1st Conference on Research Data Infrastructure

Securing RDM

https://doi.org/10.52825/CoRDI.v1i.275



Published: 07 Sept. 2023

# Developing Consent Tools for the Research Community at the German Human Genome-Phenome Archive (GHGA)

Andreas Bruns[1][https://orcid.org/0000-0002-1961-1691], Simon Parker[2,3][https://orcid.org/0000-0001-9993-533X], Fruzsina Molnár-Gábor[2][https://orcid.org/0000-0002-9406-2776], Eva C. Winkler[1][https://orcid.org/ 0000-0001-7460-0154], and the GHGA Consortium

[1] University Hospital Heidelberg, National Center for Tumor Diseases (NCT), Section Translational Medical Ethics

[2] Universität Heidelberg, BioQuant, Governance of Emerging Technologies in Medical Research and Health Care

[3] German Human Genome-Phenome Archive (GHGA, W620), German Cancer Research Center (DKFZ)

**Abstract.** The German Human Genome-Phenome Archive (GHGA) aims to enable the responsible sharing of human omics data for secondary research use across Germany and Europe. Informed consent is the most commonly used legal and ethical basis for processing omics data for secondary use. However, obtaining informed consent from Data Subjects can be challenging when data is to be widely shared and reused beyond the initial purpose of collection. To address these challenges, the ELSI (Ethical, Legal, and Social Implications) Group of GHGA has developed consent tools for the research community. First, we have developed a toolkit for prospective data collection, which consists of consent modules and complementary advice on how to update or create new consent forms. Second, we have created a legacy consent toolkit that can be used by researchers to assess whether the consent under which data was originally collected covers further data processing for secondary research purposes.

**Keywords:** Informed consent, Legacy consent, GDPR, Omics data, Secondary use

## 1. Introduction

The German Human Genome-Phenome Archive (GHGA) is currently establishing a federated data infrastructure allowing the secure storage of, and controlled access to, human genome and other omics data for scientific research use. So long as omics data relates to a specific identifiable person (Data Subject), it is considered sensitive personal data. Informed consent is the most commonly used legal and ethical basis for processing such data for scientific research purposes. Using consent as a legal basis has several advantages; it is a legal basis for processing both personal and special category personal data, and its implementation enables data consented for research use to be shared across European nations easily. Outside of its legal value, informed consent is a central ethical principle as consent ensures that Data Subjects are not treated as mere means to the ends of research but as participants freely choosing to contribute to research activities.

Importantly, consent is needed not only for collecting and using omics data for the primary purpose but also for sharing data for secondary research use. However, obtaining consent from data subjects for secondary use can be challenging. According to the European Union's General Data Protection Regulation (GDPR (EU) 2016/679), consent should be freely

given, unambiguous, specific, and informed. While in biomedical research with derivatives such as medical data there has been a shift away from strictly project-specific consent to so-called broad consent (consent to a broadly defined range of future research uses), it can be difficult to ensure that consent is sufficiently informed when data is to be widely shared and reused beyond the initial purpose of collection. Moreover, researchers may wish to process data collected prior to the introduction of the GDPR and may wonder whether the original consent allows data sharing and secondary use.

GHGA seeks to help address these regulatory challenges. The ELSI (Ethical, Legal, and Social Implications) Group of GHGA has developed consent tools for the research community to help researchers ensure that data is collected and shared with the consent of Data Subjects. First, we have developed a Modular Consent Toolkit for prospective data collection, which consists of consent modules and complementary advice on how to update or create new consent forms. Second, we have created a Legacy Consent Toolkit for legacy data (previously collected data) to assess whether the consent under which data was originally collected covers further data processing for secondary research purposes.

