Chapter03EthicsAndPrivacy

January 28, 2024

1 ETHICS AND PRIVACY

1.1 Introduction

The collection and analysis of health data have become integral components of healthcare delivery and research. Technologies such as electronic health records, wearable devices, and telemedicine platforms have all revolutionized the way health data is collected, stored, and utilized. These advancements promise new opportunities for improving patient care, advancing medical research, and enhancing public health interventions. However, they also raise profound ethical and privacy concerns that demand our attention and thoughtful consideration.

The ethical difficulties surrounding health data collection are multifaceted and complex. They encompass issues of consent, data ownership, data security, and the potential for misuse or misinterpretation of data. They also touch on the fundamental human rights of privacy and autonomy, which can be jeopardized when sensitive health information is collected and shared without appropriate safeguards or permissions.

Privacy concerns, on the other hand, revolve around the potential for unauthorized access, data breaches, and the misuse of personal health information. Data is often referred to as the *new gold* and the value of health data is immense, making it a prime target for cybercriminals. Furthermore, the use of health data by third parties, such as insurance companies or employers, can lead to discrimination or stigmatization.

This chapter aims to explore the ethical and privacy concerns associated with health data collection, highlighting the potential risks and proposing strategies to mitigate them. It is crucial to strike a balance between leveraging health data for the greater good and preserving the individual's right to privacy and autonomy. The future of healthcare may be determined by our ability to navigate these ethical considerations with wisdom, foresight, and a respect for the dignity and rights of every individual.

1.2 The misuse of health data

The misuse of health data is a concern to healthcare and research, with the potential to infringe upon individual and community rights, and to erode trust in medical institutions. One of the most notable cases of such misuse involves the Havasupai Tribe of Arizona, a case that has since become a touchstone in discussions about informed consent, genetic research, and the rights of indigenous peoples.

In the early 1990s, researchers from Arizona State University (ASU) approached the Havasupai Tribe, a Native American community living in the Grand Canyon, to conduct a study on the tribe's high prevalence of Type 2 diabetes. The researchers collected blood samples from tribe

members, assuring them that the samples would be used to investigate the genetic underpinnings of the diabetes epidemic within their community.

However, without obtaining further consent, the researchers used the blood samples for additional studies, investigating genetic markers for schizophrenia, metabolic disorders, and tracing the tribe's geographical origins. When the Havasupai Tribe discovered this in 2003, they felt betrayed and violated, as the additional research conducted was not only beyond the scope of the initial consent but also challenged the tribe's cultural and spiritual beliefs.

The Havasupai Tribe sued the Arizona Board of Regents and the ASU, alleging lack of informed consent, misuse of blood samples, and violation of their civil rights. The case was settled in 2010, with ASU agreeing to pay $700,000 in damages, return the blood samples, and provide other forms of assistance to the tribe.

The case highlighted several ethical issues in health data collection and research. Firstly, it underscored the importance of informed consent, not just as a legal requirement, but as a moral obligation to respect individuals' autonomy and cultural beliefs. Secondly, it emphasized the need for transparency in research, with researchers fully disclosing their intentions and potential uses of collected data. Lastly, it brought to the fore the potential for harm when health data is misused, not only in terms of legal repercussions but also in damaging trust between researchers and communities.

The Havasupai Tribe case serves as a reminder of the ethical and privacy concerns surrounding health data collection. It shows the need for stringent safeguards, transparency, and respect for the rights and cultural beliefs of individuals and communities. This case offers valuable lessons in the importance of maintaining the delicate balance between scientific advancement and ethical responsibility.

One unintended consequence of the judgement in this case has been a perceived reluctance to conduct studies involving minorities or vulnerable communities. By excluding such people from research, researchers may be missing out on important insights that could benefit them.

Task: Find another example of the misuse of health data. Give details about the event and discuss the ethical issues involved. What could have been done to prevent this from happening?

1.3 Breaches in health data security

The security of health data has become a paramount concern. With the increasing digitization of health records, the risk of data breaches has escalated, posing threats to privacy and security. One of the most significant breaches in healthcare history occurred in 2015, involving Anthem Inc., one of the largest health insurance companies in the United States. This incident serves as a reminder of the vulnerabilities inherent in our digital health infrastructure.

