

Nuffield Council on Bioethics
The linking and use of biological and health data consultation

A personal submission from Dr John Saunders, chair, Royal College of Physicians (RCP) Committee on Ethical Issues in Medicine

Q1: Do biomedical data have special significance?

The key term here is, of course, “biomedical”. How this is construed largely determines the response to the question. Fundamental physiological data – the genetic profile of an individual, for example – are the result of determination, not of freedom; of scientific law and not of choice; of the effect of things around us that we cannot alter and not of things around us that we have willingly (or even wilfully) chosen to do. Expressing this differently they are data that come from the way things are, not the way we might make them to be; from the analysis of structures, not from morality. We cannot be blamed for them: nobody chooses their parents; nobody has responsibility for their own genes. At the most basic level, there is no *reason* why we should be concerned about privacy of our genetic data, any more than we should be concerned about the fact that we exist. It’s just the way things are. On this basis, there appears to be no justification for concern about privacy on the same level as those characteristics that result from what we do, what we choose, what could be otherwise, what reflects upon us as moral beings with practical reason (in that Kantian sense).

However that is to construe “biomedical” very narrowly. On a wider basis, the term may include any number of acquired features: the results of laboratory or radiological investigations that concern not merely what we are born with and what must come to pass, but also those features that we could have avoided: the positive HIV or hepatitis B test, that imaginary gene for criminality, the data that tell us something about our behaviour, of the sort of person that we are. In one sense, these are reflections of social values: they reflect features of our personalities and behaviours that we would prefer others not to know – we drink too much, we have sexual relationships etc that for any number of reasons we do not want others to know: we wish to avoid disapproval, judgment, condemnation. And we fear the control over us that knowledge about us gives to others to infringe our autonomy. Even those things for which we have no responsibility (let us imagine that a blood test for homosexuality existed), may give others power over us (including the power of blackmail) if social values are such as to condemn those things that the data reveal.

Biomedical data then seem to have no special significance in themselves, but do appear to have a very considerable significance in almost any social context – more perhaps in a rigidly controlled society (let us say a Stalinist or Taliban one). The significance of biological data appears to be chiefly a function of its social context, not in itself.

Is this special? It seems not. Biomedical data have significance in the same way as data about our criminality, intelligence, social interactions and a thousand other things which interest commerce, government, science and, of course, the nosy and prurient. In a thought experiment in which we knew or could know everything about everybody, we might cease to be interested in

many human peccadilloes as we realise that they are part of the spectrum of human being and experience.

Are biomedical data, however construed, more sensitive than other data and, if so, is that a contingent or an intrinsic feature as appears to be suggested in the consultation? The obvious comparisons are those data that are not biomedical but about which we have strong beliefs, in many quarters, about privacy. A good example might be financial data. Many of us believe that a healthier society might be one in which all tax returns were publically available and where, implicitly, incomes were no longer a closely guarded secret for large numbers of people (obviously, for many these are publically available through national pay scales.) Yet we do, in fact, have considerable reservations about our financial data. How much we earn and spend, and where it comes from and to where it goes carry huge social significance; and rather like the point made about biomedical data above, such information enables approbation or condemnation or all points in between. On the one hand there could be benefits (shall we say the possibility of less extreme variations or greater generosity to others), but on the other hand such freely available information could lead to greater envy and social strife. It may be in the interests of us all that privacy about our finances is maintained.

It is hard to see any difference between our financial data and biomedical data. Anonymised financial data may be of huge value in social policy, choice about taxation and the public weal and so on. Similarly it is not clear that subsets of data offer any distinct differences either. (As an aside, it is not clear what the word “ethically” adds to the phrase “ethically important benefits” – surely a benefit is an ethical term, implying an ethical judgement).

Privacy concerns go further than the influence or power of others over us. The mythical London taxi driver who responds to the passenger’s parting remark of “have a nice day” with the response, “I’ll bloody well have the sort of day I want” is perhaps merely asserting that he wants to be left alone, not advised or wished about by someone else. Don’t tell me you are praying for me or offer to do so might be a more serious and sensitive example. Individualism asserts itself against social interaction or community membership.

However, we cannot opt out of society: we are part of it, even the most private, solitary, solipsistic human being. All of us are genetically related to someone else (and, theories of cloning aside) to two others at least. We survive because others care for us before we can care for ourselves. There are social as well as genetic reciprocities. The analysis of the “I” is a profound philosophical conundrum to whom this information belongs: and ‘belongs’ is probably not the right concept here. Whatever our jurisprudential concepts, some theists may suggest that nothing in our bodies ‘belongs’ to us: but that is to stray from the practical conclusions that must order a secular society. Genetic data that has implications for others should, on the face of it, be shared if significant benefits (or harms) would otherwise accrue to a third party. In practice, such possibilities are new and unfamiliar to many of us: and consequently we are suspicious of data sharing. An open process of consent is probably the best way to address this, with some proviso for an over-riding judgment by a professional where the harms really are of life changing magnitude.

