(Genomic) Data Sharing, Privacy, and Academic Research

Amy L. McGuire, JD, PhD

Baylor College of Medicine

CENTER FOR MEDICAL ETHICS & HEALTH POLICY

Genomics Data Sharing Policies and Privacy

- History: rapid public release of all generated sequence data

1991 NHGRI and DOE data release policy
1996 Bermuda Principles
2000 NHGRI policy extension
2003 Ft. Lauderdale Principles
2003 NHGRI policy

- Developed in the context of large-scale sequencing studies (HGP, HapMap)
  - Primary purpose: create a reference dataset
  - Cost efficient
  - Promotes scientific utility

Examples of Existing Open Access Databases

- Human Genome Project
- HapMap
- 1000 Genomes
- Personal Genomes
  - PGP, Venter, Watson
- Human Microbiome Project

NIH GDS Policy: non-human genomic data should be made publicly available no later than date of initial publication

Human Genomic Data Sharing and Privacy

- OHRP (2008): research using de-identified biospecimens and data ≠ human subjects research because not “readily identifiable”

2004: “Specifying DNA sequence at only 30 to 80 statistically independent SNP positions will uniquely identify a single person.”

2008: Can uniquely identify an individual from pooled or aggregate DNA data.

“Surnames can be recovered from personal genomes by profiling short tandem repeats on the Y chromosome (Y-STRs) and querying recreational genetic geneology databases.”

Policy Response (part 1): Shift to Restricted Databases (dbGaP)

- NIH Genomic Data Sharing (GDS) Policy
  - Prospective studies: NIH expects specific language about sharing genotype and phenotype data in the ICD
  - For retrospective studies: Submitting institution ensures data sharing is not inconsistent with informed consent
    - Role of RACs and other governance mechanisms

Policy Response (part 2): NPRM

Biospecimen: secondary research using de-identified biospecimen = human subjects res ; need broad consent

Data: secondary r identified data ≠ human subject

WHY?

Best policy p obably lies somewhere in the middle

Complicated exclusions and exemptions

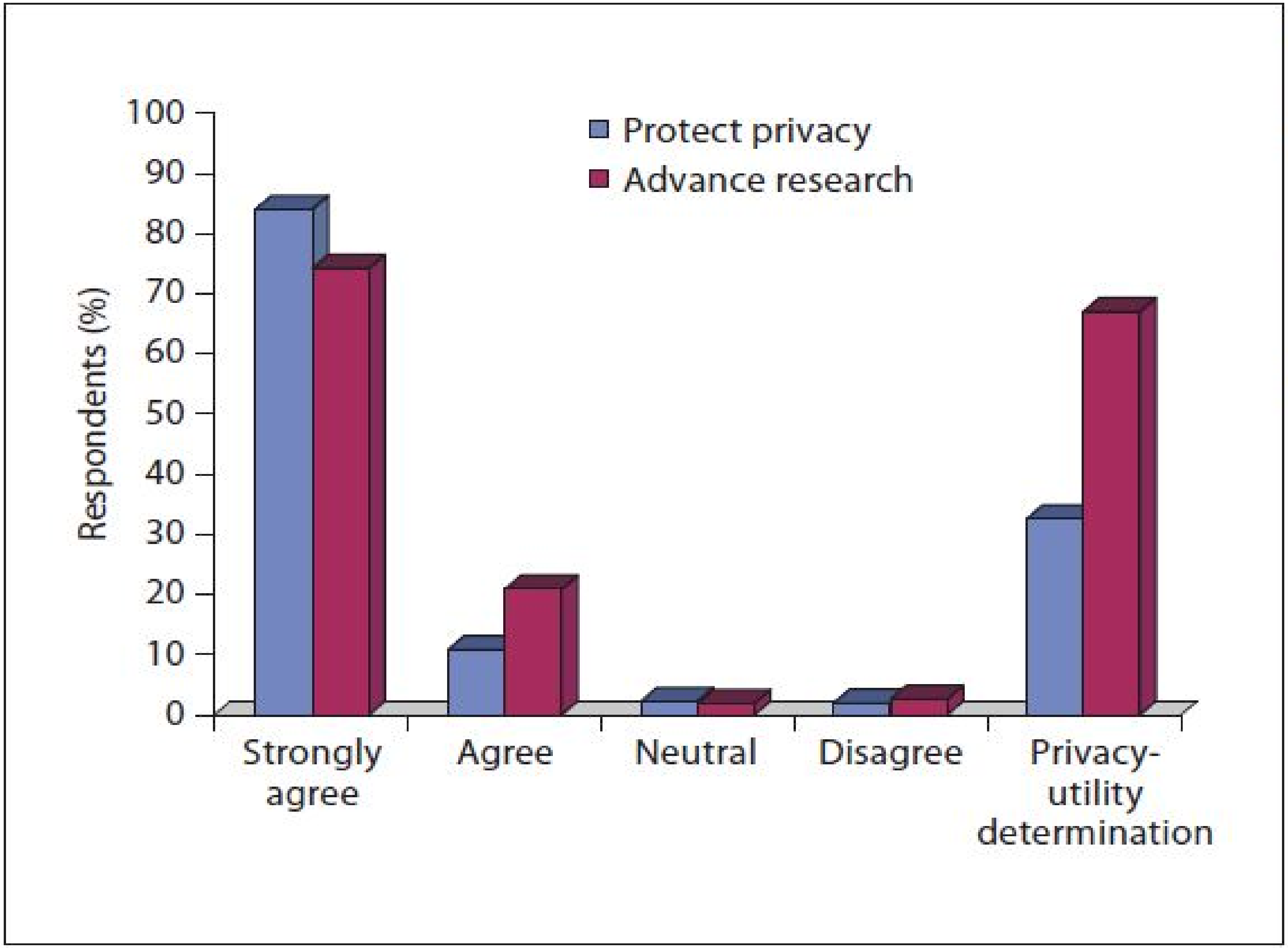
will make compliance a challenge

Consent for Data Sharing (DNA and clinical data)

