

# Returning raw genomic data: rights of research participants and obligations of health care professionals

As the number of requests for raw genomic data increases, policies and protocols reflecting the perspectives of clinicians, patients, researchers and research participants are required

**N**ext generation sequencing technologies generate vast quantities of genomic data and can identify significant molecular markers associated with disease development and its response to treatments. Much attention has focused on the use of such data for research purposes, and on return of clinically significant findings to research participants via their clinicians.

Increasingly, patients are being recruited to research studies utilising genomics, particularly in oncology. Using whole genome sequencing, researchers may identify significant molecular markers giving rise to the potential for more targeted treatment options. The long established division between clinical care and research is becoming increasingly blurred as clinicians have a role in recruiting patients to research studies and then translating genomic results into clinical care. The scope of their obligations to individuals as both research participants and patients requires analysis, particularly in the context of what information individuals are entitled to.

## Requests for raw genomic data

The literature identifies increasing requests from patients and participants for their raw genomic data,<sup>1,2</sup> and we are aware of national and international genomic precision medicine studies where such requests have been made. Raw genomic data comprise genomic sequence data before annotation and interpretation (Box 1). Although it has no direct clinical utility, further interpretation may reveal information of value.<sup>3</sup> Individuals may want to access their raw genomic data for a variety of reasons that may or may not be associated with further health choices or treatment, and feelings of strong ownership and entitlement.<sup>4</sup>

In practice, raw genomic data may be provided on a case-by-case basis to those who request it. A number of studies have indicated that patients, research participants and relevant health professionals support the return of raw genomic data to patients and research participants upon request.<sup>4</sup> It has been predicted that the number of requests for raw genomic data is likely to increase exponentially as concepts of personalised medicine, health care autonomy and perceptions of the right to possess one's personal genomic data become more prominent.<sup>5</sup>

This article highlights arguments for a right of access to raw genomic data, the obligations of clinicians in mediating return, and features of a process that would support the ethical return of such data.

## Approaches to returning genomic data

The National Statement on Ethical Conduct in Human Research provides that researchers are not expected to return raw genomic data to study participants.<sup>6</sup> In contrast, Genomics England is obliged to return genomic data to participants in its 100,000 Genomes Project if requested.<sup>7</sup>

Do Australian study participants have any right to obtain genomic data upon request? Clinicians who recruit their patients to genomic research studies are often provided with a summary report interpreting clinically relevant results they may choose to share with patients.

Genomic data fall within the scope of privacy legislation. The *Privacy Act 1988* (Cth) provides that individuals have a right to access a copy of their personal information from the agency or organisation that holds it. Section 6 (1) of the Act defines personal information as "information or an opinion about an identified individual, or an individual who is reasonably identifiable". The Federal Court has stated that "a determination of whether the identity can reasonably be ascertained will require an evaluative conclusion".<sup>8</sup> The key issue then is whether raw genomic data identify an individual, and this is a subject of ongoing discussion. In Europe, the General Data Protection Regulation has been interpreted as providing a right to participants and patients to have their raw genomic data released to them (there are exemptions within the context of scientific research).<sup>5</sup> A recent Canadian article considers it likely that clinical laboratories will soon be obliged to provide access on legal grounds even where data was generated for research purposes.<sup>9</sup>

Irrespective of any legal right of return, there are strong ethical reasons why raw genomic data should be returned to patients enrolled in a research study. The return of raw genomic data respects the autonomy of participants and the personal meaning and value that genomic information has for them. Access to raw data can benefit participants by providing hope and options for further interpretation. Even if ultimately unfounded, providing raw data to those who request it reciprocates the participants' contribution to research and builds empowerment.<sup>9</sup>

## Implications for clinicians

Clinicians who treat patients and recruit them to genomic research studies will increasingly field requests for raw data. Patients might approach

Jane L Nielsen<sup>1</sup>

Carolyn Johnston<sup>1</sup>

Tracey O'Brien<sup>2</sup>

Vanessa J Tyrrell<sup>3</sup>

<sup>1</sup> University of Tasmania, Hobart, TAS.

<sup>2</sup> Kids Cancer Centre, Sydney, NSW.

