Health data research in New Zealand: updating the ethical governance framework

Angela Ballantyne, Rochelle Style

ABSTRACT
Demand for health data for secondary research is increasing, both in New Zealand and worldwide. The New Zealand government has established a large research database, the Integrated Data Infrastructure (IDI), which facilitates research, and an independent ministerial advisory group, the Data Futures Partnership (DFP), to engage with citizens, the private sector and non-government organisations (NGOs) to facilitate trusted data use and strengthen the data ecosystem in New Zealand.

We commend these steps but argue that key strategies for effective health-data governance remain absent in New Zealand. In particular, we argue in favour of the establishment of: (1) a specialist Health and Disability Ethics Committee (HDEC) to review applications for secondary-use data research; (2) a public registry of approved secondary-use research projects (similar to a clinical trials registry); and (3) detailed guidelines for the review and approval of secondary-use data research. We present an ethical framework based on the values of public interest, trust and transparency to justify these innovations.

Demand for data

Both the amount of data and demand for access to use this data for research are growing. These trends are apparent for health data and for other data collected by government, including welfare and social services data, justice, migration, education and taxation data. Better use of data underpins the New Zealand Government's 'social investment approach', which aims to increase the use of public sector data to drive innovation, save costs and better target services to people in need. It seeks to do this by applying rigorous, evidence-based investment practices backed up by big data.¹

In the health context, secondary-use research involves the use of clinical data for purposes other than that for which the data was collected and which is therefore generally outside the original patient consent. This involves re-using, sharing and linking data to better understand New Zealand's health needs. This often involves access to initially identifiable health data for quality control purposes (ensuring the data is correct and complete). Researchers seeking to ensure accuracy and consistency of data from various datasets will require access to patient identifiers such as a National Health Index (NHI) number or personal patient information such as name and birthdate. There is increasing interest in establishing patient registries across the health sector and linking between registries, and from registries to other Ministry of Health (MOH) data. For example, research carried out by the New Zealand Treasury investigated the impact of eight different health conditions on the employment rates and incomes of working-aged New Zealanders.² The study used the IDI, linking data from Inland Revenue, the MOH (including data on public hospital admissions and discharges, outpatient and emergency department treatments, pharmaceuticals dispensed, laboratory tests and enrolments at primary health organisations) and the Ministry of Social Development (MSD).

Broader context

There is increasing interest in the establishment of an expert national data ethics committee to oversee all significant data use in New Zealand, not just health data.
The collection and subsequent use of data must be undertaken with care. This has been recently highlighted by MSD’s ‘Data-for-funding’ policy, which required non-governmental organisations (NGOs) providing social services to disclose to MSD individual client-level data. The Privacy Commissioner’s report into this policy found the proposed data collection was excessive and inconsistent with privacy principles; and there was a lack of transparency about the purposes for which the data was being collected and for which it might subsequently be used.¹ Roll-out of this data sharing program was delayed in April 2017 due to a security flaw in the information-sharing portal. The MSD situation provides a sober lesson regarding the need for data collection and use to consider both the interests of data subjects and stakeholder communities, as well as the needs of researchers and government.

The Integrated Data Infrastructure (IDI) includes de-identified microdata (information at the level of individuals) from government agencies, NGOs and surveys such as the Census. Researchers can access this data for economic, social or policy-related research. As of April 2017, there were 112 research projects listed by Statistics New Zealand which are using, or have used, data in the IDI, including projects for benefits and social services (12), business and employment (24), education and training (15), families and households (17), health and safety (24), housing (7), justice (4) and travel and migration (9).² Use of the IDI data requires approval from the Government Statistician according to the 5 Safes framework: safe people; safe projects; safe settings; safe data; safe output. While the ‘safe projects’ framework requires the Government statistician to sign off all research proposals and to ensure they are in the public interest, there is no independent ethics committee review of projects, no community input into determining what uses of data would be in the ‘public interest’ and no requirement for review of projects by Māori/iwi. We suggest that our proposals for the governance of health data could provide a useful pilot for such an expert data ethics committee.

In this paper we focus on the secondary-use of data in the health context and suggest ways in which the public interest may be served and trust preserved through expert oversight, public engagement and transparency. We propose an expert health-data research ethics committee, a public register of health-data research and expanded ethical guidelines for secondary-use.

