
Sharing research data to improve public health: A joint statement by funders of health research

Introduction

Recent advances in information technology have revolutionised science - providing new opportunities for researchers to share data and build on one another's work. Informatics and the ability to mine large datasets and combine them with information from many other sources presents a huge potential to advance developments in public health. The importance of data sharing in advancing health is becoming increasingly widely recognised, and has been strongly endorsed by the H8 group of global health organisations.

In some research fields – such as genetics and physics – data sharing is well-established and has accelerated the progress of research and its application for the public good. In public health research, however, while research collaborations are growing more common, the sharing of data is not yet the norm, even within the scientific community.

Much of the data collection that could improve public health research is expensive and time-consuming. As public and charitable funders of this research, we believe that making research data sets available to investigators beyond the original research team in a timely and responsible manner, subject to appropriate safeguards, will generate three key benefits:

- Faster progress in improving health
- Better value for money
- Higher quality science

Each funding institution will work within its own legal and operational framework, and we are committed to working towards these goals together. We intend to establish joint working groups where appropriate. We call on governments and other actors that generate routine health service statistics and other types of public health data to adopt a similar approach.

This Statement establishes guiding principles and desired goals. It recognizes that flexibility and a variety of approaches will be needed in order to balance the rights of the individuals and communities that contribute data, the investigators that design research and collect and analyse data, and the wider scientific community that might productively use data for further research.

The Joint Statement of Purpose

Vision

We, as funders of health research, intend to work together to increase the availability to the scientific community of the research data we fund that is collected from populations for the purpose of health research*, and to promote the efficient use of those data to accelerate improvements in public health.

Principles

Funders agree to promote greater access to and use of data in ways that are:

Equitable: Any approach to the sharing of data should recognize and balance the needs of researchers who generate and use data, other analysts who may want to reuse those data, and communities and funders who expect health benefits to arise from research.

Ethical: All data sharing should protect the privacy of individuals and the dignity of communities, while simultaneously respecting the imperative to improve public health through the most productive use of data.

Efficient: Any approach to data sharing should improve the quality and value of research and increase its contribution to improving public health. Approaches should be proportionate and build on existing practice and reduce unnecessary duplication and competition.

Goals

While we recognize that progress may be gradual as we develop mechanisms and resources consistent with these principles, we aim to work in concert to achieve the following:

Immediate goals:

Data management standards support data sharing

Standards of data management are developed, promoted and entrenched so that research data can be shared routinely, and re-used effectively.

Data sharing is recognized as a professional achievement

Funders and employers of researchers recognize data management and sharing of well-managed datasets as an important professional indicator of success in research.

Secondary data users respect the rights of producers and add value to the data they use

Researchers creating data sets for secondary analysis from shared primary data are expected to share those data sets and act with integrity and in line with good practice - giving due acknowledgement to the generators of the original data.

Longer Term Aspirations:

Well documented data sets are available for secondary analysis

Data collected for health research are made available to the scientific community for analysis which adds value to existing knowledge and which leads to improvements in health.

Capacity to manage and analyse data is strengthened

The research community, particularly those collecting data in developing countries, develop the capacity to manage and analyse those data locally, as well as contributing to international analysis efforts.

* The statement does not currently cover routinely-collected clinical data, or public health data to which the signatories of this statement contribute no funding.

Published work and data are linked and archived

To the extent possible, datasets underpinning research papers in peer-reviewed journals are archived and made available to other researchers in a clear and transparent manner.

Data sharing is sustainably resourced for the long term

The human and technical resources and infrastructures needed to support data management, archiving and access are developed and supported for long-term sustainability.

Appendix 1: Development of the Joint Statement

In recent years, many of the major funders of public health research have adopted policies on data sharing. International institutions and major medical journals have also made declarations about the importance of increasing access to public health data.

In 2008, the World Health Organization and the Wellcome Trust initiated discussions about the development of a common Code of Conduct on the sharing of data of public health importance. The aim was to develop principles and targets that would meet the needs of researchers and science funders alike, while providing the flexibility for individual institutions to respond to different circumstances.

The Wellcome Trust commissioned an analysis of obstacles to data sharing, which concluded that obstacles could be grouped into four major areas: incentive-related, capacity-related, ethical and technical. Together with the WHO, the Wellcome Trust convened a meeting held in London on October 6th, 2008 to discuss potential ways of overcoming these obstacles. Epidemiologists and data managers from several continents gathered with a number of representatives from governments, international organizations and major funders of public health research. Together, they suggested core principles for a Code of Conduct.

A draft version of the Code was then presented at the Global Ministerial Forum on Research for Health, held in Bamako in November 2008. Further discussions were held with groups interested in the implications of the Code throughout 2009. These included editors of major medical journals, and the funders and consumers of health research grouped in the H8. The Hewlett Foundation commissioned a paper looking specifically at data sharing issues from the point of view of foundations interested in capacity development in developing countries. The Lancet, the British Medical Journal, the New England Journal of Medicine, PLoS Medicine and Nature have all recently published commentaries calling for greater sharing of data of importance to public health and Annals and the BMJ now require authors to provide a statement on what data, if any, are available to other researchers, and under what conditions.

The H8, which groups WHO, a number of other United Nations agencies and several major funders of public health research, in January 2010 published a paper on data sharing which committed to develop a Code of Conduct in data sharing. In April 2010, the World Bank announced that it would make its data base open access, and would adopt disclosure as the default position for both data and documents relating to lending.

The informal working group developing this Code has been in active discussion with all of these partners. It has also been talking to groups such as MalariaGen and INDEPTH that have been involved in increasing access to data generated by their research partners, and those such as UK Data Archive that have experience with developing infrastructure for sharing. Their contributions and experiences have been used to inform a revised Code of Conduct, which was discussed at a high level meeting of research funders and partners from the research community in Foggy Bottom, Washington DC on May 25th, 2010. Meeting participants felt that action to enable more effective re-use of research data was urgently needed, but legal and institutional constraints required a more flexible framework. This Joint Statement is the result of those discussions. Signatories are committed to engaging with other funders of health research (especially those based in developing countries) as well as with other members of the research community in order to invite greater participation in discussion, faster implementation of agreed measures and more effective progress towards our common purpose.