Information Security and the "Privacy Broker"

Michael Douma
Institute for Dynamic Educational Advancement, USA

Eduard J. Gamito
University of Colorado Health Sciences Center, USA

INTRODUCTION

The term "privacy broker" describes a concept developed in 2004 to address important privacy issues in public health research. The privacy broker provides a combination of an information "firewall" between the research team and the study participants' personal data, to protect the identity of study participants. It also provides a "masquerading" feature so that participants can be indirectly tracked or contacted, which allows for the collection of scientifically useful information over time.

BACKGROUND

Over the past several hundred years, research involving humans as subjects has resulted in countless exciting discoveries that have greatly benefited humanity. Advances like the discovery of penicillin to fight infection or the vaccine that brought polio under control, and nearly all modern-day treatments for cancer, diabetes, heart disease, and so forth, would not have been possible without research on humans. Unfortunately, among the many bright moments in research history, some dark episodes have occurred. For example, in the infamous Tuskegee Syphilis Study that took place from 1932 to 1972, researchers withheld available treatments from poor African American men with syphilis so that the progression of this terrible disease could be studied. Many of the men in the study suffered and died needlessly as a direct result of the researchers' unethical behavior. While the Tuskegee study is an extreme example of the abuses that have occurred, and is considered the worst case of research abuse in U.S. history, many other abuses have occurred and continue to occur to this day. In recent years, research abuses have been less extreme and much more rare than in the past. Most of these abuses have pertained to improperly informing study participants of known risks to participating in research. In some cases, researcher negligence has resulted in the release and misuse of the personal health information of research participants.

To address these problems, a series of laws and regulations have been put into place over the years to protect humans participating in research studies. One of the most recently implemented laws is known as the Health Insurance Portability and Accountability Act (HIPAA) of 1996. This law, the first phase of which covered privacy issues, was implemented in 2003. The HIPAA Privacy Rule sets standards to protect information related to health care. Specifically, it regulates health care information that can be linked with a person.

The Problem

While HIPAA is an important law that is needed to protect study participants, its restrictions have created barriers to the successful implementation of some research studies with human subjects (Feld, 2005; Erlen, 2005; Kaiser, 2004; O'Herrin, Fost, & Kudsk, 2004). For example, many studies that require long-term follow-up to collect data from study participants have become more difficult to implement or have become completely unfeasible. This is because, in order to follow up with a participant, a researcher needs to record the participant's contact information, which falls under the purview of HIPAA. Many researchers in this situation can comply with HIPAA by obtaining written consent (a signed subject consent form) to collect, store, and use the participant's personal information.

This is not an onerous requirement for a researcher who is based at one institution and who has his or her study subjects come in for clinical appointments on a regular basis. However, in many research scenarios, such as public health research that often entails the

Table 1.

STUDY ID	GENDER	AGE	CANCER TYPE	COUNTY
1	Male	82	Prostate	Jefferson
2	Female	74	Breast	Los Angeles

study of entire populations, this requirement can be an insurmountable barrier.

What is Personally Identifying Information?

Some obvious personal identifiers are names, addresses, and dates of birth. However, some other potentially identifying information may not be as obvious and may identify an individual depending on the setting. Table 1 provides a fictional example of research data for a pain management study for men and women with cancer.

Research subject number one lives in a rural area. He is one of six men in the county who have been recently diagnosed with prostate cancer and he is the only 82-year-old still living in that area. In this case, the study participant's age, cancer type, and county of residence constitute personal identifying information and should not have been collected without written consent of the study subject. On the other hand, study subject number two is a 74-year-old woman with breast cancer living in Los Angeles, California. Her age, cancer type, and county of residence would be much less likely to identify her. Table 2 provides a list of potentially identifying information that may fall under the purview of HIPAA.

A Research Example

As an example, consider a scenario where researchers want to measure changes in tobacco use (both cigarettes and chewing tobacco) in rural areas throughout the U.S. The researchers have developed educational materials that are intended to inform users of the increased health risks associated with tobacco use. The materials also include suggested strategies for tobacco use cessation.

Table 2. Additional examples of personally identifying information

Name(s), including nicknames and initials
Geographic location other than state (address, city, ZIP, county)
Date(s) of treatment, appointments, tests, hospitalizations (can use year only)
Month or date of birth (year of birth can be used unless over 89)
Social Security number
Phone number
Fax number
E-mail address
Medical Record number
Health Plan number
Account numbers
Certificate/license numbers
URL
IP address
Vehicle identifiers
Device ID
Biometric ID
Full-face/identifying photo
Any other unique identifying number, characteristic, or code