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

Presentation at Cambridge Second Annual Big Data, Biomarkers and Diagnostics World Congress
Philadelphia, PA
17 May 2016
Medical Progress: From Superstitions to Symptoms to Signatures
Molecular Diagnostics and Precision Medicine: PanOmics Profiling and Mapping the Disruption of Molecular Networks in Disease

(Epi)Genomics

Proteomics

Molecular Pathways and Networks

Network Regulatory Mechanisms

ID of Causal Relationships Between Network Perturbations and Disease

Patient-Specific Signals and Signatures of Disease or Predisposition to Disease
PanOmics Molecular Profiling and Precision Medicine

- a new molecular taxonomy for disease
- intellectual foundation for improved diagnostic accuracy and rational therapy selection
- profiling individual variation in disease risk, patterns of disease progression and therapeutic responses
- mapping the diversity and dynamic range of disease-associated alterations in the architecture of molecular signaling (information) networks
Analytical and Clinical Validation Protocols to Demonstrate the Utility of Molecular Profiling in Precision Medicine

- Multiplex disease biomarkers and molecular variants
- Massive data: volume, velocity, variety, veracity
- Evidentiary standards for regulatory qualification
- Clinical outcomes, value and reimbursement
- Clinical utility and adoption
Analytical and Clinical Validation Protocols to Demonstrate the Utility of Molecular Profiling in Precision Medicine

**DEMONSTRATING VALUE**

**DRIVE ALTERED CLINICAL PRACTICE TO IMPROVE OUTCOMES AND/OR LOWER COST**

**REDUCING RISK: OPTIMIZING PROSPECTIVE ACTIONS TO SUSTAIN HEALTH VERSUS REACTIVE RESPONSES TO DISEASE EPISODES**

- multiplex disease biomarkers and molecular variants
- massive data: volume, velocity, variety, veracity
- evidentiary standards for regulatory qualification
- clinical outcomes, value and reimbursement
- clinical utility and adoption
Precision Medicine:
The Dangers of Reductionism
and Ignoring Biological Complexity

Genes For ....

The Over-Simplified Perspective That
While Exome-and Whole Genome-sequencing
Will Reveal the Full Etiology of Disease Pathogenesis
The Current Reductionist, Simplistic Obsession with Genome Sequencing as the Dominant Platform for Precision Medicine

Revealing Pathology With Genome Sequencing
Informatics Challenges to Advance Regulatory Evaluation of NGS-Based Diagnostics: The Imperative for Analytical and Interpretation Standards

Science Translational Medicine, 22 April 2016, 8,335,335ps10, 2
The Exabyte-Zettabyte World Awaits for panOmics Profiling of Large Scale Populations*

- 2015 estimated 3.6 petabases raw sequence data (all species)
- c.7 month current doubling time of archived sequences
- Projected 100 million to 2 billion human WGS by 2025
- Projected production 1 exabase and 1 zettabase of sequence/year by 2020 and 2025 respectively
- Additional need for multiple sequence datasets: (epi)genome, RNA universe, microbiome)

*Adapted from: Z.D. Stephens et al (2015) PLOS Biology 1002195
Genome Sequencing Alone Will Not Suffice: The Need for Deep Phenotyping

Phenome-Association Data (PheWAS): Integration of panOmics Profiling with Clinical Disease Patterns and Treatment Outcomes

Understanding the Complex Interplay Between PanOmics, Environment and Behavior
Individual Variation, (Epi)Genome Complexity and the Challenge of Genotype-Phenotype Predictions

Junk No More: Pervasive Transcription
- alternate transcription/translation/(co)splicing
- SNPs, CNVs
- pseudogenes
- indels, SVs
- ncRNAs
- phasing
- epistasis
- imprinting
- silencing
- miRNAs/ceRNAs/circRNAs

recognition of (epi)genome organizational and regulatory complexity

Cell-specific Molecular Interaction Networks

Perturbed Networks and Disease
Molecular Diagnostics and Precision Medicine: Mapping The Topologies and Dynamics of Molecular Signaling (Information) Networks

- “health”
- homeostasis
- subclinical disease
- graded threshold states
- overt clinical disease
- diverse phenomes

Emergence (E)

E₁, E₂,..., Eₙ

network topology

state shifts
The Challenge of Translation of Burgeoning panOmics Data Into Clinically Relevant (Actionable) Knowledge

- Data
- Reliability and Robustness
- Biological Insight
- Clinical Utility
The Virtuous Circle of Data on Population Health and Individuals in Driving Precision Medicine

1. Large Scale Population Data Profiles
2. Pattern Analysis to ID Subgroup/Individual Profiles
3. Correlation of Subgroup/Individual Patterns with Disease Progression/Rx Outcomes
4. Guidelines/Best Practices for Precision Medicine
5. Continued Data Capture and Analytical Refinement

The circle flows in the following order: Large Scale Population Data Profiles → Pattern Analysis to ID Subgroup/Individual Profiles → Correlation of Subgroup/Individual Patterns with Disease Progression/Rx Outcomes → Guidelines/Best Practices for Precision Medicine → Continued Data Capture and Analytical Refinement → Large Scale Population Data Profiles.
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
Untethering Healthcare From Fixed Clinical Facilities: Extending the Care Space

The Majority of Events Affecting Individual Health Occur Outside of Formal Interactions with Health Systems

Wearables, Remote Sensors, Mobile Devices and Real-Time Monitoring of Risk, Health Status and Treatment Compliance
Invasion of the Body Trackers: Wearables, Sensors, the Internet of Things and the Focus on Prevention and Wellness

quantified self

telemedicine

grey technologies and treatment compliance

in-home support and reduced readmissions

wireless monitoring of implantable devices

IoT: escalating device connectivities
An Apps-Based Information Economy in Healthcare