In February 2015, Anthem Inc. disclosed that it had been the victim of a sophisticated cyberattack, resulting in the largest health data breach to that date. The hackers gained access to the personal information of nearly 78.8 million individuals, including names, birth dates, social security numbers, addresses, and employment information. Notably, no medical information or credit card data was reported to have been compromised.

The breach was attributed to a sophisticated external cyber attacker, with investigations suggesting state-sponsored actors. The hackers used a spear-phishing email to infiltrate Anthem's system, demonstrating the vulnerability of even large, well-resourced organizations to such attacks.

The Anthem breach had far-reaching implications. It exposed millions of individuals to potential identity theft and fraud, causing significant distress and concern. Furthermore, it highlighted the vulnerabilities in health data security and the potential scale of damage that can be inflicted by a single breach.

In the aftermath of the breach, Anthem agreed to a series of settlements. In 2017, it agreed to a \$115 million settlement to cover credit monitoring services for affected individuals, marking the largest settlement ever for a data breach. In 2018, Anthem settled with the U.S. Department of Health and Human Services' Office for Civil Rights for a record \$16 million, acknowledging the potential violation of the Health Insurance Portability and Accountability Act (HIPAA).

The Anthem data breach serves as a cautionary tale. It underscores the critical importance of robust cybersecurity measures and the need for continuous vigilance and adaptation in the face of evolving threats. The Anthem case reminds us of the pressing need to safeguard the privacy and security of health data, to maintain public trust, and to protect individuals from the potentially devastating consequences of data breaches.

Task: Find another example of a breach in health data. Describe the event, including the number of people affected and the consequences. What could have been done to prevent this from happening?

1.4   Bias in health data

In health-related research, bias can be a silent and difficult to detect intruder, skewing results, perpetuating inequalities, and potentially causing harm. One example of bias in health data involves the pulse oximeter, a common medical device used worldwide to measure oxygen levels in the blood. Recent studies have revealed that this seemingly innocuous device may harbor significant racial bias, highlighting the urgent need to address such biases in health data and technology.

Pulse oximeters are small, clip-like devices that are typically placed on a patient's finger to measure oxygen saturation in the blood, a vital parameter in many medical and surgical conditions. They are widely used in hospitals, clinics, and homes around the world. However, this correspondence in the New England Journal of Medicine in 2020 revealed a significant racial bias in the device's readings.

The study found that pulse oximeters were nearly three times more likely to miss hypoxemia, or low blood oxygen levels, in Black patients compared to White patients. This discrepancy was attributed to the fact that the algorithms used by pulse oximeters to calculate oxygen saturation were primarily based on data from light-skinned individuals. The darker skin pigmentation in Black individuals can absorb more light, potentially leading to overestimation of oxygen levels by the device.

The implications of this bias are profound. Misleading pulse oximeter readings can lead to delayed or inappropriate treatment, potentially resulting in serious harm, particularly in conditions like COVID-19 where oxygen levels are a critical determinant of disease severity.

This case serves as a reminder of the potential for bias in health data and technology. It shows the importance of diversity and representation in health data collection and algorithm development. Furthermore, it highlights the need for rigorous testing and validation of medical devices across diverse populations.

The pulse oximeter controversy illuminates the hidden biases that can lurk within health data and technology. It serves as a call to action to us all to actively seek out and address such biases.

Ensuring the fairness and inclusivity of health data is not just an ethical imperative, but a necessity for delivering safe and effective healthcare to all.

Task: Find another example of bias in health data. Describe the event and the consequences. What could have been done to prevent this from happening?

1.5 Anonymization of health data

Data anonymization or **data de-identification** is a process of protecting private or sensitive information by erasing or encrypting identifiers that connect an individual to the stored data. The primary goal of data anonymization is to make it impossible or at least extremely difficult to identify the individuals to whom the data belongs, while still allowing the data to be useful for analysis or research.

