

‘As long as it’s used for beneficial things’: An investigation of non-Māori, Māori and young people’s perceptions regarding the research use of the Aotearoa/New Zealand Integrated Data Infrastructure (IDI)

ABSTRACT

The Aotearoa/New Zealand Integrated Data Infrastructure (IDI) is a national database containing a wide range of data about people and households. There is limited information about public views regarding its use for research.

A qualitative study was undertaken to examine the views of forty individuals, including those of Māori ethnicity and young people. Data were analysed using Braun and Clarke’s method of thematic analysis.

Six themes emerged: 1) Conditional support for the use of the IDI; 2) Concerns regarding the misuse of data; 3) The importance of privacy; 4) Different views regarding consent for use of data for research; 5) Desire for access to personal data and research results; and 6) Concerns regarding 3rd party and commercial use.

Although there is reasonable support for its use for research, there is more work to be done to ensure ethical use of IDI data via improved consent, privacy and researcher training.

INTRODUCTION

Advances in digital access and technology over the past few decades have led to the generation and storage of vast amounts of data from electronic public service records, websites, applications, blogs, organisational memberships and even twitter hashtags at an institutional, national and international level. Large datasets, sometimes referred to as ‘Big Data’, may contain diverse information from heterogeneous sources, produced at high frequency. They can be examined to reveal patterns, trends, and correlations that have relevance for personal or population wellbeing (Lenca et al., 2018). The use of such data may be associated with perceived and actual benefits and risks (Custers & Uršič, 2016).

Safeguards against misuse, such as the General Data Protection Regulation (GDPR) in the European Union (Mourby et al., 2018) and national privacy laws, are trailing behind technological possibilities (Rubinstein, 2012; Valentine, Scampion, & Taylor, 2020).

The Aotearoa/New Zealand Integrated Data Infrastructure (IDI) is a large national database first established in 2011 that contains a wide range of administrative and survey data about people and households, linked at an individual level (Milne et al., 2019; Statistics New Zealand, 2017a). Managed by Statistics New Zealand, the IDI includes (i) administrative data from government departments, such as health and education, as well as non-government sectors and (ii) survey data, including the New Zealand census (Figure 1) (Statistics New Zealand, 2017a). Records are probabilistically linked, usually by name, date of birth, and sex, and held in a secure environment (Atkinson & Blakely, 2017; Milne et al., 2019).

The use of IDI data is governed by Statistics New Zealand using the ‘Five Safes’ and Ngā Tikanga Paihere frameworks (Statistics New Zealand, 2017). Data can be accessed by approved researchers for projects in the public interest, such as policy and interventions research and analysis of life outcomes of population cohorts over time. Data and results must

be aggregated and anonymised according to Statistics New Zealand protocols, and all results are checked for confidentiality by Statistics New Zealand prior to their release from the secure environment (Statistics New Zealand, 2020). Legal requirements to protect IDI data include New Zealand's *Statistics Act 1975*, *Privacy Act 2020* and *Tax Administration Act 1994* [9] and there are further Statistics New Zealand policies and protocols that govern the use of the IDI (Statistics New Zealand, 2020). Regular privacy impact assessments undertaken by Statistics New Zealand also provide a systematic evaluation of the benefits and risks associated with integrating data from multiple sources (Statistics New Zealand, 2017b).

While linked administrative and survey data have been used for research purposes in Aotearoa/New Zealand for many years, the IDI makes these data more readily available and at lower cost. This opens the door to a wider range of researchers, who have greater potential to access large amounts of data covering a growing number of government sectors and aspects of society. As a result, research using the IDI is increasing (Bowden et al., 2019; Bowden et al., 2020; Charania, Paynter, Lee, Watson, & Turner, 2018; McLeod et al., 2015; Shackleton, Broadbent, et al., 2018; Shackleton, Milne, et al., 2018) and is expected to grow exponentially. On the one hand, while it might be argued that the safeguards to protect these data are a significant improvement on pre-IDI data use, equally it can be argued that the IDI lends itself to a higher likelihood of misuse of data, both intentional and unintentional, particularly in the absence of sufficient guidance for researchers regarding ethical and culturally appropriate data use. This is especially problematic for Māori peoples and individuals of Pacific Island ethnicity, as research findings derived from IDI data have frequently been perceived as deficit-framed and polarising for the communities they describe (Kokaua et al., 2020). To date, only one government-commissioned investigation and one independent study have examined the views of people from Aotearoa/New Zealand on the

inclusion of census data in the IDI (A. Davidson et al., 2015; Gulliver, Jonas, McIntosh, Fanslow, & Waayer, 2018) and none have explicitly examined people's views regarding the use of IDI data for research. Additionally, given the potential for cultural and age-related variations in perspective (Milberg, Burke, Smith, & Kallman, 1995) and the country's commitment to Te Tiriti o Waitangi (the contract between the British Crown and Māori peoples) (Palmer, 2008), there has not been any attempt to specifically understand the views of people of Māori ethnicity or young people when considering whether existing IDI research-related legal protections, policies and processes are adequate.

