Digital Technologies are Transforming the Behavioral and Social Sciences into Data Rich Sciences

William Riley, Ph.D.
NIH Associate Director for Behavioral and Social Sciences Research
Director, Office of Behavioral and Social Sciences Research
Research Methods in a Data Poor Environment

• Priority is on prospective design and data collection

• Limited data collection opportunities
  • Predominately cross-sectional or minimally longitudinal designs

• Unable to assess or control myriad confounds
  • Control confounds via randomization

• And once the study is completed . .
Research Methods in a Data Rich Environment

- Temporally Dense
- Computational
- Predictive vs. Causal
A Brief History of a Data Rich Science: Meteorology

- Local, limited measurement
- Leverage communications technologies (telegraph) to connect data across sites
- Set standards for data integration
- Continued leveraging of technical advances in measurement and communication

Result: Rich, integrated data that is computationally modeled to explain and predict phenomena

Is it possible for health behavior research to become a data rich science?

"Nearly all the grandest discoveries of science have been but the rewards of accurate measurement." Lord Kelvin, 1872
Previous State of Behavioral Measurement
Dawn of a Data Rich Behavioral Science

• Rapidly accelerating technology development
  • Ecological Momentary Assessment (EMA) methods improved and delivered on cell phones
  • Capture of digital traces from daily interactions with technology
    • Social media
    • Call data records
    • Consumer sensors
  • Sensors that can passively and continuously monitor health risk behaviors in context
    • Physical activity sensors
    • Smoking sensors
    • Sun exposure sensors
    • Environmental exposure sensors
    • Dietary intake sensors (sort of)
• Applications of computational modeling and new statistical modeling approaches that provide the analytic capabilities for intensive longitudinal (temporally dense) data.
Ecological Momentary Assessment

Archival Big Data Sources in the Behavioral Sciences

“Digital Breadcrumbs” (Pentland, MIT)

• Behavioral Data gleaned from consumer-based data sources
  – Social Media (Twitter, Facebook)
  – Internet Searches (Google)
  – Cell phone Use (# calls and texts)
  – Cable Box Data (hours of TV)
  – Auto Black Box data
    (miles driven, seat belt use)
Sensor Technologies
Emerging Technologies and Assays for Adherence Monitoring

Xhale SMART “breathalyzer” for GRAS drug taggants

Drug (metabolite) concentrations via hair samples or dried blood spots

Proteus pill microchips and sensor

GlowCaps
But We Need More than Technology and Precise Assessment

• Data Access – especially from proprietary systems
• Common Ontologies/Taxonomies and CDEs
• Data Infrastructures (Registries, Repositories, Distributed Networks, APIs)
• Broad Consent
• And a Culture of Data Sharing where the knowledge base, not the publication, is the currency of research
Computational Modeling

WHAT WE COULD DO WITH A DATA RICH KNOWLEDGE BASE
Framingham Heart Study Social Network in 2000

Obesity

Happiness

Modeling of Tobacco Taxes on Smoking Prevalence
Cobiac et al., Tobacco Control, 2015
A Step Toward a Data Rich Cohort

THE U.S. PRECISION MEDICINE INITIATIVE
“And that’s why we’re here today. Because something called precision medicine … gives us one of the greatest opportunities for new medical breakthroughs that we have ever seen.”

President Barack Obama
January 30, 2015
A New Initiative on Precision Medicine

Francis S. Collins, M.D., Ph.D., and Harold Varmus, M.D.

"Tonight, I’m launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes — and to give all of us access to the personalized information we need to keep ourselves and our families healthier."

— President Barack Obama, State of the Union Address, January 20, 2015

The proposed initiative has two main components: a near-term focus on cancers and a longer-term aim to generate knowledge applicable to the whole range of health and disease. Both components are now within our reach because of advances in basic research, including molecular biology, genomics, and bioinformatics. Furthermore, the initiative
Precision Medicine Initiative

“The initiative will encourage and support the next generation of scientists to develop creative approaches for detecting, measuring, and analyzing a wide range of biomedical information – including molecular, genomic, cellular, clinical, behavioral, physiological, and environmental parameters.”
Collins and Varmus, NEJM, 2015

Behavioral and environmental measurement tools to:
• better characterize disease processes and treatment outcomes
• assess not only disease states but also the physical, mental, and social functional status
• monitor behavioral (e.g. smoking/diet) and environmental (e.g., particulate matter, social isolation) exposures that contribute to disease and that interact with genetic influences on disease and treatment, and
• provide potential behavioral and environmental predictors of treatment response beyond that obtained from genetics alone.
PMI: National Research Cohort “All of Us”

