Ethics, Privacy, etc.

Peter Szolovits
6.872/HST.950
Treatment of Human Subjects: The Belmont Report 1979

Ethical Principles and Guidelines for the Protection of Human Subjects of Research

- Balancing (societal) benefits vs. (individual) risks
- History of abuses
  - Nazi “experiments” ⇒ Nuremberg code
- Tuskegee syphilis study
Nazi Medical Experiments

- Freezing / Hypothermia
- Genetics
- Infectious Diseases
- Interrogation and Torture
- Killing / Genocide
- High Altitude
- Pharmacological
- Sterilization
- Surgery
- Traumatic Injuries

A cold water immersion experiment at Dachau concentration camp presided over by Professor Ernst Holzlöhner (left) and Dr. Sigmund Rascher (right). The subject is wearing an experimental Luftwaffe garment.


© unknown. All rights reserved. This content is excluded from our Creative Commons license. For more information, see http://ocw.mit.edu/fairuse.
Tuskegee Syphilis Experiment

- 1932-1972 experiment to study natural progression of disease
- 399 African-American sharecroppers w/syphilis
- failed to treat even after penicillin was shown to be an effective treatment in 1940’s

http://en.wikipedia.org/wiki/Tuskegee_syphilis_experiment
Practice & Research

• The term “practice” refers to interventions that are designed solely to enhance the well-being of an individual patient or client and that have a reasonable expectation of success.

• The term “research” designates an activity designed to test an hypothesis, permit conclusions to be drawn, and thereby to develop or contribute to generalizable knowledge.

• Research and practice may be carried on together when research is designed to evaluate the safety and efficacy of a therapy. ... if there is any element of research in an activity, that activity should undergo review for the protection of human subjects.

From the Belmont Report
Basic Ethical Principles

- Respect for Persons
- Beneficence
- Justice

From the Belmont Report
Respect for Persons

• Each person is an autonomous agent, capable of deliberation about personal goals and of acting under the direction of such deliberation

• Persons with diminished autonomy are entitled to protection: e.g., children, physically or mentally disabled, prisoners.

• Requires *Informed Consent*
  • Adequate information
  • Voluntary participation

From the Belmont Report
(Informed Consent)

- Study involves research, purpose of research, duration, procedures, what is experimental?
- Foreseeable risks and discomforts
- Possible benefits to participants or others
- Alternative procedures that might be beneficial
- How confidentiality will be maintained
- For research involving more than minimal risk, what compensations and treatments may be available, and where to get further information
- Participation is voluntary; no penalty for refusal

http://www.hhs.gov/ohrp/policy/consent/index.html
Beneficence

- Do no harm
  - one should not injure one person regardless of the benefits that might come to others
  - minimize risk to participants
- Maximize possible benefits
  - to society
  - but, research subjects may not benefit directly
- Some tradeoffs are unavoidable

From the Belmont Report
Justice

• Varied views of equal treatment
  • equal share
  • individual need
  • individual effort
  • societal contribution
  • merit

• Select participants fairly
• Distribute benefits fairly

From the Belmont Report
Enforcement: The Common Rule

- Applies to all US Government funded projects involving human subjects
- Institutional Review Boards (IRB) review and must approve all such proposed research; responsible to protect subjects
  - yearly review of research protocols, informed consent, training of researchers, etc. Criteria of Belmont Report.
  - expedited review for research involving “no more than minimal risk”; consent may be waived
  - exemptions for educational research, food quality research, and retrospective research on public or de-identified data
- IRB’s also responsible for protection of confidentiality
- MIT’s IRB is the Committee on Use of Humans as Experimental Subjects (COUHES)

http://www.hhs.gov/ohrp/policy/consent/index.html
Privacy vs. privacy
Protecting…

• What?
  • Privacy
    • Individual’s desire to limit disclosure of personal information
  • Confidentiality
    • Information sharing in a controlled manner
  • Security
    • Protecting information against accident, disaster, theft, alteration, sabotage, denial of service, …

• Against what?
  • “Evil hackers”
  • Malicious insiders
  • Stupidity
  • Information Warfare
Privacy

