

# Applying the Reason Model to enhance health record research in the age of ‘big data’

Rajan Ragupathy, Vithya Yogarajan

There is considerable international interest in applying powerful data research techniques (‘big data’) to health records.<sup>1</sup> This will become increasingly viable in New Zealand as previously siloed records held by private primary providers and public secondary institutions become accessible at the regional level—such as through the Midland Clinical Portal—and perhaps eventually even at the national level. Such research could, for instance, uncover the real-world safety and effectiveness of new health technologies, identify health disparities or optimise application of international guidelines in New Zealand.

However, such research opportunities will only be viable if the public can trust that patient confidentiality will be protected. Similarly, providers would rightly be wary of making records available for research if doing so might breach legal and ethical obligations, such as those in the Health and Disability Consumers’ Code of Rights and the Health Information Privacy Code. We propose here a framework—drawing from patient safety methods in healthcare and ‘big data’ safeguards already in use in New Zealand—for protecting the confidentiality of health records in research.

Many health professionals will be familiar with the Reason Model of patient safety. This model posits that any safeguard against patient harm will inevitably have holes in it, due to system design limitations and human factors. These holes are constantly opening and closing, creating opportunities for an error to bypass the safeguard. It is therefore necessary to have multiple adaptive safeguards and continuous learning to prevent an error reaching the patient.<sup>2</sup>

Health professionals may be less familiar with the safeguards used by Statistics New

Zealand’s Integrated Data Infrastructure (IDI).<sup>3</sup> The IDI allows ‘big data’ techniques to be applied to linked datasets that contain information about (among other things) New Zealanders’ income, certain health interventions, educational outcomes, housing, use of Government benefits and social services, and interactions with the criminal justice system. Multiple safeguards—known as ‘the five safes’—protect the confidentiality of this information.<sup>4</sup> Each of the ‘safes’ are detailed below, along with their applicability to protecting confidentiality in health record research.

## Safe people

Researchers are referee checked and must sign a declaration of secrecy under the Statistics Act 1975 before accessing data in the IDI.<sup>4</sup> This can be replicated in the health sector through confidentiality agreements, employment contracts and professional codes of practice. However, as cases of health professionals inappropriately accessing patient records (‘employee browsing’) have shown, such measures may not fully protect against human nature.<sup>5</sup>

## Safe projects

Statistics New Zealand requires the Government Statistician (or a delegated person) to sign off all research projects as being in the public interest, and not compromising individual privacy.<sup>4</sup> However, there is currently no single standardised pathway for signing off data research projects in the health sector. The National Ethics Advisory Committee’s Ethical Guidelines for Observational Studies provide guidance for investigators and institutional review boards, but allow considerable scope for determining whether projects need to be reviewed by the Health and Disability Ethics Committee (HDEC).<sup>6</sup> It has also been argued

that the current HDEC review process is tailored for interventional research, and may not be ideal for considering the risks of data research.<sup>7</sup> Institutional review boards such as those at universities or DHBs may face similar or even greater limitations. We strongly endorse calls to strengthen the review process with expertise in data science, a public registry of projects and a detailed pathway for approval.<sup>7</sup>

### Safe settings

Statistics New Zealand uses a secure Data Lab environment to release data to researchers, with designated non-networked computers from which Statistics New Zealand staff control the release.<sup>4</sup> It may not be financially or logistically feasible to replicate this in the health sector. However, a corresponding safeguard may include a set of universally agreed encryption and cybersecurity protocols that researchers must be validated against before accessing health records. (Encryption keys should be kept secure and regularly updated to conform to current encryption best practices). If data is to be stored 'in the cloud' at any point in this process, there should be transparency about where the cloud servers are located, who else may be accessing the data and (in cases where the servers or cloud providers are based overseas) which nations may be able to assert jurisdiction.

