

# Ethical considerations regarding the use of pooled data from electronic health records in general practice

Kylie Vuong, Rowena Ivers,  
Sally Hall Dykgraaf, Michael Nixon,  
Gail Roberts, Siaw-Teng Liaw

## Background

It is important that research conducted in general practice is ethical and technically robust. Electronic health records (EHRs) have high potential to contribute to research and policy that benefits both individual patients and the broader community. General practitioners are increasingly approached to make EHRs available for research, but they also have ethical and legal obligations to ensure personal information remains confidential and is handled safely.

## Objective

The aim of this article is to discuss the ethical considerations associated with the pooling of general practice EHRs for research.

## Discussion

There are associated benefits and potential risks to patients and practices when using pooled EHR data for research. Careful consideration and judgement are required to ensure ethical requirements are met, with particular attention to informed consent, data quality and confidentiality.

**MOST GENERAL PRACTICES** in Australia use electronic health record (EHR) systems to manage patient information. The range of variables recorded in the EHR may include patient demographics, pre-existing conditions, immunisations, pathology results, prescriptions, reasons for encounters and clinical assessments. General practitioners (GPs) have a role as custodians of EHR data.<sup>1</sup> While the primary use of EHRs is to inform and enhance patient care, their detailed longitudinal nature, recorded in electronic format during routine interactions at the point of care, makes the secondary use of EHR data an attractive data source for researchers and policy makers.<sup>2,3</sup>

Pooled data from general practice EHRs provide an insight into general practice workload and the delivery of clinical care, with large sample sizes. The comprehensive nature of general practice enables the study of outcomes, including those that may be difficult to capture in other settings, from across the entire patient journey. The data collection costs associated with using pooled EHRs are also less than those of other data collection methods, particularly methods involving individual consent.<sup>4,5</sup> However, there are difficulties related to data quality and,<sup>6</sup> especially for pooled data, interoperability.<sup>7</sup>

In Australia, as in the UK and USA, there has been substantial investment in

pooling EHRs.<sup>5</sup> For example, in 2011, National Prescribing Service (NPS) MedicineInsight was funded by the Australian Government to establish a national general practice database. It de-identifies general practice EHRs at the source, then extracts longitudinal data from general practice information systems using third-party extraction tools such as the GeneRic Health Network Information for the Enterprise (GRHANITE)<sup>8</sup> and cdmNET.<sup>9</sup> By October 2018 NPS MedicineInsight had grown to include 662 participating general practices and 2.3 million regular patients.<sup>10</sup> It has since been used in a range of research and quality improvement projects, ranging from pharmaco-epidemiological studies to large cohort studies.<sup>11,12</sup> In another example, the Lumos program – a collaboration between NSW Health, Primary Health Networks and general practices – de-identified data from general practice EHRs that are linked to other New South Wales health and registry data.<sup>13</sup> From 1 August 2019, many general practices in Australia have been submitting de-identified data to Practice Incentives Program Eligible Data Sets with their local Primary Health Networks,<sup>14,15</sup> and may be familiar with the extraction process.

Enhanced system capabilities have allowed the types of research studies possible with EHR data to include opportunities for pragmatic interventional

studies with randomisation within the database at the point of care.<sup>4,5</sup> For example, when a potentially eligible patient visits, the GP may see a computer-generated flag on the screen linked to the participant information statement and consent form.<sup>16,17</sup> In these interventional studies, consenting patients may be randomised by the trial to different treatment groups, and outcomes may be collected from the EHRs.<sup>17</sup> There are also other opportunities for using EHRs to facilitate recruitment by identifying study eligibility criteria in the EHR, and for data collection, by linking questionnaires to EHRs to provide additional data or adding additional coded fields to the EHR to evaluate outcomes.<sup>5</sup>

