A Web-Based Interface for Communication of Data Between the Clinical and Research Environments without Revealing Identifying Information[1]

Peyton H. Bland, PhD, Gary E. Laderach, BS, Charles R. Meyer, PhD

Rationale and Objectives. Recent health care policies and regulations have affected the manner in which patient data—especially protected health information (PHI)—are handled in both the clinical and research settings. Specifically, it is now more challenging to obtain de-identified PHI from the clinic for use in research while adhering to the requirements of this new environment.

Materials and Methods. To meet this challenge, we have devised a novel web-based interface that facilitates the communication of data (eg, biopsy results) between the clinic and research environments without revealing PHI to the research team or associated research identifiers to the clinical collaborators. At the heart of the scheme is a web application that coordinates message passing between the researchers (in general, the requesters of de-identified PHI) and clinical collaborators (who have access to PHI) by use of a protocol that protects confidentiality.

Results. We describe the design requirements of this communication scheme and present implementation details of the web application and its associated database.

Conclusions. We conclude that this scheme provides a useful communication mechanism that facilitates clinical research while maintaining confidentiality of patient data.

Key Words. PHI; protected health information; HIPAA; DICOM; anonymization; de-identification.

Recent years have seen significant changes in the management of patient data in the medical setting. These changes have been heavily influenced and, in some cases, dictated by regulations such as the rather broad mandates of the Department of Health and Human Services. Specifically, certain provisions of Health Insurance Portability and Accountability Act, as detailed in Department of Health and Human Services 45 CFR Parts 160 to 164 (1)—notably, the Privacy Rule and the Security Rule (2)—have greatly affected the manner in which medical staff and medical centers operate with respect to both clinical practice and research. Further, privacy practices in some hospitals and medical centers are influenced by policies of The Joint Commission (formerly the Joint Commission on Accreditation of Healthcare Organizations). On a local level, clinical research facilities are influenced by local institutional review boards (IRB). All of these are concerned with the proper handling of protected health information (PHI), which is any information about health and health care that can be linked to an individual. Confidentiality of PHI must be maintained for both legal and ethical reasons, and the adherence to appropriate standards is mandatory for facilities that receive federal

Acad Radiol 2007; 14:757–764

[1] From the Department of Radiology, University of Michigan Health System, BSRB, Room AC51, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200 (P.H.B., G.E.L., C.R.M.). Received November 3, 2006; accepted February 17, 2007. Supported in part by NIH/NCI 1P01 CA87634, Automated 3D Registration for Enhanced Cancer Management; and NIH/NCI 1U01 CA91099, Lung Image Database Consortium. **Address correspondence to:** P.B. e-mail: bland@umich.edu

doi:10.1016/j.acra.2007.02.015

funding. Clearly, a new set of procedures is needed to facilitate research while protecting PHI in this new environment.

As an example of PHI, consider the results obtained from a radiologic examination. Images from these exams are commonly made available in DICOM format (Digital Imaging and Communications in Medicine [3]). This widely used set of standards specifies not only the encoding and storage of image data, but also specifies how meta-data relating to the patient, the imaging equipment, the imaging exam, etc., are to be stored along with the image data. The meta-data (the DICOM attributes, often referred to collectively as the "DICOM header") are integral with the image data and thus cannot be lost or separated from it, a clear advantage to researchers. However, even though the DICOM standard has been a boon to researchers, it comes with a price: In the current environment of confidentiality practices, the information stored in the meta-data—much of it PHI—must be dealt with properly.

Continuing with this example, consider the transfer of DICOM data files into the research environment. One way of complying with confidentiality regulations is to de-identify the DICOM attributes that contain PHI. (This process is sometimes referred to as "anonymization.") In our lab, de-identification is performed automatically by use of a custom-designed DICOM receiver that processes DICOM image files as they come into the lab via the network. Some DICOM attributes are removed entirely by this process (or have their values set to null), whereas others are given new values. In the latter case, the original values are generally stored in an encrypted, protected database table along with the new values in such a way that the original values can be reassociated with the data when needed and when properly authorized. This process is described in more detail in the following sections.

Concerns with PHI, however, go beyond DICOM image files. The primary focus of this article is not the transfer of de-identified DICOM images; instead, the primary focus is the communication of PHI that, in general, originates or exists independently of images (and the PHI that may be contained in them), such as clinical data from the patient interview, physical examination, reports from pathology and surgery, and so on. To obtain local IRB approval, researchers must observe proper practices in handling and using these other types of PHI. Because much radiologic research involves the use of this information along with image data, the question arises how this information can be accessed and handled while maintaining confidentiality. Access to PHI typically results from direct communication between the researcher and a clinical collaborator—often via ad hoc or informal mechanisms. These conventional mechanisms do not guarantee protection of PHI nor can they reliably conceal the identity of patients from the research team.

One side effect of de-identifying DICOM images as described previously is the increased difficulty in matching de-identified clinical PHI with research results that are derived from or related to the de-identified clinical data. At the same time, however, this separation works in the favor of increasing the integrity of the research because it prevents the researcher from being unduly influenced by knowledge of patient information, clinical data, and diagnostic outcomes. For example, knowledge of a patient's diagnosis could influence how research analysis results might be interpreted or used. However, if de-identified PHI and research results are matched in a blinded fashion, integrity can be maintained.

In summary, successful conduct of research often requires the transfer of de-identified patient-related information from the clinical environment to the researchers while maintaining confidentiality and integrity of the information. To fulfill this requirement, we have designed a novel scheme built around a web application that we call "Web-CAP" (web-based communications application for PHI). Web-CAP is designed as a tool to assist in the process of communicating de-identified PHI between the hospital and research environments by implementing a secure, well-defined, controlled protocol that improves on ad hoc procedures sometimes used to manage PHI. Patients whose PHI is being communicated through the web interface have either given informed consent or an IRB waiver has been obtained by the project's investigators. Web-CAP is not intended to circumvent any mechanisms or procedures that are in place for the protection of PHI; final authority for the manner in which PHI is managed in a medical center resides with the local IRB.[1] Instead, proper use of Web-CAP and the systems that surround it facilitate the adherence to federal and local established regulations and thus potentially assist in obtaining IRB approvals by demonstrating to an IRB that proper safeguards for handling PHI are in place.

[1] It is assumed that the local IRB has authorized the clinical collaborators (eg, clinical data technicians) using Web-CAP to have access to PHI and that it has authorized the research personnel using Web-CAP to have access to de-identified PHI. In this manuscript, the terms "clinical collaborator" and "researcher" imply those who have been properly authorized by the local IRB.