



Ethics in the Era of Big Data

G. Owen Schaefer¹

Published online: 18 June 2019

© National University of Singapore and Springer Nature Singapore Pte Ltd. 2019

While health researchers have long relied on data to develop insights and innovations, it is generally recognised that in recent years, there has been a substantial shift—we no longer talk about just data, but *big* data. The big data era is characterised not only by the very large quantity of data that is being generated but also by the depth of detail the data contain and the ability of researchers to rapidly transmit, link and analyse the data. In this context, perennial ethical concerns such as obtaining valid informed consent and respecting privacy take on a new character, and it is not clear if existing approaches to these ethical issues are still fit for purpose. Yet if new (or at least revised) approaches are needed, what should they look like? How should they be formulated, and by whom?

The Science, Health and Policy-relevant Ethics in Singapore (SHAPES) initiative is currently engaged in responding to such challenges. SHAPES, funded by Singapore's National Medical Research Council and housed by the Centre for Biomedical Ethics, Yong Loo Lin School of Medicine, National University of Singapore, conducts analyses and research on ethical issues of relevance to those involved in the practice and oversight of biomedical researchers. The degree of enthusiasm as well as concern over the expanding role of big data in contemporary research has naturally led SHAPES to focus on big data as one of its core areas of investigation.

In March 2018, SHAPES convened a symposium on big data ethics to develop an understanding of the most important and under-addressed issues in the area. Authors of the papers published in this special section were present at that symposium, to discuss and deliberate over their perspectives. And indeed, all three papers present important ideas that lie at the forefront of these debates.

Abdul Aziz and Mohd Yusof (2019) critically appraise existing approaches of broad or blanket consent, where participants engage only once in a consent process while research continues for years, expanding into areas that could not be anticipated at the outset. They argue that broad and blanket consent does not live up to Malaysian legal principles, and the ethical norms that underpin them, by failing to convey appropriate

✉ G. Owen Schaefer
owen_schaefer@nus.edu.sg

¹ Science, Health and Policy-relevant Ethics in Singapore (SHAPES), Centre for Biomedical Ethics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

information and control that subjects deserve. The new paradigm of dynamic consent, where online tools give participants information to and some degree of control over the use of data about them, is offered as an approach to address these issues.

Erwin Tantoso and colleagues (2019) suggest that current practices overemphasise privacy at the expense of public interest while obscuring under-examined issues like the ethical ramifications of failure to ‘clean’ data properly. They offer an important perspective of practitioners—bioinformaticians who are motivated to develop knowledge that will benefit patients but frustrated by the perception that existing oversight mechanisms inhibit that goal. Of course, the social value of science is not guaranteed—data research is only as valuable as the quality of the data that is used, and so there is an ethical imperative to ensure the quality and integrity of data that is used.

And while much of the current discourse is focused on clinical or research data, Gwendolyn Gilbert, Chris Degeling and Jane Johnson (2019) explore how disease surveillance data must be considered in ethical deliberations as well. Public health ethics and research ethics are increasingly becoming intertwined, as the data generated in surveillance is of substantial value beyond the immediate purpose for which it is gathered. But how to ensure it is used responsibly? Their paper ends with a series of difficult questions on the use of surveillance data that are prompts for further ethical analysis and deliberation.

One common thread that emerged was that existing paradigms of consent and anonymisation are inadequate to address the challenges of health research in the big data era. As such, careful deliberation over appropriate systems of use, management and oversight of data is needed by a variety of stakeholders, including researchers, regulators, data access committee members and institutional leaders.

In light of this, SHAPES has convened an international working group to develop a framework for deliberation over ethical issues of big data in health research, spurred on by discussions at the March 2018 workshop from which this special section derives. The aim of the framework is not to itself provide recommendations or directions concerning appropriate courses of action in relation to specific health data research. The divergence of contexts in which and purposes for which data is being used makes a one-size-fits-all approach untenable. Instead, the framework will outline a series of guiding values to be taken into account when stakeholders deliberate over the relevant issues and suggest a systematic approach to those deliberations. Moreover, the framework will apply those values to a series of particular topical areas (such as open data access or real-world evidence), highlighting for stakeholders relevant issues at play and how reasonable deliberation might occur over them.

As of this writing, the framework is still under development, and it is expected to be released later on in 2019. But while we hope for it to be a valuable resource, it is still only one part of a much larger conversation over big data ethics. That conversation, aided by articles like those that appear in this special section, will continue to shape policy and practice around the world.

References

- Abdul Aziz, Mohammad Firdaus, and Aimi Nadia Mohd Yusof. 2019. Can dynamic consent facilitate the protection of biomedical big data in biobanking in Malaysia? *Asian Bioethics Review* 11 (2) 1-14. <https://doi.org/10.1007/s41649-019-00086-2> .

- Gilbert, Gwendolyn, Chris Degeling, and Jane Johnson. 2019. Communicable disease surveillance ethics in the age of big data and new technology. *Asian Bioethics Review* 11 (2) 1-14. <https://doi.org/10.1007/s41649-019-00087-1> .
- Tantoso, Erwin, Wing-Cheong Wong, Wei Hong Tay, Joanne Lee, Swati Sinha, Birgit Eisenhaber, and Frank Eisenhaber. 2019. Hypocrisy around medical patient data: Issues of access for biomedical research, data quality, usefulness for the purpose and omics data as game changer. *Asian Bioethics Review* 11 (2) 1-14. <https://doi.org/10.1007/s41649-019-00085-3>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.