This study was conceptualized to try and address some of these gaps. A qualitative study was undertaken to examine the views of people from Aotearoa/New Zealand, including those of non-Māori and Māori ethnicity, and young people, on the secondary use of health-related and other administrative data within the IDI for research.

METHODS

The study was conceptualized by all seven authors and undertaken at a large city hospital in Auckland, Aotearoa/New Zealand between April and October 2019. Forty individuals were recruited from Auckland District Health Board outpatient clinics via opportunistic means (they were informed about the study by clinic staff during these visits). Eligible participants were aged 16 years or over (and thereby able to provide individual consent, without any need for paired parental consent, as per ethics committee regulations), had any health condition, and sufficient English to understand the material being discussed. Purposive sampling was employed to ensure that 10 participants were young people aged 16 to 18 years and that a further 10 were Māori. All experimental protocols were approved by the National Health and

Disability Ethics Committee. All methods were conducted as per ethics committee requirements (NZ HDEC approval number: 18/NTB/241/AM01). Participants were provided with oral and written information about the study and time to consider whether they wished to participate, following which both oral and written consent were obtained. Interviews lasting for 30-60 minutes were conducted in person at the hospital or via telephone by two female, non-clinical, paid research assistants of non-Māori ethnicity, both of whom had previous qualitative research experience and had no prior relationship with study participants (HK and PR). These interviewers followed guidelines established in conjunction with the Māori clinician on our team (TS) who also oversaw the analysis and interpretation of the Māori data. Additional cultural support was verbally offered to, but not used by any, Māori participants. Participants received a NZ\$20 gift voucher for their time.

Semi-structured interviews were conducted on a single occasion for each participant using a pre-specified and pilot-tested template (Appendix 1). Participants were provided with verbal and written explanations about the IDI (Appendix 2). Data were audio recorded, transcribed and manually examined by three authors (HT, HK and TS) for key themes and supporting statements using Braun and Clarke's method of thematic analysis (Braun, Clarke, & Terry, 2014). Transcripts were not returned to participants for review. Coding discrepancies were addressed by consensus. Following analysis of themes from the entire sample, comparative analysis of themes from non-Māori participants, Māori participants and young people was undertaken to discern any differences between these subgroups. Demographic information was tabulated. No other quantitative analysis or statistical procedures were undertaken. The paper was drafted and reviewed by all authors.

RESULTS

Participant demographics

Participant demographics are outlined in Table 1. Of the 54 individuals approached to take part, 40 agreed to do so. Reasons given by others were: ‘not interested’ (N=10; 71%), ‘too busy to participate’ (N=1, 0.7%), ‘want to focus on physical healthcare’ (N=1; 0.7%) and ‘none’ (N=2; 1.4%). Planned numbers of individuals of Māori ethnicity and young people were interviewed. Over 60% of participants were female.

Key findings

Six themes were identified from the information provided by participants, namely: 1) Conditional support for the use of the IDI, including for research; 2) Concerns regarding the misuse of IDI data; 3) The importance of privacy; 4) Different views regarding consent for use of data for research; 5) Desire for access to personal data and the results of research; and 6) Concerns regarding 3rd party, commercial use. Themes and supporting examples are summarised in Table 2. No clear differences in themes were identified in this limited sample between Māori, young people and other participants.

DISCUSSION

This study aimed to understand the views of the people of Aotearoa/New Zealand, including those of Māori ethnicity and young people, regarding the secondary use of IDI data for research purposes. Our findings demonstrate local acceptance of the legitimate use of the IDI for research that is in the public interest. Adequate data de-identification, regulation by an approved authority and provision of consent prior to collection and/or use are preferred, while third party and commercial data use are less supported. Despite international literature suggesting that preferences may differ by age and ethnicity (Buckley, Murphy, & MacFarlane, 2011; Caulfield, Rachul, & Nelson, 2012; Eloranta & Auvinen, 2015; Grande, Mitra, Shah, Wan, & Asch, 2013; Jagsi et al., 2017; Kim, Sankar, Wilson, & Haynes, 2017; King, Brankovic, & Gillard, 2012; Lofters, Shankardass, Kirst, & Quiñonez, 2011; Page, Manhas, & Muruve, 2016; Riordan et al., 2015; Trinidad et al., 2010), we did not identify any clear differences between the views of young people, Māori and other Aotearoa/New Zealand adults. Potential reasons for this are explored below.