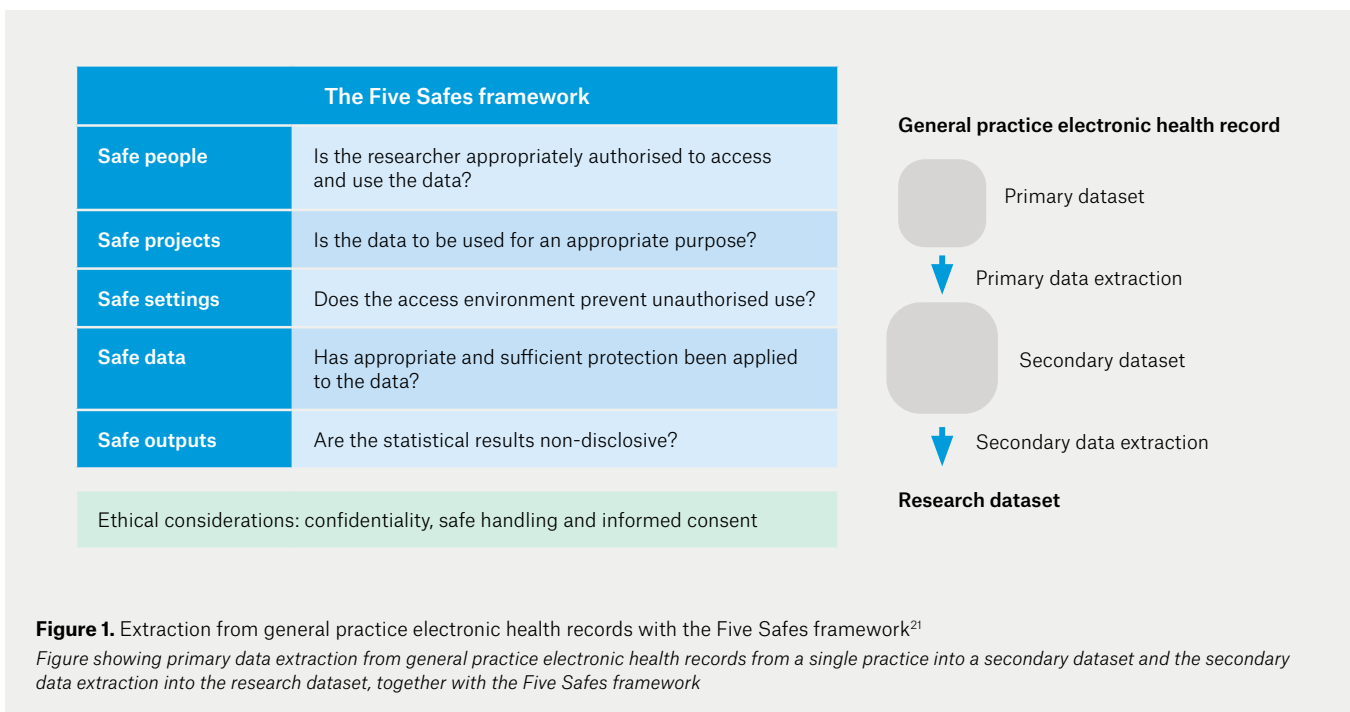
While some GPs have a dual role as both the custodian of EHR data and researcher, increasingly, practices are approached to make EHRs available for research and quality improvement. GPs have ethical as well as legal responsibilities to ensure patient medical records are confidential, stored and managed securely, and fit for purpose.<sup>18</sup> In this article, we aim to discuss the ethical considerations associated with the use of general practice EHRs, individually and pooled, for research.

In Australia, standards for the conduct of human research, including the use of health records and other personal materials, are set by the *National statement on ethical conduct of human research* (National Statement).<sup>19</sup> The National Statement encourages research that is based on ‘the values of respect, research merit and integrity, justice, and beneficence’, with careful consideration, judgement and appraisal of its unique context.<sup>19</sup> The potential benefit of the research needs to be greater than its risks for the research to proceed ethically.<sup>20</sup> The National Statement stipulates careful consideration of the benefits and risks of research as well as informed consent, and we will use it to guide our discussion.

**Balancing the potential benefits and risks**

Risk is ‘a potential for harm, discomfort or inconvenience’.<sup>19</sup> In using pooled EHRs for research, the risks to patients and GPs include potential psychological, social, economic and legal harms that may be associated with breaches of confidentiality when a person or organisation is re-identified.

The Five Safes framework has been adopted by several Australian and international agencies as a multidimensional structure to minimise risk (Figure 1).<sup>21</sup> Researchers should have the knowledge and skills to use the data in an appropriate manner (safe people). The research project should deliver public benefits, with its design meeting the stated objectives (safe project). Procedures to minimise the risk of potential breaches of confidentiality to patient participants in EHR-based research may include the removal of patient identifiers at source, encryption, and allocation of a unique identifier before data storage (safe data). High levels of security should be ensured through data management standards and infrastructure, and restriction of lower-level geographical information is important to prevent unintended identification (safe setting). The collection, storage and handling of EHR data should be in accordance with Australian privacy legislation.<sup>22,23</sup> Human research ethics committees (HRECs) routinely review storage plans for de-identified data, considering security, legislation, data backup, period



of storage and plans for destruction of the data at conclusion of the project. Final checks should also be conducted on all research outputs to further minimise confidentiality risks (safe outputs).

