REVIEW

'Encore, encore!' – secondary use of health data for research: A practical guide to POPIA

A Edgcumbe, LLM; **D Thaldar,** PhD

School of Law, University of KwaZulu-Natal, Durban, South Africa

***Corresponding author**: A Edgcumbe (edgcumbea@ukzn.ac.za)*

Health data hold immense potential for advancing medical science and informing public health strategies; however, their secondary use for research purposes has been poorly exploited, partly due to uncertainty and compliance concerns brought about by the Protection of Personal Information Act (POPIA). Researchers often assume that re-consent is the only option for secondary use. Yet POPIA provides several lawful grounds for processing special personal information, which includes health data, without necessarily obtaining re-consent. The present article serves as a practical guide for researchers navigating POPIA's requirements for the secondary use of health data for research. A systematic approach is recommended: first, researchers should establish whether the original consent encompasses the intended secondary use. If it does not, they should determine which lawful grounds for processing special personal information best suit their context and can be relied upon. Options include obtaining re-consent, relying on a research purpose, or seeking authorisation from the Information Regulator. In addition, researchers may consider de-identifying the dataset, which would exempt it from the ambit of POPIA. Situations where pseudonymisation might achieve a similar outcome are also unpacked. By taking proactive, well-documented steps, researchers can better capitalise on existing health datasets to advance vital health research in South Africa.

S Afr J Bioethics Law 2025;18(3):e3350. https://doi.org/10.7196/SAJBL.2025.v18i3.3350

Picture a South African research group at the forefront of cutting-edge health research. Years of meticulously collected data on patient outcomes, disease patterns and genetic predispositions have created a treasure trove of information capable of answering critical questions, informing public health strategies, and driving innovation in healthcare. Yet, despite the value of these datasets, the researchers face a significant hurdle. Can data collected for one study now be re-used for secondary research without violating South Africa's Protection of Personal Information Act (POPIA)?[1]

These concerns are not unfounded. POPIA imposes stringent requirements on processing special personal information, which includes health data. The research group grapples with questions, such as: What constitutes 'specific' consent under POPIA? Is re-consent required? Can strategies like de-identification or pseudonymisation help to bypass certain legal obligations? Without clear answers, ground-breaking research risks being stalled.

Ethical oversight of health research in South Africa is decentralised, with institutional Health Research Ethics Committees (HRECs) having the discretion to develop and amend their own guidelines, as permitted under Section 73(2)(a) of the National Health Act 61 of 2003 (NHA). While these committees operate under the broad guidance of the National Health Research Ethics Council (NHREC), their standards are not necessarily uniform and can vary across institutions. In contrast, the legal framework governing the secondary use of health data is uniform across South Africa and does not allow for institutional-level variation. However, given the relative novelty of POPIA, much of the existing literature remains focused on conceptual discussions of its application to health data, often addressing only isolated aspects. This article seeks to fill that gap by providing a comprehensive analysis of the legal provisions relevant to the secondary use of health data and, crucially, offering practical guidance to help researchers navigate these requirements effectively.

Conceptual clarity: Health data under POPIA

POPIA establishes two governance levels for managing personal data: one for personal information and an additional, stricter level for special personal information. Health data fall under both. This dual classification requires researchers to navigate both tiers of compliance effectively.[2-4]

The secondary use of health data is classified as 'further processing' in POPIA. Further processing is any operation or activity concerning personal information for purposes **other** than those for which it was **initially** collected.[5]

For instance, a clinic collects patients' personal information to provide and manage their patients' medical treatment. Later, the clinic decides to use this information for a secondary purpose: to determine the most common health complaints at the clinic. As the new use (determining the most common ailments) differs from the original purpose for which the personal information was initially collected (managing patients' treatment), this secondary use constitutes **further** processing.

