

Research Ethics in Big Data

Marilyn J. Hammer, PhD, DC, RN

The ethical conduct of research includes, in part, patient agreement to participate in studies and the protection of health information. In the evolving world of data science and the accessibility of large quantities of web-based data created by millions of individuals, novel methodologic approaches to answering research questions are emerging. This article explores research ethics in the context of big data.

Hammer is the director of research and evidence-based practice in the Department of Nursing at Mount Sinai Hospital in New York, NY.

No financial relationships to disclose.

Hammer can be reached at marilyn.hammer@mountsinai.org, with copy to editor at ONFEditor@ons.org.

Keywords: big data; research guidelines; patient autonomy

ONF, 44(3), 293–295.

doi: 10.1188/17.ONF.293-295

George Orwell wrote the novel *1984*, in which every individual was under constant surveillance in a futuristic world. Now, 33 years past the setting of that novel, people are under considerable surveillance far exceeding that described in the story line of *1984*. Well beyond being captured on security cameras at every turn and occasional drones, we actively disclose where we are, what we are doing, what we like and dislike, and who we are connected to with virtually every tap of our e-devices. The amount of data created daily through technology has become so exponential (Mittelstadt & Floridi, 2015) that it is deemed an important commodity and has been termed *data capital* (MIT Technology Review Custom, 2016). We have essentially become our own best historians in real time, leaving a highly accessible data trail. These “digital footprints” are being used en masse as a rich source of research data (Bietz et al., 2015). All these data elements are freely available without our consent, so where does this leave us in terms of research ethics?

The Impetus for the Ethical Conduct of Research on Humans

Research ethics guidelines were initially based on the horrific treatment of human participants without their consent, such as in the Tuskegee experiment and

in the atrocities of World War II. Historic information about the establishment of guidelines for the ethical conduct of research were detailed in a prior research ethics article (Hammer, 2016). Keeping in mind the historic inflections that inform the current research guidelines provides a solid, although sometimes narrow, framework with which research studies are conducted today. In brief, based on the Nuremberg trials, the 1947 Nuremberg Code, which introduced new research ethics principles, was written (National Institutes of Health, n.d.), followed by the Declaration of Geneva in 1948 (Fischer, 2005). Orwell’s novel was published the following year but not with research ethics in mind. Current adherence to the *Belmont Report* holds researchers to the highest standards of ethics to ensure not only that study participants are treated respectfully and that studies provide more benefit than potential risk, but also that personal health information is protected (Office for Human Research Protections, 1979, 2016).

The protection of health information to prevent individuals from losing health insurance or employment, or incur other life losses because of personal health information disclosures became central with the 1996 Health Insurance Portability and Accountability Act (HIPAA) (Colorafi & Bailey, 2016). HIPAA was strongly embraced.