

Sharing Genomic (and Other) Data: The Perspective of the Kaiser Permanente Research Program on Genes, Environment and Health

Catherine Schaefer, PhD



Secretary's
Advisory
Committee on
Genetics, Health
and Society

February 4, 2010

Washington, DC

Division of Research

Overview

- Kaiser Permanente Northern California (KPNC) is developing a comprehensive resource for research on genetic and environmental influences on health
- The resource will link data on 500,000 members of KPNC, including
 - Comprehensive clinical and administrative data from electronic medical records (EMR)
 - Participant surveys
 - Environmental exposures based in GIS database
 - Genetic, biomarker, and environmental data from biospecimens

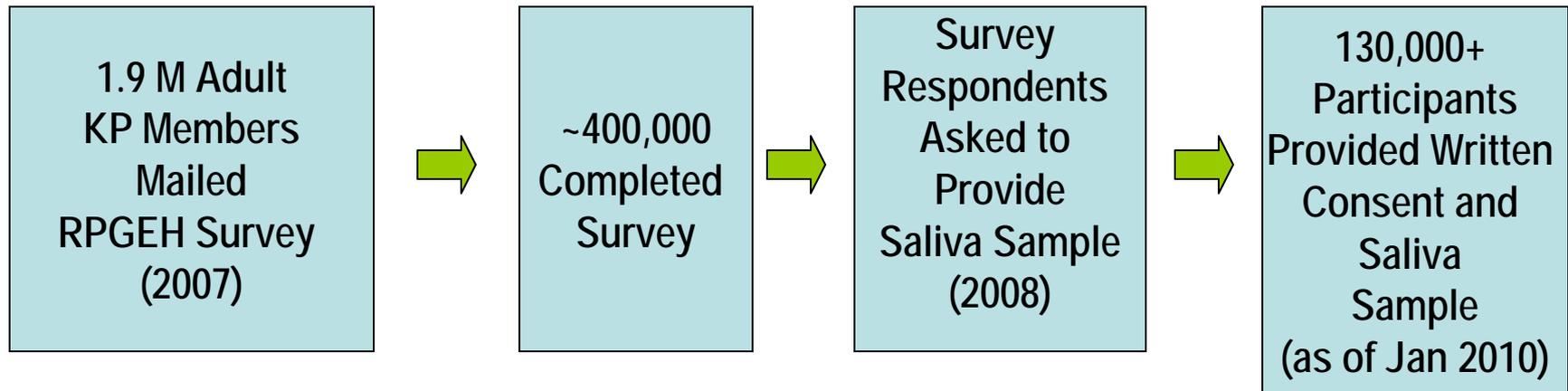
- The aims are to enable scientists to conduct research on genetic and environmental influences on
 - **Disease susceptibility**
 - **Disease course, prognosis, and outcomes**
 - **Response to treatment (pharmacogenetics)**
- Conduct research to translate findings into improvements in medical care and public health
- Conduct research on the ethical, legal and social implications of genetic research and the use of genomic information in medical care

A Brief History of the RPGEH

- Initial funding in 2005-2007
 - Engaged KPNC membership through focus groups, internal communications, and media
 - Organized Community, Scientific, and Bioethics Advisory Panels
 - Organized EMR data by disease groups to facilitate research

- Development of the RPGEH has been funded by Kaiser Permanente and grants from the Robert Wood Johnson Foundation, the Wayne and Gladys Valley Foundation, and The Ellison Medical Foundation.

Development of the RPGEH Cohort (survey & samples)



RPGEH survey includes questions regarding demographics, health history, family history, smoking, alcohol, diet, physical activity, and reproductive history and health (available in English, Spanish, and Chinese).

Key Current Activities

- Continuing enrollment and sample collection
 - 200K by year-end 2010; 500 K by year end 2013
- Collection of blood samples using KPNC clinical infrastructure
- Continuing work on funded GWA studies: multi-ethnic study of bipolar disorder; study of prostate cancer among African Americans
- Collaborations portal and access review committee receiving applications in 2010

Key Current Activities

- Genome-wide genotyping of 100,000 participants by year-end 2011, funded by GO grant from NIH
 - Measurement of telomeres
 - Average age is 65; resource for study of age-related diseases, healthy aging, and longevity
 - Accessible through dbGaP and collaborations with RPGEH
 - May require re-consent for deposit of data in dbGaP

Considerations for Data Sharing

- **RPGEH is a unique resource**
 - Large, diverse and generally representative of population
 - Comprehensive, continuously updated EMR enables excellent phenotypic characterization and follow-up
- **KP sees that RPGEH can make an important contribution and wants to ensure that the best and broadest use is made, consistent with commitment to members**
- **KP/RPGEH perspective on data sharing is shaped by commitment to members; we are invested in them and they determine KP future**
 - Trust in KP by our members enables us to do research
 - Over 50% of first 100K participants in RPGEH have been members for more than 20 years

Data Sharing

■ Broad consent

- No restrictions on kinds of health data that can be studied
- Health information can be updated from EMR
- All studies must be IRB-approved

■ Use and share genomic data for research only

- Research results will not be placed in medical record
- Participation is confidential
- Genomic data will not be returned to individuals or their providers
- Consent form does include option to contact participant if information is developed that has significance for health
- Participant may withdraw and ask that sample be destroyed
 - How do we ensure these commitments are met when data are used through public database

Data Sharing

- **Significant concern about data sharing through federal database such as dbGaP**
- **Community Advisory Panel, focus groups, and some survey respondents are concerned that the government may “take” and misuse data**
 - **Building of other federal DNA databases increases perceived vulnerability of NIH database**
 - **Use of DNA to deny treaty rights or label immigrants is community concern**
 - **Perception that KP/ RPGEH storage and control of data give participants better recourse and control**
 - **Concern that data may be misused to stigmatize vulnerable group**

Sharing Phenotypic Data

- In most research contexts, sharing genomic data means sharing phenotypic data, since interest is in linking genomic data with health
 - **Must also consider factors affecting sharing other types of data (e.g., re-identification and discrimination or stigma)**
 - **Health plans with EMRs have huge investments in these data**
- Quality of phenotypic data is critical to best use for discovery and characterization
- Concern that data may be misused to stigmatize or denigrate a vulnerable group
 - **Best use of data depend on knowledge of system that generated the data --**
 - **Important to involve Investigators who know the data**

RPGEH Key Personnel

Catherine Schaefer, PhD, Director

Neil Risch, PhD, Co-Director

Mary Henderson, MBA, MPH, Managing Director

Sarah Rowell, MPH, Research Operations

Larry Walter, MA, Informatics

Carol Somkin, PhD, Lead of ELSI and Community Cores

Stephen Van Den Eeden, PhD, Environmental Exposures
Core

Charles Quesenberry, PhD, Biostatistics Core

Lisa Croen, PhD, Biorepository Design Core

Larry Kushi, ScD, Access and Collaborations Core