1. Do biomedical data have special significance?

Any kind of data is different from another in terms of size, detail, identifiability and how precious it might be to the data subject, controller or processor and, to this extent, this is a potentially misleading question as one of the most salient attributes of any piece of data is the circumstances under which it was acquired. Obviously this is linked to the dynamic nature of consent - what one agrees to when desperately seeking treatment may be different to what one agrees to when recovered or met further adversity etc. It is hard to consider the ethical implications without thinking about the risks inherent to the data subject / subject of care by the mere existence of biomedical data relating to them and their use. There will always be special significance for people about whom the data are being collected - processors must remain sensitive to that and should be in a position to reassure data subjects that even though there are cases when they will not be identified, they are sensitive to the nature of the data and that it relates to something that is very personal, which was originally collected under a solemn promise of keeping confidence.

Additionally, context always matters, but again this comes down to risks - to the individual, the processors and continued use of such data for research. It also depends on what is meant by "special protection" - does this entail special legislation and procedures? Or is this a means of applying additional controls over and above what already exists for wider sensitive information? The protection measures must all fit within a framework of update and implementation, and should not move them outside the realms of protection that already exist - arguably, the need for special protection indicts the existing provisions. Measures must reflect the sensitivity and preciousness that the data subject applies to the information and the availability that ethically approved research and public health policy projects need in addition to the legal requirements. This is especially pertinent in the case of genomics data - this is unique to the individual and even though one can remove names, addresses, NHS numbers and so forth, this does not remove the uniquely identifying nature of this data and the ease with which it could be attributed to an individual.

2. What are the new privacy issues?

The main issue is about risk to privacy and how this is changed with the arrival of big data sets and biomedical linkage. This is what needs to be considered, so that people can make an informed decision about whether they want to opt out or not, or express a view if they will not be explicitly permitted to do so. Linking across large numbers of databases makes anonymity even more of a myth rather than a feature. There is also very little by way of articulated actual benefit to the

individual or the society in which they live: this is where much work and engagement needs to be done. Nobody has really meaningfully articulated specific interests or benefits brought by big data, and there is plenty of literature on the balance of public good versus individual rights - big data does not change this, it makes it harder to ask people's permission to process their information. One must ask - if people's data are being used in ways that they don't know about and will not affect them, what is the point of them being used in the first place, and how does this fit with changing understandings of consent and real implementation of consent as a dynamic, relational matter?

Social media etc are probably making users think more about them as concepts. That being said, I'm sure we've all shared things that in hindsight we might have thought better of at the time! But what of the privacy of others? Is there really a shift where privacy is being violated or whether feelings might be hurt and trust eroded? There is a lack of education in this area. Just as we were taught at school not to do drugs, perhaps we need more awareness in all education levels from primary school through to Masters level about how to proceed safely online. But if we are being asked to provide a means to allow people to feel comfortable sharing personal information, we refer back to empowering individuals to make informed decisions so that what we mean by autonomy is made explicit and therefore requiring full disclosure of risks in a way not dissimilar to that increasingly being required when patients are offered treatment.

If we were to treat data as property, this would bring a plethora of legislative instruments into the fray, though they are probably already there anyway. This would affect the ethics surrounding their use, however, and raise issues all of its own.

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

Larger scale computing resources are being constructed in terms of storage capacity and processing power, and a host of software tools with varying degrees of reliability and provenance are being deployed. The issue that is coming to the fore relates to data dictionaries and classification of the information assets that are being held, as well as the legal and ethical issues. Additionally, science (not to mention surveillance) could have more access to information resources that they might not have a legal, ethical or reasonable right to access, so the technological and scientific methodology needs to bear this in mind… This should also not preclude other funding opportunities that don't necessarily need data and assets on such a scale but are nevertheless expensive to run. In terms of biomedical data, the benefit to individuals seems to be getting lost in the excitement - where such data is concerned, there must be a demonstrable benefit to the specific individuals about whom the information is collected. But there definitely needs to be some expectations management here - additionally, the definition of big data must remain inclusive: it's not just about accessing millions of people's records, high performance computing methods should also be employed to benefit one person.

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

That's a very good question - we're not sure the research community have articulated this particularly well and we should be careful that traditional techniques should not be excluded or overshadowed by this approach. Is there an issue here about reinforcing the paramountcy of adherence to the ethics of 'good' science? It is clear though that the beneficiaries must be the people about whom the biological information has been collected, and the benefits of population scale studies have to be linked back to individual benefit as well as a basis for sound public health policy decisions and conclusions are about population scale studies.

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

Data completeness and availability - the only barrier that should exist for medical practitioners should be based on the patient's wishes and legitimacy and need to access records, not because the information about an individual is spread across several different system that don't interoperate with one another. Moreover, where practice falls short of the ideal isn't there a requirement for patients to be told that the practitioner attending them may not have access to full information. Too often we hear of scenarios where a patient fails to mention something to a second or third practitioner because (s)he has already provided the information. The issue of transparency is one of ethics, quality and safety in an area which is ripe for both professional and patient/public education.

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

Not sure about opportunities, but impacts would certainly relate to point and purpose, as well as risks to privacy and indeed confidentiality - we must not forget that the original collection of biomedical data has happened under a solemn promise of keeping confidence, there is no fundamental right to this information for researchers or public health studies, but it is permitted. This makes uses outside healthcare provision, research and public health even more privileged and the precise purposes need to be identified.

7. What legal and governance mechanisms might support the ethical linking of biomedical data?

The legal bases for the processing of information should not be changed in reaction to a fad or buzzword technology. There are wider issues regarding the processing of information, but certainly in the UK we have seen a very lenient set of prosecutions in general regarding privacy violations - we don't need more legislation, we do need more consideration of what is currently available, as well as a less liberal interpretation of privacy law by the judiciary. More legislation without a realistic set of processes for its implementation is meaningless and also, arguably, disingenuous and harmful to public confidence. How linked biomedical data fit within this framework is not clear, but there is certainly a bigger issue in general that linkage does not change.

The real issue is how to apply sensible yet responsible governance rules to linked data - if it is being made more identifiable by virtue of having been linked, then more will need to be done in terms of access, use and dissemination within an ethical framework that has appropriate security controls applied. Data Safe Havens are a step in the right direction, but curators of such resources should perhaps treat the resources as identifiable and go over and above prescribed requirements to provide as much assurance as possible.