However, in some instances, confidentiality concerns remain because it may be possible to re-identify participants or link participant information from different sources, particularly in small samples or in cases with rare conditions. The confidentiality of the patient, GP, general practice staff and the general practice needs to be protected at all times. HRECs consider situations where there is potential for misuse of data – for example, with vulnerable people who ‘may have an increased likelihood of being wronged or of incurring additional harm’,<sup>24,25</sup> such as children or people with impaired intellectual function. It is important that analyses of pooled EHR data are not used or presented in ways that might exacerbate inequalities or perpetuate social injustice – risks that have been identified by other uses of ‘big data’, especially in commercial settings.<sup>26</sup>

As large repositories of real-world data from EHRs become increasingly used, the FAIR Guiding Principles have been developed to promote best practice among data custodians. Data should be Findable, Accessible, Interoperable and Reusable to promote transparency and reproducibility of data-driven research.<sup>7</sup> Additionally, data custodians, whether a GP with a single EHR system or large aggregated data systems such as NPS MedicineInsight, have a responsibility to ensure that the data are fit for purpose; that is, that they meet benchmarks for completeness, conformance and plausibility, and that they comply with the research merit and integrity requirement.<sup>27,28</sup>

### Informed consent

In most research, participant consent is voluntary and based on sufficient information about the research, including potential benefits, risks and impacts on participants such as costs to patients or increased workload to clinicians, with opportunities for participants to ask questions and discuss the information

together with others if they wish to.<sup>19</sup> Participant information should be conveyed in a timely, meaningful and accessible way to meet the needs of a diverse Australian population, encompassing participants, including patients, providers and general practices. People do not need to provide any reason for non-participation. Participants are entitled to withdraw from research at any stage without any consequences.

However, in large-scale projects using pooled EHR data, seeking such explicit consent may not be feasible or practical. In these circumstances, for low-risk research with a high potential benefit, an opt-out approach or a waiver of consent for research may be considered by an HREC. Opt-out consent models require that reasonable attempts are made to explain the research in a way that is timely, meaningful and accessible, with opportunities to decline or withdraw participation, and an adequate plan to protect confidentiality. An opt-out approach must also align with State and Territory, Commonwealth and international laws.<sup>19</sup> Participating general practices should display waiting room posters with information on using EHRs for research, contacts for further information and instructions to decline or withdraw by completing an opt-out form that is available from the practice reception or online.<sup>29,30</sup> It is possible to remove de-identified data retrospectively if consent is withdrawn. In both explicit and presumed consent with opt-out approaches, consent is an active process with three main components, including capacity, adequate information and opportunities to consider the information.<sup>31</sup>

### Insurance

GPs may wish to review what data is being extracted from the practice EHRs and to consider opt-out procedures for individual patients. GPs should check their personal and practice medical indemnity to ensure coverage for research participation. Research projects that are sponsored by universities or other external agencies often

carry additional insurance. It is prudent to review any insurance arrangements in the formal agreement to participate in the research.

### Conclusion

There are benefits of using EHRs for research as well as potential risks to patients, GPs and the practices. GPs, in their role as custodian of EHR data, should ensure there are processes for informing patients – in a timely, meaningful and accessible way – about pooling EHRs, its uses and opt-out procedures (Box 1). Careful consideration and judgement are required to ensure ethical requirements are met, with particular attention to confidentiality, data quality and informed consent.

### Key points

- Pooled general practice EHR data provide an insight into general practice workload and the delivery of clinical care, and have high potential to contribute to research and policy that is relevant to the community.
- GPs have ethical as well as legal responsibilities to ensure patient health records are confidential, including that they are handled safely, with

#### Box 1. Issues for general practitioners to consider when approached by researchers with a request for electronic health record (EHR) data

1. Is there ethics approval from a National Health and Medical Research Council-approved ethics committee?
2. To what extent is the ethics approval dependent on an opt-in or opt-out approach to informed consent?
3. How will the EHR data be used and handled?
4. How will patient, general practitioner, general practice staff and practice confidentiality be ensured?
5. Is there adequate insurance coverage for research participation?
6. Does the research team have the skills to undertake this research?
7. Are the potential benefits of the research greater than the risks?

good governance and provenance arrangements for secondary use.

