Discussion 3 Ethics Behind Data-driven Medicine

Discussion Topic:

1. Context
In national healthcare systems, great weight is put on the possibilities of large data collections and "big data" for generating economic growth, enhancing medical research, and boosting health and wellbeing in totally new ways. This massive data-gathering and usage is justified by the moral principle of improving health. The imperative of health thus legitimizes data collection, new infrastructures, and innovation policy. It is also supported by the rhetoric of health promotion. New arrangements in health research and innovations in the health sector are justified, as they produce health, while the moral principle of health also obligates individual persons to pursue healthy lifestyles and become healthy citizens.

2. Discuss
Please discuss the following points:

- With data-driven medicine, comment on the impacts on privacy and autonomy when contrasted with the moral principle of health.
- Suggest a plausible ethical vulnerability when using data-driven medicine and a potential remedy for such a vulnerability.

My Post:

Hello class,
As the discussion context notes, large-scale data collection is often justified to drive economic growth, advance medical research, and improve health and well-being. This suggests that the "imperative of health" can be used as a moral argument to legitimize building data sets of the population's health information and manipulating this information to promote innovation, with the assumption that more data means better health outcomes.
However, this logic can pressure individuals to share private health information for the sake of accessing care or health benefits.

Impacts on privacy and autonomy when contrasted with the moral principle of health
Data-driven medicine can be used for earlier diagnosis, more precise care treatment, and better epidemiology. This usage can be very beneficial for society; however, when seen through the lens of Kantian ethics, which is a human-centric ethical framework that places human existence and capacity at the center of systems, it can be argued that when privacy is eroded, autonomy and dignity erode as well (Ulgen, 2017). This erosion may cause individuals to lose their ability to choose what parts of their private data are exposed and when. To put it bluntly, being unable to protect one's own private data is like being in a continual state of exposure, "comparable to having no choice but to walk down the street naked," (O'Keefe & Brien, 2023, p.52). In other words, through large-scale health data collection, which is a component of data-driven medicine, one's data is constantly exposed, effectively removing choice over which parts and when one's own data can be used; ultimately, choice is what disappears.

On the other hand, an ethical consequentialist approach, "the basic intuition that what is best or right (for today) is whatever makes the world best in the future" (Sinnott-Armstrong, 2023), will

argue that the beneficial health outcomes outweigh the privacy costs. However, this is precisely the kind of trade-off that can rationalize serious human rights violations, that is, if the perceived benefits are large enough. In essence, this approach limits an individual's choice by arguing that the future needs of the many outweigh the rights of the one. However, an individual's ability to choose is a fundamental right. For example, even with the maximizing benefit (to all) and minimizing harm (to individuals) approach, individuals are treated as inputs into a health-optimization system, such as algorithmic predictive modeling. Immanuel Kant (the father of Kantian ethics) argued that morality is based on duty, not on the consequences of actions (Seikh, 2025). Meaning that an ethical framework for data-driven medicine cannot be based solely on good outcomes, but it must also account for whether the system treats private individual data as being part of a person rather than as a means, just an input, to improve health systems. Not doing so converts patients into a product that can be used or manipulated, such as enabling secondary uses (eligibility screening or risk scoring). In a sense, the harms are not only violations of privacy; they are also injuries in the form of loss of autonomy and dignity, and in certain situations, they can undermine the very health outcomes that data-driven medicine systems were created to maximize.

A plausible ethical vulnerability when using data-driven medicine, and a potential remedy for such a vulnerability

A plausible ethical vulnerability when using data-driven medicine is function creep in combination with re-identification, which can result in discrimination against individuals/populations and private data misuse. For example, even de-identified datasets (lacking individual identification) do not prevent harm; as shown by the work of Sweeney (2000) on re-identification, simple “quasi-identifiers” (like ZIP code, date of birth, and sex) can identify a large portion of people, and be used to link an health record back to a real individual, resulting in expose diagnoses, prescriptions, or health condition. Additionally, this knowledge can be repurposed for profiling, behavioral prediction, advertising, or even used for decision-making by insurance and employment. This reproposing of data can be defined as “data” function creep. Data function creep can be defined as data collected for one specific purpose getting gradually reused for new purposes, which deviates from the original intended, disclosed, or consent.
A potential remedy is to use a Privacy-by-design combined with a patient choice mechanism approach. An approach that combines:

- Governance limits the usage of the data (how much of and how many times the data can be used), overseas who can use the data, and sets boundaries on how or for what purpose the data can be used.
- Technical safeguards implement a minimal collection approach and stronger privacy techniques dictating what is needed, reducing personal identifications, and, when possible, use privacy-protecting techniques such as data-controlled access and data de-identification.
- Autonomy mechanisms that implement dynamic consent and ongoing data operation control by moving away from a forever consent form to a more dynamic consent approach, where the patients update consent forms as risks and data usage (operations) change over time.

-Alex

References:

O'Keefe, K., & Brien, D. O. (2023). Chapter 02: Introduction to ethical concepts and frameworks. *Data ethics: Practical strategies for implementing ethical information management and governance*. Kogan Page.

Seikh, R. (2025). Digital privacy and human dignity: Revisiting Kantian ethics in the age of surveillance [PDF]. *International Journal of Novel Research and Development, 10*(8), a160–a166. https://www.ijnrd.org/papers/IJNRD2508019.pdf

Sinnott-Armstrong, W. (2023). *Consequentialism*. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2023/entries/consequentialism/

Sweeney, L. (2000). *Simple demographics often identify people uniquely* (Data Privacy Working Paper No. 3) [PDF]. Carnegie Mellon University. https://dataprivacylab.org/projects/identifiability/paper1.pdf

Ulgen, O. (2017, October 31). *Kantian ethics in the age of artificial intelligence and robotics*. QIL QDI. https://www.qil-qdi.org/kantian-ethics-age-artificial-intelligence-robotics/