Waving away waivers: an obligation to contribute to ‘herd knowledge’ for data linkage research?

Owen M Bradfield
University of Melbourne, Australia

Abstract
In today’s online data-driven world, people constantly shed data and deposit digital footprints. When individuals access health services, governments and health providers collect and store large volumes of health information about people that can later be retrieved, linked and analysed for research purposes. This can lead to new discoveries in medicine and healthcare. In addition, when securely stored and de-identified, the privacy risks are minimal and manageable. In many jurisdictions, ethics committees routinely waive the requirement for researchers to obtain consent from data subjects before using and linking these datasets in an effort to balance respect for individuals with research efficiency. In this paper, I explore the ethical justification for using routinely collected health data for research without consent. I conclude that, not only is this morally justified but also that data subjects have a moral obligation to contribute their data to such research, which would obviate the need for ethics committees to consider consent waivers. In justifying this argument, I look to the duty of easy rescue, distributive justice and draw analogies with vaccination ethics.

Keywords
Autonomy, data linkage, consent, public health, research ethics

Corresponding author:
Owen M Bradfield, Law and Public Health Unit, Centre for Health Policy, Melbourne School of Population and Global Health, University of Melbourne, 207 Bouverie Street, Carlton, VIC 3053, Australia.
Email: o.bradfield@student.unimelb.edu.au
Introduction

For much of the last 50 years, medical research primarily involved testing potentially harmful interventions on humans. However, when research is limited to linking and then analysing pre-existing data, different risks are created. This raises the ethical question of whether consent to link pre-existing data is necessary, particularly since requiring consent may hinder research, distort results and create novel privacy risks. In this paper, I focus on the secondary use of pre-existing de-identified public sector health data in linkage research, where linkage is defined as the process of matching records from multiple datasets that belong to the same individual (Winglee et al., 2005). I will argue that consent from data subjects is not required before linking pre-existing data, when considering the potential benefits and risks involved. I will also argue that citizens have a conditional moral responsibility to make their data available for such research because maximal benefit from this type of research is realised only when datasets contain information about sufficiently large numbers of data subjects. We each have an individual and collective responsibility to contribute to the ‘herd knowledge’ that these datasets yield. This paper adds to the literature by exploring novel arguments for this position, based on the rescue principle, distributive justice and vaccination ethics.

I will make several assumptions to limit the scope of my arguments. First, I will only consider linkage research involving administrative data that has been stripped of identifying information. I will not consider research involving access to descriptive content data, such as clinical records. Second, I will only consider data linkage research projects that are of sufficient scientific merit and social benefit to have received Human Research Ethics Committee (Institutional Review Board) approval. Third, while attempts should always be made to obtain some form of consent from data subjects prior to data collection, this paper only addresses the question of what to do with potentially useful data that has already been collected without consent. As such, I will not consider the relative merits of different types of consent models, such as broad consent, opt-out consent, dynamic consent or meta-consent (Teare et al., 2021). Finally, I will only consider public sector health data collected by democratic governments for use in publicly funded research. Important concerns about the nefarious use or sale of health information collected by private corporations such as Google and Facebook, or health insurers, are beyond the scope of this paper and described elsewhere (Véliz, 2020).

Why consent to linkage research is unnecessary

Sensitive information about individuals is routinely generated, collected and compiled into large datasets by governments and healthcare providers without consent. Examples include the notifications of births, deaths, cancers, infectious diseases
and hospital admissions. Digitisation has accelerated the capacity to extract, manipulate, aggregate, share and analyse this information. This creates innovative research opportunities that is supported in many jurisdictions by open data policies (Andreu-Perez et al., 2015). For example, the Australian Government has promised to transform public sector data sharing from ‘aversion and avoidance’ to ‘transparency and confidence’ by allowing researchers to access de-identified data for research with clear and direct public benefits that minimises privacy risks (Anton, 2020).

**Benefits and risks of data linkage research**

The compilation of datasets pertaining to entire populations over many years facilitates the efficient investigation of rare diseases, rare outcomes and under-researched populations. Findings drawn from these datasets can be statistically powerful (Corrada et al., 2014). Linkage between datasets can uncover new information, without exposing research participants to potentially harmful experiments or interventions. In some circumstances, this is the only morally acceptable way of evaluating ‘real world’ outcomes in groups, such as pregnant women, for whom interventional research may be too risky (Nguyen and Barshes, 2010). However, data linkage is not itself without risks.

