A Status Report on Biomarkers:
biomarkers that are not; standards that are forgot;
and the need to transform the system we’ve got!

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: National Biomarkers Development Alliance
JW Marriott Camelback Inn, Scottsdale, Arizona
December 13, 2012
Slides available @ http://casi.asu.edu/
Healthcare: An Expensive Menu Without Prices

Managing the Demands of an Aging Society and Chronic Disease Burden in an Era of Economic Constraint

Shift From a “Do More, Bill More” Healthcare System to Managing Individual Risk to Improve Health Outcomes and Control Cost

Sustainable Health: Societal (Economic) and Individual (Wellness)
Medical Progress:
From Superstitions to Symptoms to Signatures
Mapping The Molecular Signatures of Disease: The Intellectual Foundation of Rational Diagnosis and Treatment Selection

- **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**
Mapping Causal Perturbations in Biological Networks in Disease: Defining the Molecular Taxonomy of Disease

iOmmics Profiling to Identify Disease Subtypes (+ or - Rx Target)

Altered Network Structure and ID of Molecular Targets for MDx and/or Rx Action

Modeling Information Flow in Biological Networks
Biomarkers, Disease Subtyping and Targeted Therapy: Companion Diagnostics – the Right Rx for the Right Disease (Subtype)

- Her-2+ (Herceptin)
- EML4-ALK (Xalkori)
- KRAS (Erbitux) (Vectibix)
- BRAF-V600 (Zelboraf)
Molecular Profiling of Disease and the Trajectory for Data-Intensive Healthcare

earlier detection of disease
disease subtypes and rational Rx
monitoring of health status
predisposition risk
integrated, continuity of care
optimized resource allocation

profiling

analysis

actionable information and knowledge-driven decisions
The Integrative Personal Omics Profile (iPOP)

Population Datasets
“Big N”

iPOP – an individual’s ‘omics inventory
“N=1”
Technology Convergence and the Trajectory for Molecular Medicine

- informatics, statistics, mathematics, modeling
  - big data: analytics and infrastructure

- iPOP
- systems-based processes

- biomedicine

- engineering
  - devices
  - sensors
  - m.Health

- computing

- optimized decisions/actions
  (professional and personal)

- comprehensive connectivity (networks)
  - integrated data
  - continuity of care
  - risk management and maximizing wellness
Translating Molecular Profiling into Clinical Medicine and Healthcare Delivery

- what is the best biospecimen for the intended use?
- how will the biospecimen be profiled and when?
- who will conduct the profiling?
- what will be reported?
- how will profiling information be used to support a clinical decision and/or clinical decision support system?
- how will MD/HCPs be trained/supported in interpretation of complex profiling data?
- how will profiling results be incorporated into EHR/EMR and/or aggregated into larger databanks?
- how will profiling assays be regulated and reimbursed?
disturbing low reproducibility of biomarker publications

- poor access to rigorously annotated biospecimens from stringently phenotyped sources

- insufficient control of pre-analytical parameters and variable analytical standards

- idiosynchratic, ‘lab-specific’ analytical methods

- ‘small N’ studies lacking statistical power

- chaotic incompatible data reporting formats and poor dbase interoperability

- pressure to publish and poor compliance with funding agency/journal policies on open data sharing

- failure to work to (or understand) industry and regulatory standards
Access to High Quality Biospecimens, Biobanks and DNA Repositories: An Obligate Prerequisite to Productive Validation of Putative Causal Disease Markers

- requisite scale and stringent QA/QC standards
- academic anecdotes and wasted investment
The Dismal Productivity of Biomarker R&D

Legacy of Failure to Embrace Multidisciplinary Expertise and Adopt Stringent QA/QC Processes

The Complexity of Biomarker Discovery, Validation and Clinical Adoption is Comparable to (Bio)Pharmaceutical R&D

In Common With R&D for Drugs and Vaccines
Success Demands a Systems-Based Approach
Building Large Scale, Standardized Resources for Biomarker Research

- rigorously phenotyped/matched/consented disease and normal specimens
- biobanking: leadership and national policies to create a vital research resource
- standardization of pre-analytical and analytical methods
- standardized data ontologies and formats for large scale datasets/federated databanks
Blood-Based Biomarkers