# 2. The GHGA Consent Tools

## 2.1. The Modular Consent Toolkit

The Modular Consent Toolkit, which contains consent modules and complementary guidance on how to use them, can be used to update or create new consent forms to enable data sharing for secondary research use and is available as a white paper on Zenodo under open access [1]. The consent modules explain the relevant processes around data sharing and secondary use, using comprehensible yet legally appropriate language, and are specifically designed to enable broad future research use and archival of the data being collected.

There is a total of four consent modules, i.e., text blocks that can be integrated into consent form text. All modules are available in German and in English. (1) The central module, the Data Sharing Module, explains what happens when data is securely archived and made available to researchers for secondary research use under conditions of controlled access. (2) The De-Identification Module provides further information on the process of removing direct identifiers and the security status of de-identified data. (3) The Controlled Access Module elaborates on the process of how researchers can request and be granted access to data. (4) Finally, the Consent Options Module allows Data Subjects to record their consent decision regarding the secondary research use of their data. Following the consent modules and guidance on how to use them, the white paper also contains a section on how researchers are to evaluate consent forms that have been updated using these consent modules to ensure that these are consistent and coherent.

## 2.2. The Legacy Consent Toolkit

The Legacy Consent Toolkit addresses the issue of legacy data and can be used to assess whether the consent under which data was originally collected allows further data processing for secondary research purposes. It can be accessed via the GHGA Website [2]. The Toolkit does not provide formal legal advice on whether legacy consent is a legitimate legal basis for processing but is instead used to provide guidance.

The structure of the Legacy Consent Toolkit is designed to ascertain whether the proposed secondary processing increases the risk to the rights and freedoms of the Data Subjects in a manner unforeseen to the Data Subjects at the time they gave their consent. To do so, it is necessary to compare the primary and proposed secondary processing. In the first stage of the assessment, the information content at the time of the secondary processing is compared to that originally collected. In the second stage, the purpose of the primary and secondary

processing is compared. In the third stage, the person originally permitted to process the data is compared to the proposed secondary processor. In the final stage, an assessment is performed to understand whether the secondary processing increases the risk to the Data Subjects from a data protection perspective.

## 3. Conclusion

While our consent toolkits provide a service to researchers and their institutions hoping to ensure consent for the sharing of omics data for scientific research use, they may also function as a measure to protect the interests of patients and other groups of Data Subjects. So long as consent remains the basis for data sharing for secondary research use, it is important that Data Subjects are put in a position to make informed decisions about the further use of their data. Our toolkits aim to help Data Subjects gain a better understanding of the implications of secondary use as well as to ensure that it is properly assessed whether further use is covered by their original provision of consent.

## Data availability statement

The submission is not based on any data.

## Underlying and related material

None.

## Author contributions

Andreas Bruns and Simon Parker have shared first authorship for this abstract (“Writing – original draft”, “Writing – review and editing”). Fruzsina Molnár-Gábor and Eva C. Winkler have supervised the research on which this abstract reports (“Supervision”). The GHGA Consortium functions as collective co-author on all GHGA publications and has acquired funding for this research (“Funding acquisition”).

## Competing interests

The authors declare that they have no competing interests.

## Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of GHGA – The German Human Genome-Phenome Archive (www.ghga.de, Grant Number 441914366 (NFDI 1/1)).

## Acknowledgement

The authors would like to thank the GHGA Team for valuable feedback on the Consent Tools.

## References

1. Andreas Bruns, Anna Benet-Pages, Jan Eufinger, Holm Graessner, Oliver Kohlbacher, Fruzsina Molnár-Gábor, Simon Parker, Christoph Schickhardt, Oliver Stegle, & Eva

Winkler, “Consent Modules for Data Sharing via the German Human Genome-Phenome Archive (GHGA),” Zenodo, July, 2022, doi: 10.5281/zenodo.6828131, https://zenodo.org/record/6828131#.ZEKD3i0Rp-V (21/04/2023).
2. GHGA, “Legacy Consent Toolkit.” GHGA Website. https://www.ghga.de/resources/legacy-consent (21/04/2023).