**Current New Zealand regulatory framework**

The National Ethics Advisory Committee’s Ethical Guidelines for Observational Studies (NEAC Guidelines) permit some uses of health data without independent ethical review.³ Linking data for observational epidemiological studies is allowed without consent, provided that identity is only disclosed for linking purposes (NEAC guideline 8.11); and linking for audit and related activities needs no justification to an ethics committee provided that it is part of high-quality health care (NEAC guideline 8.12).

Access to potentially identifiable patient information for research generally requires approval from an approved research ethics committee. There is limited instruction available in the NEAC Guidelines for when ethics committees can grant access to identifiable health information:

NEAC guideline 6.43 provides that:

Access to identified or potentially identifiable data for research (without consent) may be justifiable when:

a) obtaining consent would cause either:

• unnecessary anxiety
• prejudice the scientific value of the study; or
• it is impossible in practice due to the quantity or age of the records; and

b) there would be no disadvantage to the participants or their relatives or to any collectivities involved; and

c) the public interest in the study outweighs the public interest in privacy.⁴

The current regulatory approach is insufficient in a number of ways. Health and Disability Ethics Committees (HDECs) and
other research ethics committees primarily review interventional or observational clinical research and some have questioned whether such committees have sufficient relevant expertise to review sophisticated data linkage projects. The current regulatory framework also lacks specificity. For example, there is no definition of ‘public interest’ or guidance on when consultation with Māori is required. And finally, the review system is fragmented, making it near impossible for the public, media or the research community to track who is accessing patient data and for what purposes.

Current regulations to guide the ethical use of health data continue to focus predominantly on consent. Individual-level interests such as control of health information, confidentiality, privacy and personal disadvantage are well-articulated. Collective values such as public interest, solidarity, trust, equity and participation have been comparatively underdeveloped in the data governance literature and especially in research regulations and guidelines. This is true both in New Zealand and internationally. Consider the focus of NEAC’s primary guidelines on such matters, Guideline 6.43—it is first and foremost about consent. Public interest and potential disadvantage to communities are appealed to here but there is no definition of these concepts or guidance regarding how they should be interpreted.

A ‘consent’ approach is insufficient in the era of population-level research and big-data projects. Many current and future uses of health data—including population-level data analytics, predictive risk modelling and diverse data-sharing arrangements—make patient consent impractical, if not impossible. A framework that focuses primarily on consent is therefore inadequate to provide substantive guidance on which uses of health data are acceptable.

To resolve these challenges we propose the use of a framework based on three core values which should guide the secondary-use of patient data. We present three suggestions for developing the governance framework for health data research in New Zealand that is guided by these three values.

Core values

In standard clinical research, patient consent plays a major role in legitimising the research. As we have noted, there are a number of circumstances in which patient consent is not currently required for the secondary-use of health data. However, other values can legitimise research in the absence of consent: public interest, trust and transparency.

1. Public interest

The core value that should guide secondary use of health data is public interest. Ethics committees are acting as stewards for a public resource and need to ensure that those accessing the resource are likely to produce knowledge that will benefit the public. The current ethical guidelines in New Zealand require that research be in the public interest, and recognise that public interest includes both research and privacy interests, but do not define the concept further. In relation to data use, public interest (also sometimes referred to as public or common good) relates to collective interests such as national research priorities, equity and public access to the research results. We would argue for example that a public interest requirement is met when research addresses neglected diseases, health conditions with high social cost and/or sources of health inequality between different groups. Public interest may be undermined by research that involves the risk of surveillance of specific populations, individual or group discrimination, stigma and predictive privacy risk (where privacy invasions occur through inference rather than direct collection of personal data). The relative presence of these risks in a research project can decrease the public interest.

In a pluralistic society with diverse conceptions of wellbeing, a community engagement process can help answer the question of which data uses are in the public interest. Community engagement is especially important in the absence of individual consent.

Indigenous communities are increasingly calling for sovereignty over their data, including genetic material. In New Zealand, Te Mana Raraunga (the Māori Data Sovereignty Network) is advocating for Māori data...
sovereignty, arguing that data innovations are occurring in the absence of a robust Māori data governance partnership that is representative, enabling and provides clear lines of accountability back to Māori/Iwi. Te Mana Raraunga have argued that Māori data (including data about Māori, data used to describe Māori collectives and data about Te Ao Māori) are a living tāonga and should be subject to Māori governance. How to operationalise these sovereignty claims and how to ensure that data use serves the interests of Māori are complex questions that require more attention across the academic, government and private sectors.