- wearables and continuous sensors (individual, populations)
- theoretical rationale but integration of data with EHR platforms poses numerous challenges
  - lack of developer access to high quality healthcare data to validate app platforms
  - cross-platform standardization and application programming interfaces (APIs)
  - regulation: accuracy, reliability, security and privacy
  - reimbursement (developers and healthcare providers)
- FDA focus on apps that transform phone/tablet into a regulated medical device
- renewed FTC interest on apps making unsubstantiated claims
Untethering Healthcare From Fixed Clinical Facilities: Monitoring of Health Status, Compliance and Risk Behavior

- every encounter (clinical and non-clinical) is a data point
- every individual is a data node
- every individual is a research asset
- every individual is their own control
The Principal Forces Shaping The Evolution of Precision Medicine

- Engineering and device-based medicine
  - wearables
  - sensors
  - smart implants
  - remote health monitoring
  - telemedicine
  - robotics

- Information-based healthcare
  - m.health/e.health
  - data- and evidence-based decisions and Rx selection

- Molecular (precision) medicine
  - panOmics profiling
  - analysis of disruption in biological networks

- Outcomes-based healthcare and sustainable health

BIG DATA

- New value propositions, new business models and services
Now Comes the Hard Part!
Integration of Diverse Datasets in Biomedical R&D and Healthcare

(Epi) Genomics | Proteomics | Rx Targets | HTS | Rx SAR

Clinical Trials | Dx Profiling | Mobiles and Wearables | EMR | Outcomes

FDA | Circular Genome Map | Wearable Technology | Robotic Assay | Value Analysis
HELL IS THE PLACE WHERE NOTHING CONNECTS — T.S. ELIOT
Silos Subvert Solutions: Protecting Turf and Sustaining the Status Quo

WELCOME TO BIOMEDICAL RESEARCH AND PATIENT MEDICAL RECORDS
The Troubled State of Too Many Data Sets in Biomedical R&D and Healthcare Delivery

- sloppy science (the reproducibility problem)
- statistics (underpowering, overfitting of small N sample sets profiled by large N panOmics feature sets)
- silos (data tombs)
- sharing (what’s that?)
- semantics (limited use of common ontologies)
- standards (incompatible data formats and dbase inter-operabilities)
- safety (source of medical errors, problematic patient ID and match across different records)
Critical Challenges in the Generation and Analysis of Robust, Large Scale Data in Biomedical R&D and Healthcare

- scale (exabyte and zettabyte and beyond)
- structure (80% unstructured; need for NLE methods)
- speed (latency, inadequate infrastructure and few fat pipes)
- storage (cost)
- security (healthcare records most frequently attacked category in 2015)
- surveillance (privacy, consent, data ownership)
- states (compliance with patchwork of US state-laws; EU data directive)
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 Security and Privacy

Data Science and Data Scientists
The Emergence of Big Data Changes the Questions That Can Be Asked

Isolated Data

Complex Networked Data

Complex Computational Data
The Pending Era of Cognitive Computing and Decision-Support Systems: Overcoming the “Bandwidth” Limits of Human Individuals

- 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
Advanced Computing and Artificial Intelligence: The Rise of ‘Learning Machines’ in the Analysis of Massive Datasets and Decision Algorithms
Data-Driven Knowledge, Intelligence and Actionable Decisions

- changing the nature of discovery
  - unbiased analytics of large datasets (patterns, rules) vs. traditional hypothesis-driven methods

- changing the cultural process of knowledge acquisition
  - large scale collaboration network, consortia and open systems versus individual investigators and siloed data

- changing the application of knowledge
  - increased quantification, big data analytics and decision-support systems

- changing education, training, research and care delivery
Technology Acceleration and Convergence: The Escalating Challenge for Professional Competency, Decision-Support and Future Medical Education Curricula

Data Deluge

Cognitive Bandwidth Limits

Automated Analytics and Decision Support

Facile Formats for Actionable Decisions
“I Can’t Let You Do That Dave”
Living in a World Where the Data Analytics and Interpretation Algorithms Are Obscure to the End User

- ceding decision authority to computerized support systems
- culturally alien to professionals in their expertise domain but they accept in all 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?
“DNR”

- Denial
- Negativity
- Resistance
The Evolution of Precision Medicine: Dependence on Formation of a New Comprehensive Healthcare Data Ecosystem

Incrementalism versus Disruptive Innovation

Squeezing Savings from Outmoded Processes and Business Models versus Radical Data-Driven Shifts in Diagnostic Medicine and Patient Care to Improve Outcomes/Reduce Cost
The Evolution of Precision Medicine: Data-Intensive Healthcare, Data Analytics and Computational Decision Systems

Precision Medicine: Managing Individual Risk

Population Health: Pattern Analytics for Risk/Outcomes Management

Real Time Health Monitoring: IOT and Sensor Networks

Big Data: Interpretation and Intelligence at Ingestion

Escalating Data Complexity: Machine Learning, Decision Support
The Evolution of Precision Medicine: Data-Intensive Healthcare, Data Analytics and Computational Decision Systems

- Precision Medicine: Managing Individual Risk
- Population Health: Pattern Analytics for Risk/Outcomes Management
- Real Time Health Monitoring: IOT and Sensor Networks
- Big Data: Interpretation and Intelligence at Ingestion
- Escalating Data Complexity: Machine Learning, Decision Support

New Patterns of Technology Convergence, Evolution and Adoption

- New Knowledge Networks
- New Participants
- New Organizational Models

Opportunity Space