<sup>3</sup> Children's Cancer Institute, Sydney, NSW.

jane.nielsen@utas.edu.au

## 1 Raw data file types typically released

### Whole genome sequence data

- FASTQ file: a FASTQ file is a text file that contains the sequence data from the clusters that pass filter on a flow cell; it stores both a biological sequence (usually nucleotide) and its corresponding quality scores
- Binary Alignment Map (BAM) file: a BAM file is an alignment file in a compressed binary format that is machine readable only; it contains information about sequenced reads (typically after alignment to a reference genome)
- Variant Call Format (VCF) file: a VCF file is a text file format that contains information about variants found at specific positions in a reference genome; it consists of meta-information lines, a header line, and then data lines (each data line contains information about a single variant)

### RNA sequencing data

- FASTQ file
- BAM file

### Other data

- A copy of a multidisciplinary tumour board research report (PDF file)
- Other pipeline specific processed data which vary according to the purpose of the pipeline (in the form of text or other file types)
- A letter explaining the release of data and the file types released

Clinicians considering returning raw genomic data may have valid concerns about their obligations to advise their patients about use of third-party interpretation (TPI) services (eg, Promethease [<https://promethease.com/>], Luna DNA [<https://www.lunadna.com/>]). TPI services analyse raw genomic data further, aiming to identify genetic markers or therapeutic targets for diseases not already identified through a research project. TPI services accept genomic data files from individuals and subject them to proprietary bioinformatic analyses that may not be clinically validated.<sup>11</sup> While many of the TPI service websites describe themselves in terms of literature retrieval and genome exploration, users may perceive reports generated as medical information. Many of these websites recommend their customers discuss results with their health care provider.<sup>10</sup>

The desire to protect patients from potential harms through utilising TPI services should not extend to denying people access to their raw data, propagating paternalism, and diminishing patient choice. Rather, fair and transparent processes for return of raw data should be explored, and we are involved with organisations that are proceeding with this work.

## Processes for return of raw genomic data

Concerns voiced regarding return of raw genomic data centre on the resource implications for health care providers, genetic services and health care systems.<sup>1,3,9</sup> Genetic counsellors are likely to be called upon to play a greater role in managing expectations of what genomic data can reveal.<sup>1</sup> More generally, the return of research results will require considerable

their clinician to seek further or more detailed interpretation where initial data analysis has yielded no concrete actionable finding. They may also request the information in order to share it on open access research platforms, or simply to take possession of something they believe they “own”,<sup>10</sup> or with which they identify in terms of ancestral ownership.

## 2 Recommendations for inclusion in protocols for the return of raw genomic data

Issue	Recommendation
Consent*	In order to avoid re-consenting patients, development of future policies should recommend broad consent as a mechanism to provide patients with information that feedback of genomic data is possible, <sup>12</sup> but that notes the limitations of the data in its uninterpreted form, <sup>5</sup> particularly that the data may not meet clinical standards. <sup>9</sup>
Protection of researchers	It should be made clear to patients recruited into research studies that research groups that generate data are not able to provide interpretation of those data beyond the study objectives.
Protection of clinicians	Clinicians who collect samples for the purpose of genomic analysis are not in a position to provide interpretation of resulting data. A statement that patients accept responsibility for subsequent uses of the data is recommended; good practice would dictate that such advice be provided in written form. <sup>5</sup> The ease of identifiability of genomic data should also be communicated, and transfer of data accompanied by appropriate advice. <sup>5</sup>
Protection of patients: interpretation by TPI services†	Raw genomic data returned by clinicians should be accompanied by a statement that interpretations provided by TPI services should be treated with caution and not treated as health advice, given the potential for false negative and false positive findings. Some recognition of the risks inherent in further use of the data is also likely to be prudent, particularly where use may include utilisation of TPI services, other research groups, or clinicians. This may necessarily include warning of over-reliance on or reassurance from false positives or negatives. <sup>14</sup>
Protection of close relatives	Patients should be advised of the potential sensitivity of genomic information for close relatives, <sup>2</sup> and privacy concerns should data be more broadly disseminated. Patients should be directed to observe adequate privacy thresholds and be made aware of limitations that subsequent use of the data might generate. <sup>9</sup> Processes should also be implemented to ensure data are able to be traced and re-verified after transfer. <sup>9</sup>
Protection of institutions	Institutions involved in sharing of genomic and health-related data should work together in generating uniform guidelines and standards that protect and promote the confidentiality, integrity and availability of data and services, and the privacy of individuals, families and communities whose data are shared.

