The Academy of Medical Sciences

Submission to the British Academy and Royal Society's call for evidence on data governance
November 2016

Summary

- The UK's outstanding medical research base is underpinned by access to data, with significant opportunities presented by increasing capabilities in the collection, linkage and analysis of data.
- The medical research community has been active in exploring data governance, particularly in relation to sharing of health data. The recently published National Data Guardian's 'Review of Data Security, Consent and Opt-outs' outlined steps for establishing a clear governance framework for sharing health data in the UK.
- Secure systems with appropriate safeguards are important to allow reliable and secure data access and linkage whilst protecting individual privacy. There are many examples of good practice in establishing such systems for health data, including the Scottish Health Informatics Programme and the Clinical Practice Research Datalink.
- Ensuring high levels of patient consent will facilitate the collection of high-quality, comprehensive datasets. This is essential as incomplete datasets have the potential to compromise the robustness and validity of research outputs. Extensive engagement with the public, clinicians and wider stakeholder base – using exemplars such as the Million Women Study and UK Biobank – to build transparency and trust around data sharing is key in order to achieve high levels of consent. This engagement will ensure that there is a clear understanding of how, and why, health data might be used for research, the value of such data, and the systems in place to protect data safety. The UK patient data taskforce, initiated by Wellcome, is anticipated to play a key role in facilitating this engagement.

Introduction

The Academy of Medical Sciences promotes advances in medical science and campaigns to ensure that these are translated into healthcare benefits for society. Our elected Fellowship comprises some of the UK's foremost experts in medical science, drawn from a diverse range of research areas, from basic research, through clinical application, to commercialisation and healthcare delivery.

The UK's outstanding research base is underpinned by access to data, which is essential for a large proportion of medical and healthcare research. It is therefore important that any data governance model continues to facilitate access to data for research and support this excellent research base. The medical research community has been active in exploring data governance, particularly in relation to sharing of health data, and the Academy has itself addressed this area through its reports on 'Personal data for public good: using health information in medical research', 'A new pathway for the regulation and governance of health research' and 'Improving the health of the public by 2040', amongst other work.[1,2,3]

Therefore our response focuses on health data, and explores some of the opportunities and challenges in creating an overarching governance framework for use of these data for research.

Opportunities presented by developments in data

As outlined in the Academy's report 'Improving the health of the public by 2040', advances in digital technology are continually opening up ever-increasing volumes of quantitative and qualitative data from a range of health and non-health sources.[3] The nature of 'data' is itself changing, with data increasingly seen as text, image, video and sound, and with many new forms inevitably on the horizon. Changes to data access and ownership are also taking place, with the volume of data held by commercial organisations, for example, dwarfing that held by public bodies; a difference that is set to increase.

The ability for researchers to utilise these rapidly increasing volumes of data is a vital component of any strategy to facilitate research for societal benefit. Enabling this will require a shift in our approach, as a society, to the use and sharing of data for research.

Historically, population data have been the foundation upon which some of the great achievements in medical and health research have been built. Large, representative datasets including those held by the NHS, Government departments, non-Governmental organisations, researchers and the private sector, as well as data generated by individuals, provide extraordinary power to understand the full spectrum and complex interactions of the broad range of factors that drive population and individual health. Combined with increasing computing capability, this will offer unprecedented opportunities to:

- Understand the distribution and determinants of health and disease.
- Explore competing risks and the relative contributions of environmental, behavioural, biological and genetic factors on health and interventions to improve health.
- Develop population-level interventions and diagnostics alongside personalised care, healthcare services and prevention, and evaluate their effectiveness, potentially in real time and at relatively low cost.
- Model future scenarios for non-communicable and infectious disease outcomes.
- Develop early warning and real-time systems for emerging health risks.

Capitalising on this opportunity for medical research will require the development of integrated systems using increasing capabilities in linking various datasets. Organisations such as the Farr Institute are working to find ways to build a better infrastructure for this health data linkage. The development of electronic health records also provide a valuable opportunity to integrate primary and secondary care data with data from across many different care pathways and services in – and indeed outside of – the NHS. With the benefits of having a single national health and social care system, there is potential for the UK to lead the way with health data linkage.

Finally, to avoid vulnerability, such systems must be adaptable to the constantly changing data landscape and to the players within it. It must also operate within a careful ethical and governance framework based on public interest.

Governance requirements

Any data governance model must support and enhance the UK's outstanding research ecosystem, facilitating and strengthening data sharing and linkage with appropriate safeguards. It is essential that any model also facilitates the collection of high-quality, robust datasets, and does not unnecessarily impede access to data for research.

Secure systems for data access

At a meeting on 'Data safe havens' hosted by the Academy, the Medical Research Council and Wellcome, participants considered different models of secure environments for handling data.[4] It was noted that whilst there are many challenges to providing data security, risks can be managed in a number of ways. These include segregation of sensitive data, minimising the movement of data between different locations, effective coding and anonymisation processes of identifiable data, developing agreed criteria for maintaining data safety, and robust recording and archiving of data usage and access.

Some participants noted that access to data should only be allowed for 'approved' researchers.[4] The Confidentiality Advisory Group (CAG) at the Health Research Authority, for instance, currently provides safeguards for access to identifiable patient information by reviewing requests for use of identifiable patient information in medical research where consent cannot be sought, under Section 251 of the NHS Act.

It was also noted at the meeting that appropriate penalties for misuse of data should be put in place to incentivise best practice and accountability in handling of data and to minimise negligent, or even malicious, use. This, however, should go hand in hand with processes and culture that facilitates appropriate data stewardship. The development of a sector-wide training and accreditation programme, directed at individuals and institutions, was suggested.

