Novel Biomarkers Associated with Future CVD Events

L. Kristin Newby, MD, MHS, FACC, FAHA
Professor of Medicine, Division of Cardiology
Duke University Medical Center

Disclosures

- Research contracts: NIH, PCORI, Amylin/Bristol Myers Squibb, MURDOCK Study, Google Life Sciences, GlaxoSmithKline, Sanofi
- Consulting: BioKier, CardioDx, Cubist Pharmaceuticals, Roche Diagnostics, DemeRx, MedScape/TheHeart.org, Philips Healthcare, NIH, Merck & Company, Inc.
- Organizations: Society of Cardiovascular Patient Care, Journal of the American Heart Association
- Full listing see w*ww.dcri.duke.edu/research/coi.jsp*

Increasing Emphasis on "Novel Biomarkers" of Future CV Events

The Fundamental Premise

- Health and disease reflect the complex interplay of biology, environment, socioeconomic and cultural influences, and time
- Characteristics that result in transitions between health and disease or herald the coming of these transitions, are poorly understood and only superficially characterized

Increasing Emphasis on "Novel Biomarkers" of Future CV Events

- Therapies (drugs, devices and behavioral interventions) are developed to treat all patients with the same clinical diagnosis – "one size fits all"
- Many therapies work in only a fraction of the patients for which they are prescribed
- Over 100,000 people die annually from medical errors or adverse events from therapy
- So, although people are living longer and are more functional than ever before, we have room for significant progress in getting the right treatment to the right person ("precision" medicine)

Combining Clinical and Molecular Data Will Redefine How We Manage Diseases

- Quantify risk
  - Predict death and disability
- Establish diagnoses earlier
- Prevent disability by treating earlier
- Use healthcare resources strategically
  - Stratify treatment based on expected benefit and risk
- Molecular data must provide incremental information to readily available clinical features and testing results

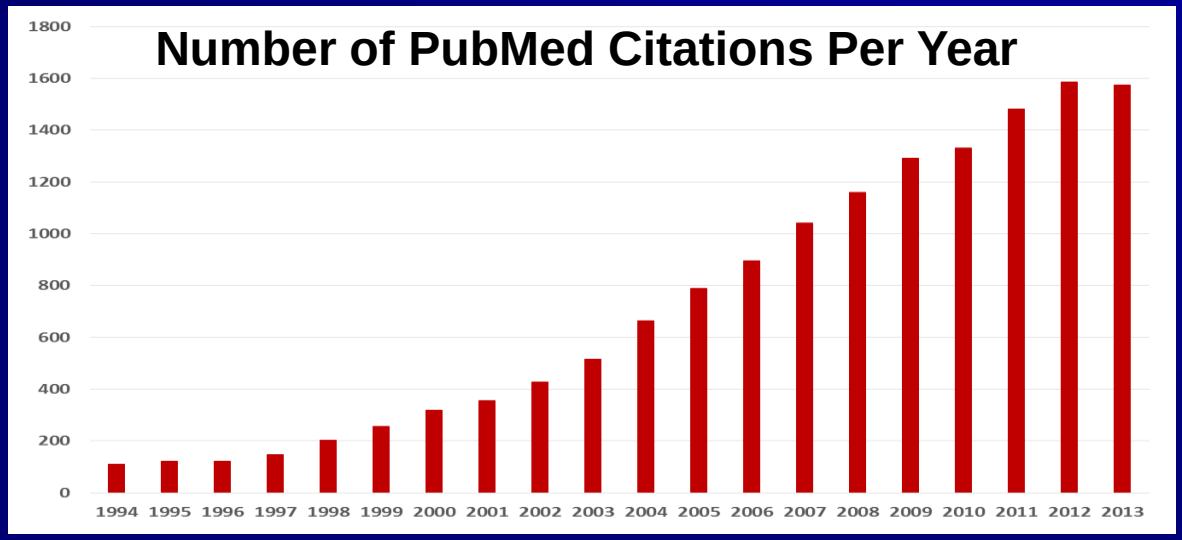
Table 3. Phases of Evaluation of a Novel Risk Marker

1. Proof of concept—Do novel marker levels differ between subjects with and without outcome?
2. Prospective validation—Does the novel marker predict development of future outcomes in a prospective cohort or nested case-cohort/case-cohort study?
3. Incremental value—Does the novel marker add predictive information to established, standard risk markers?
4. Clinical utility—Does the novel risk marker change predicted risk sufficiently to change recommended therapy?
5. Clinical outcomes—Does use of the novel risk marker improve clinical outcomes, especially when tested in a randomized clinical trial?
6. Cost-effectiveness—Does use of the marker improve clinical outcomes sufficiently to justify the additional costs of testing and treatment?

