

When big data can bring big gains to society

Why be wary about sharing health details when we give so much to tech giants?

The Observer Domestic edition · 22 apr. 2018 · Sonia Sodha,

Half of us born after 1960 will be told we have cancer at some point in our lives. Virtually no one will go through life untouched by the disease, whether as a sufferer, a survivor or supporter. So every year, millions of us lace up running shoes, bake cakes and cultivate moustaches in memory of loved ones to raise the cash needed for cancer research.



But money is not the only critical ingredient in developing new cures. An important constraint on the pace of progress is the lack of large datasets that contain valuable information on how patients are responding to different types of treatment.

So it is hugely significant that Tessa Jowell has pledged to make a different sort of donation to cancer research. The former culture secretary has won plaudits for her bravery in speaking out about her experience of suffering from an aggressive brain tumour and her calls for greater collaboration in the fight against cancer. She has now announced she will become the first person to donate her health records to a new global cancer database being set up by an Australian philanthropist. Cancer researchers have hailed it a gamechanger: it will enable them to look at what people responding well to particular cancer treatments have in common so as to refine and develop new treatments.

It might be difficult to recall at a time when concerns about data privacy have never been more heightened, but 10 years ago, big data was being talked about as the panacea to a range of social ills. Like nudge theory, its potential was unquestionably overhyped, but there is no doubt that in relation to healthcare, our personal data has not just life-changing but lifesaving potential. Observational data from the real world has paved the way for

huge medical breakthroughs over the years, including establishing the links between thalidomide and birth defects and between smoking and cancer.

But the power of patient data remains cripplingly underharnessed as a result of privacy concerns. Five years ago, NHS England announced it would be setting up a single, opt-out database – care.data – bringing together millions of patient records. The data would be made available in “pseudonymised” form (with personal information such as name and address removed) to medical researchers in the public and private sectors.

There are, however, privacy risks to such an undertaking that can’t be eliminated. Even when data is pseudonymised, there’s a small risk people can be identified if the information were matched with data from other datasets, for example those held by insurance companies. The public accounts committee said that last year’s NHS ransomware attack was made possible by a lack of investment in cybersecurity – which doesn’t inspire confidence.

The NHS sent every household a leaflet to explain the endeavour that “raised more questions than it answered”, according to the independent health data guardian. It certainly wasn’t clear about what sort of safeguards might be adopted to ensure patient data would not get into the wrong hands.

So the resulting public outcry was perhaps inevitable and resulted in the whole thing being scrapped a couple of years ago. The tragic price? Potential delays to medical advances that could help us find new cures for diseases such as dementia and cancer.

Today, we are awash with examples of how our personal data could potentially be put to malign use. Insurers would love to get their hands on our health data so they can charge us premiums based on how likely we are to become ill. A social media giant such as Facebook is able to use the massive amounts of data it holds on us, including our likes, our private messages and our reading habits, to try to predict how we will react to commercial and political advertising and push us targeted marketing accordingly – a capability its developers are continuously honing.

And companies can use data to evade regulation and help them to break the law. The New York Times last year revealed how Uber had been using passenger data to systematically deceive local law enforcement officials.

Too often, data experts throw up their hands at the way in which we apparently undervalue the privacy of our personal data. The saga over care.data shows us how far from the truth this is: people can get very animated about the privacy of our data when we don’t trust the people charged with keeping it safe and the benefits are either poorly explained or feel obtuse or too far into the future.

But there is a tragic asymmetry in our response to care.data and to the big tech companies.

Revelations about the reams of personal data that Facebook and Google hold about us, including which events in your calendar you actually attended and which apps you use and when, haven’t resulted in masses of people closing their accounts. We seem to be less anguished about under-regulated, global tech giants holding deeply personal information about us than an arm’s-length NHS body.

That asymmetry isn't surprising – it's just yet another product of our irrationality. We trade away our data for short-term benefits such as the convenience of an app, but the risks make us more reluctant to hand it over in service of long-term collective benefits such as improved medical care and new cures for diseases.

If we are to beat diseases such as cancer and dementia, that irrationality is something we need to overcome. Tessa Jowell has taken the first step. Hopefully some of us will follow. Perhaps in 10 years' time, we'll see people running races and baking cakes asking not just for cash sponsorship – but for friends and family to donate their health data.