2. Trust

There is widespread agreement on the importance of trust in medical research. Studies have shown that trust is fundamental to the successful conduct of research and is conditional. The Data Futures Partnership (DFP) has emphasised that maintaining trust is vital to ongoing data innovation and has warned that public trust in the data-use ecosystem is tenuous and, once lost, may be hard to restore. Trust is multi-dimensional, complex and contextual. Nevertheless, common qualities of trustworthy institutions include:

- integrity—the institution is fair and just;
- dependability—the institution will do what it says it will do; and
- competence—the institution has the ability to do what it says it will do.

It is essential to ensure that the public has adequate trust in the governance of secondary-use research, especially given the absence of individual consent. Both researchers and research ethics committees must demonstrate trustworthiness. Much of the trust that study participants place in researchers is associated with the reputation of researchers as a professional group, with associated standards of conduct and systems of independent oversight and review.

We believe that HDECs in New Zealand have built a reputation for integrity and dependability. We question however whether committees primarily designed to review interventional clinical research have the necessary competence to review complex data research. Ethical issues relating to the secondary-use of data for health research raise unique concerns which differ substantially from those raised by interventional research. Data research moves inquiry away from familiar categories of research harm, such as physical pain or psychological distress, to other categories such as surveillance, discrimination and stigma. Rapid advances in data research involve both a change in scale of the analytic tools—speed, capacity, continuous generation—as well as a change in the relationality, flexibility, re-purposing and de-contextualisation of data.

Internationally, calls are being made for “algorithmic literacy, transparency and oversight” because of the challenges of algorithmic biases. Algorithms reflect the biases of programmers and dataset, for example regarding race, gender and other variables related to social justice. Algorithms are biased towards what their writers understand to be ‘normal’. For example, a search for images of ‘professor’ results in pictures of white males, and to find images of women or people of colour, the search algorithm requires the user to include ‘woman professor’ or ‘Asian professor’; thereby reinforcing the assumption that a real professor is white and male. Overseeing data research requires specific expertise in information technology, computer science and topics such as data security, algorithms and data privacy. Appropriate oversight also requires specialist understanding of the risks data-harnessing may pose, including security failures and the wide-ranging ramifications for privacy, even absent security breaches. Ethical research with data therefore requires review by a committee which includes data specialists as well as lay members.

3. Transparency

Key features of transparency include visibility, accessibility and honesty. New Zealand has a strong track record of transparency in the context of HDEC ethical review—all meetings are open to the public and minutes are published on the website. This degree of transparency is not reflected in other countries.
Transparency helps facilitate community engagement because it allows other researchers and members of the public to see what is being done with patient data and by whom. In this way, transparency can help ensure accountability, secure trust and maintain the social licence for data research in the absence of patient consent.

Our three core values—public interest, trust and transparency—are consistent with the New Zealand government’s commitment to transparency, including the Declaration of Open and Transparent Government, which proclaimed that high value public data must, inter alia, be open, trusted and readily available. Our core values are also consistent with the four principles developed by the DFP of value, inclusion, trust and control.12

**Data governance needs**

In light of these core values, we propose three recommendations for revising health-data governance in New Zealand.

1. **Guidance**

   We propose developing expanded guidance on the ethical uses of patient data in the absence of consent. This could consist of new independent guidelines or additional chapters within existing NEAC guidelines. The guidance should focus on defining the criteria for public interest and articulating the potential benefits and harms regarding secondary-use data research. Expanded guidance should consider use of patient data by academic researchers, NGOs, commercial or for-profit entities, public-private partnerships and government agencies. Research justified on the grounds of public interest should also be made publicly accessible and any power of commissioning agencies to limit the publication of results should be considered.

2. **Independent expert review**

   We propose establishing a specialist HDEC that reviews applications for secondary-uses of health data. The committee membership should include expertise in computer science, information technology, data ethics, privacy, as well as patient and community advocates; and should follow current HDEC policy and the HRC Guidelines regarding lay, gender and Māori representation.18,19

   An alternative model would be to increase the data expertise on the existing four HDECs. However, given there are only eight members per HDEC, it would be hard to achieve the degree of expertise we advocate across all four committees, without either displacing existing categories of expertise or increasing the size of the committees. Our preference is therefore for a specialist HDEC committee but we acknowledge there are different ways of achieving the desired expertise during the review process.