- obvious appeal of low cost, minimally invasive sampling and ease of repeat sampling
- DNA banking and WES/WGS of germ line variants and disease predisposition/PGX profile
- current major knowledge gaps
  - lability of different molecular species in different pre-analytical process protocols (intended use setting)
  - biomarker(s) abundance, dynamic range and intra- and inter-patient variation
  - detection threshold(s) for different analytical platforms
  - relationship between stage of disease progression, (sub)clinical phenotype and intravascular biomarker shedding kinetics
new informed consent provisions in an era of Omics profiling and DNA banking

- broad (future proof) versus narrow (explicit) research investigations
- flexibility to address personal preferences

dynamic consent: e.consent tools and regular updating
- EnCoRe, Indivo, PrivateAccess

separate informed consent for sample banking for PGx testing versus MDx and PDx and clinical trial stratification?

the ‘incidentalome’ and recontact criteria
are the phenotypes and profiled molecular pathways of cell lines and 2D cell cultures so unrepresentative of the situation in vivo to render them irrelevant and pose blind avenues for diagnostic/therapeutic discovery?

can the biology of metastasis and MDx biomarkers for metastasis and new Rx targets be elucidated by the analysis of non-metastatic cells?

how representative are circulating tumor cells (CTC) of the metastatic phenotypic and the extravagant inter-patient and intra-patient heterogeneity of cancer cell phenotypes?
massively parallel multiplex assays (global profiling)
automated, miniaturized, high throughput
novel test formats and data formats
escalating data volume, velocity and variety (V3)
complex deconvolution of low signal to noise signatures
big data, analytics and infrastructure
Implications of the Underlying Distribution of the Dataset and Selection of Classification Algorithms

- large scale, molecular datasets differ fundamentally from traditional biological and clinical datasets
- traditional clinical/epidemiological datasets
  - small number of variables (p) tracked across a proportionately larger number (n) of samples (n>p)
- molecular datasets
  - the high dimensionality problem: number of variables measured per sample ($10^5$ – $10^9$ analytes/WGS) far exceeds the typical number of samples ($10^1$ – $10^2$) (p>>n)
- without detailed understanding of the biology of cellular network regulation impossible to know whether any or all of the measured features (genes/proteins) are related in a linear or non-linear fashion
  - difficult to demonstrate non-linear relationship in incompletely-characterized biological systems
Standards, Complexity and New IRB Competencies
Quality happens only when someone is responsible for it
Sophisticated Project Management Systems and Automated Workflow Systems

- Visual modeling of multi-step processing and analysis
- Automated documentation: ID, tracking, and provenance
- Integration of disparate software tools, edits, and provenance
- Automated translation to required formats
- Longitudinal data integration (from discovery to EHR/EMR)
- Regulatory compliance and inspection audits
Improved Conduct and Reporting of Biomarker Validation Studies and QI Standards for Biobanking, Analysis and Publication

- CONSORT
- AMSTAR
- AGREE
- STROBE

- REMARK
- BRISQ
- EGAPP
- STREGA
- MIAME (loc.cit)

- Biospecimen Reporting for Improved Study Quality
- Best practices: NCI, OECD, UK, Australia, Canada
- Int. Society for Biological and Environmental Repositories

Reinforcement by Funding Agencies and Journals of Reporting Formats and Open Data Submission
IHE-LAW
A Major Advance for Integration of Diagnostic Laboratory Automation, Information Systems and Electronic Health Records

- partnership between Integrating the Healthcare Enterprise (IHE) and In Vitro Diagnostics Connectivity Consortium (IICC)

- IHE-LAW (Laboratory Analytical Workflow) standard
  - uniform IT connectivity standards for LIS, automation systems, middleware, CPOE and EHR
  - use of ISO HL7 messaging

- participation of leading instrument manufacturers
  - Abbott, BD, Beckman Coulter, BioMerieux, Ortho, Roche, Siemens

- projected final standard in collaboration with The Clinical and Laboratory Standards Institute (CLSI) in 2013
Will Low Cost Whole Genome Sequencing Change Everything?
Current Chokepoints and Challenges in Adoption of Personal Omics Profiling Data for Clinical Decisions

