Karmela Krleza-Jeric[1], Iain Hrynaszkiewicz[2]

ENVIRONMENTAL SCAN OF REPOSITORIES OF CLINICAL RESEARCH DATA: HOW FAR HAVE WE GOT WITH PUBLIC DISCLOSURE OF TRIAL DATA?

Executive Summary (November 2012)

Project outline

There is growing interest in sharing of research data, as data sharing is expected to accelerate research and increase accountability. Many opportunities for data sharing are enabled by the Internet and electronic medical record keeping but some barriers to data sharing still remain. In clinical research data sharing is expected to decrease the burden on research subjects, enabling more efficient and reliable research; promote innovations and entrepreneurship; and facilitate knowledge synthesis, discovery and translation that could ultimately inform clinical decision making. Sharing of data underlying published research is an increasingly important consideration for peer-reviewed journals, and some journals require researchers to state their data sharing policy in published articles.

However, the field of clinical trials has been slow to adopt a culture of data transparency. Previous research indicates that there are practical, legal, and ethical considerations that must be addressed to enable public disclosure of data.

Data repositories are key elements for achieving broad data sharing but in clinical research there is no repository for raw data which covers a variety of medical disciplines. Furthermore, to our knowledge there is no wide agreement on the standards, best practices and essential features of clinical data repositories.

This project aims to address these gaps in knowledge and produce comprehensive information on the features and practices of existing repositories which have common interests, or potential interests, in sharing and public disclosure of clinical trials data.

The methodology of this study includes reviewing existing resources that catalogue information of data repositories, such as Databib (http://databib.org/), literature review, analysis of websites of repositories, and engagement of relevant stakeholders – such as interviews with repository managers.

We aim to capture the methods of existing repositories for public disclosure of clinical data, and non-public forms of data sharing, such as the unique and persistent identification systems for datasets; the license; use or other agreements employed by the repositories; sustainability (business) models. We aim to understand how they have addressed these issues and summarize what are considered good practices. We also plan to gather information on how repositories define raw data and meta-data; data formats and standards; methodology of data preparation; privacy; standards of quality control; policies and terms of data inclusion and access to data for

[1] University of Ottawa, Ottawa group, CMJ, kkrlezaj@uottawa.ca
[2] BioMed Central Ltd, iain.hrynaszkiewicz@biomedcentral.com

re-use; system architecture. We will also identify features that encourage data sharing across geographical and domain boundaries, such as networks of repositories (federated) vs. centralized repository, and collaborations with other stakeholders in clinical research data including journals and publishers.

Expected impact and knowledge translation plan

Results and findings of this study will be made publicly available and used to inform the elaboration of study on development of methodology for data sharing of clinical trials.

The outcomes of this research will also inform development of methodologies of public disclosure of data, as well as standards and guidelines for data repositories involved in the public disclosure of participant level datasets of clinical trials.

Furthermore, the outcomes of this research may encourage collaboration between repositories on areas of common interest, such as appropriate funding models, protecting individual privacy and foster collaboration between journals, publishers and data repositories to help enhance the reliability and connectedness of the scientific literature.

Transparency

In the spirit of transparency we shall share the development of this study on the Ottawa group website (http//ottawagroup.ohri.ca) and via the BioMed Central http://blogs.biomedcentral.com/bmcblog/. We are inviting you to follow, comment, suggest additions, and contribute to this study by communicating with us via these websites as well as directly.

Funder: COPE- Committee on Publication Ethics