At the basic level, POPIA allows for the secondary use of personal information so long as the secondary use is compatible with the 'specific, explicitly defined and lawful purpose' (section 13) for which it was initially collected (section 15). Where the information is used for 'historical, statistical or research purposes and the responsible party ensures that the further processing is carried out solely for such

purposes and will not be published in an identifiable form' (section 15(3)(e)), such secondary use will **not** be deemed **incompatible** with the original purpose.

In other words, this provision allows researchers to utilise existing datasets containing personal information for research purposes without needing to re-obtain the participants' consent where:

(i) the personal information was originally collected for a specific, explicitly defined and lawful purpose, and

(ii) the responsible party ensures:

a. the secondary use is carried out solely for research purposes, and

b. it will not be published in an identifiable form.

This approach is often referred to as POPIA's 'research exception'. However, it is important to emphasise that this exception does not override the stricter requirements of section 27, which governs the processing of **special** personal information. In the context of health research, participant data, in most cases, qualify as special personal information.[5] This means that researchers will typically need to comply with the more stringent provisions applicable to special personal information.

Firstly, as a starting point, section 26 provides that the processing of special personal information is **prohibited** except in the limited circumstances outlined in section 27. Notably, POPIA does not provide for the secondary use of special personal information, so researchers will need to rely on one of the lawful grounds listed in section 27 to process the health data as special personal information for secondary use. Re-consenting participants is one option; however, it may not always be practical or feasible, in which case researchers may rely on another lawful ground specified in section 27.

It is a common misconception that POPIA mandates consent for all processing activities. In reality, POPIA provides researchers with additional legally valid grounds for processing special personal information beyond consent, offering flexibility in complying with the Act and, thereby, enabling critical research activities to continue.

Navigating POPIA's requirements for secondary use

As health data used in research are generally classified as special personal information, it is essential that researchers navigate the stricter requirements applicable to such data. To achieve compliance for the secondary use of health data for research purposes, researchers should follow a systematic process (outlined in Flowchart 1):

1. **Evaluate the original consent:** Determine whether the original consent encompasses the intended secondary use.
2. **Explore alternative lawful grounds under POPIA:** If the consent does not encompass the intended secondary use, consider alternative grounds such as:
   - obtaining re-consent
   - relying on a research purpose, or
   - seeking authorisation from the Information Regulator.
3. **Consider de-identification and pseudonymisation:** Explore whether de-identifying the dataset can exempt it from POPIA compliance. Additionally, assess whether pseudonymisation may, in certain limited cases, achieve a similar outcome. (The precise meanings of 'de-identification' and 'pseudonymisation' are addressed in paragraphs 3.1 and 3.2 below.)

This article examines each step and, in particular, how POPIA applies to health data as special personal information. First, we begin by clarifying the requirements for consent and exploring alternative grounds for lawful processing under POPIA. Next, we analyse the conditions under which data may be exempted from POPIA's provisions through de-identification. Finally, we consider the legal position of pseudonymised data.

Options for POPIA compliance: A step-by-step approach to the legal processing of health data as special personal information

1. Assess existing consent

Section 27 (1)(a) provides that the prohibition on the processing of special personal information does not apply if the 'processing is carried out with the consent of a data subject'. Therefore, the starting point for a researcher intending to use health data for a secondary research purpose is to assess whether the initial consent obtained allows for the intended secondary use. If the original consent explicitly permits the secondary use, this serves as a valid legal basis for the use of special personal information under section 27(1)(a), allowing the researcher to proceed. However, if the original consent does not extend to the new purpose, the researcher must explore alternative lawful grounds for the secondary use of health data under POPIA.

For example, a university researcher wants to study trends in diabetes management using patient records from a local clinic. If the clinic originally obtained patient consent stating that the patient's personal information could be used for future medical research, this consent may allow the researcher to process the health data under section 27(1)(a) of POPIA. However, if the original consent was only for direct patient care, the researcher would need to re-consent the patient or identify another lawful basis for secondary use. (These are discussed below.)