TPI = third-party interpretive. \* Currently, patient consent processes do not inform patients of the possibility of accessing genomic data. † Interpretations of raw genomic data by TPI services vary in quality and veracity, and their operation is unregulated. ♦

work and cross-discipline collaboration to clarify clinicians' obligations in relation to the return of raw genomic data, and to assist them in managing patient expectations and best practices for return.<sup>12</sup> Logistical questions associated with appropriate storage and transfer mechanisms for large data files,<sup>9</sup> and the integration of genomic data into e-health records will also need to be navigated to ensure interoperability and ease the significant burden of transferring data.<sup>1,11,13</sup> Box 2 provides further recommendations that would form the basis for protocols for returning raw genomic data.

## Conclusion

This brief exposition of the issues for clinicians confronted with requests for raw data generated through participation in research highlights the myriad issues arising from a simple request by a patient. Although there is little doubt that ethical grounds compel the release of data, there are associated burdens and challenges that must be addressed. Further attention to overcoming these challenges is essential, necessitating collaborative, institutional development of policies and protocols reflecting the broad perspectives of clinicians, patients, researchers and research participants. The resource implications of establishing infrastructures for return should be addressed by health systems nationally and globally.

**Acknowledgements:** This project is supported by the Medical Research Futures Fund (Genomic Health Missions stream; grant number N76758). We acknowledge the contributions of other team members towards conception of this project.

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

**Competing interests:** Vanessa Tyrrell is Vice President of the Human Genetics Society of Australasia and a member of the NSW Health Genomics Strategy Translational Medicine Committee.

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

© 2022 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-NonCommercial](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

- 1 McGuire A, Caulfield T, Cho M. Research ethics and the challenge of whole genome sequencing. *Nat Rev Genet* 2008; 9: 152-156.
- 2 Chad L, Szego M. Please give me a copy of my child's raw genomic data. *NPJ Genom Med* 2021; 6: 15.
- 3 Nelson S, Bowen D, Fullerton S. Third-party genetic interpretation tools: a mixed-methods study of consumer motivation and behavior. *Am J Hum Genet* 2019; 105, 122-131.
- 4 Middleton A, Wright CF, Morley KI, et al. Potential research participants support the return of raw sequence data. *J Med Genet* 2015; 52: 571-574.
- 5 Schickhardt C, Fleischer H, Winkler E. Do patients and research subjects have a right to receive their genomic raw data? An ethical and legal analysis. *BMC Med Ethics* 2020; 21: 7.
- 6 National Health and Medical Research Council, Australian Research Council, Universities Australia. National Statement on Ethical Conduct in Human Research 2007 (updated 2018). Canberra: Commonwealth of Australia, 2018. <https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018#block-views-block-file-attachments-content-block-1> (viewed May 2022). [
- 7 Genomics England. Participant data requests under the GDPR and Data Protection Act 2018. <https://www.genomicsengland.co.uk/understanding-genomics/data/participant-data-requests/> (viewed Mar 2022).
- 8 *Privacy Commissioner v Telstra Corporation Limited* [2017] FCAFC 4.
- 9 Thorogood A. APPLaUD: access for patients and participants to individual level uninterpreted genomic data. *Hum Genomics* 2018; 12: 7.
- 10 Badalato L, Kalokairinou L, Borry P. Third party interpretation of raw genetic data: an ethical exploration. *Eur J Hum Genet* 2017; 25: 1189-1194.
- 11 Swaminathan R, Huang Y, Miller K. Transferring exome sequencing data from clinical laboratories to healthcare providers: lessons learned at a pediatric hospital. *Front Genet* 2018; <https://doi.org/10.3389/fgene.2018.00054>.
- 12 Budin-Ljøsne I, Mascalzoni D, Soini S, et al. Feedback of individual genetic results to research participants: is it feasible in Europe? *Biopreserv Biobank* 2016; 14: 241-248.
- 13 Narayanasamay S, Markina V, Thorogood A, et al. Genomic sequencing capacity, data retention and personal access to raw data in Europe. *Front Genet* 2020; 11: 303.
- 14 May T, Nakano-Okuno M, Kelley WV, et al. Return of raw data in genomic testing and research: ownership, partnership, and risk-benefit. *Genet Med* 2020; 22: 12-14. ■