Hlatky, et al Circulation 2009

The State of the Field

- Hundreds of protein biomarker associations with cardiovascular risk and outcomes appear in peer reviewed literature each year



- Reported associations are often in isolation
  - of other putative biomarkers of risk
  - of full consideration of clinical predictors of risk

Circulation
Cardiovascular Genetics

Original Article

Simultaneous Consideration of Multiple Candidate Protein Biomarkers for Long-Term Risk for Cardiovascular Events

Sharif A. Halim, MD, MHS; Megan L. Neely, PhD; Karen S. Pieper, MS;
Svati H. Shah, MD, MHS; William E. Kraus, MD; Elizabeth R. Hauser, PhD;
Robert M. Califf, MD; Christopher B. Granger, MD; L. Kristin Newby, MD, MHS

Background—Although individual protein biomarkers are associated with cardiovascular risk, rarely have multiple proteins been considered simultaneously to identify which set of proteins best predicts risk.

Halim SA, et al. Circ Cardiovasc Genet 2015

Novel Biomarkers Must Provide Incremental Information

MURDOCK Survival Model Top 10 predictors

	HR	X2	C-index
Age		129	0.792
(per 5y >60)	1.24 (1.18-1.30)		
(per 5y up to 60)	1.09 (1.01-1.18)		
Modifed Charlson	1.31 (1.23-1.39)	83	
BUN		57	
(per 5U >20)	1.11 (1.07-1.15)		
(per 5U up to 20)	1.22 (1.10-1.38)		
RDW (per 1%)	1.12 (1.08-1.16)	43	
Weight (per 10kg to 80)	0.78 (0.72-0.84)	43	
HR (per 5 bpm to 80)	1.10 (1.06-1.14)	26	
Female	0.66 (0.56-0.77)	26	
WBC (per 1000)	1.05 (1.03-1.07)	22	
Chest pain freq	0.86 (0.80-0.92)	20	
QTc (per 20 msec)	1.08 (1.04-1.11)	19	

Also: EF, DBP, Hbg, SBP, diabetes, Duke index, creatinine, smoking, LVH, afib/flut, sodium, CHF severity, LBBB

MURDOCK Death/MI Model Top 10 Predictors

	HR (95% CI)	X^2	C-index
Age (per 5y >60)	1.21 (1.17-1.26)	99	0.746
Modified Charlson	1.29 (1.22-1.36)	90	
RDW (per 1%)	1.12 (1.08-1.16)	45	
Duke CAD SI (per 10 u)	1.06 (1.04-1.08)	38	
Weight (per 10kg to 80)	0.81 (0.76-0.87)	37	
BUN (per 5U >20)	1.08 (1.04-1.12)	33	
Diabetes	1.34 (1.19-1.51)	22	
Female	0.73 (0.64-0.85)	18	
Cigarette smoking	1.27 (1.13-1.42)	15	
QTc (per 20 msec)	1.06 (1.03-1.09)	14	

Also: HR, CP frequency, SBP, hemoglobin, DBP, WBC, EF, afib/flutter, LBBB, CHF severity, creatinine

Note: model is stratified by presentation type (AMI, outpatient or hospitalized patient) due to violation of proportional hazards assumption

N=6667 individuals with or at risk for CAD

Incremental Contribution of Biomarkers

- **53 biomarkers**
  - Literature reports of at least modest association with CV outcomes
  - Expert panel review and marker referral
  - Inflammation and atherosclerosis, thrombosis, myocardial necrosis, endothelial dysfunction and extracellular matrix remodeling, hemodynamic distress and metabolism
- **Marker assay platform**
  - MesoScale Discovery
  - Luminex
- **Intra-assay CV <20% for 88% of assays**

Statistical Methods

- **Penalized logistic regression using Elastic Net method**
  - Penalty on size of estimate coefficients, shrinking estimated coefficients of nonimportant variables to 0
  - Allows for individual consideration of correlated variables
- **3 models fit, each with 5-fold cross validation**
  - Protein biomarkers alone
  - Protein biomarkers in clinical model
  - Protein biomarkers or clinical variables
- **500 bootstrap samples**
  - Strong evidence: variable selected in $\geq$85%
  - Moderate evidence: variable selected in $\geq$70 and <85%

Biomarkers with Incremental Contribution

	Model 1		Model 2		Model 3	
	% of 500 bootstrapped samples selected	Odds Ratio*	% of 500 bootstrapped samples selected	Odds Ratio*	% of 500 bootstrapped samples selected	Odds Ratio*
ICAM-1	87.4	1.83	80.2	–	88.4	1.79
MMP-3	88.4	1.23	74.7	–	89.2	1.19
sCD40L	91.6	1.16	90.0	1.15	97.4	1.20
NT-proBNP	99.8	1.21	62.3	–	91.0	1.10
IL-6	94.2	1.22	78.6	–	93.2	1.19
IGFBP-2	99.8	1.30	82.2	–	97.6	1.17
Age (per 5 yr)	–	–	100	1.26	94.8	1.26
RDW	–	–	100	1.19	97.6	1.23
NYHA Class	–	–	100	1.17	85.2	1.12
Baseline hbg	–	–	100	0.91	86.2	0.91
Diabetes	–	–	100	1.60	85.6	1.53

