Big Data and the Evolution of Precision Medicine

Dr. George Poste
Chief Scientist, Complex Adaptive Systems Initiative and Del E. Webb Chair in Health Innovation
Arizona State University
george.poste@asu.edu
www.casi.asu.edu

Michigan Institute for Data Science (MIDAS)
Ann Arbor, MI
October 6, 2015
Medical Progress:
From Superstitions to Symptoms to Signatures
Precision Medicine: Understanding the Organization of Complex Molecular Networks in the Health-Disease Continuum

(Epi)Genome

Cell- and Organ-Specific Molecular Information Networks

The Phenotype and Individual Variation

Health

Instructional Code

Disease-Induced Network Changes

panOmic Individual Profiling and Optimum Care Decisions

Disease
Precision Medicine: Managing Individual Health Risk

Managing Active Disease
- right diagnosis
- right treatment(s)
- right time
- right follow-up and compliance

Sustaining Health (Wellness)
- disease predisposition
- risk exposure
- risk mitigation

Reactive

Proactive
Precision Medicine: The Complexity of Genotype-Phenotype Relationships

The Need for Deep Phenotyping

Genome Sequencing Alone Will Not Suffice

Understanding the Complex Interplay Between PanOmics, Environment and Lifestyle
Establishment of Causal Relationships Between Alterations in Molecular Networks and Disease Risk, Disease Progression and Intervention Outcomes

- Large Scale Population Profiling
- Integrated Molecular, Clinical, Environmental and Lifestyle Data
- Pattern Analysis of the Health-Disease Continuum
- Multi-parameter Individual Risk Profile

- University of Michigan Health System
- The Fourth Paradigm: Data-Intensive Scientific Discovery
- The Molecular Taxonomy of Health and Disease
- Optimized Individual Care and Health Risk Reduction

Population Health Management
Big Data and Data Science

Big Data and Data Science
The Evolution of a Data-Driven Health Ecosystem: Systematic Integration of Diverse Data Sets for Population Health Analytics

Continuity of Care Record: From Womb to Tomb

Behavior

Environment
AORTA (Always On Real Time Access): Continuous Monitoring of Health Status

- majority of events affecting an individual’s health occur outside of healthcare facilities
- new technologies and real-time, remote monitoring of health status and treatment compliance
  - wearables, sensors, social media
- new patterns of consumer/patient interaction with the healthcare system and healthcare professionals (“expanded touch points”)
- progressive evolution of a seamless blend of online and physical services for clinical care and individual health risk management
Real Time Remote Health Monitoring and Chronic Disease Management

Lifestyle and Fitness

Information for Proactive Health Awareness (Wellness)
Telemedicine: Diagnostics, Robotics, and Remote Monitoring of Health
Implantable Devices and Wireless Monitoring (and Modulation)

next-generation miniaturized power sources

security and hacker protections
Gray Technologies and Aging in Place: Independent But Monitored Living for Aging Populations

- Rx compliance
- Cognitive stimulation
- In home support and reduced readmissions
- Reduced office visits
Mobile Apps, Wearables, Sensors and Continuous Monitoring

- who sets the standards?
- who integrates and interprets the data?
- who pays?
- who consents?
- who owns the data?
Social Spaces and Individual Behavior Become Quantifiable

- who knows why people do what they do?
  - the fact is that they do!
- these actions can now be traced and measured with unprecedented precision
- with sufficient data, the numbers reveal increasingly predictable behavior and individual risk patterns
- rapid growth in new business opportunities in multiple sectors including healthcare
- new ethical and legal issues
  - consent and data ownership
  - privacy, surveillance, security
**Population Health Research and Precision Medicine:**
**Blurring the Boundaries Between Research and Clinical Care**

- every individual is a data node
- every encounter (clinical and non-clinical) is a data point
- every individual becomes a research asset
“Do you solemnly swear to have no involvement in your own care?”
The Rise of Consumerism in Healthcare

Provider Performance, Pricing Transparency Plus Choice

UX: User Experience

“liquid expectations”
- positive consumer experience in one domain generates expectations of similar convenience/value in other domains
The Principal Forces Shaping Biomedical R&D and Healthcare Delivery

- Engineering and device-based medicine
  - Wearables
  - Sensors
  - Smart implants
  - Remote health monitoring
  - Telemedicine
  - Robotics

- Molecular (precision) medicine
  - PanOmics profiling
  - Analysis of disruption in biological networks in disease

- Information-based healthcare
  - M.health/e.health
  - Data- and evidence-based decisions and Rx selection

- BIG DATA
  - Outcomes-based healthcare and sustainable health
  - New value propositions, new business models and services
The Worst Supply Chain in Our Society is the Health Information Supply Chain

- no area of the economy (15-20% GDP) generates as much information as the health sector yet uses it so poorly
- fragmented, disconnected data (data tombs)
- incompatible data formats as barriers to data integration
- incomplete and inaccurate data
The Worst Supply Chain in Our Society in the Health Information Supply Chain

- slow transition from paper to electronic systems
- inadequate information on behavioral and environmental influences
- poor data protection at rest and in transit
- legislative barriers to data transfer based on well intentioned privacy protections
- EMR vendor barriers to facile data exchange
The Painful Evolution of Electronic Medical Records (EMRs)