There are several examples of incomplete epidemiological data due to suspicions of data use. Scientifically this can distort conclusions. There is no simple way to address this. Full anonymisation may prevent identification (anonymise and throw away the ‘key’) and that may be better than failure to proceed. (This technique was used, for example, in the national tonsil archive for CJD). What is not acceptable is the ability to withdraw data once submitted because of suspicion that research might deliver an unpalatable conclusion. Pressure groups do indeed sometimes have strong beliefs about what the answers should be. (There is a crude parallel here with data fabrication by investigators).

Q2. What are the new privacy issues?

The introduction in the consultation paper offers an excellent summary of the current state of affairs and of the tensions that require resolution.

The sheer volume of data that are now available raise concerns that are similar in concept but different in scale. There is still a belief that individual anonymity can be achieved by ‘anonymity’: that is, by not naming. In practice, deductive disclosure may be easier than many patients believe when giving consent to, for example, case reports. Rare diseases may reflect rare genes and the description of the one lead to the deduction of the other. The power of technology to analyse data is vastly greater than even a few years ago. Identification of individuals is far easier. The volume of data and the power to analyse it therefore makes possible a greater privacy threat.

It is hard to construe ‘public interest’ as other than an ethical judgement. We might think that ‘public benefit’ implies peace (or, more minimally, an absence of war). Yet there are just wars. We may also wish to debate whether public benefit extends to the unborn: and, if so, how far into the future. Does public benefit apply to a public that does not exist and may never exist? More cynically, some may (indeed, do) ask whether it extends beyond the lifetime of a parliament. And in turn, to take the question further, we need some theory of the state itself: how far its role extends – a subject of ongoing political philosophy and debate. Should the state have a positive role in encouraging the flourishing of its citizens, a wider role towards all mankind, or a minimal one of basic laws securing its borders – and none in education, health or social welfare. This working party will surely have to state some basic position on this to make sense of what it recommends. But certainly genetic data has implications for social well being and, as we are organised at present, the state is the ultimate guardian of the public interest – and preferably a democratic state, however we construe *that*.

Public interest, almost by definition, is not private interest. As a multi-millionaire I have a private interest in retaining my wealth; the public interest may be a more egalitarian society with fairer wealth distribution. Public and private interests clash. I may want the best for me – in healthcare, income, education, power. Public interest may demand curtailing my choices and freedoms. Such issues are the traditional tensions between left and right, between community and individual concerns. Different societies will reach a consensus at a different point and one that will change over time. There is no formula to set this consensus, no line that can be drawn without debate through this grey area. The only assertion that requires continual emphasis is that

the failure to be able to draw a rigid line through a grey area is not an argument for dismissing the importance of drawing such a line. Or failing to distinguish between the (necessarily) arbitrary and the merely arbitrary. It is a fact of life that the lines that we most importantly draw are those that are arbitrary, yet not merely arbitrary: between, shall we say, youth and old age, between wealth and poverty, between compulsion and liberty. Exactly the same principle applies in determining the line between the precedence of public interest in biomedical data over individual and minority group interests. In practice, UK society has been able to accommodate most minority group interests without a high level of anguish or harm. Discrimination (unjustified discrimination that is) and stigmatisation have not been major issues: a statement that should signal the need for continuing vigilance and concern, rather than one of complacency.

People's data may harm them in the hands of others. Even democratic well intentioned polities make mistakes. Those harms may occur whether the persons concerned know of their data being collected or not. It is not an argument against data collection, but in favour of clearly stated reasons for its collection even if acknowledging that important subsequent uses may arise. These should at least be subject to some form of ethical regulatory review. The term "affect them" in the consultation seems to imply "affect them personally or individually": there is an important distinction here between affecting the individual personally or as affecting that individual as a member of society.

The implications of treating biomedical data as property deserve a detailed discussion. Such a status is at odds with (non) ownership of bodies and concepts of 'stewardship'. At present, this writer would be reluctant to express a firm view. Once again conceptual clarity is important: what does ownership imply, if ownership lies at the core of the concept of 'property'?

Q3 What is the impact of developments in data science and information technology?

The issues set out in the explanation of this question are almost entirely empirical and not conceptual: 'what is the evidence' is an empirical question. The incentives and interests driving research forward and the barriers that resist advances are complex. Given the potential of such data to uncover valuable new knowledge, it seems justifiable to ensure that research about the factors driving or limiting such research should themselves be the subject of scientific study. 'Big data' has a journalistic ring to it: as such it may serve to interest and involve the public. It therefore is a term that serves a useful purpose. Of course, expectations are likely to be excessive in the public mind especially – this is so often the way that science attracts interest. Good information requires promotion that is informative, assists enthusiasm and general interest, but is also ethically appropriate.

Q4 What are the opportunities for, and the impacts of, use of linked biomedical data in research?

The opportunities appear huge. The collection and analysis of data cost time, money and effort whether the costs derive from purely academic centres or from commercial ones. The legitimate interests in 'intellectual property' by academic centres is similar to the interests of commercial companies. It has to be understood by those who cooperate in agreeing to submission of data that these costs arise and that all data cannot be immediately available. This does not mean, of

course, that none should. There is a line to be drawn between data that are genuinely needed to remain secret for justifiable commercial concern, as opposed to data that could (& should) be made available in the interests of openness and transparency: not least, so that those whose data may have contributed can see the benefits that might be accruing.