- GPs should ensure there are processes for informing patients about pooling EHRs, their uses and opt-out procedures.
- Careful consideration and judgement are required to ensure ethical requirements are met, with particular attention to confidentiality and informed consent.

## Authors

Kylie Vuong FRACGP, PhD, MPH, Associate Professor in Primary Care and Academic Lead in General Practice, School of Medicine and Dentistry, Griffith University, Qld

Rowena Ivers FRACGP, PhD, MPH, Academic Leader, Community Based Health Education, Graduate School of Medicine, University Of Wollongong, NSW  
Sally Hall Dykgraaf RN, PhD, Senior Lecturer & Research Lead, Rural Clinical School, College of Health & Medicine, Australian National University, ACT  
Michael Nixon OAM, Th Dip, MBBS, FRACGP, Senior Medical Officer, Galangoor Duwalam Primary Healthcare Service, Hervey Bay, Qld

Gail Roberts RN, MA, Senior Project Officer – Research and Ethics, The Royal Australian College of General Practitioners, East Melbourne, Vic

Siaw-Teng Liaw PhD, FRACGP, FIAHSI, FACMI, FAIDH, Emeritus Professor & Head, WHO Collaborating Centre (eHealth), UNSW, Sydney, NSW

Competing interests: KV, RI, SHD, MN and STL receive sitting fees as members of The Royal Australian College of General Practitioners (RACGP) National Research and Evaluation Ethics Committee. RI is a member of the *Australian Journal of General Practice* Editorial Advisory Committee, NSW Council and RACGP Expert Committee – Quality Care.

Funding: None.

Provenance and peer review: Not commissioned, externally peer reviewed.