2. Explore lawful grounds for secondary use under POPIA

If the initial consent did not include consent for the intended secondary use, researchers may rely on several other lawful grounds for processing special personal information. The most relevant options for researchers are obtaining re-consent, relying on a research purpose, or seeking approval from the Information Regulator.

2.1 Re-consent

If re-consent is successfully obtained, it serves as a valid legal basis for the secondary use of data under section 27(1)(a). However, researchers should assess the feasibility of re-consenting participants as opinions on the appropriateness of re-consent vary.[6,7]

While an examination of this debate is beyond the scope of this article, it is important to note that researchers seeking to re-consent participants will need to weigh up the various ethical and practical considerations, including the time, cost and effort required to recontact participants, as well as the potential impact of refusals on the health data cohort, such as introducing selection bias, which may potentially compromise the validity of the dataset.[6,7]

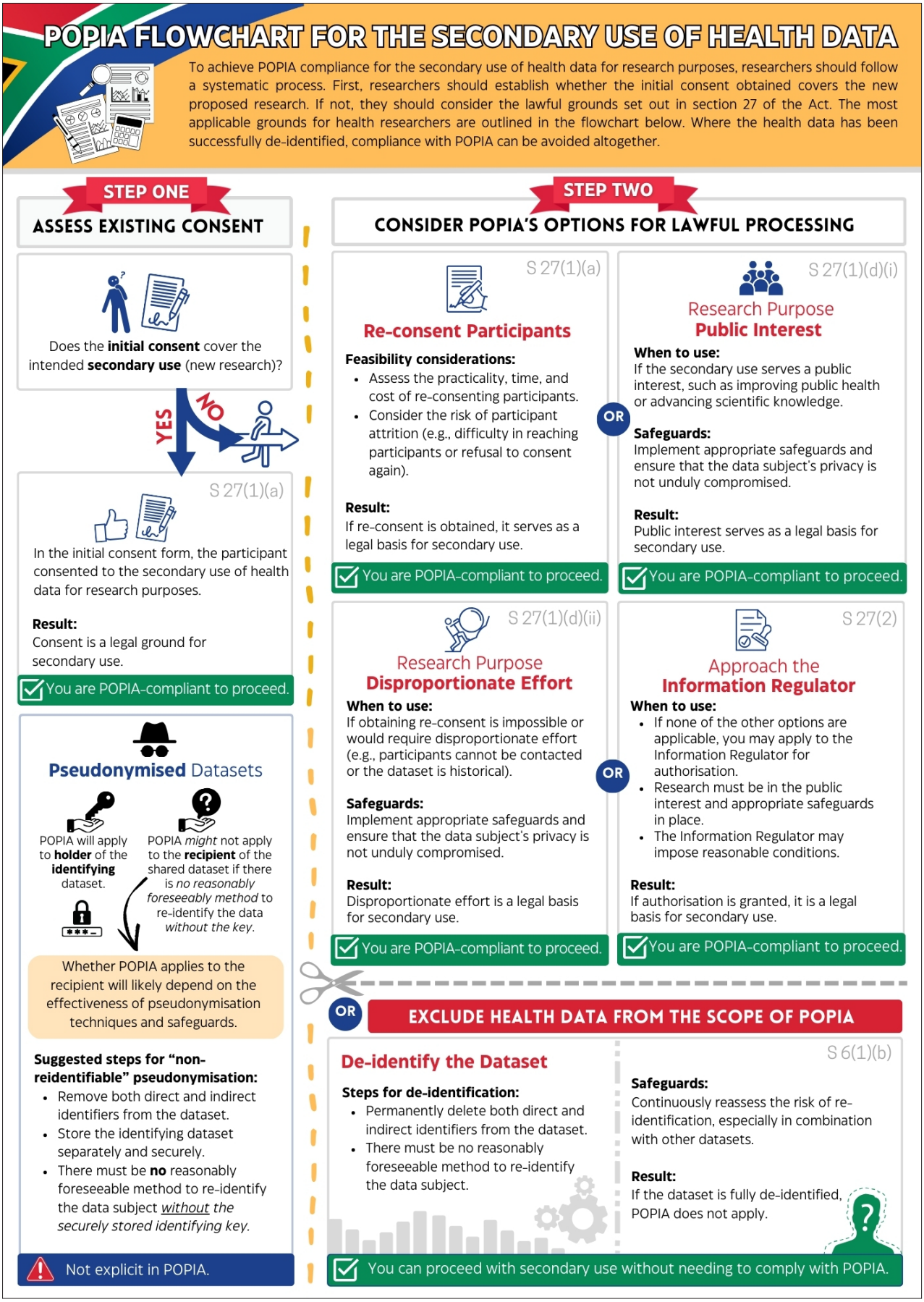

Fig. 1. POPIA flowchart for the secondary use of health data.

2.2. Research purposes

Section 27 permits the processing of special personal information **without** consent if the processing is conducted for, inter alia, **research** purposes, provided that:

(a) such research serves a public interest, and the processing of the personal information is essential for the research; **or**

(a) obtaining consent is impossible or would require disproportionate effort.

In cases (*a*) and (*b*) above, researchers must ensure that the processing does not disproportionately infringe on the privacy of the data subject and must provide 'sufficient guarantees' to protect individual privacy (section 27(1)(d)). This means that if researchers can demonstrate that the research serves a public interest **or** that obtaining consent is impossible or impractical, they can proceed **without** re-consenting participants, provided that robust privacy safeguards are implemented. These safeguards must align with generally accepted information security practices and industry-specific regulations (section 19), which may include measures such as pseudonymisation, encryption, regular audits, and access controls to prevent unauthorised disclosure.[8]

