

September 2015

Response submitted by:

Hilary Burton
hilary.burton@phgfoundation.org

Lucia von Bredow
lucia.vonbredow@phgfoundation.org

The Big Data Dilemma Inquiry

The PHGF welcomes the opportunity to contribute to this consultation. In light of our specific sectoral experience, our response focuses on genomics in the NHS. The following broad issues are considered in our response: infrastructure challenges and solutions; the role of big data in precision medicine; and legal considerations and appropriate safeguards for data sharing.

In particular, we would like to highlight recommendations 12, 13 and 15 from our Realising Genomics in Clinical Practice report (reproduced below), which we have used as the basis for our recommendations.

Recommendation 12: A secure, comprehensive, accessible NHS Database is urgently required that can underpin ongoing genomic sequence interpretation, improve clinical outcomes and support the needs of clinical services. This nationally accessible database should be considered an integral part of NHS genomic testing services and will need to be resourced. Any initiative should be long-term and sustainable.

Recommendation 13: Deposition of data into the secure NHS Database needs to be (i) mandated through enhanced service specification, accreditation, and commissioning and (ii) supported by NHS England policies. Any compulsory data sharing must be consistent with existing regulatory frameworks, and address potential concerns about safeguarding privacy and identifiability.

Recommendation 15: Systems and legal processes need to be put in place to allow the contents of the NHS Database to be shared more widely outside the NHS. In order to address proposed legislative changes, the optimal method of establishing a firm legal basis for sharing identifiable patient data beyond the clinical care of the patient would be to seek routine appropriate consent. This will contribute to building public trust.

In particular, we would like to highlight recommendations 12, 13 and 15 from our Realising Genomics in Clinical Practice report (reproduced below), which we have used as the basis for our recommendations.

The majority of patients support the sharing of their data (with appropriate safeguards) on the basis that by sharing their data they will benefit others in a similar position to themselves

RESPONSE

The opportunities for big data, and the risks

We would like to draw attention to the following risks and opportunities:

There is a need for an infrastructure capable of sharing and learning from health-related big data, and without any inappropriate regulatory barriers which might restrict the potential benefits that the scale of the data affords. This needs to be balanced by safeguards to ensure that data is handled in a manner that protects patient confidentiality and enables patients to opt out of having their data shared in ways that they do not expect (e.g. outside of the NHS, or for purposes other than health).

As part of the 'mutuality contract' of health care provided through the NHS, there should be an expectation for individuals to share their anonymised data within the NHS. Individuals should be encouraged to share their data with non-NHS research bodies, specifically pharma, but there should not be the same expectation. The NHS Constitution now enshrines these expectations.

The majority of patients support the sharing of their data (with appropriate safeguards) on the basis that by sharing their data they will benefit others in a similar position to themselves. It hinges on the fact that their receipt of effective, evidence-based healthcare depends upon others having previously made their data available for research, both within the NHS and in the pharmaceutical sector.

The NHS is in a unique position internationally to use health data in order to innovate and enable new therapies to become available. This is especially the case for patients with rare diseases, for whom the rarity of their condition and paucity of data has until now restricted the potential for research into their condition and the development of therapies. Very few health systems are based on a nationalised system and so have the potential to generate and share anonymised data from such a large population.

Genomics is yielding big data as NGS (Next Generation Sequencing) technologies become part of clinical practice. Within health care, genomics is probably the field that produces the largest quantity of big data. We are in the very early stages of generating genomic big data in clinical practice, and so the time is very apposite for ensuring that the potential for the responsible use of this data is maximised.

If used in a coherent fashion across the NHS, and shared with the research community within the NHS and in pharma, this could enable more personalised approaches and the development of more effective health care: the right treatment is provided to the right person, rather than the one size fits all paradigm that causes iatrogenic illness as well as uses scarce resource in an inefficient manner.

There is a need for an infrastructure capable of sharing and learning from health-related big data, and without any inappropriate regulatory barriers which might restrict the potential benefits that the scale of the data affords.

If pharma in the UK is to be profitable it needs to be able to develop drugs and therapies based on the new paradigm of precision medicine.

In addition, if collated and sufficiently integrated, big data presents opportunities to better understand and identify a range of other health-related solutions, such as for AMR (Antimicrobial Resistance) – to develop vaccines and improve detection of disease outbreaks.

Big data is necessary for the development of precision medicine and to ensure that precision medicine is able to yield the benefits that it has the potential to deliver. Genomic and other -omic based medicine (transcriptomics, glycomics, proteomics, metabolomics etc) can transform health care, but in order to deliver this transformation, it will need to keep refining its understanding of the way that the healthy body functions and of disease mechanisms.

Q: Whether the Government has set out an appropriate and up-to-date path for the continued evolution of big data and the technologies required to support it

The potential of genomic data will remain untapped unless the challenges to data collation and sharing are addressed. These challenges include: technical and practical barriers – the lack of a dedicated and sustainable infrastructure, legal uncertainty (around sharing human genetic data), and the fact that the wider infrastructure is not conducive to collecting data (it often involves laborious non-digital pathways: for an example please see our **briefing on Phenotyping Patients for Genomic Diagnostics**). There is also a need for clear standards so that data that is aggregated is interoperable.

It is our view that a proportionate approach is required, taking account of the purposes of data sharing, linkage, applicable safeguards, potential benefits and risks. Potential barriers to effective implementation include a lack of harmonised approaches between different agencies.

Government intervention through NHS England could support the sharing of data by requiring such sharing as part of commissioning guidance. NHSE could ensure that commissioners make this a contractual requirement, although ensuring that those who want to opt out of sharing their own data outside of the NHS are allowed to do so.

Q: Where gaps persist in the skills needed to take advantage of the opportunities, and be protected from the risks, and how these gaps can be filled

It is our view that interdisciplinary skills are required to (1) analyse big data; (2) develop algorithms to do so; and (3) develop software tools, particularly user-friendly tools to enable non-computational experts to mine data.

Developing expertise is important, but perhaps even more crucial is the need to develop general principles and consensus around what constitutes 'trustworthiness'. In a clinical context this could include developing federated models of sharing within the NHS.

Q: How public understanding of the opportunities, implications and the skills required can be improved, and 'informed consent' secured

Rather than questioning the mechanisms for achieving public understanding and achieving consent, it might be more effective to question how public trust can be facilitated. This might include meaningful public engagement when setting up data repositories and identifying safeguards: also offering opt-out, especially when data will be used for purposes beyond those of original data collection.

Q: Any further support needed from Government to facilitate R&D on big data, including to secure the required capital investment in big data research facilities and for their ongoing operation.

Further support is required to improve the infrastructure: it is the view of the PHGF that the Government should prioritise this demand. In addition, we would also like to highlight the need to focus on sustainability, as relying on research funding inevitably drives short-term solutions.

Thus while capital investment for research facilities is important, the requirement for infrastructure which is purpose-built for big-data (in the health care domain), as well as resources and support for curation and management of that data, is indispensable.

The PHG Foundation is an independent genomics and health policy think tank based in Cambridge, UK. Our mission is making science work for health.

For more information about the PHG Foundation visit
www.phgfoundation.org