- production of sequencing data outstripping interpretational capacities
- CLIA compliance, RUO materials and other regulatory requirements for clinical decisions
- confusing maze of base calling, alignment, assembly and analysis tools
- many software tools insufficiently robust and/or customized for one type of data or sequencing platform
- variation in clinical significance predictions from different algorithms using well known algorithms (SIFT, PolyPhen, LRT, MAPP, VarioWatch)
- (comparable data standardization/validation problems in large scale proteomics)
Genes For .... The Overly Simplistic and Deterministic Dangers of a Genome-Sequence Centric Perspective

The Over-Simplified Perspective That Whole Exome-and Whole Genome-Sequencing Will Reveal the Full Etiology of Disease Pathogenesis
Individual Variation, Genome Complexity and the Challenge of Genotype-Phenotype Predictions

Junk No More: Pervasive Transcription
- alternate transcription
- SNPs, CNVs
- pseudogenes
- indels, SVs
- ncRNAs
- phasing
- epistasis
- imprinting
- silencing

recognition of genome organizational and regulatory complexity

Mapping Cell-specific Molecular Interaction Networks

Perturbed Networks and Disease
Defining Human Genetic Variation

- GWAS
- Linkage Studies: Mendelian Disorders
- Epistasis
- WGS

Effect Size (Penetrance, Impact)

- high
- low

Allele Frequency

- common (>1%)
- intermediate (>1% <0.001%)
- rare (< 0.0001%)
what are the network epistatic interactions between causal and modifying genes that define expressivity, penetrance and ultimate phenotypic impact?

- genesis of a likely continuum of clinical phenotypes
- potential spectrum of subtle to severe disease generated by graded perturbations within- and between- molecular pathways and networks in the same cell type/organ
Diversity of the Human Variome and Role of Rare/Private Disease-Associated Variants

- Implications for cohort selection for profiling studies
  - Clinically relevant/medically actionable variants may be heavily weighted toward recent rare/private variants (clan/pedigree/individual)
  - Ethnic factors and study design
- Implications for genoprofiling
  - Deep sequencing (>100X coverage) of 20,000 or more individuals to link variants/variant combinations to disease phenotype(s)
- Cancer presents unique challenges
  - Extravagant scale of causal somatic mutations plus rapid progression of intra-and inter-lesional heterogeneity in advanced disease
Biomarkers for Disease Subtyping and Improved Design of Clinical Trials and Enrollment Efficiency

- imperative to reduce cost/time of high failure rate of investigational drugs
- targeted Rx and stratification of cohorts based on target BM (and/or other Rx response marker)
  - enrichment trials and adaptive trials
  - faster trials?
  - greater regulatory clarity?
- logistics: time, cost and scale of screening to ID BM+ cohort(s) (frequency dependent e.g. ALK)
- social media and accelerated enrollment
Miniaturization of Analytical Technologies

“Lab-on-a-Chip”

“Lab-on-a-Tip”

“Lab-Always On” and “Lab-On-Me”
Invasion of the Body Trackers

Individual Biosignature Profiling Via On Body: In Body (OBIB) Sensors and Devices

Remote Health Status Monitoring
Remote Health Monitoring and Chronic Disease Management

Information for Proactive Health Awareness (Wellness)

Lifestyle and Fitness

m.Health
Proactive Engagement of Patient Communities in Investigational Clinical Trials and Observational Outcomes Studies

- Collate, Annotate and Curate Clinical Trial Data with Genomic Information from the Comparator Arms of Industry- and Foundation-Sponsored Clinical Trials
- Building a Site for Sharing Data and Models to evolve better Disease Maps.

