

# Centring equity in data-driven public health: a call for guiding principles to support the equitable design and outcomes of Australia's data integration systems

We need to design data systems that hold social and health equity as a core value and desired outcome of data integration

The secondary use of health data — and data pertaining to the social determinants of health — is widely acknowledged as an underutilised yet powerful contributor to research and evidence-based policy. This is particularly the case when data pertaining to individuals are linked across multiple datasets and scaled to include large segments of the population, or even whole-of-population data, in a process known as data integration. The rise of big data, more broadly, has seen data integration increasingly systematised and embedded into government-regulated data infrastructures in order to support quality assurance, research, policy and many forms of innovation. Examples include the Australian Bureau of Statistics' data asset Multi-Agency Data Integration Project (MADIP),<sup>1</sup> the Australian Institute of Health and Welfare's (AIHW) coronavirus disease 2019 (COVID-19) register and linked dataset,<sup>2</sup> and the AIHW National Integrated Health Services Information asset.<sup>3</sup>

While this whole-of-government process affects many fields, data integration is expected to be highly significant to the future of public health. Data integration generates comprehensive and granular data, which, rigorously analysed, contribute to strong research, the design and monitoring of evidence-based policy and health systems, and the development of new medical technologies.<sup>4</sup> Linked data have recently demonstrated that education-related inequalities drive inequities in virtually all causes of death.<sup>5</sup> Others are analysing linked data to improve evidence-based policy around the social and health outcomes of early childhood experience, which carries long term health implications.<sup>6</sup> Other applications include the monitoring and improvement of prescription practices,<sup>7</sup> and the development of artificial intelligence (AI)-supported medical technologies.<sup>8</sup> In addition to universities and learned academies, the benefits of data integration are widely recognised across Australian Government agencies, so that considerable investment and reform have been underway to strengthen data integration capacities to support research and innovation, create efficiencies and private-public partnerships, and accelerate data-driven decision making within policy and routine modes of government.<sup>9</sup>

However, the uptake of big data within government carries ambiguous implications for social and health equity. A major concern globally stems from the

growing awareness that social biases can become embedded within data, and subsequently within its analyses and the algorithmic design of AI-informed technologies, including automated decision making within government.<sup>10</sup> This can be seen, for instance, in the contested ways that disability is classified, measured and acted upon by government within the National Disability Insurance Scheme.<sup>11</sup> In countries that are further advanced in data integration, such as the United Kingdom, private sector access to health data has provoked public controversy, indicating that Australia should be proactive in developing strategies to regulate and share data from and with the private sector in ways that allow innovation while also aligning with broadly shared visions of the public good.<sup>12,13</sup> The imperative to keep human decision making visible, and data systems transparent and accountable, is a major concern globally.<sup>8,10,13-15</sup> Many agree that public confidence in the secondary uses of health data requires a social license.<sup>12,13,15</sup> Such complexities stand alongside privacy considerations and concerns about cyber security, which are also substantial challenges.

In recognition of these complexities, a growing number of researchers, advocacy groups, think tanks and international governments are developing strategies to inform the development of equity-oriented data systems.<sup>16-19</sup> Although some researchers and policy makers are using data to analyse equity or working collaboratively at a project level, and despite much effort going toward data governance, Australian governments have not yet developed strategies to ensure that social equity considerations are both protected and pursued through the design of government-regulated data systems. Despite sharing an interest in the potential benefits of data integration, some groups have expressed concern that the relevant agencies have not sufficiently consulted with community stakeholders or put adequate effort into developing data strategies to safeguard and improve social equity.<sup>20,21</sup> In this article, we argue that the Australian public will be better served if we take concerted action to collaboratively design data systems that hold social and health equity as a core value and desired outcome of data integration. The equity concerns mentioned above are global issues that are inherent to big data and their changing role within government and society. We propose that a transparent, proactive approach to the design and oversight of data systems to minimise inequitable outcomes, and

Catherine Smith<sup>1</sup> 

Claire M Vajdic<sup>2</sup>

Niamh Stephenson<sup>1</sup>

<sup>1</sup> University of New South Wales, Sydney, NSW.

<sup>2</sup> Kirby Institute, University of New South Wales, Sydney, NSW.

catherine.smith3@unsw.edu.au

doi: 10.5694/mja2.51902

to capitalise on the potential for data integration to improve social and health equity. As researchers aiming to support social and health equity, we seek to look beyond risk mitigation, to call for the intentional design of data systems and governance processes that are geared towards social and health equity.