Data linkage involves the storage, transfer and analysis of sensitive information. Privacy breaches can occur. Researchers can match de-identified data with publicly available information, allowing sensitive inferences to be drawn from non-sensitive data (Townsend, 2021). More importantly, data linkage research can adversely impact minority groups, including Indigenous communities. This can occur when they are excluded from data collections or when research questions, methodologies and analysis do not appropriately reflect Indigenous priorities, values, identity and diversity. ‘Indigenous data sovereignty’ allows Indigenous people to be the primary beneficiaries of their data, knowledge and cultural heritage by entrusting Indigenous communities with control over what and how data is collected, analysed, interpreted and disseminated (Kukutai and Taylor, 2016).

To reduce data linkage risks, Human Research Ethics Committees usually require researchers to separate identifiable data (e.g. name, address and date of birth) from non-identifiable data and use a non-identifiable ‘linkage key’ to compares data from the same individual across multiple datasets (Emery and Boyle, 2017). Also, in Australia, data linkage often only occurs within accredited data linkage integrating authorities (Australian Government, 2020) where researchers must consider the privacy impacts of funded research and securely store data on encrypted drives and adopt stringent data management plans (Christen et al., 2020). Together with growing understanding of Indigenous data sovereignty, these
measures have proven largely effective in mitigating the risks, with few complaints or breaches registered (O’Keefe and Connolly, 2010).

**Benefits and risks of abrogating consent**

Central to research ethics is the requirement that human research participation should follow a free and voluntary choice made by potential participants based on sufficient understanding of risks posed by participation (World Medical Association, 2013). The compulsory acquisition of data for future use is potentially coercive because it removes the choice of citizens not to yield their data (Wertheimer, 1996). This moral cost is worth carefully considering, given the salience afforded to personal autonomy and privacy in liberal democracies (Herring and Wall, 2017). Uncritically abrogating informed consent may be ‘morally hazardous’ (Boyd, 2007) because privacy rights protect people from unsolicited surveillance of, and intrusion upon, their private lives. Moreover, enforcing research participation could undermine public trust if research governance inadequately protects subjects from risks they want to avoid (Carter et al., 2015). As a result, some citizens might avoid accessing health services or withhold information from health providers, resulting in suboptimal health outcomes.

Nevertheless, requiring consent for linkage research may be unfeasible and may even create new risks. First, it would be almost impossible to obtain consent from the millions of data subjects whose data is contained within some large datasets. The time and cost expended could curtail important and time-critical research if entire research budgets were spent obtaining consent rather than analysing and interpreting data. It might also be impractical to obtain consent when data from smaller datasets has been collected over a prolonged period because data subjects may be uncontactable if their name or contact details have changed since enrolment, or if they have died.

Second, requiring consent can reduce inclusion of sufficient numbers of research subjects through low response rates and attrition, leading to sub-optimal or incomplete linkage (Bohensky et al., 2010). In the case of research involving rare diseases or rare outcomes, it can also skew results, causing ‘consent bias’. This occurs when those who consent to research participation systematically differ from those who do not in ways that make those studied unrepresentative of the population (Mazzali and Duca, 2015). While there is debate about its frequency and impact (Ploug, 2020), it likely reduces the reliability and validity of research findings to some extent.

Third, the information used in data linkage research already exists, which means that declining consent to its use does not guarantee secure collection or storage. Most importantly, reidentifying anonymous data subjects for the purpose of
obtaining consent aggravates the risk of privacy breaches to de-identified individuals because it exposes their identity to researchers.

**Balancing private interests and public benefits**

Clearly, linking data without consent entails balancing autonomy infringements with potential healthcare benefits (Gostin and Wiley, 2016). Many moral frameworks converge on the premise that coercive policies may justify curtailing individual freedoms in pursuit of anticipated social benefits (Department of Health, Education, and Welfare; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 2014). Even small benefits may allow for interventions that are relatively costless (Childress et al., 2002), provided they are effective, necessary and proportionate (Bradfield and Giubilini, 2021). In this case, the re-use of data without consent is both necessary and effective in maximising participation to increase the size, power and potential benefits of large public datasets. It is also proportionate because, even though this imposes coercive pressure on large populations, the individual impact is usually limited to loss of informational control that can be mitigated and is less ‘central to one’s self-governance’ than breaches of bodily integrity (Ballantyne and Schaefer, 2018).