   The Privacy Commissioner has spoken about the possibility of having an independent body to promote the ethical and safe use of data,20 as has the DFP which noted there is no independent trusted forum for an inclusive conversation on data use.21 Specialist ethics review committees are already part of the New Zealand ethical landscape—there is an Ethics Committee on Assisted Reproductive Technology and a Gene Technology Advisory Committee.

   There are four national HDECs and, as of 1 May 2017, none have a member described as having expertise in computer science or statistics; although some of the clinical research members will likely have expertise in data analytics due to their personal research programs.22 A specialist data ethics committee could develop expertise in judging issues such as public interest and potential population-level harms from the misuse of data, and may also provide greater consistency and predictability in the review process.

   Building on the HDEC model makes sense as this: (1) would prevent the siloing of a health data ethics committee outside the broader health research ethics review structure because Chairs of the HDECs meet regularly and discuss procedural and policy matters; and (2) leverages existing secretarial support services at the Ministry of Health. In addition there would be the advantage of a shared line of public accountability, statutory empowerment and a common regulatory framework.

3. **Registry**

   Third, we propose the development of a New Zealand data research registry that is public, searchable and based on the structure and content of clinical trials registries such as the Australian New Zealand Clinical Trials
Register (ANZCTR). Registration should be a condition of receiving access to health data without consent. We have argued above that one strength of the New Zealand research ethics ecosystem is its relative transparency, but more needs to be done.

While HDEC minutes are published online, the applications are not categorised other than to indicate the title of the study, the principal investigator, the sponsor and the ‘clock-start-date’. To identify secondary-use data research studies, interested parties must read the full minutes of each meeting. The lists of applications in HDEC Annual Reports include a description of whether the study is observational or interventional, but there is no further categorisation and, at the time of writing, the Annual Reports are only available until 2013. The absence of any detailed and consistent coding of health research applications makes it difficult to determine how many data projects are occurring using patients’ health data, which agencies or researchers are accessing the data, what sorts of research questions are being pursued and where and when research results are being published. A core justification in the NEAC guidelines for approving access to patient information without consent is ‘public interest’. We argue that public interest requires public transparency. A registry similar to the ANZCTR with settled and mandatory criteria with respect to data content, quality and validity, the assignment of unique identification numbers and advance searching capabilities would dramatically improve the transparency of secondary-use data research and its accessibility. A public registry is especially important in relation to tracking data research (as opposed to other observational research) because much data research does not have specific patient consent. A registry therefore provides a mechanism for the public to see who has access to their data and for what purposes.

Conclusion

The governance framework for secondary-use health data research in New Zealand is piecemeal and underdeveloped. These deficiencies cannot be remedied by the established clinical research framework, which is not fully suited to dealing with the complexities raised by secondary-use data research. Our proposals to establish a specialist HDEC and a data research registry, combined with specific and expanded guidelines, would provide a robust governance framework for the secondary-use of health data, reflecting the core values of public interest, trust and transparency.

Summary of key arguments

• The governance framework for secondary-use health data research in New Zealand is piecemeal and underdeveloped.
• Data use should be governed by the values of public interest, trust and transparency.
• An ethical framework focused on consent and individual control provides insufficient guidance for population and big data studies.
• We need expanded guidelines on the ethical uses of health information without consent that focus on collective benefits and harms.
• A specialist health data ethics committee would provide expert oversight and improve consistency of review.
• A registry of secondary-use health data research would increase accountability and transparency regarding the use of patient health data.
Competing interests:
Angela Ballantyne and Rochelle Style are current members of the Health and Disability Ethics Committees in New Zealand.

Author information:
Angela Ballantyne, Department of Primary Health Care and General Practice, University of Otago, Wellington; Rochelle Style, Independent Researcher, Wellington.

Corresponding author:
Dr Angela Ballantyne, Department of Primary Health Care and General Practice, University of Otago, Mein St, Wellington 6242.
angela.ballantyne@otago.ac.nz

URL:

REFERENCES:

17. Personal communication to Angela Ballantyne from research ethics experts in Australia, United States, Sweden, Amsterdam and the United Kingdom.


22. Details of the members of each of the four HDECs are available from: http://ethics.health.govt.nz/about-committees