This process is crucial in fields like healthcare or social science research, where sensitive data needs to be used for studies without compromising the privacy of the individuals involved. Anonymization techniques can include data masking, pseudonymization (replacing identifiers with pseudonyms), data swapping (interchanging data between records), or adding noise to data. Below are eight techniques that can be used to achieve de-identification of data. In many cases, the original gatekeepers of the data must maintain a key to re-identify the data if necessary. This is especially required if the data is to be used for research and analytical purposes.

1. **Data Masking or Redaction:** This involves replacing identifiable information with random characters or removing it entirely. For example, a patient's name *John Doe* might be replaced with *XXXX XXXX* or removed completely.

2. **Pseudonymization:** This technique replaces private identifiers with artificial identifiers or pseudonyms. For example, a patient's social security number might be replaced with a random number that serves as a pseudonym.

3. **Data Swapping or Shuffling:** This involves interchanging values of attributes among the records in the dataset to prevent identification of individuals. For example, the ages of patients in a dataset might be shuffled so that they no longer correspond to the correct individuals.

4. **Generalization:** This technique replaces precise data with more general information. For example, a patient's exact age of 37 might be replaced with an age range of $30 - 40$.

5. **Noise Addition or Perturbation:** This involves adding random noise to the data to obscure individual values. For example, a patient's weight of 150 pounds might be reported as anywhere between 145 and 155 pounds.

6. **Data Aggregation or K-anonymity:** This technique involves grouping the data in such a way that each group contains at least k records, thereby making it difficult to single out individuals. For example, instead of reporting individual incomes, patients might be grouped into income brackets, with each bracket containing at least k patients.

7. **Synthetic Data:** This involves creating a new dataset that mimics the statistical properties of the original data but doesn't contain any actual individual's information. For example, a synthetic dataset might be generated that has the same distribution of ages, weights, and health conditions as the original dataset, but none of the synthetic records correspond to real patients.

8. **Differential Privacy:** This is a more advanced technique that adds a specific amount of noise to queries against the dataset, ensuring that the inclusion or exclusion of a single database item does not significantly change the probability distribution of the query's outcome. For example, if a researcher queries the number of patients with a certain condition, the response might be slightly altered by a random amount to protect individual privacy.

It is important to note that no anonymization technique can guarantee complete privacy. There's always a risk, known as **re-identification** or **de-anonymization**, where anonymized data can be linked back to the individual, especially when combined with other data sources. Therefore, robust strategies and ongoing vigilance are required to protect privacy in the era of big data.

Task: Find an example of anonymization of health data. Describe the technique used.

1.6   Consent in health data collection

The collection of health data plays a pivotal role in advancing medical research, improving patient care, and informing public health policies. However, this practice also raises significant ethical concerns, chief among them being the issue of consent. Ensuring informed consent in health data collection is not merely a legal requirement but an ethical imperative that respects individual autonomy and privacy.

1.6.1   The importance of consent

Consent is a fundamental principle in healthcare and research. It is an expression of respect for the autonomy and dignity of individuals, acknowledging their right to control what happens to their bodies and personal information. In the context of health data collection, consent ensures that individuals are aware of and agree to the collection, storage, and use of their health data.

1.6.2   Informed consent

Informed consent goes a step further, requiring that individuals are provided with sufficient information to make an educated decision about their participation. This includes information about the purpose of the data collection, how the data will be used, who will have access to it, the potential risks and benefits, and the individual's right to withdraw their consent at any time.

1.6.3   Challenges in Obtaining Consent

Obtaining informed consent for health data collection is not without challenges. The complexity of data use in the digital age can make it difficult for individuals to fully understand the implications of their consent. Furthermore, the use of broad or blanket consent, where individuals agree to unspecified future uses of their data, raises concerns about whether such consent can truly be informed.

Moreover, in certain situations, obtaining consent may not be feasible or practical, such as in the use of large-scale anonymized datasets for research or public health surveillance. In these cases, alternative safeguards, such as robust de-identification and strict data access and use policies, are crucial to protect individual privacy.