Scheduling and Billing

Compliance

Real Time Data and Decisions
The Pending Zettabyte Era
1,000,000,000,000,000,000,000,000

Managing Big Data in Biomedicine is Not a Simple Extrapolation from Current Practices

Current Institutional Structures and Competencies Are Ill-Prepared for Pending Disruptive Change
• human genome sequencing data doubling every 7 months
• projected 1 exabyte/year within 5 years
• projected 100 million to 2 billion human genomes sequenced by 2025
• data storage needs of 2 to 40 exabytes (@30x coverage)

• YouTube 1-2 exabytes for video storage
• Twitter 1-17 petabytes/year
• Square Kilometre Array 1 exabyte/year
The Unavoidable Data-Intensive Evolution of Healthcare: Major Challenges Ahead

PB and TB Data Streams

Ontologies and Formats for Data Integration

Longitudinal Data Migration and Inter-operable Dbases

New Data Analytics, Machine Learning, NLP Methods

Infrastructure, Storage and Privacy

Data Science and Data Scientists
“Digital Darwinism”: Stark Selection Pressures Will Create Haves and Have Nots

- Growing imbalance between different end user populations and their ability to embrace large data scale and complex analytics.
- Institutions unable to access and analyze large data sets will suffer ‘cognitive starvation’ and relegation to competitive irrelevance.
- Understanding the structure of information and its productive application/customization will emerge as a critical institutional competency.
  - “Intelligence at ingestion”
The Big ‘N’ Challenge in Making Precision Medicine a Reality: Building the ‘Data Commons’

- development of a robust molecular taxonomy for the health-disease continuum will require comprehensive data capture and pattern analysis of multiple features
  - panOmics, clinical, risk exposure, life style
  - longitudinal continuity
- required scale will transcend the population cohort(s) available in all but the largest healthcare providers/payors
- new models for open data sharing and meta-analysis
The Big ‘N’ Challenge in Making Precision Medicine a Reality: Building the ‘Data Commons’

- urgent need for new policies and incentives for data sharing and open infrastructure (international?)
- how to integrate proprietary databases into an open infrastructure
- privacy and security: is individual de-identification illusory?
Computational- and Data-Enabled Science

- Bigger Data and Better Questions
- Data Science: Thinking More Deeply About Data and Knowledge Generation
- Big Data and Data Science Will Generate Destabilizing and Disruptive Knowledge
The Pending Era of Machine Intelligence and Cognitive Systems: Overcoming the “Bandwidth” Limits of Humans

- limits to individual expertise
- limits to our multi-dimensionality
- limits to our sensory systems
- limits to our experiences and perceptions
- limits to our objective decision-making

“helping the slow brain catch up with the fast machine”
Data Deluge

Technology Acceleration and Convergence: The Escalating Challenge for Professional Competency, Decision-Support and Future Medical Education Curricula

Automated Analytics and Decision Support

Cognitive Bandwidth Limits

Facile Formats for Actionable Decisions
Living in a World Where the Data Analytics and Interpretation Algorithms Are Obscure to the End User

- ceding decision authority to computerized support systems
- resistance and push back in a MD-centric culture
- culturally alien to professionals in their expertise domains while they accept machine-based decision-support in many other aspects of their lives
- who will have the responsibility for validation and oversight of critical assumptions used in decision tree analytics for big data?
  - regulatory agencies and professional societies (humans)?
  - machines?
Big Data: Changing the Intellectual Framework for Discovery and Knowledge Acquisition

Defining An Optimum Balance

- Multi-disciplinary: team-based, systems-focus, big data sets
- Reductionist: Individual investigator-centric, single discipline datasets

Hypothesis-driven research

Unbiased datasets and new analytics for pattern mining
• changing the nature of discovery
  – hypothesis-driven versus unbiased analytics of large datasets (patterns, rules)
• changing the cultural process of knowledge acquisition
  – large scale collaboration networks, open systems versus individual investigators and siloed data
• changing knowledge content
  – increased quantification and complexity
  – integration of diverse data streams
• changing the cognitive and intellectual competencies for knowledge-intensive competitiveness in multiple domains
• changing education, training and research
The Evolving Data-Intensive Healthcare Ecosystem

- Technology convergence
- Computing and automation
- Life sciences and medicine
- Sensors, robotics
- BIG DATA
  - Population Data
  - Precision Medicine
  - Data Science
- Connectivity, continuity and consumerism
  - Social media
  - Patient engagement
  - Life style metrics

- Services integration (systems)
- Analytics for actionable Information and improved outcomes (value)
- The expanded care space (individuals)
CHANGE
is good
you go first
Leveraging the Potential of Precision Medicine Will Require **PROFOUND CHANGES** in the Organization and Proficiency of Healthcare Services

- Seamless integration of complex, diverse and dynamic data for real-time monitoring of health status and risk management
- Shift from reactive episodic care encounters to increasingly proactive risk mitigation
- Progressive shift from management of overt disease to sustained wellness and continuity in care

**technology**
- Integration of molecular, clinical and lifestyle data

**competencies**
- Infrastructure, data science, new clinical skills

**efficiency**
- Continuity in care

**incentives**
- New financial reward systems