Therefore, while consent buttresses individual autonomy and maintains trust in public health research, it ought not be required for linkage of data already collected because: (1) the risk of using de-identified information is minimal; (2) the potential benefits to others are significant; and (3) requiring consent can create novel privacy risks and frustrate the realisation of the full potential of large datasets. Indeed, not only is consent not required, but there is a moral obligation on citizens to contribute their data to research.

**An obligation to contribute data to research**

The established view that research participation is voluntary and supererogatory is increasingly being challenged by the view that citizens have a moral duty to participate in some research (Stjernschantz Forsberg et al., 2014). In this section, I will argue why citizens have a moral obligation to contribute their data to linkage research.

**Duty of easy rescue**

According to the rescue principle, when the actions required by an individual to avoid possible harm to others are sufficiently easy and costless, an individual has a moral obligation to act. Peter Singer conceives this duty through his much-cited shallow pond example:
. . . if I am walking past a shallow pond and see a child drowning in it, I ought to wade in and pull the child out. This will mean getting my clothes muddy, but this is insignificant, while the death of the child would presumably be a very bad thing (Singer, 1972).

The moral duty to rescue is sculpted by the urgency of the situation, the consequences of inaction and the exertion required by the potential rescuer to prevent those consequences (Bradfield, 2021). Data contribution entails a relatively small cost to individuals. Although the individual impact of a privacy breach may be significant, the likelihood is small. Likewise, when Indigenous and other minority groups affected by research projects are involved in the planning, conduct, analysis, interpretation and publication of linkage research, group harms can also be mitigated. On the other hand, data contribution also entails a potentially large benefit to others in terms of knowledge gained that may save future lives or prevent serious suffering. Some have therefore argued that a ‘duty of easy rescue’ grounds an individual’s duty to contribute data to research (Porsdam Mann et al., 2016).

However, there are two criticisms of applying this principle to the duty to contribute data. First, unlike some potentially life-saving interventional research, data linkage research is not ‘an instrument for rescuing sudden victims of accident’ (Ploug, 2020). Instead, its aim is to prevent harm to potential future patients. Second, there is no guarantee that data linkage research will uncover useful information that will prevent future harm. Even if the information is useful, there is often a ‘lengthy, complex and unpredictable path’ (Ploug, 2020) from scientific discovery to improvements in healthcare.

These criticisms weaken, but do not invalidate, the argument that a moral obligation to contribute data could be grounded in a duty of easy rescue. Some research may be urgently required to address issues of pressing global significance, such as the need to understand, trace and prevent COVID-19 transmission (Bhattacharya et al., 2021).

**Distributive justice**

A stronger argument in favour of a duty to contribute data to linkage research is grounded in the principle of fairness. In his seminal work on the ‘duty to research’, John Harris argues that those who benefit from medical research without contributing to it are free riding because they benefit from others’ efforts without reciprocating (Harris, 2005). Being a free rider is unfair and people have a moral obligation not to act unfairly. Therefore, fairness requires that we contribute data to research from which we will benefit.

Brassington disagrees with Harris’ fairness argument. He asserts that there is no free riding problem if people who do not participate in research have paid for medical benefits they receive through taxes, medical insurance premiums or other
means (Brassington, 2007). However, while Brassington’s response addresses the question of free riding in relation to the end benefits of research, it does not solve the free riding problem with respect to the research discovery process, which may lead to those medical benefits. Applying these opposing views to vaccination, if individuals fail to vaccinate themselves, but society at large is sufficiently vaccinated, then unvaccinated individuals benefit from herd immunity (Kim et al., 2011). On the one hand, Harris might say that this is a form of free riding because unvaccinated individuals benefit from the indirect protection afforded to them by those who are vaccinated. On the other hand, Brassington might say that unvaccinated individuals are not free riding because their taxes paid for the vaccines given to others. However, neither position is entirely correct because achieving herd immunity requires, amongst other things, payment of taxes and vaccination.