• Will comprise:
  – >1 million U.S. volunteers
    • From HPOs
    • From Direct-to-Volunteer efforts
• Participants will be:
  – Centrally involved in design, implementation
  – Able to share genomic data, lifestyle information, biological samples – all linked to their electronic health records
• Will forge new model for scientific research that emphasizes:
  – Engaged participants
  – Open, responsible data sharing with privacy protections
PMI COHORT PROGRAM
MAJOR COMPONENTS

• Data and Research Support Center (DRC)
• Biobank
• Participant Technologies Center (PTC)
• Healthcare Provider Organizations (HPOs)
• Regional Medical Centers
• Community Health Centers (Federally Qualified Health Centers)
• VA Medical Centers
PMI COHORT PROGRAM DATA

• The Program will start by collecting a limited set of standardized data from sources that will include:

  • Participant provided information
  • Electronic health records
  • Physical evaluation
  • Biospecimens (blood and urine samples)
  • Mobile/wearable technologies
  • Geospatial/environmental data

• Data types will grow and evolve with the science, technology, and participant trust.

• Tiered approach (not all data from all participants)
Concept is **not** entirely new:

- Prescription Eyeglasses
- Blood Transfusions
Precision Behavioral Interventions

• History
  – Project Match – Alcohol Abuse Treatments
  – Internet Tailored Interventions (Expert Systems)

• Beyond Tx A better than Tx B
  – For whom (tailored, personalized, precision)
  – In what context and at what time (JITAI, EMI)
  – In what combination and sequence (MOST, SMART)

• To Achieve Precision Behavioral Medicine
  – Identify more robust moderators and mediators
  – Reliable and intensive longitudinal data
  – Conceptual models that guide when, in what context, and for whom to deliver intervention strategies
“More than Genes, Drugs, and Disease”

News from the NIH: potential contributions of the behavioral and social sciences to the precision medicine initiative

William T. Riley, PhD, Wendy J. Nilsen, Ph.D., Teri A. Manolio, M.D., Ph.D., Daniel R. Masys, M.D., Michael Lauer, M.D.

Keywords
Precision medicine, Tailored interventions, Personalized medicine, Mobile health, Health informatics, Pharmacogenetics, Cohort studies, Behavioral risk factors, Environmental risk factors

At this year’s State of the Union address, the President announced a new $215 million Precision Medicine Initiative in the 2016 budget that will pioneer a new model of patient-empowered research and mobile/wireless technologies make now an opportune time for a large precision medicine cohort initiative.

NIH PRECISION MEDICINE WORKSHOP
To initiate planning of a large precision medicine cohort that could fully leverage these advances in genomics, cohorts, informatics, and mobile/wireless technologies, the NIH hosted a workshop on February 11–12, 2015. This workshop was attended in person by approximately 80 invited
Why Mobile Technologies?

Better Characterize Phenotypes and Outcomes
Why Mobile Technologies?

Better Characterize Treatments
Why Mobile Technologies?

Assess Treatment Predictors beyond Genetics
Why Mobile Technologies?

Intensively Measure Behavioral and Environmental Risk Factors
Why Mobile Technologies?

Fully Engage Participants as Partners
OBSSR Strategic Plan

Office of Behavioral and Social Sciences Research
Scientific Priority 1: Improve the Synergy of Basic and Applied Behavioral and Social Science Research

- Objective 1.1: Identify and encourage promising basic behavioral and social sciences research (bBSSR) with strong potential for applied translation relevant to health.
- Objective 1.2: Facilitate greater bidirectional interaction between basic and applied BSSR researchers to facilitate the translation of basic and applied behavioral and social sciences research.
Scientific Priority 2: Enhance the Methods, Measures, and Data Infrastructures to Encourage a More Cumulative Behavioral and Social Sciences

- Objective 2.1: Encourage data integration and replication in the behavioral and social sciences
- Objective 2.2: Facilitate the development and testing of new measurement approaches
- Objective 2.3: Expand the repertoire of methods available to social and behavioral researchers
Scientific Priority 3: Facilitate the Adoption of Behavioral and Social Science Research Findings in Health Research and Practice

- Objective 3.1: Encourage research that studies mechanisms and interventions in context
- Objective 3.2: Enhance the relevance and scalability of social and behavioral interventions
- Objective 3.3: Foster collaborations with agencies and entities that utilize and/or deliver social and behavioral research findings, and evaluate systemic and policy changes that facilitate or impede adoption of effective approaches
THANK YOU

WILLIAM.RILEY@NIH.GOV