

## ***AMA Code of Medical Ethics***

### ***3.3.4 Research Handling of De-Identified Patient Data***

Within health care systems, identifiable private health information, initially derived from and used in the care and treatment of individual patients, has led to the creation of massive de-identified datasets. As aggregate datasets, clinical data takes on a secondary promising use as a means for quality improvement and innovation that can be used for the benefit of future patients and patient populations. While de-identification of data is meant to protect the privacy of patients, there remains a risk of re-identification, so while patient anonymity can be safeguarded it cannot be guaranteed. In handling patient data, individual physicians thus strive to balance supporting and respecting patient privacy while also upholding ethical obligations to the betterment of public health.

When clinical data are de-identified and aggregated, their potential use for societal benefits through research and development is an emergent, secondary use of electronic health records that goes beyond individual benefit. Such data, due to their potential to benefit public health, should thus be treated as a form of public good, and the ethical standards and values of health care should follow the data and be upheld and maintained even if the data are sold to entities outside of health care. The medical profession's responsibility to protect patient privacy as well as to society to improve future health care should be recognized as inherently tied to these datasets, such that all entities granted access to the data become data stewards with a duty to uphold the ethical values of health care in which the data were produced.

As individuals or members of health care institutions, physicians should:

- (a) Follow existing and emerging regulatory safety measures to protect patient privacy.
- (b) Practice good data intake, including collecting patient data equitably to reduce bias in datasets.
- (c) Answer any patient questions about data use in an honest and transparent manner to the best of their ability in accordance with current federal and state legal standards.

Health care entities, in interacting with patients, should adopt policies and practices that provide patients with transparent information regarding:

- (d) The high value that health care institutions place on protecting patient data.
- (e) The reality that no data can be guaranteed to be permanently anonymized, and that risk of re-identification does exist.
- (f) How patient data may be used.
- (g) The importance of de-identified aggregated data for improving the care of future patients.

Health care entities managing de-identified datasets, as health data stewards, should:

- (h) Ensure appropriate data collection methods and practices that meet industry standards to support the creation of high-quality datasets.

- (i) Ensure proper oversight of patient data is in place, including Data Use/Data Sharing Agreements for the use of de-identified datasets that may be shared, sold, or resold.
- (j) Develop models for the ethical use of de-identified datasets when such provisions do not exist, such as establishing and contractually requiring independent data ethics review boards free of conflicts of interest and verifiable data audits, to evaluate the use, sale, and potential resale of clinically derived datasets.
- (k) Take appropriate cyber security measures to seek to ensure the highest level of protection is provided to patients and patient data.
- (l) Develop proactive post-compromise planning strategies for use in the event of a data breach to minimize additional harm to patients.
- (m) Advocate that health- and non-health entities using any health data adopt the strongest protections and seek to uphold the ethical values of the medical profession.

There is an inherent tension between the potential benefits and burdens of de-identified datasets as both sources for quality improvement to care as well as risks to patient privacy. Re-identification of data may be permissible, or even obligatory, in rare circumstances when done in the interest of the health of individual patients. Re-identification of aggregated patient data for other purposes without obtaining patients' express consent, by anyone outside or inside of health care, is impermissible.

*AMA Principles of Medical Ethics: IV*