### Safe data

Statistics New Zealand de-identifies data before releasing it to researchers. However, de-identification of health records may be more challenging due to the wealth of potentially identifying information in health records, especially those of a longitudinal nature. The United States Health Insurance Portability and Accountability Act (HIPPA) Privacy Rule sets what is arguably the gold standard in health record de-identification,

and lists 18 different categories of personal information that must be removed before a record is considered de-identified.<sup>8</sup> The HIPPA requirements could be considered equivalent to the standard in the Health Information Privacy Code that information does not identify an individual.<sup>9</sup> This makes the de-identification of records in New Zealand equally challenging. Manual de-identification of health records requires specially trained staff who also have medical knowledge, is laborious and is potentially very expensive. Automated de-identification is still an unsolved problem and has not yet achieved the 95% threshold needed (across all 18 HIPPA categories) for a record to be considered de-identified.<sup>10,11</sup>

### Safe outputs

Statistics New Zealand sets standards for research outputs to ensure individuals cannot be identified. This could be replicated by universally agreed standards for outputs that researchers must agree to before gaining access to the data. It is important here to consider not just the final outputs (such as reports and publications) but also intermediate outputs (such as data collated into a spreadsheet for analysis). The accidental release of such a document could result in a major privacy breach.<sup>12</sup>

In summary, there are a number of potential hurdles to be overcome to fully unlock the potential of health record research in New Zealand. Instead of either putting such research in the 'too-hard' basket, or alternatively allowing a set of ad-hoc and potentially variable standards to develop, we urge cautious progress through a set of universally agreed safeguards that benefit patients, providers and researchers. These also concord with proposed ethical frameworks for data research: public interest, trust and transparency.<sup>7</sup>

**Competing interests:**

Nil.

**Author information:**

Rajan Ragupathy, Clinical Trials & Research Pharmacist, Pharmacy, Waikato District Health Board, Hamilton; Vithya Yogarajan, Graduate Assistant & Research Student, Department of Computer Science, University of Waikato, Hamilton.

**Corresponding author:**

Rajan Ragupathy, Pharmacy, Waikato District Health Board, Private Bag 3200, Waikato Mail Centre, Hamilton 3204.

rajan.ragupathy@waikatodhb.health.nz

**URL:**

<http://www.nzma.org.nz/journal/read-the-journal/all-issues/2010-2019/2018/vol-131-no-1478-13-july-2018/7624>

**REFERENCES:**

- Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA*. 2013 Apr 3; 309(13):1351–2.
- Reason J. Human error: models and management. *BMJ*. 2000 Mar 18; 320(7237):768–70.
- Statistics New Zealand. Integrated Data Infrastructure. [Online]. [Last accessed 07 May 2018]. Available at: [http://archive.stats.govt.nz/browse\\_for\\_stats/snapshots-of-nz/integrated-data-infrastructure.aspx](http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx)
- Statistics New Zealand. How we keep IDI and LBD data safe. [Online]. [Last accessed 07 May 2018]. Available at: [http://archive.stats.govt.nz/browse\\_for\\_stats/snapshots-of-nz/integrated-data-infrastructure/keep-data-safe.aspx](http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/keep-data-safe.aspx)
- Office of the Privacy Commissioner. The temptation of inappropriate employee browsing. [Online]. [Last accessed 07 May 2018]. Available at: [http://www.privacy.org.nz/assets/Uploads/](http://www.privacy.org.nz/assets/Uploads/CCDHB-Grand-Round-7-April-2016.pdf)
- National Ethics Advisory Committee. 2012. Ethical Guidelines for Observational Studies: Observational research, audits and related activities. Revised edition. Wellington: Ministry of Health.
- Ballantyne A, Style R. Health data research in New Zealand: updating the ethical governance framework. *N Z Med J*. 2017 Oct; 130(1464):64–71.
- United States Department of Health & Human Services. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. [Online]. [Last accessed 07 May 2018]. Available at: <http://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html>
- Office of the Privacy Commissioner. Comparison paper on health privacy laws. [Online]. [Last accessed 07 May 2018]. Available at: <http://www.privacy.org.nz/news-and-publications/books-and-articles/comparison-paper-on-health-privacy-laws-2/>
- Stubbs A, Filannino M, Uzuner Ö. De-identification of psychiatric intake records: overview of 2016 CEGS N-GRID shared tasks track 1. *J Biomed Inform*. 2017 Nov 1; 75:S4–18.
- Stubbs A, Kotfila C, Uzuner Ö. Automated systems for the de-identification of longitudinal clinical narratives: Overview of 2014 i2b2/UTHealth shared task Track 1. *J Biomed Inform*. 2015 Dec 1; 58:S11–9.
- KMPG & Information Integrity Solutions Pty Ltd. Independent review of ACC's privacy and security of information. [Online]. [Last accessed 07 May 2018]. Available at: <http://www.privacy.org.nz/assets/Files/Media-Releases/22-August-2012-ACC-Independent-Review-FINAL-REPORT.pdf>