This approach aligns with the Department of Health's *South African Ethics in Health Research Guidelines: Principles, Processes and Structures*, 3rd edition, 2024 (NDoH 2024).[9] The Guidelines adopt the position that where re-consent is simply not feasible for the secondary use of data, and where the envisaged research is of 'important social value' – in other words, where it serves a 'public interest' – a Health Research Ethics Committee (HREC) may waive the requirement for reconsent, provided the research poses no more than **minimal risk of harm** to the research participants (NDoH 2024, 4.1.5 (e)). Similarly, where a link to the identifiers exists but is not accessible to the research team, reconsent is **not** required, so long as the research is unlikely to place any individual, family or community at social, psychological, legal or economic risk of harm (NDoH 2024, 4.1.5 (c)). This reflects the principle that, where research serves the public interest and strong safeguards are in place to mitigate the risk of harm, the ethical requirement for consent may justifiably be waived.

2.3. Other lawful grounds

POPIA also allows for the processing of special personal information that has been publicly disclosed by the data subject (section 27(1)(e)). Furthermore, when none of the options already outlined is feasible, researchers can apply for authorisation from the Information Regulator. The Information Regulator may permit the processing of special personal information, subject to reasonable conditions, where it is in the public interest and there are sufficient safeguards in place to protect the data subject's privacy (section 27(2) and (3)). However, there is currently little guidance or precedent on how this authorisation process operates in practice.

3. Take-homes for scientists seeking to be POPIA-compliant

As outlined, when the initial consent does not cover the new research purpose, researchers have several lawful alternatives under POPIA to process special personal information. Re-consenting participants provides a clear legal basis but may be impractical owing to cost or participant attrition, especially with historical data. In such cases, the disproportionate effort exception can be relied upon, provided that robust privacy safeguards are implemented. Public interest grounds, such as research aimed at improving public health or advancing knowledge,[10] may also justify secondary processing if strong privacy measures are in place. If these options are not applicable, researchers can seek authorisation from the Information Regulator.

