giving them access to the other details in that record.³ Nothing apart from those answers is extracted, and the threats to confidentiality (however theoretical) are avoided.

Consent to share data involves trust, and the ill fated care.data project five years ago collapsed partly because of patients' mistrust about how their data might be used.⁴ Concern hasn't been voiced on a similar scale this time, but that may be because so few people know about GDPR. Our patients have the choice to opt out—and we have a duty to let them know about it.

Competing interests: See www.bmj.com/about-bmj/freelance-contributors.

Provenance and peer review: Commissioned; not externally peer reviewed.

- 1 NHSX. COPI notice—frequently asked questions. <https://www.nhs.uk/covid-19-response/data-and-covid-19/information-governance/copi-notice-frequently-asked-questions/>.
- 2 NHS Digital. General Practice Data for Planning and Research (GPDPR). <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/general-practice-data-for-planning-and-research>.
- 3 OpenSafely. <https://www.opensafely.org/>.
- 4 Godlee F. What can we salvage from care.data? *BMJ* 2016;354:i3907doi: 10.1136/bmj.i3907.