Greater clarity and harmonisation of guidance and terminology

Consistency across different legislative and policy frameworks, as well as in governance and operational structures, will be essential if we are to capitalise on the full potential of data generated and collated from across different disciplines and sectors. At present, there are numerous sources of guidance on access and standards for health data, with many different bodies involved, such as the Information Commissioner's Office, the National Information Board, the Health Research Authority's CAG and NHS Digital. However, there is no single authoritative voice or source of guidance, which can cause confusion for those trying to navigate this landscape.[2] It would be helpful to consolidate and align guidance where possible, and for further clarity on how the various bodies involved work with one another.

In general there is a key challenge posed by terminology and vocabulary used in communications around data use and governance. There are several terms used for different forms of patient data, such as 'identifiable', 'pseudonymised', 'de-identified' and 'anonymised', amongst others, and it is essential that common definitions are established for these terms to ensure transparency for the system users, clinicians and patients.

Public and wider stakeholder engagement

As identified in 'Personal data for public good: using health information in medical research', health data can be regarded differently, and sometimes as particularly sensitive or private, when compared with other types of data.[1,5] Recent events have led to some erosion of public trust around sharing of health data including NHS England's care.data initiative. Therefore public, clinician and wider stakeholder engagement around use of health data in research is critical. This should help individuals to understand how, and why, health data might be used in research. Clear communication around the value of data sharing and its contribution towards the health and social care system, as well as transparency around data access, will help to build trust around sharing

health data. The recently established UK patient data taskforce initiated by Wellcome will play a valuable role in this engagement.[6]

Consent for data use

There is a risk to medical research if high numbers of patients do not share health data, as this could compromise the robustness and validity of research. A study comparing the care given to affluent and deprived women with breast cancer, for instance, demonstrates how challenges in obtaining consent can introduce bias.[2] At the start of this particular study, patient consent was not required for the review of medical records, but the requirement for consent was introduced later in the study process. Comparing the findings of the original study to a reanalysis of the second smaller dataset of patients who consented showed that the second dataset missed one of the key research findings: that more women from deprived areas, compared with those from affluent areas, presented with locally advanced or metastatic tumours. In addition to averting misleading research findings, access to comprehensive health data is also important to provide the public with equal opportunities to participate in research and to facilitate identification of eligible patients for recruitment into certain studies. Evidence shows that only a small number of patients do not wish to receive direct invitations to participate in research.[1]

There are many nuances to seeking consent for use of health data in line with requirements of relevant regulation and guidance such as the Data Protection Regulation. Seeking and obtaining consent for research can have significant cost implications and be impracticable in some cases. For instance, when seeking re-consent from participants to use data where contact details may be outdated or when seeking consent may cause inconvenience, distress or harm. The Department of Health is currently looking to implement a clear, single framework for governance of health data through a newly proposed consent model. This will be a positive step towards a more transparent, navigable system for data governance, replacing the confused and often opaque systems currently in place where patient opt-outs are interpreted differently across the country. The recent publication of the National Data Guardian's 'Review of Data Security, Consent and Opt-outs' in healthcare outlined next steps for establishing a UK governance framework for sharing health data, proposing an associated 'opt-out' model.[7] The Academy responded to the consultation on this governance model.[8]

Some examples of best practice in data governance and access

Some examples of current systems that successfully allow secure access to linked, anonymised patient data include the **Scottish Health Informatics Programme** and the **Clinical Practice Research Datalink** (CPRD). The **UK Biobank** provides a further example of a successful mechanism for sharing of health data with a long-term follow-up and consent model built into the system.

The **Million Women Study** is a national study of women's health involving more than one million UK women, where disease is monitored through self-reporting, follow-up and record linkage. This study provides a good example of where high-quality, comprehensive data has been used in large research studies, with successful patient engagement and communication. In addition, learnings can be taken from the extensive patient engagement carried out through the **100,000 Genomes Project**, which has also created a secure data governance system for storage and access of sensitive patient data including genomic profiles.

This response was prepared by Liberty Dixon (Policy Officer) and was informed by the Academy's previous policy work and consultation responses. For further information, please contact Liberty Dixon (Liberty.Dixon@acmedsci.ac.uk; +44(0)20 3141 3222).

Academy of Medical Sciences
41 Portland Place
London, W1B 1QH
+44(0)20 3141 3200

info@acmedsci.ac.uk
www.acmedsci.ac.uk

Registered Charity No. 1070618
Registered Company No. 3520281

[1] Academy of Medical Sciences (2006). *Personal data for public good: using health information in medical research. http://www.acmedsci.ac.uk/download.php?f=file&i=13206*
[2] Academy of Medical Sciences (2011). *A new pathway for the regulation and governance of health research.* http://www.acmedsci.ac.uk/download.php?f=file&i=13646
[3] Academy of Medical Sciences (2016). *Improving the health of the public by 2040.* http://www.acmedsci.ac.uk/download.php?f=file&i=37428
[4] Academy of Medical Sciences (2014). *Data in safe havens.* http://www.acmedsci.ac.uk/viewFile/53eb4d247ef80.pdf
[5] Wellcome (2013). *Summary Report of Qualitative Research into Public Attitudes to Personal Data and Linking Personal Data.* https://wellcome.ac.uk/sites/default/files/wtp053205_0.pdf
[6] https://wellcome.ac.uk/news/independent-patient-data-taskforce-announced
[7] National Data Guardian for Health and Care (2016). *Review of Data Security, Consent and Opt-Outs.* https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF
[8] Academy of Medical Sciences (2016). *Response to the Department of Health's consultation on the National Data Guardian for Health and Care's Review of Data Security, Consent and Opt-Outs.* http://www.acmedsci.ac.uk/viewFile/57dfa1b898cb9.pdf