- Right to be let alone; e.g.:
  - snooping on Dan Quayle by J. Rothfeder
  - “outing” of Arthur Ashe (HIV), Henry Hyde (adultery)
  - celebrity medical problems (Tammy Wynette, Nicole Simpson)

- … applies mostly to known individuals
Privacy in obscurity

• Right to remain unknown

• Correlation among pervasive databases:
  • census
  • marketing
  • health
Confidentiality

- Use and sharing of information by multiple users at many institutions
- Should be controlled by coherent policy
- Enforced by appropriate technology

- E.g., who may use results of your life insurance physical exam, for what purposes?
Legitimate Concerns
(some may be ameliorated by ACA)

- Difficulty getting insurance
  - “Individual insurers may deny you coverage based on your medical history if it includes:
    - Use of prescription drugs to treat anxiety, depression or a physical condition, including Ativan, Klonipin, Paxil, Prozac, Serzone, Zoloft, Xanax and Wellbutrin.
    - Counseling for anxiety, depression, grief or an eating or sleep disorder. Even if you briefly sought counseling as a way to cope with the Sept. 11 terrorist attacks, you could be denied individual health insurance, according to researchers with Georgetown's Health Privacy Project.” (MSN, March 9, 2004)
  - Medical Information Bureau
    - Data on all applicants for private life insurance in past 7 years
Additional Legitimate Concerns

- When employer pays insurance premiums, you may lose your job
  - Self-insured companies
  - Small employers facing “experience rated” policies
- Non-employment discrimination based on health
  - Adoption
  - Politics
- Social stigma
Security

- Integrity of data
  - No unauthorized modifications
  - No “dropped bits”
- Availability
  - Natural disaster
  - Adversary attack
  - Inadequacy of backup, fail-over
- Enforcement of confidentiality policies
De-Identification
Identifiable

- **HIPAA**: Name, address, phone number, fax number, email address, URL, IP address, social security number, medical record n., health plan n., account n., certificate/license n., vehicle id, device id, biometric id, full-face photo, date of birth, zip code, gender, race, profession
  - “any other unique identifying number, characteristic, or code”
  - “actual knowledge that the information could be used … to identify”
- Patterns of doctor visits, immunizations, etc.
  - identifiable by inference
  - depends on knowledge and abilities of data user
- Small bin sizes lead to identifiability
  - Aggregate data into larger bins
    - dob => age
    - 3 digits of zip code
Sweeney’s Cambridge

- 1997 Cambridge, MA voting list on 54,805 voters
  - Name, address, ZIP, birth date, gender, …
- Combinations that uniquely identify:
  - Birth date (mm/dd/yy) 12%
  - BD + gender 29%
  - BD + 5-digit ZIP 69%
  - BD + 9-digit ZIP 97%
- Unique individuals
  - Kid in a retirement community
  - Black woman resident in Provincetown
Problem of “other information”

- Governor Weld’s data found in Mass “de-identified dataset”
- Dates you visited a health care provider (over a lifetime) are probably unique
- Can be used to re-identify you if someone has both de-identified data and other data that link to identifiers
- Genetics makes this immensely more problematic
  - Think Gattaca
Danger of Re-identification

Figure 18 Release using de-identification

Figure by Sweeney, Latanya. "Computational disclosure control: A primer on data privacy protection." *Massachusetts Institute of Technology*, 2001.
Protection via generalization

Figure by Sweeney, Latanya. "Computational disclosure control: A primer on data privacy protection." Massachusetts Institute of Technology, 2001.
Computational Disclosure Control

- Make sure data cannot be traced back to a set of size \(< n\)
  - Generalization
  - Suppression of unique combinations
  - Account for leakage from what has been suppressed; e.g., back-calculating from aggregate statistics
- How to estimate “external information”?
- **Every** release becomes more external info.
Methods of Generalization/Suppression

- Underlying problem (find minimal generalization/suppression to achieve a level of anonymity) is NP-hard (Vinterbo)
- Mainly heuristic search over space of possible generalizations/suppressions
  - Scrub, Datafly, µ-Argus (Netherlands), k-Similar
- Lasko: spectral anonymization
  - Build a model of data that captures the n-th order statistics of the distribution
  - Synthesize “fake” patients from that distribution