Paying taxes without also being vaccinated, or vice versa, does not amount to free riding because a partial contribution has been made. Instead, it amounts to a discounted ride, which may seem unfair from the perspective of the person paying full fare. By analogy, linkage research requires, amongst many other things, payment of taxes to fund medical research and contribution of data to that research. Hence, it could be seen as unfair for citizens to equally benefit from the knowledge gained from data linkage research when they have contributed unequally to the attainment of that knowledge. However, there are flaws with this argument too. Although citizens have an obligation to do all that is reasonably within their power to maximally contribute to data linkage research, there are sound reasons for allowing some people to contribute less. For example, people with chronic illness or disability who cannot work may be exempted from paying taxes. Likewise, minority groups and Indigenous communities may legitimately mistrust how their data will be used. Hence, fairness requires citizens to contribute equally, not in absolute terms but within the limits of their resources and capacity.

Another reason why data contribution might be fair is because the amount of health data generated by individuals should be directly proportional to the amount of healthcare they receive as patients. Patients who receive more clinical care generate comparatively more health information that is available for research use. These people theoretically stand to benefit the most from the data generated and collected from others that can improve treatment of medical conditions they have. However, this argument assumes that access to health services is equitable, which we know is often not the case. Those who generate the most data may not always receive the most healthcare and this is problematic because information they have given up may not be fairly exchanged for care. Therefore, this argument only holds if everyone receives the quality and quantity of care that is commensurate with how sick they are. Moreover, we must also protect vulnerable groups from potential exploitation. For example, minority groups and Indigenous communities have not always benefited from public health research (Walter et al., 2021). Many
Indigenous and minority ethnic communities have suffered significant cultural harm and discrimination from the inappropriate use of public health data, poor quality data and erroneously implied causality. Nevertheless, fairness can still be achieved if these communities are afforded control over how linked data is analysed and interpreted.

While there are limitations with the fairness argument, routinely collected health data is a public resource and, to the extent possible, should be used in ways that maximises public benefit. The curation and generation of these datasets requires public expenditure and the information contained therein provides a platform for answering important questions that can drive public healthcare improvements.

**Contributing to ‘herd knowledge’**

I propose a third and novel basis for justifying a moral obligation to contribute data to research that draws upon arguments used to justify a moral obligation to vaccinate people in pursuit of ‘herd immunity’. Like individual vaccination decisions, individual decisions about data contribution affect both the individual and society. ‘Herd immunity’ describes the indirect protection from infectious diseases afforded to unvaccinated people when a sufficient proportion of the population is vaccinated (Randolph and Barreiro, 2020). By analogy, only when enough people contribute their data to large data linkage studies can the results emanating from large datasets be statistically significant, meaningful, reliable and applicable, even to those who do not contribute their data to the research from which they are benefitting.

I suggest the term ‘herd knowledge’ to describe the benefit that flows to society when potentially powerful information is generated from research using sufficiently large datasets. Like herd immunity, the pursuit of herd knowledge may ground an individual moral obligation to contribute data because it requires the cooperation of a sufficiently large number of people for potential collective benefits to be realised. If, like vaccination, there is a collective moral obligation to realise herd knowledge, then each of us has the individual moral responsibility to contribute to data linkage research, all things considered.

When a sufficient portion of the population is vaccinated, vaccinating more people usually contributes little more to herd immunity and individual vaccination risks outweigh the collective benefits of herd immunity (Giubilini, 2019). Even though viruses are dynamic and the threshold for herd immunity can shift, the incremental benefits in vaccinating above the threshold for herd immunity may be limited. Likewise, while collecting and analysing ever more data can produce more powerful, representative and granular results, this will likely produce diminishing returns from subsequent additions to a dataset. Therefore, just as herd immunity generally occurs only up to a threshold level of vaccination coverage,
there is also likely to be a ceiling effect when it comes to the benefits of contribution of data to linkage research. Above that ceiling, the focus is likely to change from maximising recruitment numbers to focusing on who is included and excluded from the dataset.

However, there are some important disanalogies between the putative moral obligations to pursue herd immunity and herd knowledge. First, while the failure to vaccinate can create a direct harm to others by facilitating viral transmission, no such direct harm exists when we fail to contribute our data. Research using a small dataset may yield valuable results, but research using a larger dataset may yield more valuable results. The failure to fully realise the benefits of existing information by not consenting to its linkage does not create tangible harm. However, a person could still be individually blameworthy for failing to contribute data to potentially beneficial research because it is a failure to ‘do her part in a collective action’ (Giubilini, 2019) that could derive knowledge and information about the treatment of medical conditions.