These options are lawful and compliant pathways for the secondary use of health data; however, another option is to remove the data from POPIA's scope entirely through de-identification—a topic explored in the next section.

Options to exclude health data from POPIA

3.1. De-identification

One strategy to mitigate privacy risks and, at the same time, bypass POPIA's requirements is to de-identify the dataset.[11] This approach may be particularly beneficial when seeking to share health data with other researchers for secondary use, as it eliminates the need for ongoing compliance with POPIA's privacy provisions and so facilitates secondary use and sharing far more easily. Under section 1 of POPIA, 'de-identify' is defined as removing or altering any information that:

(a) identifies the data subject

(b) can be used or manipulated by a reasonably foreseeable method to identify the data subject

(c) can be linked by a reasonably foreseeable method to other information that identifies the data subject.

To fall outside the scope of POPIA, the information must be **irreversibly** de-identified; in other words, it must not be possible to re-identify it in future (section 6(1)(b)). Where a reasonably foreseeable method exists that can be used to re-identify the data subject, either directly or indirectly, the information continues to be subject to POPIA's provisions.

To help illustrate this point, consider the following example: A hospital researcher is studying HIV treatment outcomes using patient hospital records. To de-identify the dataset, a researcher removes names, identification (ID) numbers and contact details. However, the dataset still includes birth dates, geographic locations and rare medical conditions.

To comply with POPIA's de-identification standard, the researcher must:

- ensure all direct and indirect identifiers are either removed or sufficiently altered, and
- assess whether any existing method could be used to re-identify individuals, including linking it with another dataset.

In this example, a reasonable researcher would recognise that external datasets (e.g. voter rolls or criminal records) containing birth dates and locations could be used to match and re-identify individuals.[12] Therefore, the dataset would continue to fall within the ambit of POPIA.

It is important to note that this approach differs from the European Union's General Data Protection Regulation (GDPR), which incorporates a risk-based assessment that considers the likelihood of re-identification (Recital 26). In contrast, under POPIA, the assessment is not about the risk or likelihood of re-identification but rather (*i*) whether a method *exists* to re-identify the data and (*ii*) whether **a reasonable researcher** placed in the position of the researcher de-identifying the data would **reasonably** foresee that such a method exists.

In our example, a reasonable researcher would recognise that linking datasets is an existing method that **can** be used to re-identify the data subjects. Retaining details, such as date of birth and geographical details, may make this possible. As this constitutes a reasonably foreseeable method of re-identification, the dataset remains subject to POPIA's provisions.

Meeting POPIA's de-identification standards is not merely about whether someone **would** re-identify the dataset, but whether they **could**. Of course, the **likelihood** of an existing method being used may factor into a court's assessment of **reasonable foreseeability** – but this has not yet been judicially settled.

Therefore, to remain compliant, the researcher should stay informed about evolving de-identification and re-identification techniques. Documenting the chosen de-identification approach and justifying why it meets POPIA's standards is also critical. This is important because the Act does not prescribe what de-identification techniques must be employed; POPIA is chiefly concerned with the outcome: data must be de-identified to the extent that they **cannot** be re-identified. Researchers are tasked with selecting the most appropriate methods based on the specific use case and context. Where data are successfully and permanently de-identified, they are removed from the ambit of POPIA, allowing them to be more easily shared and utilised for secondary use.

However, the greatest challenge for researchers is not just choosing the method but maintaining the utility of de-identified data,[3] particularly when working with high-dimensional health data.[13] There may also be a need to retain the original dataset, for instance where hospital records are used or where it is necessary to recontact the research participant. Where de-identification is not a realistic – or desirable – option, researchers may want to share pseudonymised datasets as an alternative. But what does POPIA have to say about this?

