The Broad Ethical Perspectives of Data Sharing

Wael Al-Delaimy MD, PhD
Chief, Division of Global Health
University of California, San Diego
Data Sharing

• Research-based
• Clinical effectiveness and quality improvement
  – Prevention
  – Diagnosis
  – Treatment
  – Research to advance practice
  – Teaching and training
Data Sharing

• Risk-benefit balance
  - Risk of loss of privacy and confidentiality
  - Risk of infringing on the autonomy of subjects
  - Risk of unethical doubt manufacturing and conflict of interest
  - Benefit of health advancement and prevention
  - Benefit of Cost-effectiveness
  - Better research and study design in pooling data
  - Sustaining disciplines such as environmental epidemiology
Data Sharing

Research Data type

• Questionnaire/written data
  – Re-analyses
  – New hypothesis

• Biomonitoring/Biobank data
  – Exposure assessment (new, re-analyze old)
  – Genetic/Omics assessment (new, re-analyze old)

• Linking datasets to create a new data set
  – Gene X Environment,
  – Medical records X Environmental exposures
## Personal Identifiers

<table>
<thead>
<tr>
<th>Name</th>
<th>Vehicle ID</th>
<th>ZipCode/address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phone</td>
<td>License number</td>
<td>Medical Record</td>
</tr>
<tr>
<td>Fax</td>
<td>IP Address #</td>
<td>Health Plan number</td>
</tr>
<tr>
<td>email</td>
<td>Account #</td>
<td>age</td>
</tr>
<tr>
<td>SS</td>
<td>URL</td>
<td></td>
</tr>
<tr>
<td>Biometrics</td>
<td>Device identifiers</td>
<td></td>
</tr>
<tr>
<td>Face image</td>
<td>Other Unique identifiers</td>
<td></td>
</tr>
</tbody>
</table>
Personal Identifiers

- All geographical subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes,
- except for the initial three digits of a zip code, if:
  1. The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and
  2. The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000.
The Players

• The researcher
  – The original investigator
  – The secondary investigator

• The subject and community

• The organization
  – Funding organization
  – The owner/storing organization
Researcher

• Original PI
  – Hypothesis
  – Appropriate design to answer the hypothesis
  – Consent
  – Ownership
  – Additional benefit from data sharing if involved
  – Burden and cost of data mining or maintenance
  – Transparency and not obstructive
Researcher

• Secondary PI
  – Hypothesis driven or post-hoc analyses
  – Can the original design appropriately address the question?
  – Conflict of interest (litigation, commercial)
  – Consent of original subjects for this analyses
  – Sharing of costs
  – Level playing field: reciprocal data sharing for private entities
Subject

– Volunteered their data and do not own it
– Autonomy to decide which study uses their data
– May not have consented if they were asked for an open consent
– Promised results and feedback
– May not be reached to be re-consented
– Concerned about privacy and confidentiality when changing hands
The Community

- Requires involvement if identifiable
- Indigenous populations (Havasupai Indians)
- Stigma
- Suffer Environmental Injustice
- Involvement in the original study approval process
- Can delay any data sharing
- Expects feedback and results
Data Sharing Organizations

- HMO, e.g. Kaiser, Medicare (HIPAA, Data Linkage)
- State and County Departments of Health (IRB, State Regulations, inefficiency)
- Registries (HIPAA, Bureaucracy, inefficiency)
- Federal Government (EPA, CDC, NCHS) (Bureaucracy and politics, WB)
- Universities (Industry influence, IP, Revenue generating)
- Industry (Conflicts of interest)
- NIEHS (Funding)
Data Sharing Access

• Data users
  – Environmental sciences need it the most
    • Climate change
    • Water contamination
    • Desert storms

• Data protectionists
  – Growing identity theft
  – Business of selling personal information

• Data obstructionists
  – Cell phone companies
  – Faulty regulations
Faulty Regulations

Pesticide levels among farmworkers

- UCSD IRB
- County Department of health
- Local community organization
- Farmworkers coalition
- Binational border health environmental task force
- California Department of Pesticide Regulation
Opposite Sides of the Table

• Data Owner
  - Requiring Zip codes for a publically available data set

• Data Seeker
  - Re-contacting pre-consented participants
International Setting

• Reconciling differences in regulations
• Lack of local IRBs and untrained researchers
• Transfer of data across borders
• The human genome
  – Quality; accessibility; responsibility of funders, generators, and users; security; transparency, accountability; integrity
• Data Sharing maximize global public benefit

Knoppers et al 2011
Re-Analyses Guidelines

• Protecting the Public’s interest
  – Cooperation of original authors, declaring conflicts of interest, independent advisory board created, agree on the hypothesis, proposal, results published regardless

• Protecting the Rights of Subjects
  – Respect privacy and consent for re-analyses

• Protecting the Right of the original and re-analyzing authors
  – Data ownership, open communication with original author, opportunity to comment before publication, allow original authors to publish first, providing funding

• Funding agencies establish guidelines on storage and access of data for secondary analyses

Neutra et al 1996
Public Health Surveillance Data

• Cancer Registries, Birth and Death Records
• Abuse of HIPAA and other Federal regulations
• More restrictive than anytime before
• De-identified linkage of data denied
• VA health data not reported to registries
• Negative impact on public health research and Practice
• Credit Agencies vs Health scientists
• Scientific community and professional organizations need to act

Wartenberg & Thompson 2010
Data Sharing Ethical Challenges

- Inconsistency across IRBs and states
- Requiring multiple IRB approvals
- Verification of conflicts of interest
- Overcoming ownership obstacles
- Publication credit of Original vs Secondary PI
- General reluctance of participants for open consent
- Prioritizing the use of finite biobank samples
Concluding Remarks

• The NIEHS leverage as a funding agency
• Oversight of data sharing
• Provide funding for the process
• Sub committee to review ethical aspects
  – facilitates local IRB approval
  – Address conflicts of interest
  – Synchronize with the original data design and consent process
  – Data exchange and linkage
Concluding Remarks

• Risk Stratification checklist
• Pro-data users
• Research ethics training
• Unethical conduct of research happens
• The researcher bears the ultimate moral responsibility towards the study subjects and the integrity of his/her work