CHAPTER FIFTEEN

A WIN: WIN FOR DATA ACCESS:
BALANCING PUBLIC GOOD WITH PRIVACY CONCERNS

Professor Fiona Stanley AC

INTRODUCTION

This chapter suggests that the current urgent issues facing modern societies demand the best information and knowledge from which decisions can be made. This is vital for governments at all levels, non-government organisations and researchers whose work is used by those making decisions and policy. Such information is commonly available but rarely used, linked, re-used and analysed intelligently to inform such decision-making. As many problems are global, finding, sharing and analysing such data in robust national and international collaborations are essential activities. Such problems include environmental degradation, climate change, global pandemics, increases in obesity and mental ill health, overpopulation and city planning, water, security, crime and youth unrest.

A recent report entitled From Data to Wisdom, prepared for the Prime Minister’s Science, Engineering and Innovation Council (PMSEIC) made several recommendations to put Australia in a strong position to both monitor and analyse these pressing problems internally and to be at the international table, when appropriate, to participate in planning and evaluating global threats.2

1 FAA, FASSA, MSc, MD, FFPHM, FAFPHM, FRACP, FRANZCOG, Hon DSc, Hon DUniv, Hon FRACGP, Hon MD, Hon FRCPCH; Director, Telethon Institute for Child Health Research; Professor, School of Paediatrics and Child Health, The University of Western Australia.

One major issue in population data linkage is the balance between using individual health records on the total population for important public good activities, while at the same time ensuring that such private information is kept confidential. The rationale for using such data includes obtaining accurate and unbiased assessments of risks of disease and the effects of medical care. A win:win process to allow access and to protect privacy that has been developed and used in Western Australia for over 30 years is described below.

RATIONALE FOR DATA SHARING AND e-RESEARCH – THE PMSEIC REPORT

The science data challenges facing Australia and all other countries include:

1. The exponential increase in data assets and how they can be turned into knowledge and wisdom;
2. The lack of data in some vital areas;
3. The increasing diversity of data (from images to languages);
4. The vulnerability of data (as the data age or the software to read them become obsolete);
5. The lack of capability in data management;
6. Missed opportunities to collaborate (which is why e-research is so vital, particularly for Australia to enable better use of data and to overcome our relative isolation);
7. Impediments to discover, preserve, share and re-use data (collected and kept in silos so many people collect even more data without knowing what is already available – the need for good longitudinal data for such things as climate change make the discovery of such information vital);
8. Lack of relevant skills (such as mathematical biologists or biological/genetic mathematicians);
The vision which the Working Group had for Australia is:

Australia is managing increasing volumes and complexity of data to enhance our country’s scientific, economic and social prosperity and to protect it from threats.

The key data issues which we included in the report were digitisation, capture, preservation, storage, discoverability, integration, interoperability, sharing, re-use, accessibility (for users), security and privacy.

The international community has clear policies about open access and data sharing with OECD (www.oecd.org), ICSU (www.icsu.org) as well as Europe and the Americas having clear philosophies and guidelines encouraging open access and wider use of research data because productivity and quality will be increased. The Australian Bureau of Meteorology summarised it well:

… foregoing proprietary rights to data and making them freely available actually benefits the individual as well as the community at large …

The recommendations from the report are included as an appendix (Appendix 1) to this chapter. In addition to those about repositories and changing the culture about sharing, access and collaboration, recommendation eight clearly outlined the need to allow researchers to access and link individual data on populations, while protecting privacy.

POPULATION DATA AND RECORD LINKAGE

Record linkage brings together records from different sources relating to the same individual. It is used for administrative purposes, case management and investigation of crimes etc, and also for describing population trends and characteristics in important areas such as the health and well being of the population. It is in the context of maternal, child, adolescent, family and Aboriginal health and development that my group has the most experience in using individual records and linking them together to answer important questions aimed at improving outcomes.
Our interest in obtaining record-linked data began in the late 1970s when there were major concerns about thalidomide and hence other drugs causing birth defects and also whether the introduction of aggressive neonatal intensive care and resuscitation of premature babies would lead to increases in brain damage. We established registers of both birth defects (1980) and cerebral palsies (1977) and then linked them back in to birth registrations and perinatal data to create the Maternal and Child Health Research Data Base. Not only did the linkage enable us to study the patterns and causes of all major perinatal and paediatric problems for the whole population, we also confirmed that intensive care did increase cerebral palsy rates in preterm babies, that birth asphyxia was a rare cause of cerebral palsy and commenced