3.2. Pseudonymisation

First, the term 'pseudonymisation' does not appear in POPIA. This is unfortunate, as it is a technique widely used by researchers to protect the identity of participants. Pseudonymisation involves replacing identifying information in a dataset with a code (a pseudonym), while maintaining a separate dataset (the key) that links the original identifiers to their codes. This technique enhances privacy while allowing re-identification when necessary. For this reason, it is frequently applied in the early stages of clinical trials[16] and in longitudinal studies, where maintaining a connection to the original identifiers is essential.[17]

Recognising its importance, the Academy of Science of South Africa (ASSAf) provides guidance on pseudonymisation in its POPIA Compliance Framework for Researchers and Research Institutions (POPIA Compliance Framework). The Framework identifies pseudonymisation as the 'default' approach in high-risk research and requires researchers to document reasons if it is not used. In the Framework, pseudonymisation is defined as:

> Pseudonymisation means that personal information is processed in such a way that the personal information can no longer be attributed to a specific research participant without the use of additional information, provided that such additional information is kept separately, confidential and secure from unauthorised access.[14]

This definition closely mirrors that of the European Union's GDPR.[15] However, because POPIA does not expressly regulate pseudonymisation, its application to pseudonymised datasets must be interpreted with care. What is clear is that under POPIA, pseudonymised health data remain classified as 'special personal information' as long as the institution retains the key. This is because the key enables re-identification, creating a 'reasonably foreseeable method' of linking the data to individuals and keeping them within POPIA's definition of 'personal information'. Consequently, pseudonymised data remain subject to the full scope of POPIA's requirements and protections – at least in the hands of the research institution. The ASSAf Framework adopts the same position: pseudonymisation is a crucial privacy safeguard, but it does not remove health data from POPIA's scope. The legal position of pseudonymised datasets becomes more complex, however, in the context of data sharing – an issue we address next.

3.3. Sharing pseudonymised datasets

A critical legal question arises when a research institution shares a pseudonymised dataset with another institution while withholding the key: Is the dataset still 'personal information' in the hands of the recipient institution? The answer is not straightforward.

Under POPIA, identifiability depends on context: a dataset may constitute personal information for one institution but not for another. For instance, if a pseudonymised dataset is transferred to a recipient institution that has no reasonably foreseeable means of re-identifying individuals and no access to the key, it may be classified as 'de-identified' information under POPIA for that institution. However, this determination hinges on various factors, including whether the recipient institution could obtain additional information that would enable re-identification.

As mentioned earlier, POPIA does not mandate the use of a specific de-identification method but instead focuses on the outcome: personal information must be 'de-identified to the extent that it cannot be re-identified again' (section 6). Institutions may use various techniques, including pseudonymisation, to achieve this standard of non-identifiability. However, the test is stringent: the pseudonymisation must eliminate the possibility of re-identification by any reasonably foreseeable method for the data to fall outside POPIA's scope. Moreover, the question is not whether the recipient **would** attempt to re-identify individuals – but rather whether they **could**.

Therefore, whether the pseudonymised dataset falls outside the scope of POPIA for a **recipient** depends on whether the pseudonymisation rendered the data incapable of re-identification by any reasonably foreseeable method. The answer determines one of two outcomes:

- If the recipient lacks the key and any reasonably foreseeable means of re-identification, the dataset is non-personal information in their hands, and POPIA does not apply.
- If the recipient can access the key or re-identify individuals through other methods, such as linking with external datasets, the dataset remains personal information in their hands, and POPIA continues to apply.

Additionally, for a pseudonymised dataset to fall outside POPIA's scope, the institution providing the dataset must implement appropriate safeguards to protect both the identifiable data it retains

and the key. POPIA requires responsible parties to ensure adequate security measures against unlawful access to personal information (section 19). Any method or safeguard must, therefore, be sufficiently robust to prevent re-identification through unauthorised access or other means. The context will determine what safeguards are appropriate and effective.

This may be a promising avenue for researchers who still need to maintain an original dataset; however, it must be emphasised that there is currently no judicial or regulatory guidance on how POPIA applies to pseudonymised data. The legal position, therefore, remains unsettled.