We call for the key government agencies involved in data integration to initiate the collaborative and consultative development of guiding principles for social and health equity in integrated data systems. Such principles could support infrastructure developers, data regulators, data custodians, data users, and the public to identify how data systems could be designed and governed to support social equity. The guiding principles can look beyond the project level to the systems level and aim to influence the cumulative impact of data integration across projects and programs and through time. This is important to safeguard the longer term and cumulative equity outcomes of data integration — a task that is beyond the remit of ethics committees. Guiding principles may support transparent and accountable decision making, and help to ensure that all involved (systems designers, data custodians, data stewards, analysts, and policy makers drawing on their work) are aware of the sensitivities, ethical complexities, and social equity implications of data integration. They may help to ensure that the right data are collected in culturally safe ways and with more consistency across jurisdictions, and that inequities are captured in ways that reflect social complexity.<sup>22,23</sup> The very process of collaboratively developing the guidelines may produce transparency by engaging a broader range of expert stakeholders and the public, and ensuring that the groups most affected by social and health inequities are given decision-making roles in developing the data systems that will affect their communities.<sup>18,22,23</sup> It will be particularly important to ensure that our data ecosystem affirms and facilitates Indigenous Data Sovereignty.<sup>19,20</sup>

These principles should be collectively developed by key stakeholders and representatives of community and consumer groups, ideally through a process that is led by a peak government agency involved in data regulation such as the Office of the National Data Commissioner (ONDC), although the Office of the Australian Information Commissioner (OAIC), the AIHW, the Australian Bureau of Statistics, and the state-level data agencies (such as Data NSW) also have regulatory functions and expertise to contribute. Whitehead<sup>24</sup> states that: “Equity policy requires a genuine commitment to decentralizing power and decision-making, encouraging people to participate in every stage of the policy-making process”. Ideally, equity will be considered and factored into data systems at all stages of the data life cycle: i) at the level of the data and the enduring integrated datasets that are created; ii) within the design of systems and infrastructures; iii) within legislation and governance; iv) in assessing and monitoring the public interest uses of data; and v) to guide the equitable distribution of the benefits of data integration over the longer term.

A growing literature examines the relative merits of various data governance models (eg, data-sharing

pools, data cooperatives, public data trusts, data commons, or personal data sovereignty systems) — framings that serve to orient data governance and to guide norms around the uses of data.<sup>25</sup> International efforts to develop guidelines on equitable data systems, big data and the social determinants of health, and on health equity more broadly offer rich resources to support the development of Australian guiding principles for equity in the era of big data.<sup>16–18,24</sup> Such resources can inform the collaborative development of guiding principles that are grounded in the specifics of the Australian health system, government, and data ecosystem, and in a vision of data integration’s public good as articulated by and meaningful to the Australian public. Conversations about the overarching purpose of data systems need to take place in a public forum.<sup>14,15,17–19</sup> In addition to epidemiologists, data analysts, consumer groups and legal experts, a burgeoning number of researchers in Australia are examining the social implications of AI, automated decision making, big data, and the regulation of technology, work which can support the equitable design of our data systems.<sup>8,11,12</sup> The Indigenous Data Sovereignty movement, disability rights activists, and other data-related movements have vital expertise to bring to this conversation and to the design of equitable data systems.<sup>11,19</sup>

Data integration is rapidly evolving and backed by significant political will, and a policy window exists to call for the development of such guiding principles, to address equity considerations in substantive ways at a system level, and to ensure that an equity-oriented lens informs all stages of the data life cycle. Guiding principles around equity might also support fair accountability if and when there is a controversial use of data and will help to ensure that decisions are not entirely contingent on the views of individuals. Collaboratively identified equity principles can work to steer data integration towards greater social equity. Although advances in data integration certainly present opportunities to improve public health, we should seek to achieve social and health equity not only through innovative research but by building data systems and a data policy environment that hold equity as a core value and aspired outcome of data integration.

**Acknowledgements:** This research was supported by the Australian Government through the Australian Research Council’s Discovery Projects funding scheme (DP200100062). The funders had no role in the design, data collection, analysis or interpretation, reporting or publication of the article.

**Open access:** Open access publishing facilitated by University of New South Wales, as part of the Wiley - University of New South Wales agreement via the Council of Australian University Librarians.

**Competing interests:** No relevant disclosures.

**Provenance:** Not commissioned; externally peer reviewed. ■

© 2023 The Authors. *Medical Journal of Australia* published by John Wiley & Sons Australia, Ltd on behalf of AMPCo Pty Ltd.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