---

the birth defects studies that led to the confirmation that periconceptual folate prevented spina bifida and related defects.4

Figure 1 shows how the original linkage has now grown to include a number of additional databases, with the capacity to link in drugs, immunisations, hospitalisations, cancers and mental health problems. These data on the total population which come from a variety of agencies, statutory and vital statistical collections and special registers and studies provide WA with the one of the most comprehensive tools for monitoring, evaluating and investigating health and health services in the world.5 Most of the data are collected for administrative purposes and are brought together for re-use to answer specific questions about disease occurrence or health service quality or effects. A list of projects which have been done using data linkage in WA is available at www.populationhealth.uwa.edu.au/__data/page/63033/projects_1995-2003.pdf.

The advantages of record linkage are:

1. Large sample sizes, no exclusions and hence unbiased data for analysis;
2. Cheap compared with studies which trace individuals, seek consent and collect information directly;
3. Valid and reliable data are obtained on sensitive issues so that individuals are not upset by the research;
4. Survey burden on populations is reduced (particularly relevant to some population subgroups such as Aboriginal who ‘have been researched to death’ as one of our Aboriginal researchers noted);
5. Fast effective linkage technology is now available;
6. Privacy can be protected by technology and protocols;
7. Better data for policy, planning and evaluation; and
8. Administrative data sets are used and improved.

---

Of course we are limited by what is available in the data and privacy issues still need to be addressed.

**HOW CONSENT CAN LEAD TO BIAS**

Consent is essential for all research involving the direct participation of individuals such as questionnaires, interviews, blood or other tissue sampling and clinical trials. But not all research requires consent and with population data the seeking of consent is neither feasible, cost-effective nor scientifically valid.

Non-participation in surveys where consent is sought comes mostly from an inability to trace the person and rarely from refusal to participate. Most people participating in studies conducted in our institute do so because they are altruistic and wish to help others.\(^6\) Non-participation is variable, unpredictable and can result in significant bias. Bias is defined as the distortion of the true relationship between exposure and outcome due to flaws in either study design or analysis. People who do not participate in studies can be very different from those who do and hence the analysis can be biased. An example is HIV status where anonymous testing of all blood samples in a population shows much higher rates than when consent is sought. This could result in very different services being developed.

The magnitude and direction of bias from seeking consent in epidemiological studies is unpredictable, not quantifiable, may well explain the differences in risks between studies (e.g. coffee is good for you one day and bad the next!) and provides poor information for health services and epidemiological research.

Examples of significant bias in studies which sought consent include a large meta-analysis of breast cancer and termination of pregnancy and the Canadian stroke register.\(^7\) In both of these the data obtained from seeking consent and interviewing patients produced such biased data as to be clinically and scientifically useless. The stroke register spent over

---


$500 000 over two years on research nurses whose only job was to seek consent from the patients and their carers. A no-consent register would have obtained 100% of patients and been useful to study the natural history of the disease, whether early treatments worked and what their long term impact would have been – all vital questions to help improve the health care system.

HARMONISING PRIVACY AND ACCESS: CAN WE HAVE A WIN-WIN?

In Australia and many other countries, legislation allows access to individual data and record linkage without consent under certain conditions which are governed by ethics committees and medical research funding agency guidelines (see www.nhmrc.gov.au). Ethics committees are guided to allow these activities when it is impractical to seek consent.

Both in Australia and the UK there is a trend towards serving the privacy lobby agenda at the expense of allowing data access for the public good. This has reduced the amount of record linkage and use of health records for research considerably in UK. One of the main barriers is the interpretations of the legislation by ethics committees. There seems to be a poor public (and ethics committee) awareness of the methods of such research and its value to society.

The WA Data linkage protocols we have developed over many years provide linked databases to researchers who have approval from ethics committees, who have an approved protocol with the Data Linkage Unit committee and approval by the Confidentiality of Health Information Committee. The preparation of the linked data base is done in two stages – the identified information is linked without any of the clinical or sensitive information attached and then the linked information without any identifiers is given to the researchers. Analyses are done on data sets that cannot identify individuals. We believe this is a win:win.