Previous studies have indicated that members of the public and healthcare staff have limited knowledge about the types of data held in national databases (Stone, Redsell, Ling, & Hay, 2005). The findings of our study suggest there is room for improvement in public education about databases such as the IDI. We agree with Aitken et al. that it would be wise for data regulators such as Statistics New Zealand to invest in clear national messages about the potential benefits and downsides of data sharing for research and to create opportunities for public deliberation, akin to the UK ‘Understanding Patient Data’ website (<https://understandingpatientdata.org.uk/>) (Aitken, Cunningham-Burley, & Pagliari, 2016), rather than merely controlled dialogue. This would lead to a more ‘balanced healthcare data access system’ as proposed by Avram et al. (Avram et al., 2018) in which individuals have

fuller understanding about the use of their data and public participation in the development of programmes for health, education and other societal issues as proposed by Dorey et al. (Dorey, Baumann, & Biller-Andorno, 2018). Such conscious “tightrope-walking”, as recommended by Robertson et al. (Robertson et al., 2016) includes finding the balance between carrying out research using datasets that would benefit the public interest, while safeguarding the privacy of individuals and data security, and maintaining adequate data governance, transparency and public support. Additional measures may be needed to ensure that people of different ages, races, ethnicities and abilities are provided with appropriately tailored and culturally relevant information regarding the secondary use of their data to genuinely increase healthcare equity (Fry, Spriggs, Arnold, & Pearce, 2014; Jagsi et al., 2017; Riordan et al., 2015; Zarate et al., 2016). In the case of Māori peoples, this may include consideration of the impact of individual decisions on whanau, hapu and iwi (extended family and communities).

Altruism seems to be the primary motivation for participants in our study being willing to share data for research. This is consistent with the findings of UK and US researchers (Bond, Ahmed, Hind, Thomas, & Hewitt-Taylor, 2013; Damschroder et al., 2007). It is also consistent with the previous findings of a Statistics New Zealand-commissioned report into public perceptions of general use the IDI (A. Davidson et al., 2015). Trust, privacy and consent appear to be interwoven issues in our study as in others (S. Davidson et al., 2013; Grande et al., 2013; King et al., 2012). Our study participants expressed varying levels of trust in the IDI, especially regarding the storage and use of information considered to be of a more personal nature. One multi-national study found that trust was inversely related to the degree of control individuals considered necessary (Aitken, Jorre, Pagliari, Jepson, & Cunningham-Burley, 2016) and another US study found that up to 44% of participants

wanted to be informed each time a new researcher uses their data (Bell, Ohno-Machado, & Grando, 2014). The fact that most of our participants recommended consent before data is used for research, but not separate consent for each study, may reflect their pragmatism or the general level of trust by people of Aotearoa/New Zealand in government (Cullen & Reilly, 2008). Although we did not directly examine the views of secondary users of data, we note that previous studies have also highlighted trust and consent as issues for researchers (Aitken, Cunningham-Burley, et al., 2016). Few have supported open access to data without the use of gatekeepers or governance committees. Some have acknowledged the tension between the need for consent and difficulties in obtaining it, while others have believed that study approval by ethics committees is equivalent to public consent (Asai et al., 2002). As long as there is appropriate oversight of secondary IDI data use, it may be reasonable to weigh up the administrative burden of seeking individual consent in the context of very large research studies, leading to consideration of a consent waiver or ‘opt-out’ approach (Council, 2007).

Although participants in our study did not initially appear to be concerned about who was permitted to use the IDI for research, on specific questioning, they were much less accepting of private or commercial use of their data. This echoes the findings of the OPUS report, a Canadian study in which profit-oriented industries and insurance companies were much less trusted than doctors and hospitals (Caulfield et al., 2012), two Scottish studies in which private sector organisations were regarded as predominantly profit-focused (S. Davidson et al., 2013) and pharmaceutical companies were believed to only be interested in commercial gain (Grant et al., 2013), and a US study in which commercial enterprise was less favoured than academic research (Grande, Mitra, Shah, Wan, & Asch, 2014). Concern about data misuse, expressed by some of our research participants, should be squarely addressed, and steps taken – such as those we recommend below – to enhance the trustworthiness of IDI

governance and uses of the data for research. There are numerous examples from other contexts of harmful data uses, such as the use of medical records to identify immigration offenders in the United Kingdom (Travis, 2017), that offer a cautionary tale for the Aotearoa/New Zealand context.

Previous local and international studies on the use of health data for research have consistently ignored or under-represented the views of young people (Clerkin, Buckley, Murphy, & MacFarlane, 2013; Haddow, Bruce, Sathanandam, & Wyatt, 2011; Hill, Turner, Martin, & Donovan, 2013; Parkin & Paul, 2011). One American study found that young people were generally positive about data sharing, less concerned about privacy, more aware of data breaches, preferred to be consented for use of their data and were not keen on private companies (Hill et al., 2013). An Irish study found that young people were more concerned with data leaks (Clerkin et al., 2013) and a multi-national study described a lack of clear relationship between age and acceptance of commercial access to health data (Aitken, Jorre, et al., 2016). Young people in our study described similar concerns to those overseas, demonstrating both the value of research-related benevolence and astuteness about the limitations of data anonymisation. Differences between genders have also only been peripherally explored. The few studies that exist have demonstrated either no difference between genders (Buckley et al., 2011) or males being more concerned about the implication of re-identification on employment and finances, while females were more concerned about social discomfort and embarrassment if information was disclosed to family, friends or acquaintances (Clerkin et al., 2013). Our study was not focused on evaluating such differences and as such, none were found.