- 1 Australian Bureau of Statistics. Multi-Agency Data Integration Project (MADIP) [website]. <https://www.abs.gov.au/about/data-services/data-integration/integrated-data/multi-agency-data-integration-project-madip> (viewed Jan 2023).
- 2 Australian Institute of Health and Welfare. COVID-19 register and linked data set [website]. <https://www.aihw.gov.au/reports-data/covid-linked-data-set> (viewed Jan 2023).
- 3 Australian Government Department of Health and Aged Care. Our data collections [website]. <https://www.health.gov.au/topics/health-data-and-medical-research/our-data-collections#national-integrated-health-services-information-nihsi> (viewed Jan 2023).
- 4 Smith M, Flack F. Data linkage in Australia: the first 50 years. *Int J Environ Res Public Health* 2021; 18: 11339.
- 5 Korda RJ, Biddle N, Lynch J, et al. Education inequalities in adult all-cause mortality: first national data for Australia using linked census and mortality data. *Int J Epidemiol* 2019; 49: 511-518.
- 6 Pilkington R, Montgomerie M, Grant J, et al. An innovative linked data platform to improve the wellbeing of children — the South Australian Early Childhood Data Project. [Australia's welfare 2019 data insights; Cat. No. AUS 226]. Canberra: Australian Institute of Health and Welfare, 2019. <https://www.aihw.gov.au/getmedia/75c9f692-c72c-438d-b68b-91318f780154/Australias-Welfare-Chapter-8-summary-18Sept2019.pdf.aspx> (viewed Mar 2023).
- 7 Costa JO, Bruno C, Schaffer AL, et al. The changing face of Australian Data reforms: impact on pharmacoepidemiology research. *Int J Popul Data Sci* 2021; 6: 1418.
- 8 Isbanner S, O'Shaughnessy P, Steel D, et al. The adoption of artificial intelligence in health care and social services in Australia: findings from a methodologically innovative national survey of values and attitudes (the AVA-AI study). *J Med Internet Res* 2022; 24: e37611.
- 9 Office of the National Data Commissioner. Introducing the DATA Scheme [website]. <https://www.datacommissioner.gov.au/the-data-scheme> (viewed Jan 2023).
- 10 Ananny M, Crawford K. Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability. *New Media and Society* 2018; 20: 973-989.
- 11 van Toorn G. Towards a social-relational model of digital disability classification. *Data and Society: Points*, 26 Oct 2022. <https://points.datasociety.net/toward-a-social-relational-model-of-digital-disability-classification-9a0a04ebdd10> (viewed Oct 2022).
- 12 Braunack-Mayer A, Fabrianesi B, Street J, et al. Sharing government health data with the private sector: community attitudes survey. *J Med Internet Res* 2021; 23: e24200.
- 13 Ipsos MORI, Wellcome Trust. The One-Way Mirror: public attitudes to commercial access to health data 2019; Wellcome Trust, 2020. [https://wellcome.figshare.com/articles/journal\\_contribution/The\\_One-Way\\_Mirror\\_Public\\_attitudes\\_to\\_commercial\\_access\\_to\\_health\\_data/5616448/1](https://wellcome.figshare.com/articles/journal_contribution/The_One-Way_Mirror_Public_attitudes_to_commercial_access_to_health_data/5616448/1) (viewed Mar 2023).
- 14 Redden J. Democratic governance in an age of datafication: Lessons from mapping government discourses and practices. *Big Data and Society* 2018; 5: 1-13.
- 15 Waind E. Trust, security and public interest: Striking the balance: a narrative review of previous literature on public attitudes towards the sharing, linking and use of administrative data for research. *Int J Popul Data Sci* 2020; 5: 1368.
- 16 CDC Foundation Health Equity Strategy Office; Panneton M, Hill F, Smith L. Principles for using public health data to drive equity: a guide to embedding equitable practices throughout the data life cycle. <https://www.cdcfoundation.org/data-equity-principles?inline> (viewed Mar 2023).
- 17 Health Data Principles. Transform health. <https://www.healthdatapinciples.org/principles> (viewed Dec 2022).
- 18 3-D Commission. Data, social determinants, and better decision-making for health: the report of the 3-D Commission. Boston: 3-D Commission, 2021. <https://3dcommission.health/report> (viewed Mar 2023).
- 19 Lovett R, Lee V, Kukutai T, et al. Good data practices for Indigenous data sovereignty and governance. In: Daly A, Devvit SK, Mann M; editors. *Good Data*. Amsterdam: Institute of Network Cultures, 2019; pp. 26-36.
- 20 Indigenous Data Network. Public submission: Office of the National Data Commissioner. Melbourne: IDN, 2022. <https://www.datacommissioner.gov.au/sites/default/files/2022-12/35%20-%20Indigenous%20Data%20Network%20Data%20Code%202022%20Submission.pdf> (viewed Mar 2023).
- 21 National Association of People with HIV Australia and Positive Life NSW. Submission to the new Australian Government data sharing and release legislation issues paper for consultation. Sydney: NAPWHA, 2018. <https://www.positivelife.org.au/wp-content/uploads/2020/10/plnsw-sb2018-data-sharing-release-legislation.pdf> (viewed Mar 2023).
- 22 Carroll C, Evans K, Elmusharaf K, et al. A review of the inclusion of equity stratifiers for the measurement of health inequalities within health and social care data collections in Ireland. *BMC Public Health* 2021; 21: 1705.
- 23 Sadana R, Harper S. Data systems linking social determinants of health with health outcomes: advancing public goods to support research and evidence-based policy and programs. *Public Health Rep* 2011; 126 (Suppl): 6-13.
- 24 Whitehead M. The concepts and principles of equity and health. *Int J Health Ser* 1992; 22: 429-445.
- 25 Micheli M, Ponti M, Craglia M, et al. Emerging models of data governance in the age of datafication. *Big Data and Society* 2020; 2: 1-15. ■