Guidance on Deidentification/Anonymization of DICOM images for use in a Research Setting

Duke Medicine is committed to protecting the privacy and confidentiality of sensitive electronic information and adherence to the Health Insurance Portability and Accountability Act of 1996 (HIPAA). Under HIPAA, protected health information (PHI) is individually identifiable health information that is transmitted or maintained in any form or medium by a covered entity or their business associates acting on behalf of the covered entity.

Digital Imaging and Communications in Medicine (DICOM) files may have PHI burned into the visible image (e.g., MRN, name, dates), but may also contain hidden data fields that help describe the contents of those files. This “data about the data,” which is commonly referred to as metadata, also fits the definition of PHI. This guidance clarifies Duke standard practice for how these files must be secured in a research setting. Note - this guidance is limited to work within research and does not address preparations for sharing images in a clinical, non-research setting.

Understanding Protected Health Information (PHI) is critical to being able to prepare DICOM images for sharing in a research setting. The below list outlines the 18 identifier classifications. Note that some are always included in DICOM images (dates), others may be included (name, MRN, phone and SSN), and others are rarely included (vehicle identifiers, web URLs, IP addresses).

Correct usage of the terms describing a data set, including one with images, is necessary for HIPAA adherence. Review the below terms carefully:

Removal of all PHI elements, particularly dates, may not always be practical with DICOM images as dates may be necessary to maintain the clinical utility of the images, may be necessary for serial image integrity, or may be required by recipients. Note that in a research setting, it is possible to use informed consent form (ICF), HIPAA waivers, or Contractual language to leave some elements of PHI intact as long as appropriate authorization is obtained.
PHI can typically be found in two places in DICOM images. This list uses fictitious information. It is not inclusive of all tags that might represent HIPAA PHI elements.

Multiple software programs exist to edit DICOM images, both in the visual image and in the metadata tags. Understanding both HIPAA requirements and DICOM architecture is necessary to configure the software to make sure that we achieve compliance. Below is a suggested list of steps to follow. It is strongly recommended that those unfamiliar with image management seek the support and advice of image subject matter experts at Duke rather than attempting image management themselves.

Step 1: Determine your image requirements and what elements of PHI are necessary in the visual or metadata tags for the success of the study. Based on this, determine whether data are anonymous, deidentified, limited data set or fully identified. Work with study sponsors and data recipients to understand both what they need included in the images and how the images will be sent to them. Obtain permission from Duke to send images.

1A: If image collection is prospective, include the specific data elements that will be sent in the study ICF.
1B: If image collection is retrospective and ICF is not possible, consult with Institutional Review Board (IRB) and Office of Corporate Research Collaborations (OCRC) to determine whether a HIPAA waiver, Data Use Agreement, or some other contract is appropriate. Note: it is very important that you tell either of these bodies exactly what elements of PHI will be retained in the DICOM images and who will have access to the images.

1C: In both cases, clearly list parties who will have access to images (study recipient/sponsor and any intermediary business affiliates of theirs, specifically cloud providers). Avoid use of terms like “secure” and “HIPAA compliant” with non-Duke entities unless this has been verified through review of controls.

1D: Clearly list all data storage in both the IRB documents and in the Research Data Security Plan (section 12.1 in eIRB application for non-exempt studies).

Step 2: Load images into locally installed software, configure for settings necessary to above data type (e.g. anonymous, etc) and appropriately modify the images. Save files locally for review prior to transmission outside of Duke. Using existing software or a locally installed DICOM viewer, visually inspect a subset of the images for compliance. In the visual image, specifically look for dates, MRNs, names and/or initials, ages over 89, and any other identifiers directly linked to the individual. In the DICOM metadata, look at ‘All Tags’ for the above as well as phone numbers, address, SSNs, and other identifiers.

2A: Looking at the metadata can be confusing. There are many dates and date/timestamps, some of which are directly linked to individuals, others of which pertain to dates that machines are configured. The image creator can also include information in the ‘Unknown Tag & Data’ fields which are not pre-set DICOM fields and these can include any information that the person performing the scan enters, which could include PHI. This confusion is a primary reason why those unfamiliar with image management should seek the support and advice of image subject matter experts at Duke rather than attempting image management themselves.

2B: If PHI is retained in the image that is not addressed in the ICF, waiver or contract, adjustments must be made either to the documents or the software settings until both are in concordance.

2C: Removal of identifiers must occur locally at Duke and not during transport via a cloud provider after the images have left our covered entity.

Step 3: Ensure that the transport method is encrypted if ANY PHI remains in the images (see DM Mobile Computing & Storage Device Standard). This can be accomplished in a number of ways (e.g., upload via https or sftp, transport on whole disc encrypted mobile devices).

3A: If encryption is not possible in a specific circumstance, an exemption from the Duke Medicine Information Security Office (ISO), iso@mc.duke.edu, must be obtained in advance.

Use of software should be department approved. Individuals should not search for and download free or purchased software (for deidentification, viewing, or transmitting images or other purposes) onto their computers without getting departmental approval and ensuring that the software is vetted, compliant, appropriately licensed, fit for purpose, and free of malware.

Researchers who do not have access to a departmentally implemented DICOM solution may wish to consider use of the Duke Multidimensional Image Processing Laboratory (Multi-D Lab) for image deidentification as a service. Research staff required to send imaging studies outside the institution (i.e. to a sponsor or an imaging CRO), can email the lab at: radiology-deid@notes.duke.edu or call (919) 681-0492.

This guidance will up updated and redistributed as necessary by the Duke Medicine Information Security Office in conjunction with Duke Ethics and Compliance Office and Duke Image Subject Matter Experts.