Researchers, ambitious for their careers, may wish to limit data sharing. These concerns are understandable, but data sharing for faster progress should be encouraged in the wider interests of new knowledge. Sometimes this sharing may occur across national boundaries or between different sectors (academic or commercial). These issues are new in scale. Values may vary between those involved; interests in developing this new knowledge may vary. The ethical conundrums that result may be new for many participants. This may highlight the desirability for a sound understanding of something more than the traditional ethical rules from professional bodies: an understanding of key ethical concepts and principles, a little more subtle than the widely taught ‘four principles’.

Q5 What are the opportunities for, and the impacts of, data linking in medical practice?

Healthcare and its possibilities has become more complicated. The intelligence of a human being however is unchanged and the basic endowment that we are born with will not change.

Education helps, of course. But it is increasingly difficult for individuals to make sense of many scientific realities. As lay people, expertise is lacking. Concepts of risk, for example, are poorly understood. However hard we try to explain some things, they will be inadequately understood. This is no excuse for not trying, only for appreciating limitations on simple statements of understanding and consent. In public life, trust in our institutions and in those who hold office is important. Most of us have little idea of how our taxes are used at a level that is specific, but acknowledge that we must entrust that to elected representatives. In trusting others to promote the values we have only very broadly expressed, we seem to be licensing a limited form of paternalism. This is better acknowledged than disguised. Occasional protests about the possibility of data sharing with others are probably less frequent than surprise at the lack of sharing of data that seems to be so obviously beneficial overall. Overall most patients would welcome the information that comes with risk profiling. Here is a situation where understanding and consent to giving and subsequently receiving the information are surely the right way forward: agree to the test and agree to learn its results – both separate functions.

All of us have some obligations that go beyond the law as part of our citizenship. Many of us believe that where an action requires minimal effort, then, all things being equal (they may not be), we have a moral obligation to act. Examples of this are a moral obligation to vote or to indicate one’s wishes with regard to organ donation. Neither of these two examples really cost us anything, yet both enable public goods. We should do them. Similarly, agreeing to use of our data where there is no reasonable effort required or risk to our own well being perceived, should bring forth our willing cooperation. Law prevents us refusing data in many situations: tax data are not something we can keep secret, infectious disease has mandatory controls and data collection and so on. There will be those who are against everything all of the time, whose attitude to the world in general and to authority in particular is an irrational suspicion, whose world view is of active conspiracies against the individual. Mostly they can be ignored: a small percentage will affect, usually, few outcomes. And there should be no implications for their care.

But let us not ignore the assertion that free riding is morally offensive. If we want an advance in healthcare, we must support research and where participation is minimal (such as contributing data) we should be willing to participate. There should be no threat to the doctor patient relationship if this is handled with sensitivity, openness and truth.

Q6 What are the opportunities for, and the impacts of, using biomedical data outside biomedical research and health care?

Our interests and activities can easily be identified by our purchases. We are reminded of this every time we log on to a website such as Amazon, which will inform us of the things that might interest us. Of course, we could avoid this by shopping in different places and paying by cash. But we don't: the convenience of the credit card is too enticing. By and large we accept the consequences: irritating at times if we are contacted by phone or even by post, but accepted. Buyer beware seems to be the applicable slogan. If such data were even more widely distributed and informed, shall we say, recruitment I suspect that most of us would feel it is an inevitable development of the way we live. We could avoid it but we don't bother. It is difficult to see a practical solution to this. If we have given out our data to others, perhaps we should accept the consequences. It would be possible to have codes of practice or even legislation to limit uses in some way, but the extent of such widespread information exchange is probably beyond simple controls. Laws that cannot be enforced are best avoided. Commerce may profit from accumulating data and selling it on. So did the company that developed the HeLa cell line: it spent time developing the line and it profited. That surely is what commerce does. Commerce can also benefit by developing blocking mechanisms to information transfer for individual use. But there does not seem to be an ethical imperative in this area. Nor is there an obligation to share data, if a research group has invested time and effort in collecting and analyzing it. To this degree the comparison with property or a patented product seems apt.

Q7 What are the opportunities for, and the impacts of, data linking in medical practice?

Information comes with a price: knowledge has benefits but also drawbacks. There are some things that we would rather not know about ourselves: perhaps the strong possibility of untreatable disease or the likely date of our deaths. Knowledge is not an unalloyed good. The greatest hope is for better diagnosis, better personalized treatment, better outcomes: and this, on balance, does seem the more probable outcome. Scientific progress can be inhibited by ethical concerns and rightly so; but it can also be inhibited by fear of the new itself and a certain boldness may be required. There is no reason why this should inhibit the doctor patient relationship if there is continued dialogue to maintain trust. Increasingly the doctor must take on the role of advocate in explaining and advising on the meaning and significance of data to the patient who may be fearful of its use. As for fully anonymised data, there can be no harm and, if there are no consequences from data collection, no justification for excluding their use.