L-Università ta' Malta

University Research Ethics Committee (UREC)

Criteria for Medical Image Data to be Considered Anonymous

July 2024

Definitions:

The following terms are the defined and are to be interpreted as such in this document

Medical Image Data: This document will restrict medical images to the following: Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Plain Radiography (X-Rays), positron emission tomography (PET) scans, and Ultrasound, mammography. This definition may be updated by the University of Malta Research Ethics Committee.

Anonymous Data: "information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable" (Recital 26 GDPR). In order to know whether data is anonymous, "account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly." (Recital 26 GDPR).

DICOM Standard: DICOM (Digital Imaging and Communications in Medicine) is an international standard for storing, transmitting, and sharing medical imaging information. It defines formats for medical images and associated data, ensuring interoperability between imaging equipment, healthcare systems, and networks from different manufacturers.

DICOM Data: DICOM data consists of Medical Image Data (as defined above) along with metadata that describes the patient's information, imaging parameters, and study details. This data is encapsulated in DICOM files, which can be exchanged and interpreted consistently across various healthcare environments.

Introduction:

In the context of medical research and clinical studies, the use of Medical Imaging data plays a crucial role in advancing our understanding of human health and disease. However, the rich and detailed nature of Medical Imaging Data presents significant challenges in ensuring compliance with data protection regulations, particularly the General Data Protection Regulation (GDPR) of the European Union. GDPR mandates stringent protections for personal data, requiring that any information which can directly or indirectly identify an individual must be treated with the highest levels of confidentiality and security.

In general, Medical Imaging Data is to be considered personal data. However, there are some very specific situations where Medical Imaging data can be considered to be anonymous. The rationale for considering whether Medical Imaging Data is to be considered anonymous will follow the GDPR definition of personal data in Article 4(1), where personal data is defined as any information relating to an identified or identifiable natural person. Anonymization, as per Recital 26, involves processing data in such a way that individuals are not identifiable, considering all means reasonably likely to be used to identify them.

Given the intrinsic difficulties in anonymizing medical images, and the extensive metadata contained within Digital Imaging and Communications in Medicine (DICOM) files it is imperative to establish clear and rigorous criteria to define when Medical Imaging Data can be considered non-personal. Such criteria must ensure that the data cannot be used to identify individuals, either on its own or in combination with other data.

This document outlines the criteria and considerations for the University of Malta's Research Ethics Committee subcommittee of Data Protection (UREC-DP) to deem Medical Imaging Data non-personal, ensuring compliance with GDPR while facilitating valuable medical (and related scientific) research.

Criteria:

Criterion 1: Use of Already Aggregated Image Data Without associated DICOM Metadata

The first criterion addresses situations where researchers utilise aggregated Medical Imaging Data. Aggregated data includes mean images across a population, template images, and statistical maps. These types of data are inherently non-identifiable as they represent collective information from multiple individuals, thus lacking any direct or indirect personal identifiers.

Criterion 2: Use of Individual Slices Without Associated DICOM Metadata

The second criterion involves the use of individual Medical images / slices from an individual, ensuring that these images/slices:

1. do not constitute whole volumes
2. they should be image data (e.g. png or jpg etc) that do not contain *any embedded* DICOM metadata. Such data can still have a limited amount of associated metadata but the researcher should ensure that these data are a) non personal and b) not embedded in a DICOM file
3. personal data is not "burned" into the image pixel data
4. have no obvious or rare features that can cause an individual to be singled out

This criterion is applicable when researchers utilise specific screenshots or segments of Medical Images/slices, ensuring that no identifiable features or information are present.

An important exception to this criterion is if the subject of the data has a very rare condition that can uniquely identify them. In such cases, the data should not be considered to be anonymous. The onus is on the researcher to consult with a medical professional if there are any doubts and to document the consultation. Researchers should document their decisions particularly related to whether they consider the slices of interest to have rare features or not.

Summary of Criteria:

Criterion	Image data	Meta-data
1	Population aggregate	None
2	Individual slices	None

Audit Requirements:

To ensure ongoing compliance with GDPR and to safeguard the anonymity of Medical Imaging Data UREC-DP will continue to audit studies that use anonymised Medical Imaging Data. These audits will be conducted as part of the regular ethics audit processes and, in-particular, will assess the completeness and transparency of the rationale for the consideration of such data to be anonymous.

Conclusion

The outlined criteria serve exclusively to determine whether Medical Imaging Data can be considered anonymous and, therefore, exempt from requiring review by the UREC-DP subcommittee. Research involving non-anonymous data remains permissible; however, it must be conducted under stringent technical and organisational safeguards to ensure GDPR compliance. UREC will continue to monitor developments in this area, proposing methodologies to streamline data handling processes for complex medical imaging data as necessary.

The following best practices should be followed, in particular when full anonymisation of data is not possible:

1. **Remove Direct Identifiers and Handle Indirect Identifiers:** Whenever possible and practical, all direct identifiers such as names, ID numbers, and exact dates that can be linked to an individual are to be removed from Medical Images. Furthermore, due consideration should also be given to the removal and/or handling of indirect identifiers (quasi-identifiers), such as location data or unique medical information that could be combined to re-identify an individual.
2. **Apply De-identification Techniques:** Use techniques like data masking, pseudonymization, or encryption to protect sensitive information within the medical images.
3. **Strive for Irreversibility:** De-identification should be irreversible, making it impossible to re-identify the individual by any means.
4. **Evaluate Re-identification Likelihood:** Consider the likelihood of re-identification in the given Medical Image dataset and apply measures to minimise this risk.
5. **Maintain Data Utility:** While de-identifying, ensure that the data remains useful for the intended research purpose.
6. **Follow DICOM Standards:** Stay informed about the latest de-identification techniques and legal requirements to ensure ongoing compliance. For DICOM images, follow the specific guidelines for de-identification and re-identification as outlined by the current version of the DICOM standard.