Second, the benefits of vaccination are primarily individual and secondarily collective, whereas in the case of data contribution, the potential benefits are primarily collective, but can also be individual. Although those who contribute more data may have more to gain because they utilise more health services (as outlined above), this individual benefit is unlikely to be sufficiently direct or contemporaneous to act as any kind of strong incentive for individuals to contribute data. It is precisely this absence of personal incentive, coupled to the potential social benefits of data research, that justifies the compulsory acquisition by the state of citizen’s data for use in data linkage research without the individual’s consent.

Third, the achievement of herd immunity has been used to justify mandatory vaccination, but not the forced vaccination of individuals against their will (Bradfield and Giubilini, 2021). On the contrary, I argue that the achievement of herd knowledge can be used to justify the analysis of collected information without any consent. The difference lies in the nature of the autonomy infringements at stake. While enforced vaccination invades bodily integrity and can give risk to physical harm, the infringement suffered by individuals when collected data is linked and analysed without consent is limited to one of informational control.

**Conclusion**

Recent empirical studies (Kalkman et al., 2019) show widespread community support for the use of de-identified routinely collected health data in research. Therefore, most people seem to largely recognise that health data research can lead to new treatments and better patient care and they support the use of health data in linkage research without consent, provided data is de-identified (Xafis, 2015). In this paper, I have argued that data linkage research can transform
healthcare at relatively low cost to individuals. Therefore, when balancing the benefits and risks of data linkage research, it is morally permissible to use and link people’s data without their consent. Moreover, I have shown that the general population also has a moral obligation to contribute their data to such research endeavours based on arguments of beneficence, fairness and the attainment of ‘herd knowledge’. However, even if linkage of large datasets can uncover important information at low cost to individuals, robust systems of research governance remain necessary to ensure privacy impacts are minimised and Indigenous data sovereignty is protected.

Acknowledgements
Dr Bradfield is the recipient of a 2020 Fulbright Future Scholarship.

Funding
Dr Bradfield is the recipient of a 2020 Fulbright Future Scholarship. All articles in Research Ethics are published as open access. There are no submission charges and no Article Processing Charges as these are fully funded by institutions through Knowledge Unlatched, resulting in no direct charge to authors. For more information about Knowledge Unlatched please see here: http://www.knowledgeunlatched.org.

ORCID iD
Owen M Bradfield https://orcid.org/0000-0002-8955-7432

References
Andreu-Perez J, Poon CC, Merrifield RD, et al. (2015) Big data for health. IEEE Journal of Biomedical and Health Informatics 19(4): 1193–1208.
Anton D (2020) Modernising government data sharing. Available at: https://www.datacommissioner.gov.au/media-hub/modernising-government-data-sharing (accessed 5 June).
Australian Government (2020) A guide for data integration projects involving Commonwealth data for statistical and research purposes. Available at: https://statistical-data-integration.govspace.gov.au/roles-and-responsibilities/integrating-authorities#:~:text=The%20integrating%20authority%20is%20the,collect%20the%20statistical%20data%20project.&text=In%20addition%2C%20the%20governance%20framework%20is%20accredited (accessed 5 June).
Ballantyne A and Schaefer GO (2018) Consent and the ethical duty to participate in health data research. Journal of Medical Ethics 44(6): 392–396.
Bhattacharya A, Collin SM, Stimson J, et al. (2021) Healthcare-associated COVID-19 in England: A national data linkage study. medRxiv.
Bohensky MA, Jolley D, Sundararajan V, et al. (2010) Data linkage: A powerful research tool with potential problems. BMC Health Services Research 10(1): 346–347.
Boyd KM (2007) Ethnicity and the ethics of data linkage. BMC Public Health 7(1): 318–323.
Bradfield OM (2021) Shining a light also casts a shadow: Neuroimaging incidental findings in neuromarketing research. Neuroethics. Epub ahead of print 12 April 2021. DOI: 10.1007/s12152-021-09463-x.
Bradfield OM and Giubilini A (2021) Spoonful of honey or a gallon of vinegar? A conditional COVID-19 vaccination policy for front-line healthcare workers. *Journal of Medical Ethics* 47(7): 467–472.

Brassington I (2007) John Harris’ argument for a duty to research. *Bioethics* 21(3): 160–168.

Carter P, Laurie GT and Dixon-Woods M (2015) The social licence for research: Why care data ran into trouble. *Journal of Medical Ethics* 41(5): 404–409.

Childress JF, Faden RR, Gaare RD, et al. (2002) Public health ethics: Mapping the terrain. *The Journal of Law Medicine & Ethics* 30(2): 170–178.