**Correspondence to:**  
k.vuong@griffith.edu.au

## References

- Hodgkins AJ, Mullan J, Mayne DJ, Boyages CS, Bonney A. Australian general practitioners' attitudes to the extraction of research data from electronic health records. *Aust J Gen Pract* 2020;49(3):145–50. doi: 10.31128/AJGP-07-19-5024.
- Häyrynen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *Int J Med Inform* 2008;77(5):291–304. doi: 10.1016/j.ijmedinf.2007.09.001.
- Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol* 2017;106(1):1–9. doi: 10.1007/s00392-016-1025-6.
- Muller S. Electronic medical records: The way forward for primary care research? *Fam Pract* 2014;31(2):127–29. doi: 10.1093/fampra/cmu009.
- Denaxas SC, Morley KI. Big biomedical data and cardiovascular disease research: Opportunities and challenges. *Eur Heart J Open* 2015;1(1):9–16. doi: 10.1093/ehjqcco/qcv005.
- Britt H, Miller G, Bayram C. The quality of data on general practice – A discussion of BEACH reliability and validity. *Aust Fam Physician* 2007;36(1–2):36–40.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016;3:160018. doi: 10.1038/sdata.2016.18.
- University of Melbourne. GRHANITE™ Health Informatics Unit. Carlton, Vic: Health and Biomedical Informatics Centre, 2011. Available at [www.grhanite.com/about/background.html](http://www.grhanite.com/about/background.html) [Accessed 2 August 2021].
- Precedence Health Care. Digital Health Solutions for Integrated Care. Melbourne, Vic: Precedence Health Care, 2021. Available at <https://precedencehealthcare.com> [Accessed 2 August 2021].
- Busingye D, Gianacas C, Pollack A, et al. Data resource profile: MedicinesInsight, an Australian national primary health care database. *Int J Epidemiol* 2019;48(6):1741–1741h. doi: 10.1093/ije/dyz147.
- Schaffer AL, Busingye D, Chidwick K, Brett J, Blogg S. Pregabalin prescribing patterns in Australian general practice, 2012–2018: A cross-sectional study. *BJGP Open* 2021;5(1):bjgpopen20X101120. doi: 10.3399/bjgpopen20X101120.
- Havard A, Manski-Nankervis JA, Thistlethwaite J, et al. Validity of algorithms for identifying five chronic conditions in MedicinesInsight, an Australian national general practice database. *BMC Health Serv Res* 2021;21(1):551. doi: 10.1186/s12913-021-06593-z.
- Correll P, Feyer A-M, Phan P-T, et al. Lumos: A statewide linkage programme in Australia integrating general practice data to guide system redesign. *Integrated Healthcare Journal* 2021;3(1):e000074. doi: 10.1136/ihj-2021-000074.
- Australian Government Department of Health. Practice Incentives Program Quality Improvement Incentive guidelines. Canberra, ACT: DoH, 2019.
- Australian Government Department of Health. Practice Incentives Program Eligible Data Set Data Governance Framework. Canberra, ACT: DoH, 2019.
- van Staa T-P, Goldacre B, Gulliford M, et al. Randomised Evaluations of Accepted Choices in Treatment (REACT) trials: Large-scale pragmatic trials within databases of routinely collected electronic healthcare records. *Trials* 2011;12(Suppl 1):A104–A104. doi: 10.1186/1745-6215-12-S1-A104.
- van Staa TP, Klungel O, Smeeth L. Use of electronic healthcare records in large-scale simple randomized trials at the point of care for the documentation of value-based medicine. *J Intern Med* 2014;275(6):562–69. doi: 10.1111/joim.12211.
- Avant Mutual. Medical records. Sydney, NSW: Avant Mutual, 2019. Available at [www.avant.org.au/Resources/Public/Medical-records-the-essentials](http://www.avant.org.au/Resources/Public/Medical-records-the-essentials) [Accessed 2 August 2021].
- National Health and Medical Research Council, the Australian Research Council and Universities Australia. National statement on the ethical conduct of human research. Canberra, ACT: Commonwealth of Australia, 2018.
- Liaw ST, Tam CW. Ethical research or research ethics? *Aust Fam Physician* 2015;44(7):522–23.
- Australian Bureau of Statistics. Five Safes framework. Belconnen, ACT: ABS, 2021. Available at [www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1160.0Main%20Features4Aug%202017](http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1160.0Main%20Features4Aug%202017) [Accessed 10 August 2021].
- Office of the Australian Information Commissioner. Australian privacy principles. Sydney, NSW: OAIC, 2021. Available at [www.oaic.gov.au/privacy/australian-privacy-principles](http://www.oaic.gov.au/privacy/australian-privacy-principles) [Accessed 10 August 2021].
- The Royal Australian College of General Practitioners. Privacy and managing health information in general practice. East Melbourne, Vic: RACGP, 2017.
- Council for International Organizations of Medical Sciences. International ethical guidelines for health-related research involving humans. Geneva, CH: CIOMS, 2016.
- World Medical Association. WMA Declaration of Helsinki – Ethical principles for medical research involving human subjects. Ferney-Voltaire, FR: WMA, 2018. Available at [www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects](http://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects) [Accessed 14 October 2021].
- O'Neil C. Weapons of math destruction. New York, NY: Crown Publishing Group, 2016.
- Liaw ST, Guo JGN, Ansari S, et al. Quality assessment of real-world data repositories across the data life cycle: A literature review. *J Am Med Inform Assoc* 2021;28(7):1591–99. doi: 10.1093/jamia/ocaa340.
- Kahn MG, Brown JS, Chun AT, et al. Transparent reporting of data quality in distributed data networks. *EGEMS (Wash DC)* 2015;3(1):1052. doi: 10.13063/2327-9214.1052.
- NPS MedicineWise. Improving patient health outcomes. Surry Hills, NSW: NPS MedicineWise, 2019. Available at [www.nps.org.au/assets/NPS1785e\\_PatientNotification\\_A3\\_v5-font-change.pdf](http://www.nps.org.au/assets/NPS1785e_PatientNotification_A3_v5-font-change.pdf) [Accessed 14 October 2021].
- South Eastern Melbourne Primary Health Network. POLAR. Heatherdale, Vic: SEM PHN, [date unknown]. Available at [www.semphn.org.au/resources/digital-health/polar.html](http://www.semphn.org.au/resources/digital-health/polar.html) [Accessed 21 October 2021].
- Prabhu PK. Is presumed consent an ethically acceptable way of obtaining organs for transplant? *J Intensive Care Soc* 2019;20(2):92–97. doi: 10.1177/1751143718777171.

correspondence [ajgp@racgp.org.au](mailto:ajgp@racgp.org.au)