Citizen Science: The Law and Ethics of Public Access to Medical Big Data

Sharona Hoffman
Professor of Law & Bioethics
Case Western Reserve University
Obama Executive Order – May 9, 2013

- Making Open and Machine Readable the New Default for Government Information
- To the extent permitted by law, government must release its data to the public in forms that make it easy to find, access, and use.
Publicly Available Databases

**Federal & State**
- HealthData.gov
- CDC Wonder
- Chronic Condition Data Warehouse
- Healthcare cost and Utilization Project
- GenBank
- All-Payer Claims Databases

**Private**
- Dryad Digital Repository
- PatientsLikeMe
- The Personal Genome Project
Benefits of Public Access

- Scientific Discovery
- Research Cost Reduction
- Tools to help patients navigate the healthcare system
- Government transparency and public education
- Improvements in healthcare quality and public health policy
Privacy Risks

- HIPAA Privacy Rule
- State laws
- HIPAA guidance re: de-identification
- **BUT** HIPAA doesn’t cover many data holders
- Re-identification risk
- Genetic information
Other Risks

• Discrimination by employers
  • Identifiable or re-identifiable information
  • De-identified data re: “undesirable” groups

• Financial institutions

• Direct targeting of consumers
  • “How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did”

• Propagation of incorrect and harmful research results
Litigation

• Defamation
  • Oprah Winfrey beef case
• Interference with economic advantage
• Public disclosure of private facts
Recommendations

• HIPAA revision
  • Expand covered entities & create national data release & de-identification standards
  • Prohibit re-identification
• Data Release Review Boards
• Data Use Agreements, privacy training, and registries
• Consent procedures for those posting own data on public websites
More recommendations

• Revise Americans with Disabilities Act to include those predicted to have disability in future
• Citizen scientist chaperoning
  • Wikipedia model