Akin to the under-representation of young people in previous international studies of big datasets, the views of people of colour, different ethnic groups and socio-demographic

subsets have also not been well examined. When they have, results have been mixed (Jagsi et al., 2017; Spencer et al., 2016). One US study identified consistent views across three racial/ethnic groups (White, Hispanic and African American) (Grande et al., 2014), another found that those of Asian origin were more likely to share data for the purposes of health information exchange than those of other migrant cultures (Kim, Joseph, & Ohno-Machado, 2015) and a third found that African-Americans were less likely to trust that their health information in databases was safe (Saito, Shofer, Saberi, & Green-McKenzie, 2017). An Indian study described narratives of powerlessness in the face of potential exploitation of participants by researchers (Hate et al., 2015). Māori and people of Pacific Island ethnicity were under-represented in the only previous telephone-based study of public views on the secondary use of electronic health records (Whiddett, Hunter, McDonald, Norris, & Waldon, 2016). In our study, Māori participants were generally positive about sharing health data and keen on clear consent processes, but warier about misuse of data that could result in personal disadvantage or profit. Related socio-demographic factors have been found to play as much of a role as ethnicity in some studies, with three US studies identifying those who had experienced recent financial barriers to health being less supportive of secondary use of their data (Grande et al., 2013), those from ethnic/racial minority neighbourhoods being more concerned about privacy violations (Lucero et al., 2015) and those who are less educated being less likely to perceive the importance of secondary data use for research (Jagsi et al., 2017). Conversely, a recent Australian study did not find any difference between the views of people of different ages, genders or residential postcodes (Lehnbom, Brien, & McLachlan, 2014).

Strengths of this study include the inclusion of a cross-section of Aotearoa/New Zealand society, especially previously understudied groups of young people and Māori peoples. Limitations include the small sample size, restriction of participants to hospital attendees

from a single urban population and lack of collection of socio-demographic data, and lack of collection of demographic data regarding the whanau, hapu and iwi (family and communities) of Māori participants, all of which might affect the generalisability of our findings. Data generated during qualitative research are always the product of interpersonal engagement between interviewers and participants (Whiddett et al., 2016), therefore the fact that the interviewers were neither young people nor Māori are likely to have affected the nature of the information provided – in retrospect, we would have preferred to have these conducted by a Māori researcher, and this would be a clear recommendation for further research of this topic. It is possible that increasing the sample size and heterogeneity of Māori participants (including those who were at Māori schools (Kura kaupapa) and from more rural communities), asking specific kaupapa Māori questions and conducting interviews in other languages (especially the Māori language, Te Reo) may have upheld recent academic discourse about culturally-linked data sovereignty. While none of our sample mentioned this issue, we acknowledge concerns about the negative implications of culturally decontextualized data interpretation and the Māori perception of data as a taonga (treasure) (Hudson, Anderson, Dewes, Whaanga, & Roa, 2017; Kukutai & Taylor, 2016; Walter & Suina, 2019). As such, we tautoko (support) the concept that the Māori data should be subject to Māori governance, that Māori should be able to access Māori data to support their development aspirations (Hudson et al., 2016) and that ‘collective benefit, authority to control, responsibility & ethics’ (CARE) principles (Kukutai, Carroll, & Walter, 2020) should be adopted when using IDI data. Additionally, as we did not explore other socio-demographic characteristics, we are unable to comment on their possible effect on participant views and recommend that our findings are understood within this context.

CONCLUSIONS

Although there is reasonable public support for the use of national datasets such as the IDI for research, there is more work to be done to ensure that there is adequate notification or consent for the secondary use of administrative data, that privacy is sufficiently managed and that researchers are optimally supported to use the data in an ethical and culturally appropriate manner. With these measures in place, we hope that the immense potential of the IDI can be harnessed to identify important national trends, infer causal relationships and evaluate the long-term impact of interventions in a cost-effective and societally useful manner that guides future resource allocation toward building a healthier, safer and more equitable Aotearoa/New Zealand.

Best Practices

Based on our findings, we offer the following recommendations:

1. Make more user-friendly information about (i) the IDI, (ii) applicable laws including privacy legislation, and (iii) current, completed and declined applications for research readily available via a centralised public and regularly updated website such as <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/>. Ensure this information is available in a range of languages, and age-appropriate and culturally-appropriate formats.
2. Establish a more robust (joint Crown and iwi) authority that centralises data sovereignty issues for Māori people and undertakes rigorous oversight, ethical approval and review of research applications, processes and results to appease New Zealanders' concerns regarding the safe use of their data. We believe that these

responsibilities are beyond the remit of existing research ethics committees which have limited knowledge of the IDI and focus on general research procedures.