Consent is a cornerstone of ethical health data collection. It empowers individuals, fostering trust in healthcare providers and researchers. However, the complexities of modern data use necessitate a nuanced approach to consent, one that combines clear, accessible information with robust privacy

safeguards. As we navigate the evolving landscape of health data, the principle of consent remains a guiding light, reminding us of the inherent dignity and autonomy of each individual whose data we hold.

Task: Find an example of a *pro-forma* consent document for health data collection.

1.7 Institutional Review Boards

Maintaining ethical standards in health-related research is paramount. Institutional Review Boards (IRBs), also known as Ethics Committees in some countries, play a crucial role in ensuring that these standards are upheld. They serve as the gatekeepers of research ethics, protecting the rights and welfare of research participants.

1.7.1 Role and Function of IRBs

IRBs are independent committees composed of diverse members, including scientists, non-scientists, and community members, who review and monitor research involving human subjects. Their primary responsibility is to ensure that such research is conducted ethically and that the rights and welfare of the participants are adequately protected.

Before a study involving human subjects can commence, the research protocol must be reviewed and approved by an IRB. The IRB assesses the study's design, methodology, and procedures, focusing on the risk-benefit ratio, the informed consent process, and the measures in place to protect participants' privacy and confidentiality.

1.7.2 Ongoing Responsibilities of IRBs

The responsibilities of an IRB extend beyond the initial approval of research. They are also tasked with ongoing monitoring of approved studies, reviewing progress reports, and assessing any proposed changes to the research protocol. If a study poses unexpected risks or if there are instances of non-compliance with ethical standards, the IRB has the authority to require modifications or even suspend or terminate the study.

1.7.3 Challenges Faced by IRBs

Despite their crucial role, IRBs face several challenges. These include managing potential conflicts of interest, dealing with the complexities of reviewing research in rapidly evolving fields like genomics or artificial intelligence, and ensuring adequate representation and understanding of diverse populations and vulnerable groups. Addressing these challenges requires ongoing education, transparency, and a commitment to ethical principles.

Institutional Review Boards stand at the forefront of ethical research, safeguarding the rights and welfare of research participants. Their role is vital in maintaining public trust in the research enterprise and ensuring that the pursuit of knowledge does not compromise the dignity and well-being of individuals. As we continue to push the boundaries of scientific inquiry, the work of IRBs remains more important than ever, reminding us of our ethical obligations to those who make our research possible.

Task: Read more about The George Washington University Office of Human Research.

1.8 Conflicts of interest

A **conflict of interest** in public health research refers to a situation where a person or organization involved in the research has a secondary interest that could unduly influence or bias the conduct, findings, or reporting of the research. This secondary interest could be personal, financial, political, academic, or organizational.

For example, a researcher might have a financial conflict of interest if they own stock in a company that could benefit from the results of their research. Similarly, an academic conflict of interest could occur if a researcher feels pressure to produce certain results to secure promotion or tenure.

Conflicts of interest can threaten the integrity of public health research, as they can lead to biased results, loss of objectivity, and a breach of trust with research participants and the public. Therefore, it's crucial that any potential conflicts of interest are disclosed and managed appropriately to maintain the credibility and integrity of the research.

Conflicts of interest can arise in various ways.

1. **Financial Conflicts**: These are the most commonly recognized conflicts of interest, where a researcher may benefit financially from the outcomes of their research. This could be through owning stocks in a company that sponsors their research, receiving consulting fees, or obtaining patents.

2. **Personal Conflicts**: These occur when researchers have personal relationships that could influence their research. This could include relationships with colleagues, students, or family members.

3. **Academic Conflicts**: These occur when researchers have academic incentives that could influence their research, such as pressure to publish or to secure funding.

1.8.1 Implications of Conflicts of Interest

Conflicts of interest in public health research represent a significant concern, with potential implications for the integrity of research, the validity of findings, and public trust in health information. A conflict of interest occurs when a secondary interest, such as financial gain, personal advancement, or institutional benefit, could unduly influence professional judgment or actions. This essay explores the impact of conflicts of interest in public health and discusses strategies for their management.

