



UPDATE ON NHANES DNA DATA DISCUSSION

**Board of Scientific Counselors
September 18, 2008**

Jennifer Madans



SAFER • HEALTHIER • PEOPLE





Introduction

- Background:
 - DNA has been collected (1991-94, 1999-2004, 2007+) by the NHANES program and is stored for future testing
 - Program in place to make DNA available for testing; to date testing has been limited but is increasing
 - Current interest by CDC and NIH in testing 'the whole genome'
- Issue:
 - NCHS considers DNA sequence information to be an 'identifier'
 - Options under consideration to expand access

Legislative Requirements for Confidentiality--NHANES

- **Public Health Service Act – Section 308(d)**
 - **NCHS is restricted from releasing identifiable data to anyone**
 - **Unless there is explicit consent**
- **CIPSEA – beginning in 2002, heavier fines and penalties for release of identifiable data**



Proposal Process to Conduct Genetic Testing on NHANES DNA Specimens

- **Proposals accepted to conduct specific DNA tests**
- **Proposals solicited through a Federal Register notice and are reviewed by three panels**
 - **Scientific technical panel (meets twice a year)**
 - **Internal CDC panel of senior scientists**
 - **Institutional Review Board**
- **Researchers receive de-identified DNA samples for genotyping**
- **Test results are sent to NCHS where they are linked to requested NHANES data**
- **Researchers access data for analysis in the RDC on site or remotely**



Proposal Process for Secondary Analysis of Genetic Data

- Genetic data can be analyzed with the thousands of phenotypic variables available from NHANES
- Proposals for secondary data analysis are accepted three times annually and follow the usual RDC process
- Except that they are reviewed by NCHS' IRB (no review by Technical Panel)
- Researchers with accepted proposals conduct their genotype-phenotype analyses in the NCHS' RDC on site or remotely

Beyond Gene Discovery Initiative

- **CDC's National Office of Public Health Genomics (NOPHG) Initiative**
- **Estimated 1,000,000 + genetic variations**
 - **Each participant's 1,000,000 genetic variations can be done on a chip the size of a standard slide**
- **Fund raising stage (estimated \$ 20 million)**
- **Funding not included in the CDC budget for testing NHANES DNA samples**

Genome Wide Association Studies (GWAS)

- **GWAS studies compare very large numbers of genetic variations to a health outcome (phenotype) of interest**
- **Many studies use cohort or case/control participants**
- **NHANES would provide a valuable sample to do original research as well as verify results from other studies**
 - **Representative of the U.S.**
 - **Multiple race/ethnicities**
 - **‘Normal’ controls for case/control studies**

GWAS Challenges

- Large data sets
- Multiple genetic pathways/complex genetic analyses
- QC process complex
- Many GWAS methods still in development
- Institutions sometimes write their own research software
- Computationally intensive
- Disclosure risks of new methods need to be assessed



Potential Ways to Expand Access

- Institute a remote access system
- Offer access to designated agents (under review with OMB)
- Open additional research data centers
- Address consent issue:
 - Reconsent of NHANES III and NHANES 1999-2002 participants
 - Change consent for future NHANES



E-mail based Remote Access

- Email-based remote access to NHANES genotypic and phenotypic database has been put in place.
 - Part of NCHS RDC
 - Researchers can access but cannot see identifiable data
- Planned development and evaluation of GWAS applications in a remote access environment
 - Have not begun due to budget constraints.



Designated Agent Authority

- **NCHS could enter into legally-binding agreements with a limited number of outside researchers to allow the controlled, conditional release of individual-level, potentially-identifiable data with oversight by NCHS**
- **NCHS has the authority to designate agents for access to data collected post 2002.**
- **Authority to designate agents for data collected through 2002 is under review at OMB – meeting held Sept 5 with OMB legal staff.**
- **Access procedures under a designated agent authority would need to be commensurate with the confidentiality risk and sensitivity of the data and would require significant oversight**



Open Additional RDC Locations

- **Add additional NCHS locations to offer access to NCHS sensitive data sets including genetic data**
- **Plans have been finalized for an Atlanta RDC but funding has not as yet been secured**



Changes to Informed Consent

Change NHANES informed consent so that confidentiality is not promised

- **Questions raised regarding the appropriateness of this approach for NHANES**
- **Initial plan to obtain community input has been put on hold – NIH sponsored a similar activity**
- **A November workshop on collecting, storing, protecting and accessing biological data, sponsored by NIA and conducted by CPOP and CNSTAT may address consent issues**



NCHS' Activities to Expand Data Access Planned or Currently Underway

- **Continue to improve current remote access system**
 - **Additional options tailored to genetic analyses**
 - **Expand menu of commercial software packages**
- **Develop mechanisms with appropriate confidentiality protections for Designated Agents accessing data protected by CIPSEA for a limited number of projects that cannot be done on the remote system**
- **Develop access tools in the NCHS RDC(s) to address the needs of GWAS projects**
 - **Computational requirements**
 - **Need for user developed software**