Exemplars for such an authority include the Confidentiality Advisory Group in England (Aitken, Jorre, et al., 2016), the Public Benefit Panel in Scotland (Aitken, Jorre, et al., 2016), Institutional Review Boards as established in some parts of the US (Zarate et al., 2016) and the data governance board that will be set up for the secondary use of My Health Record data in Australia (Australian Government: Department of Health, 2018). Ensure that working panels include Māori and non-Māori researchers, and Māori (from multiple iwi) and non-Māori members of the public akin to the Citizen's Juries in some parts of the US (Tully et al., 2018).

3. Develop guidelines for individual researchers to improve IDI research literacy and awareness of public and cultural necessities regarding the secondary use of these data. These may be supported by a checklist to accompany IDI-related project proposals and ethics applications. Clarify the training and standards for researchers who are directly accessing IDI data, and encourage the use of regularly updated peer-reviewed data dictionaries, so that data are extracted in a safe, societally beneficent and culturally non-maleficent manner.
4. Amend all consent forms for primary data collection (e.g. hospital consent forms) used by public services, so that they explicitly mention possible linkage and secondary data use via the IDI and optimise verbal consent processes, which are preferred by Māori . Provide access to further information for those who are interested. Opt-in consent for each research project may not be relevant if there is adequate transparency regarding re-use of data.
5. Further consideration and consultation is needed before purely private and commercial use of IDI data is approved, as it appears to be less supported by public

opinion. We note that in Australia, solely commercial access to My Health Record data is prohibited, while joint public-private partnerships may be eligible (Australian Government: Department of Health, 2018), suggesting that there may be ways in which commercial interests can be aligned with public ones.

6. Public and researcher responses to these measures should be regularly evaluated so that documentation and procedures can be updated as required.

Research Agenda

Further research is needed into the views of Aotearoa/New Zealand researchers, consent processes for data sharing in diverse populations and the impact of research undertaken using the IDI. More detailed analysis of Māori preferences about IDI data use using culturally sanctioned methodology, and public knowledge regarding general IDI data sources (including markers of both strengths and deficits), data storage, data sharing and general data use would also be valuable.

Educational Implications

Our findings and recommendations are suitable for IDI researchers, ethicists and policy makers. We have shared them with Statistics NZ to convey to existing researchers and intend to disseminate them to a wider audience via publication and promotion via NZ universities and other research organisations.

Acknowledgements

We acknowledge Priya Ramakrishnan (PR) for her assistance with interviews, Anne Wilson (AW) for her assistance with the background literature review and the New Zealand Health Research Council for financially supporting Noor Aljawahiri's time on the project.

Author note

Corresponding author: Hiran Thabrew, The Werry Centre, Department of Psychological Medicine, University of Auckland, Room 340, 20-22 Park Road, Grafton, Auckland 1142, New Zealand. Email: h.thabrew@auckland.ac.nz

LIST OF ABBREVIATIONS (in order of appearance)

- IDI- Integrated Data Infrastructure
- GDPR- General Data Protection Regulation
- NZ- New Zealand
- HDEC- Health and Disability Ethics Committee
- UK- United Kingdom
- US- United States
- CARE- Collective benefit, Authority to control, Responsibility & Ethics