---


FINAL COMMENTS

There have been some recent activities and debates about these issues in UK, USA, Canada and Australia as researchers, policy makers and others realise that the pendulum may have swung too far in favour of privacy. The Australian Law Reform Commission is currently reviewing the privacy legislation and may well help to clarify these issues. The NHMRC is just about to publish a new set of guidelines which outline the Australian code for the responsible conduct of research.

In the UK Academy of Medicine report, Professor John Harris (University of Manchester) asked ‘are patients morally obliged to participate in research projects as a “mandatory contribution to public good,” particularly for those aimed at preventing serious harms and providing important benefits?’

And, at an Australian conference on harmonising privacy and access Professor Eric Meslin commented: ‘[w]e are optimistic that a win:win is possible where privacy is protected, where important health research can proceed. There is evidence that a shift in thinking is possible on behalf of ethics review committees, regulators, researchers and the public.’10

We believe that we can work towards a win:win. For this to happen we need ethics committees to understand and accept current guidelines which allow identifiable data to be used for research without consent as long as privacy issues are addressed and the rationale is acceptable. We need to develop Australian privacy and research best practice for the researchers using personal data. We need to make the public more aware of how personal records are used and how research is done and why this is of great benefit to society.

I would like to end with a quote from Professor Lawrence O Gostin:

> In the late 20th century, scholars and politicians posed a key question. “What desires and needs do you have as an autonomous rights bearing person to privacy, liberty and free enterprise?” Now it is important to ask another kind of question “What kind of community do you want and deserve

---

10 Professor Eric Meslin, Indiana University, Centre for Bioethics (2005) personal communication.
APPENDIX ONE – RECOMMENDATIONS FROM PRIME MINISTER'S SCIENCE, ENGINEERING AND INNOVATION COUNCIL (PMSEIC)

A National Strategic Framework for Scientific Data

**Recommendation 1**
That Australia’s government, science, research and business communities establish a nationally supported long-term strategic framework for scientific data management including guiding principles, policies, best practices and infrastructure.

**Recommendation 2**
That a high level expert committee be established to provide the leadership role in progressing the formation of the long-term strategic framework for scientific data management.

The National Network of Digital Repositories

**Recommendation 3**
That the necessary policy and programmes be implemented with a view to establishing a sustainable publicly funded network of federated digital repositories.

**Recommendation 4**
That the expert committee consider the development of a strategic roadmap for the implementation and evolution of the national network of federated digital repositories.

---

Data Management, Access, Sharing and Collaboration – Changing the Culture

Recommendation 5
That standards-based technologies be adopted and that their use be widely promoted to ensure interoperability between data, metadata, and data management systems, providing authentic users of the data with appropriate processes and safeguards.

Recommendation 6
That the principle of open equitable access to publicly-funded scientific data be adopted wherever possible and that this principle be taken into consideration in the development of data for science and programmes. As part of this strategy, and to enable current and future data and information resources to be shared, mechanisms to enable the discovery of, and access to, data and information resources must be encouraged.

Recommendation 7
That funding agencies offer incentives to encourage researchers and institutions to:

- Develop data management plans for each research grant application involving data collection and generation, and that standards be made freely available and widely disseminated so as to encourage best practice in data management;
- Introduce policies and practices to encourage collaboration and sharing of data across Australia’s scientific research institutions and across agencies;
- Analyse and re-use existing data.

Ensuring there are no Regulatory Impediments

Recommendation 8
That funding agencies such as the NHMRC and ARC ensure that best practices and policies are developed and followed that allow bona fide researchers to access individual population data, including and linking of data from multiple sources, whilst protecting privacy, and ensuring that ethics committees fully understand these policies and their rationale.
Recommendation 9
That in the context of developing the strategic framework for scientific data management, Australia’s intellectual property approaches be checked to ensure they do not impede the sharing of data.
In particular, it should take into account the OECD Committee for Scientific and Technological Policy guidelines on access to research data and the International Council for Science statements about the benefits of sharing data.

Skills for Data Management

Recommendation 10
That data management expertise becomes a core skill for researchers, including graduate and postgraduate science students across all disciplines, and that they receive data management training as part of their education.

Recommendation 11
That the Australian Government give early consideration to the finding of the e-Research Coordinating Committee regarding changing research behaviour, practices and skills12.

---