PURPOSE
To improve cancer-related comparative effectiveness research by better capturing data on physiological, behavioral and psychological status from research participants at home and as the go about their daily lives.
Framingham Redux!
Longitudinal Sampling of Populations to Assemble Biomarker/Biosignature Correlates of Health Outcomes

- 73,000 undergraduates
- 10% enrollment for longitudinal monitoring
- bi-annual blood profile
- progressive expansion of WES/WGS datasets
- broad and portable informed consent
- upload to PHR/EHR plus reciprocal consent for epidemiological analysis
Commensal Microbiomes: The Frenemy Within. An Additional Dimension to Biomarker Profiling

Metagenome-wide Association Studies (MGWAS)

- Immune-Mediated GI Diseases
- Type 2 Diabetes Profile Shift
- Aging, Metabolism and Fragility
- Metabolic Activation of Xenobiotics
Oversight and Regulation of Diagnostic Tests: Emerging Gaps and Ambiguities

**enforcement**

- FDA or CMS or both?

**new technologies and new analytical complexities?**

- Dx-Rx combinations and patient stratification for adaptive trials
- multiplex assays and validation of complex statistics/software
- genome sequencing (WGS, exome, transcriptome)
- RUO vs. clinical use (reagents/instruments)
- international harmonization

**new marketing issues**

- no system for post-market surveillance of Dx performance
- DTC advertising and testing
- internet sales
- social media
New Locations and New Services in Primary Healthcare

What Was

What Will Be

From: S. Burrill
Evolution of CLIA-Waived POC/PON Devices

Decentralized Primary Care

Low Cost Fieldable Units: DCs/FOBs

- paper-based assays
- hepatoxicity of retroviral therapy
- malaria
- dengue
- pre-eclampsia
- mHealth Apps
Medical Policy and Reimbursement

- ICH, FISH other MDx deemed ‘experimental’
- 14% plans classify MDx as ‘experimental’
- no umbrella coverage; separate-contract for each state
- no adverse policies for MDx
- guidance that NGS will not be reimbursed
- IHC’s not reimbursed
- no adverse policies to MDx
- no adverse policies to MDx
If You Build It, Will They Pay?
If It Isn’t Billable, It Won’t Happen!

- Will test alter patient management?
  - Reduce cost of care
  - Improve outcomes
- What additional resources/services/training are affected by test adoption?
- Perception of RCT as only ‘gold standard’
  - Narrow interpretation that discounts value of observational studies
- Payer demand for regulatory approval to be eligible for reimbursement or CED
- Mindset of ‘lab data’ as low cost (<1% total care cost) despite role in most treatment decisions (>85%)
  - Unianalyte versus multiplex tests
  - Outdated US reimbursement codes
Introducing Intellectual Property and Molecular Diagnostics

Essential Protection for Innovation and Incentive to Invest in Increasingly Complex and Expensive MDx Development and Approval?

or

Inappropriate Patenting of Naturally Occurring Molecules (Biomarkers) and Anti-Competitive Barriers to Adoption of New MDx Services in Medical Centers?
Now Comes the Hardest Part of All!

Moving Downstream Beyond Discovery:

Driving iOmnics and Molecular Medicine and IT-Centric Capabilities Into Routine Clinical Practice

The Escalating Scale and Complexity of the Data Stream
Silos Subvert Solutions: Protecting Turf and Sustaining the Status Quo
The Need for Facile, Seamless Data Exchange Formats for Large Scale Biomedical Data Systems

- research and discovery
- translation and clinical trials
- healthcare delivery
- m.health
- consumers
- payors
- regulators
- outcomes analytics
- decision support tools
- patients

consumers
Representation of Datasets: Formalisms and Abstractions

**Discovery**
- controlled vocabularies and formal ontologies
- quality and provenance checklists and open source repositories
- algorithms and transparent source code for analytical tools

**Translation and Adoption in Routine Care**
- facile exchange formats and semantic interoperability
- cross-domain harmonization/integration/migration/sharing
  - community-driven (eg. SMBL.org, BioSharing catalogue), industry-driven (eg. Pistoia Alliance), regulatory-driven (eg. CDISC), clinical (eg. HL7)
  - reimbursement (CPT, ICD) and HITECH EMR/MU
Genome Technology

October 2012

Readying Clinicians for Genomics

A flood of genomic information is poised to pour into the clinic. Are doctors ready?
“We don’t teach (medical) students how to interpret lab results or how to pick them. We’re spending 61 to 302 hours in anatomic pathology and nine hours teaching laboratory medicine. To pass anatomic pathology you’ve got to pass a test. There are no tests for lab. medicine.”