Conflicts of interest can compromise the integrity of public health research. If researchers have a vested interest in the outcome of their studies, consciously or unconsciously, it might bias their approach to the design, conduct, analysis, or reporting of the research. For instance, a researcher with financial ties to a pharmaceutical company might be inclined to design a study in a way that favors the company's drug or to interpret ambiguous results positively. This could lead to skewed findings, impacting the validity of the research and potentially influencing health policies or clinical practices based on flawed evidence.

Conflicts of interest can also erode public trust in public health research and institutions. When conflicts of interest are revealed, particularly if they were not disclosed initially, it can lead to perceptions of dishonesty or manipulation. This can undermine public confidence in research findings and decrease public engagement with health initiatives or interventions. In an era where misinformation can spread rapidly, maintaining public trust in public health research is paramount.

Conflicts of interest in public health research can have serious implications, which are summarized below.

1. **Bias in Research**: Conflicts of interest can introduce bias into research, affecting the design, conduct, and reporting of research. This can lead to distorted findings and undermine the validity of the research.

2. **Loss of Public Trust**: Conflicts of interest can erode public trust in research. If the public perceives that researchers are influenced by personal interests, they may be less likely to trust research findings.

3. **Damage to Research Integrity**: Conflicts of interest can damage the integrity of the research process and the reputation of the research community.

1.8.2 Managing Conflicts of Interest

Conflicts of interest in public health research, where a secondary interest could unduly influence professional judgment or actions, pose significant threats to the integrity of research and public trust. As such, managing these conflicts is of paramount importance. This essay explores strategies for mitigating conflicts of interest in public health research.

The first step in managing conflicts of interest is transparency. Researchers should disclose any potential conflicts of interest at the outset of a study, including financial ties, personal relationships, or institutional affiliations that could influence their work. This disclosure should be updated throughout the research process and included in any publications or presentations resulting from the research. Transparency allows others to evaluate the potential influence of the conflict and helps maintain trust in the research process.

Institutions conducting public health research should have robust policies in place to manage conflicts of interest. These policies might include guidelines for disclosure, rules about financial relationships with industry, or procedures for managing conflicts when they arise. Institutions should also provide training on these policies to ensure that all researchers understand their responsibilities.

Independent oversight can be a valuable tool for managing conflicts of interest. This might involve having a separate committee review research proposals for potential conflicts, monitor ongoing research, or review findings before publication. Independent oversight can provide an additional layer of scrutiny and help ensure that conflicts of interest do not compromise the research.

Strategies to mitigate conflicts of interest are summarized below.

1. **Avoidance:** In some cases, the best way to manage a conflict of interest is to avoid it altogether. This might involve choosing not to participate in certain research or severing financial ties with industry.

2. **Divestment:** If a financial conflict of interest exists, the researcher might be required to divest from the financial interest causing the conflict.

3. **Management Plans:** For conflicts that cannot be avoided, a management plan can be developed. This might involve regular oversight, separating decision-making roles, or other strategies to mitigate the conflict.

4. **Public Disclosure:** Disclosing the conflict of interest to the public, particularly in any publications or presentations, can help manage the conflict by allowing others to evaluate the

potential influence.

5. **Institutional Policies**: Institutions should have clear policies for identifying and managing conflicts of interest. This could include processes for disclosure, review, and management of conflicts of interest.

6. **Education and Training**: Researchers should receive education and training on conflicts of interest, including how to identify, disclose, and manage them.

1.9 Quiz questions

1.9.1 Questions

1. What is the primary goal of data anonymization?

2. What is the Havasupai Tribe case about?

3. What was the largest health data breach in history?

4. What is the pulse oximeter controversy about?

5. What is the role of an Institutional Review Board (IRB) in research?

6. What is the difference between spreadsheet software and database software for health data management?

7. What is a conflict of interest in public health research?

8. What is informed consent in the context of health data collection?

9. What is data masking in the context of de-identification of health data?

10. What is the principle of k-anonymity in the context of de-identification of health data?

11. What is the main concern about conflicts of interest in public health research?

12. What is the first step in managing conflicts of interest in public health research?

13. What is the role of independent oversight in managing conflicts of interest in public health research?

14. What is one strategy for mitigating conflicts of interest in public health research?

15. What is the potential impact of not properly managing conflicts of interest in public health research?