REFERENCES

- Aitken, M., Cunningham-Burley, S., & Pagliari, C. (2016). Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. *Science and Public Policy*, 43(5), 713-723.
- Aitken, M., Jorre, J. d. S., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Medical Ethics*, 17(1), 1-24.
- Asai, A., Ohnishi, M., Nishigaki, E., Sekimoto, M., Fukuhara, S., & Fukui, T. (2002). Attitudes of the Japanese public and doctors towards use of archived information and samples without informed consent: Preliminary findings based on focus group interviews. *BMC Medical Ethics*, 3(1), 1-10.
- Atkinson, J., & Blakely, T. (2017). New Zealand's Integrated Data Infrastructure (IDI): Value to date and future opportunities: IJPDS (2017) Issue 1, Vol 1: 105, Proceedings of the IPDLN Conference (August 2016). *International Journal of Population Data Science*, 1(1).
- Australian Government: Department of Health. (2018). *Framework to Guide the Secondary Use of My Health Record System Data*. Retrieved from [https://www1.health.gov.au/internet/main/publishing.nsf/Content/F98C37D22E65A79BCA2582820006F1CF/\\$File/MHR_2nd_Use_Framework_2018_ACC_AW3.pdf](https://www1.health.gov.au/internet/main/publishing.nsf/Content/F98C37D22E65A79BCA2582820006F1CF/$File/MHR_2nd_Use_Framework_2018_ACC_AW3.pdf)
- Avram, R., Marquis-Gravel, G., Simard, F., Pacheco, C., Couture, É., Tremblay-Gravel, M., . . . Mansour, S. (2018). Understanding the patient perspective on research access to national health records databases for conduct of randomized registry trials. *International journal of cardiology*, 262, 110-116.
- Bell, E. A., Ohno-Machado, L., & Grando, M. A. (2014). *Sharing my health data: a survey of data sharing preferences of healthy individuals*. Paper presented at the AMIA Annual Symposium Proceedings.
- Bond, C. S., Ahmed, O. H., Hind, M., Thomas, B., & Hewitt-Taylor, J. (2013). The conceptual and practical ethical dilemmas of using health discussion board posts as research data. *Journal of medical Internet research*, 15(6), e112.
- Bowden, N., Gibb, S., Thabrew, H., Audas, R., Camp, J., Taylor, B., & Hetrick, S. (2019). IDI trends in antidepressant dispensing to New Zealand children and young people between 2007/08 and 2015/16. *NZ Med J*, 132(1505), 48-61.
- Bowden, N., Gibb, S., Thabrew, H., Kokaua, J., Audas, R., Merry, S., . . . Hetrick, S. E. (2020). Case identification of mental health and related problems in children and young people using the New Zealand Integrated Data Infrastructure. *BMC medical informatics and decision making*, 20(1), 1-13.
- Braun, V., Clarke, V., & Terry, G. (2014). Thematic analysis. *Qual Res Clin Health Psychol*: 95–114. In.
- Buckley, B. S., Murphy, A. W., & MacFarlane, A. E. (2011). Public attitudes to the use in research of personal health information from general practitioners' records: a survey of the Irish general public. *Journal of Medical Ethics*, 37(1), 50-55.
- Caulfield, T., Rachul, C., & Nelson, E. (2012). Biobanking, consent, and control: a survey of Albertans on key research ethics issues. *Biopreservation and Biobanking*, 10(5), 433-438.

- Charania, N. A., Paynter, J., Lee, A. C., Watson, D. G., & Turner, N. M. (2018). Exploring immunisation inequities among migrant and refugee children in New Zealand. *Human vaccines & immunotherapeutics*, *14*(12), 3026-3033.
- Clerkin, P., Buckley, B., Murphy, A., & MacFarlane, A. (2013). Patients' views about the use of their personal information from general practice medical records in health research: a qualitative study in Ireland. *Family practice*, *30*(1), 105-112.
- Council, T. N. H. a. M. R. (2007). Qualifying or waiving conditions for consent in National Statement on Ethical Conduct in Human Research 2007 Retrieved from https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018#toc_296. (2018).
- Cullen, R., & Reilly, P. (2008). Information Privacy and Trust in Government: a citizen-based perspective from New Zealand. *Journal of Information Technology & Politics*, *4*(3), 61-80.
- Custers, B., & Uršič, H. (2016). Big data and data reuse: a taxonomy of data reuse for balancing big data benefits and personal data protection. *International data privacy law*, *6*(1), 4-15.
- Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social science & medicine*, *64*(1), 223-235.
- Davidson, A., Bowie, C., Ivory, V., Harding, A., Thomas, J., & Beetham, J. (2015). *Public Attitudes to Data Integration (Rep)*. Retrieved from <http://archive.stats.govt.nz/~media/Statistics/browse-categories/snapshots-of-nz/integrated-data-infrastructure/pub-att-data-int-2015/public-attit-data-integr-2015>
- Davidson, S., McLean, C., Treanor, S., Aitken, M., Cunningham-Burley, S., Laurie, G., . . . Sethi, N. (2013). *Public acceptability of data sharing between the public, private and third sectors for research purposes*: Scottish Government.
- Dorey, C. M., Baumann, H., & Biller-Andorno, N. (2018). Patient data and patient rights: Swiss healthcare stakeholders' ethical awareness regarding large patient data sets—a qualitative study. *BMC Medical Ethics*, *19*(1), 1-14.
- Eloranta, K., & Auvinen, A. (2015). Population attitudes towards research use of health care registries: a population-based survey in Finland. *BMC Medical Ethics*, *16*(1), 1-10.
- Fry, C. L., Spriggs, M., Arnold, M., & Pearce, C. (2014). Unresolved ethical challenges for the Australian personally controlled electronic health record (PCEHR) system: key informant interview findings. *AJOB Empirical Bioethics*, *5*(4), 30-36.
- Grande, D., Mitra, N., Shah, A., Wan, F., & Asch, D. A. (2013). Public preferences about secondary uses of electronic health information. *JAMA internal medicine*, *173*(19), 1798-1806.
- Grande, D., Mitra, N., Shah, A., Wan, F., & Asch, D. A. (2014). The importance of purpose: moving beyond consent in the societal use of personal health information. *Annals of internal medicine*, *161*(12), 855-862.
- Grant, A., Ure, J., Nicolson, D. J., Hanley, J., Sheikh, A., McKinstry, B., & Sullivan, F. (2013). Acceptability and perceived barriers and facilitators to creating a national research register to enable 'direct to patient' enrolment into research: the Scottish Health Research Register (SHARE). *BMC health services research*, *13*(1), 1-15.
- Gulliver, P., Jonas, M., McIntosh, T., Fanslow, J., & Waayer, D. (2018). Qualitative research: Surveys, social licence and the integrated data infrastructure. *Aotearoa New Zealand Social Work*, *30*(3), 57-71.