Christen P, Ranbaduge T and Schnell R (2020) *Regulatory Frameworks. Linking Sensitive Data*. Cham, Switzerland: Springer, 27–45.

Corrada E, Ferrante G, Mazzali C, et al. (2014) Eleven-year trends in gender differences of treatments and mortality in ST-elevation acute myocardial infarction in northern Italy, 2000 to 2010. *The American Journal of Cardiology* 114(3): 336–341.

Department of Health, Education, and Welfare; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (2014) The Belmont report. Ethical principles and guidelines for the protection of human subjects of research. *The Journal of the American College of Dentists* 81(3): 4–13.

Emery J and Boyle D (2017) Data linkage. *Australian Family Physician* 46(8): 615–619.

Giubilini A (2019) *The Ethics of Vaccination*. Cham, Switzerland: Springer Nature.

Gostin LO and Wiley LF (2016) *Public Health Law: Power, Duty, Restraint*. Berkeley, CA: University of California Press.

Harris J (2005) Scientific research is a moral duty. *Journal of Medical Ethics* 31(4): 242–248.

Herring J and Wall J (2017) The nature and significance of the right to bodily integrity. *The Cambridge Law Journal* 76(3): 566–588.

Kalkman S, van Delden J, Banerjee A, et al. (2019) Patients’ and public views and attitudes towards the sharing of health data for research: A narrative review of the empirical evidence. *Journal of Medical Ethics*. Epub ahead of print 12 November 2019. DOI: 10.1136/medethics-2019-105651.

Kim TH, Johnstone J and Loeb M (2011) Vaccine herd effect. *Scandinavian Journal of Infectious Diseases* 43(9): 683–689.

Kukutai T and Taylor J (2016) *Indigenous Data Sovereignty: Toward an Agenda*. Canberra: ANU Press.

Mazzali C and Duca P (2015) Use of administrative data in healthcare research. *Internal and Emergency Medicine* 10(4): 517–524.

Nguyen LL and Barshes NR (2010) Analysis of large databases in vascular surgery. *Journal of Vascular Surgery* 52(3): 768–774.

O’Keefe CM and Connolly CJ (2010) Privacy and the use of health data for research. *The Medical Journal of Australia* 193(9): 537–541.

Ploug T (2020) In defence of informed consent for health record research - why arguments from ‘easy rescue’, ‘no harm’ and ‘consent bias’ fail. *BMC Medical Ethics* 21(1): 1–13.

Porsdam Mann S, Savulescu J and Sahakian BJ (2016) Facilitating the ethical use of health data for the benefit of society: Electronic health records, consent and the duty of easy rescue. *Philosophical Transactions of the Royal Society A Mathematical Physical and Engineering Sciences* 374(2083): 20160130.

Randolph HE and Barreiro LB (2020) Herd immunity: Understanding COVID-19. *Immunity* 52(5): 737–741.
Singer P (1972) Famine, affluence and morality. *Philosophy & Public Affairs* 1: 229–243.
Stjernschantz Forsberg J, Hansson MG and Eriksson S (2014) Why participating in (certain) scientific research is a moral duty. *Journal of Medical Ethics* 40(5): 325–328.
Teare HJA, Prictor M and Kaye J (2014) Reflections on dynamic consent in biomedical research: The story so far. *European Journal of Human Genetics* 29(4): 649–656.
Townsend B (2021) The lawful sharing of health research data in South Africa and beyond. *Information & Communications Technology Law*. Epub ahead of print 12 April 2019. DOI: 10.1080/13600834.2021.1918905.
Vélez C (2020) *Privacy is Power: Why and How You Should Take Back Control of Your Data*. New York, NY: Random House.
Walter M, Kukutai T, Carroll SR, et al. (2021) *Indigenous Data Sovereignty and Policy*. Abingdon, Oxon; New York, NY: Taylor & Francis.
Wertheimer A (1996) Remarks on coercion and exploitation. *Denver Law Review* 74: 889.
Winglee M, Valliant R and Scheuren F (2005) A case study in record linkage. *Survey Methodology* 31(1): 3–11.
World Medical Association (2013) World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects. *The Journal of the American Medical Association* 310(20): 2191–2194.
Xafis V (2015) The acceptability of conducting data linkage research without obtaining consent: Lay people’s views and justifications. *BMC Medical Ethics* 16(1): 79–16.