3.4. Take-homes: de-identification and pseudonymisation

De-identification and pseudonymisation are valuable strategies for enabling the secondary use of health data. De-identification removes a dataset from the scope of POPIA by ensuring that no party has a reasonably foreseeable method of re-identification. If properly executed, de-identification renders the information non-personal and outside POPIA's regulatory framework.

Pseudonymisation, on the other hand, does not automatically achieve this outcome. A context-specific approach is crucial for correctly classifying and managing pseudonymised data. To exclude a dataset from POPIA's scope, the pseudonymisation must be sufficiently robust to eliminate any foreseeable means of re-identification. If an institution retains the re-identification key, the dataset remains personal information in its hands. For a recipient institution, the dataset is only considered de-identified non-personal information if the recipient lacks a reasonably foreseeable method of re-identification.

However, given the lack of judicial or regulatory guidance on pseudonymisation in South Africa, institutions should take a cautious approach when sharing pseudonymised datasets. The following recommendations can help ensure compliance with POPIA while facilitating responsible data sharing:

1. **Clearly document the de-identification process and pseudonymisation techniques**
   Institutions should keep detailed records of the techniques applied, the rationale for their use, and any assessments conducted to evaluate their effectiveness in preventing re-identification.
2. **Assess the recipient's ability to re-identify the data**
   Before transferring a pseudonymised dataset, institutions should conduct a context-specific assessment of whether the recipient has a reasonably foreseeable method to re-identify data subjects, including access to external datasets or other methods.
3. **Establish clear transfer conditions**
   Data-sharing agreements should state:
   - The recipient does not have access to the key.
   - The recipient is prohibited from attempting re-identification.
   - The dataset is provided in a manner that prevents foreseeable re-identification.
4. **Balance data utility against compliance requirements if necessary, identify a lawful ground for secondary use**
   Institutions must carefully balance maintaining data utility with meeting regulatory requirements. If de-identification significantly compromises data usability, researchers should explore alternative lawful grounds for processing special personal information under section 27 of POPIA.

Effective de-identification and pseudonymisation strategies enable the responsible secondary use of health data, while eliminating the need to comply with POPIA's stringent provisions for data sharing. Careful consideration of all the implications of de-identifying or pseudonymising health data is essential for researchers seeking to employ these strategies for the secondary use of health data. Given the evolving regulatory landscape, institutions should adopt a cautious approach, ensuring that their data protection practices align with both local requirements and international best practices.

Conclusion

The secondary use of health data presents immense opportunities for advancing research and improving public health. However, navigating POPIA's legal requirements is essential to ensure compliance while maximising the utility of valuable datasets. This article provides a structured roadmap for researchers, outlining key legal pathways, including assessing original consent, exploring alternative lawful grounds, and implementing privacy-enhancing measures such as pseudonymisation and de-identification.

A crucial takeaway is that re-consent is not the only option. While often assumed to be the default, POPIA provides alternative lawful grounds that acknowledge the realities of long-term research and the impracticality of tracking down participants years later. Researchers must make strategic choices: if secondary use is necessary and re-consent is impractical, they can either comply with POPIA's requirements by identifying a lawful ground for the further processing or remove it from the Act's scope by fully de-identifying the dataset. Pseudonymisation is also a promising option, provided it is implemented in a way that prevents the recipient from re-identifying data subjects – though this is yet to be judicially settled. Ultimately, determining the most appropriate option requires researchers to balance data utility with legal obligations.

As secondary research becomes increasingly vital to medical advancements, researchers cannot afford to ignore these legal considerations. By taking proactive, well-documented steps, institutions can unlock the full potential of health data while navigating regulatory requirements with confidence. The message is clear: compliance with the law and advancing health research goals are not mutually exclusive; by following the right approach, researchers can achieve both.

Declaration. None.