- Haddow, G., Bruce, A., Sathanandam, S., & Wyatt, J. C. (2011). 'Nothing is really safe': a focus group study on the processes of anonymizing and sharing of health data for research purposes. *Journal of evaluation in clinical practice*, 17(6), 1140-1146.
- Hate, K., Meherally, S., Shah More, N., Jayaraman, A., Bull, S., Parker, M., & Osrin, D. (2015). Sweat, skepticism, and uncharted territory: a qualitative study of opinions on data sharing among public health researchers and research participants in Mumbai, India. *Journal of Empirical Research on Human Research Ethics*, 10(3), 239-250.
- Hill, E. M., Turner, E. L., Martin, R. M., & Donovan, J. L. (2013). "Let's get the best quality research we can": public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. *BMC medical research methodology*, 13(1), 1-10.
- Hudson, M., Anderson, T., Dewes, T., Temara, P., Whaanga, H., & Roa, T. (2016). Tribal data sovereignty: Whakatōhea rights and interests. In T. J. Kukutai T (Ed.), *Indigenous data sovereignty: Toward an agenda* (pp. 157-178): ANU Press.
- Hudson, M., Anderson, T., Dewes, T., Whaanga, H., & Roa, T. (2017). Conceptualising Big Data through a Māori lens. In *He Whare Hangarau Māori — Language, culture & technology*. (pp. 64-73): ANU Press.
- Jagsi, R., Griffith, K. A., Sabolch, A., Jones, R., Spence, R., De Vries, R., . . . Bradbury, A. R. (2017). Perspectives of patients with cancer on the ethics of rapid-learning health systems. *Journal of Clinical Oncology*, 35(20), 2315.
- Kim, K. K., Joseph, J. G., & Ohno-Machado, L. (2015). Comparison of consumers' views on electronic data sharing for healthcare and research. *Journal of the American Medical Informatics Association*, 22(4), 821-830.
- Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. *BMC Medical Ethics*, 18(1), 1-10.
- King, T., Brankovic, L., & Gillard, P. (2012). Perspectives of Australian adults about protecting the privacy of their health information in statistical databases. *International journal of medical informatics*, 81(4), 279-289.
- Kokaua, J., Jensen, S., Ruhe, T., Camp, J., Jensen, W., Sorensen, D., . . . Richards, R. (2020). An Application of a Tivaivai Research framework to a quantitative Pacific health research project using New Zealand's Integrated Data Infrastructure: Tivaivai framework for quantitative Pacific research. *Pacific Health Dialog*, 21(5), 206-215.
- Kukutai, T., Carroll, S., & Walter, M. (2020). *Indigenous data sovereignty*. Retrieved from Copenhagen, Denmark: IWGIA:
- Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*: ANU press.
- Lehnbom, E., Brien, J., & McLachlan, A. (2014). Knowledge and attitudes regarding the personally controlled electronic health record: an Australian national survey. *Internal medicine journal*, 44(4), 406-409.
- Lenca, M., Ferretti, A., Hurst, S., Puhan, M., Lovis, C., & Vayena, E. (2018). Considerations from ethics review of big data health research: a scoping review. *PLoS one*, 13(10), e020493.
- Lofters, A. K., Shankardass, K., Kirst, M., & Quiñonez, C. (2011). Sociodemographic data collection in healthcare settings: an examination of public opinions. *Medical care*, 193-199.
- Lucero, R. J., Kearney, J., Cortes, Y., Arcia, A., Appelbaum, P., Fernández, R. L., & Luchsinger, J. (2015). Benefits and risks in secondary use of digitized clinical data:

- views of community members living in a predominantly ethnic minority urban neighborhood. *AJOB Empirical Bioethics*, 6(2), 12-22.
- McLeod, K., Templeton, R., Ball, C., Tumen, S., Crichton, S., & Dixon, S. (2015). *Using Integrated Administrative Data to Identify Youth Who Are at Risk of Poor Outcomes as Adults*. Retrieved from
- Milberg, S. J., Burke, S. J., Smith, H. J., & Kallman, E. A. (1995). Values, personal information privacy, and regulatory approaches. *Communications of the ACM*, 38(12), 65-74.
- Milne, B. J., Atkinson, J., Blakely, T., Day, H., Douwes, J., Gibb, S., . . . Teng, A. (2019). Data resource profile: the New Zealand integrated data infrastructure (IDI). *Int J Epidemiol*, 48(3), 677-677e.
- Mourby, M., Mackey, E., Elliot, M., Gowans, H., Wallace, S. E., Bell, J., . . . Kaye, J. (2018). Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. *Computer Law & Security Review*, 34(2), 222-233.
- Page, S. A., Manhas, K. P., & Muruve, D. A. (2016). A survey of patient perspectives on the research use of health information and biospecimens. *BMC Medical Ethics*, 17(1), 1-9.
- Palmer, M. S. (2008). The Treaty of Waitangi in New Zealand's law and constitution.
- Parkin, L., & Paul, C. (2011). Public good, personal privacy: a citizens' deliberation about using medical information for pharmacoepidemiological research. *Journal of Epidemiology & Community Health*, 65(2), 150-156.
- Riordan, F., Papoutsis, C., Reed, J. E., Marston, C., Bell, D., & Majeed, A. (2015). Patient and public attitudes towards informed consent models and levels of awareness of Electronic Health Records in the UK. *International journal of medical informatics*, 84(4), 237-247.
- Robertson, A. R., Smith, P., Sood, H., Cresswell, K., Nurmatov, U., & Sheikh, A. (2016). Tightrope walking towards maximising secondary uses of digitised health data: a qualitative study. *Journal of Innovation in Health Informatics*, 23(3), 591-599.
- Rubinstein, I. (2012). Big data: the end of privacy or a new beginning? *International Data Privacy Law (2013 Forthcoming)*, NYU School of Law, Public Law Research Paper(12-56).
- Saito, K., Shofer, F. S., Saberi, P., & Green-McKenzie, J. (2017). Health Care Personnel Perception of the Privacy of Electronic Health Records. *Journal of occupational and environmental medicine*, 59(6), 535-538.
- Shackleton, N., Broadbent, J. M., Thornley, S., Milne, B. J., Crengle, S., & Exeter, D. J. (2018). Inequalities in dental caries experience among 4-year-old New Zealand children. *Community dentistry and oral epidemiology*, 46(3), 288-296.
- Shackleton, N., Milne, B., Audas, R., Derraik, J. G., Zhu, T., Taylor, R., . . . Taylor, B. (2018). Improving rates of overweight, obesity and extreme obesity in New Zealand 4-year-old children in 2010–2016. *Pediatric obesity*, 13(12), 766-777.
- Spencer, K., Sanders, C., Whitley, E. A., Lund, D., Kaye, J., & Dixon, W. G. (2016). Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study. *Journal of medical Internet research*, 18(4), e5011.
- Statistics New Zealand. (2017a). Integrated Data Infrastructure. Retrieved from <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure>.
- Statistics New Zealand. (2017b). Privacy impact assessments for the IDI and LBD.

- Statistics New Zealand. (2017). How we keep IDI and LBD data safe. Retrieved from http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure/keep-data-safe.aspx
- Statistics New Zealand. (2020). Legislation, policies, and protocols.
- Stone, M. A., Redsell, S. A., Ling, J. T., & Hay, A. D. (2005). Sharing patient data: competing demands of privacy, trust and research in primary care. *British journal of general practice*, 55(519), 783-789.
- Travis, A. (2017). NHS hands over patient records to Home Office for immigration crackdown. . Retrieved from <https://www.theguardian.com/uk-news/2017/jan/24/nhs-hands-over-patient-records-to-home-office-for-immigration-crackdown>
- Trinidad, S. B., Fullerton, S. M., Bares, J. M., Jarvik, G. P., Larson, E. B., & Burke, W. (2010). Genomic research and wide data sharing: views of prospective participants. *Genetics in Medicine*, 12(8), 486-495.
- Tully, M. P., Bozentko, K., Clement, S., Hunn, A., Hassan, L., Norris, R., . . . Peek, N. (2018). Investigating the extent to which patients should control access to patient records for research: a deliberative process using citizens' juries. *Journal of medical Internet research*, 20(3), e7763.
- Valentine, N., Scampion, L., & Taylor, R. (2020). New Zealand: Significant changes to NZ's privacy act – but where is the bite?
- Walter, M., & Suina, M. (2019). Indigenous data, indigenous methodologies and indigenous data sovereignty. *International Journal of Social Research Methodology*, 22(3), 233-243.
- Whiddett, D., Hunter, I., McDonald, B., Norris, T., & Waldon, J. (2016). Consent and widespread access to personal health information for the delivery of care: a large scale telephone survey of consumers' attitudes using vignettes in New Zealand. *BMJ open*, 6(8), e011640.
- Zarate, O. A., Brody, J. G., Brown, P., Ramirez-Andreotta, M. D., Perovich, L., & Matz, J. (2016). Balancing benefits and risks of immortal data: participants' views of open consent in the personal genome project. *Hastings Center Report*, 46(1), 36-45.