	Before Debriefing (n=335)			**After Debriefing**		
	Public Release	*Restricted Release*	*No Release*	*Public Release*	*Restricted Release*	*No Release*
All Participants	83.9	6.6	9.6	**53.1**	**33.1**	**13.7**
Consent Type						
Traditional	**100.0**	0.0	0.0	62.1	31.9	6.0
Binary	84.9	0.0	**15.1**	50.9	30.2	18.9
Tiered	66.4	**19.5**	14.1	46.0	37.2	16.8

A. McGuire (PI), 1 R01 HG004333
McGuire et al. *Genetics in Medicine* 2011

Privacy-Utility Trade-Off



Oliver J et al. Balancing the Risks and Benefits of Genomic Data Sharing: Genome Research Participants’ Perspectives. *Public Health Genomics* (2012).

Factors Influencing Data Sharing Decision

- Groups statistically more likely to choose restricted data release:
  - Hispanic participants
  - Unmarried participants, including those who were divorced, widowed, separated, or never married
  - Participants with some college or a college degree
  - Participants providing parental consent
- Other patient groups may feel differently (e.g., disadvantaged populations, HIV+ patients)

McGuire AL *et al.* To Share or Not to Share: A Randomized Trial of Consent for Data Sharing in Genomic Research. *Genetics in Medicine* (2011); Robinson JO *et al*. It Depends Whose Data are Being Shared. *J Law and Biosciences* (2014)

Desire for Involvement

- The majority (86.4%) of participants reported it is important for them to be involved in the decision about whether to share their genetic information.
  - Why: RESPECT
- Opinion ≠ Policy (Ellen Clayton)
- Respect for persons versus respect for autonomy
  - Consent
  - Engagement, Transparency, Accountability

Translational Research

Proprietary Databases

The Free the Data Program

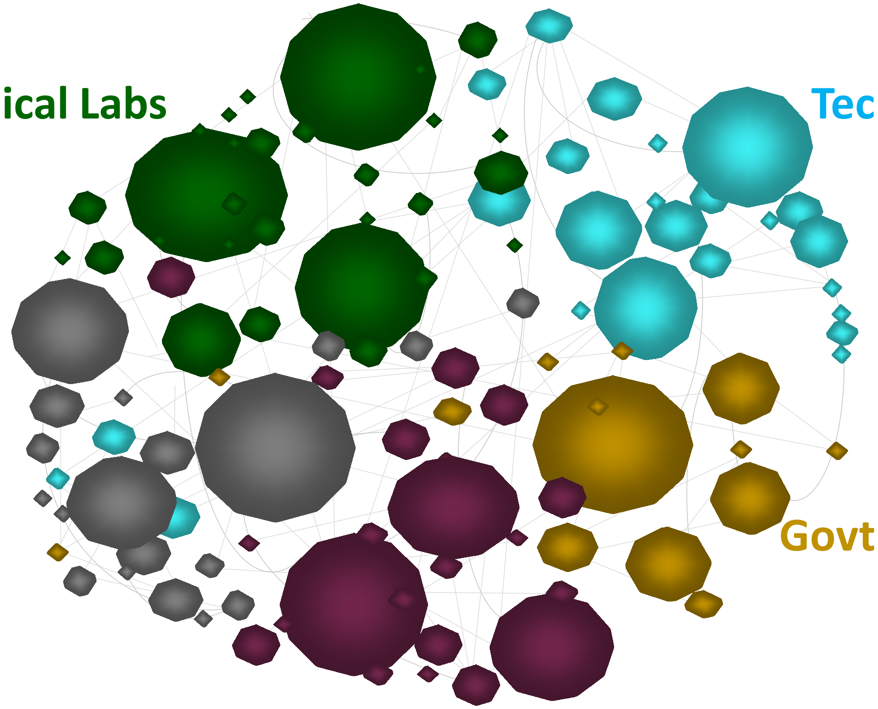
Clinical Genomic Data Sharing

- Delphi Panel with 43 experts (payers, policy makers, researchers, patient advocates, lawyers, funders)
  - Data sharing was the most important and least politically tractable policy challenge to clinical translation
- Potential policy solutions
  - Make data sharing a condition of approval/clearance, certification, or approval (FDA, CLIA, CAP)
  - Payers reimburse more for tests from laboratories that share data
  - Make data sharing a requirement for NIH funding

Moving Beyond Genomic Data…

...And Academic Research

Building the Medical Information Commons



HC Systems & Clinical Labs

Tech Companies

NGOs

Govt Programs

Research Consortia

A. McGuire, B. Cook-Deegan (Pis), R01 HG008918

What is Needed

- Harmonization of regulations and reporting requirements
- Clear ethical framework to guide policy making and decision making
  - National Commission did not anticipate these considerations
  - No framework of normative bioethics behind NRPM (Barbara Evans)

Baylor College of Medicine

CENTER FOR
MEDICAL ETHICS
& HEALTH POLICY

Thank You