Acknowledgements. The authors used ChatGPT to assist with editing and proofreading this manuscript for grammar, clarity, and readability. All intellectual content, analysis, and arguments are the original work of the authors, who take full responsibility for the accuracy and integrity of the work.

Author contributions. Both authors contributed to the conceptualisation of this work. Author 1 prepared the initial draft. Both authors contributed to the revision and approved the final manuscript.

Funding. U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. The content of this article is solely the authors' responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

Conflicts of interest. None.

1. Protection of Personal Information Act 4 of 2013 [South Africa]. https://www.gov.za/sites/default/files/gcis_document/201409/3706726-11act4of2013popi.pdf (cited 31 August 2022).
2. Thaldar D, Edgcumbe A, Donnelly DL. How to interpret core concepts in POPIA? Recommendations on the draft Code of Conduct for Research. South Afr J Sci 2023;119(7/8) https://doi.org/10.17159/sajs.2023/15062
3. Thaldar DW, Townsend BA. Exempting health research from the consent provisions of POPIA. Potchefstroom Electron Law J 2021;24:1-32. https://doi.org/10.17159/1727-3781/2021/v24i0a10420
4. Staunton C, Adams R, Anderson D, et al. Protection of Personal Information Act 2013 and data protection for health research in South Africa. Int Data Priv Law 2020;10(2):160-179.
5. Swales L. The Protection of Personal Information Act 4 of 2013 in the context of health research: Enabler of privacy rights or roadblock? Potchefstroom Electron Law J 2022;25:1-32.
6. Murdoch B, Jandura A, Caulfield T. Reconsenting paediatric research participants for use of identifying data. J Med Ethics 2023;49(2):106-109.
7. Wallace SE, Gourna EG, Laurie G, Shoush O, Wright J. Respecting autonomy over time: Policy and empirical evidence on re-consent in longitudinal biomedical research. Bioethics 2016;30(3):210-217.
8. Adams R, Veldsman S, Ramsay M, Soodyall H. Drafting a code of conduct for research under the Protection of Personal Information Act No. 4 of 2013 (with corrigendum). South Afr J Sci 2021;117(5/6).
9. National Health Research Ethics Council South African Ethics in Health Research Guidelines: Principles, Processes and Structures, 3rd ed (2024). National Department of Health, Pretoria. https://www.health.gov.za/nhrec-guidelines/
10. Thaldar D. Research and the meaning of 'public interest' in POPIA. South Afr J Sci 2022;118(3/4). https://doi.org/10.17159/sajs.2022/13206
11. Edgcumbe A, Botes M, Donnelly DL, Townsend B, Shachar C, Thaldar D. 'Potato potahto'? Disentangling de-identification, anonymisation, and pseudonymisation for health research in Africa. J Law Biosci 2025;12(1):lsae029.
12. Simon GE, Shortreed SM, Coley RY, et al. Assessing and minimizing re-identification risk in research data derived from health care records. EGEMs Gener Evid Methods Improve Patient Outcomes 2019;7(1):6 .
13. Gadotti A, Rocher L, Houssiau F, Creţu AM, De Montjoye YA. Anonymisation: The imperfect science of using data while preserving privacy. Sci Adv 2024;10(29):eadn7053.
14. Academy of Science of South Africa (ASSAf). ASSAf POPIA compliance framework for researchers and research institutions. ASSAf; 2025. https://hdl.handle.net/20.500.11911/438 (cited 24 July2025).
15. General Data Protection Regulation (EU) 2016/679, European Union.
16. Rodriguez A, Tuck C, Dozier MF, et al. Current recommendations/practices for anonymising data from clinical trials in order to make it available for sharing: A scoping review. Clin Trials 2022;19(4):452-463.
17. Kohlmayer F, Lautenschläger R, Prasser F. Pseudonymisation for research data collection: Is the juice worth the squeeze? BMC Med Inform Decis Mak 2019;19(1):178.

Received 1 April 2025. Accepted 22 September 2025.