Dr. M. Laposta MD. Ph.D.
Executive Vice-Chair of Pathology, Microbiology and Immunology
Vanderbilt Univ. School of Medicine
Member, CDC Clinical Laboratory Integration Into Healthcare Collaborative (CLIHC)
Reducing Multi-Dimensional Complexity and Diverse Data Sources to End-User Simplicity and Facile Decisions
The Imminent Arrival of the Zettabyte ($10^{21}$) Era

The Emergence of Large Scale, Integrated Data and Knowledge Networks:

Profound Consequences for Individuals, Enterprises, Infrastructure, Investment, Education and Public Policy
Managing the Impact of PB and TB-Scale Datasets on Archives, Data Centers and User Communities

- current-approaches for data-access/download not sustainable
- network bandwidth limitations on large scale transfer
- need for rapid access to data in multiple distributed archives
- most end-user local capacities lack power to process large TB/PB datasets
- looming impact of austerity budgets on infrastructure and personnel
- prospect of serious performance degradation/breakdown
Data-Driven Knowledge, Intelligence and Actionable Decisions

- changing the nature of discovery
  - hypothesis-driven versus unbiased analytics of large datasets (patterns, rules)
- changing the nature of explanation
  - statistical probabilities versus unitary values
- changing the cultural process of knowledge acquisition
  - large scale collaboration networks, open systems, social media
- changing how knowledge is analyzed and used
  - increased quantification, complex analytics and decision-support systems
- changing education and training
Cross-Domain Convergence, Complexity and Increasing Dependency on Data-Intensive Methods and New Knowledge Networks

- systems-focused, big data sets mining and analytics
- reductionist, investigator-centric datasets and hypotheses
Silos Subvert Solutions: The Slow Response of Biomedicine to Technology Convergence and Cross-Disciplinary Requirements

- Anachronistic curricula
- Institutional sclerosis and career barriers
- Inadequate cyberinfrastructure
- Limited linkage between experiment and theory and iterative refinement of models and simulations
- Predominance of investigator-centric, reductionist approaches
- Domain silos reinforced by funding policies
- Qualitative data
- Poor standardization and reproducibility

Inadequate cyberinfrastructure
New Conceptual, Methodological and Organizational Frameworks for Data-Intensive Biomedical R&D

increasing dependency on systems-based, data-intensive analytics and new knowledge networks

agile knowledge networks
new curricula and career rewards
intelligent systems
integration and analytics for large scale datasets
ontologies, semantics, facile data sharing and interoperabilities

technology and cross-domain convergence
large scale team-based projects
increased automation
quantitative data open data
● the potential economic and health benefits from biomarkers for molecular diagnostic profiling, rational treatment selection and continuous health monitoring transcend any other current category of healthcare innovation

● realization of this potential will depend not only on technological advances but equally on overcoming entrenched cultural, institutional and economic interests
realization of this potential will not be straightforward and will require:

- improved technical standards for biomarker R&D
- sophisticated integration of complex multidisciplinary expertise (from silos to systems)
- proactive (inter)national leadership to establish comprehensive resources for biobanks, cyberinfrastructure and HIX data inter-operability
- new clinical trial designs for Rx/MDx combinations
- streamlined updating of SOC guidelines to reflect disease subtypes, patient heterogeneity and predisposition risk
The Primacy of Biomarkers and Biosignatures in Charting a New Ecology for Healthcare

Sustainable Health Systems

Identification of Biological Network Dysregulation in Disease

Biomarkers and Biosignatures

Digital Medicine

Molecular Medicine
The Primacy of Biomarkers and Biosignatures in Charting a New Ecology for Healthcare

- **Digital Medicine**
- **Molecular Medicine**
- **Network Dysregulation**
- **Biomarkers and Biosignatures**

**DISRUPTIVE INNOVATION**

**NEW, SYSTEMS-BASED, END-TO-END STRATEGIES**

**TRANSFORMATIVE TECHNICAL, INSTITUTIONAL AND CULTURAL CHANGES**

**Sustainable Health Systems**
Slides available @ http://casi.asu.edu/