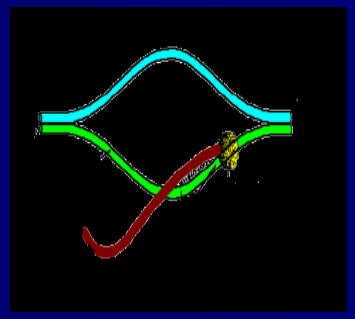
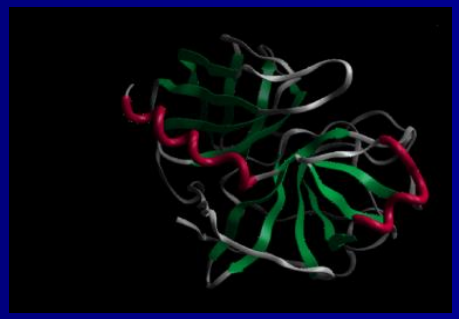
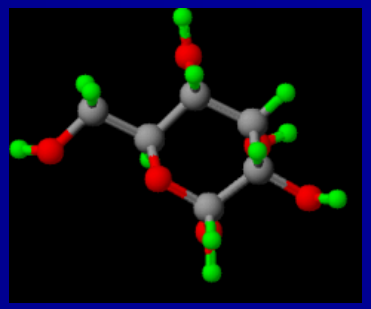
Biomarkers with Incremental Contribution

- Biomarkers selected represented inflammation and atherosclerosis, thrombosis, hemodynamic stress, metabolism, and vascular/endothelial function
- Notably absent:
  - SAA, hsCRP, TNF-α, TNF-β, IL-1β, IL-2, IL-4, IL-10, MCSF, GCSF, IL-1α, IL-1RA, IL-18, MCP-1, Lp-PLA2, E-selectin, P-selectin, LBP, RANTES (inflammation)
  - VCAM-1, ICAM-3, Apo A1, Apo B, Apo E (atherosclerosis)
  - fibrinogen, thrombomodulin, D-Dimer, PAI-1, vwF, tPA (thrombosis)
  - MMP-1, MMP-9, TIMP-1, bFGF, sFlt-1, PIGF, VEGF, PAPP-A, MPO, PDGF AA, PDGF AB/BB, OPGN (endothelial dysfunction)
  - CKMB, myoglobin, troponin I (myocardial necrosis)
  - Growth hormone (metabolism)
  - GDF-15 (growth, remodeling)

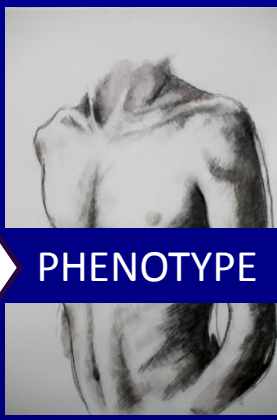
Key Points

- **"Proof of concept" for distilling the proliferating literature on protein biomarkers in cardiovascular risk assessment**
  - Assess associations of multiple, highly-correlated putative protein biomarkers with outcomes simultaneously in the context of one another and clinical information
  - Identify high priority candidates for further assessment in risk prediction, reclassification
  - Applicable to other "big p, little n" problems created with "-omics" platforms and EHRs
- **Limitations**

Genome | Transcriptome | Proteome | Metabolome

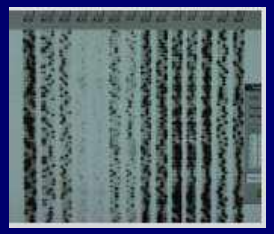
DNA N~$2x10^4$ → RNA N~$4x10^4$ ⇉ Protein N~10^6 → Modified Proteins N~10^7 —Enzymes→ Metabolites N~$5x10^3$



ENVIRONMENT

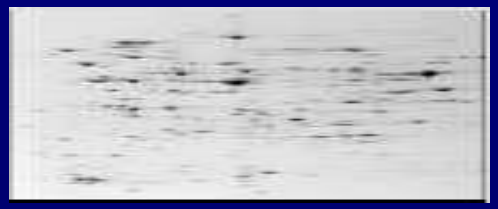
PHENOTYPE

Sequencing

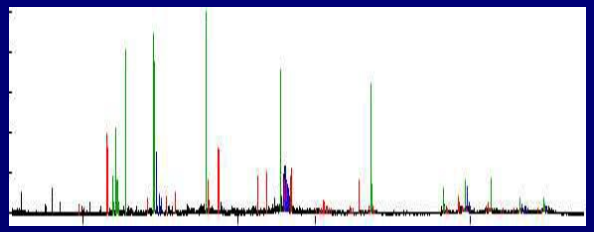


Microarray

2DGel, ELISA



GC, MS, NMR

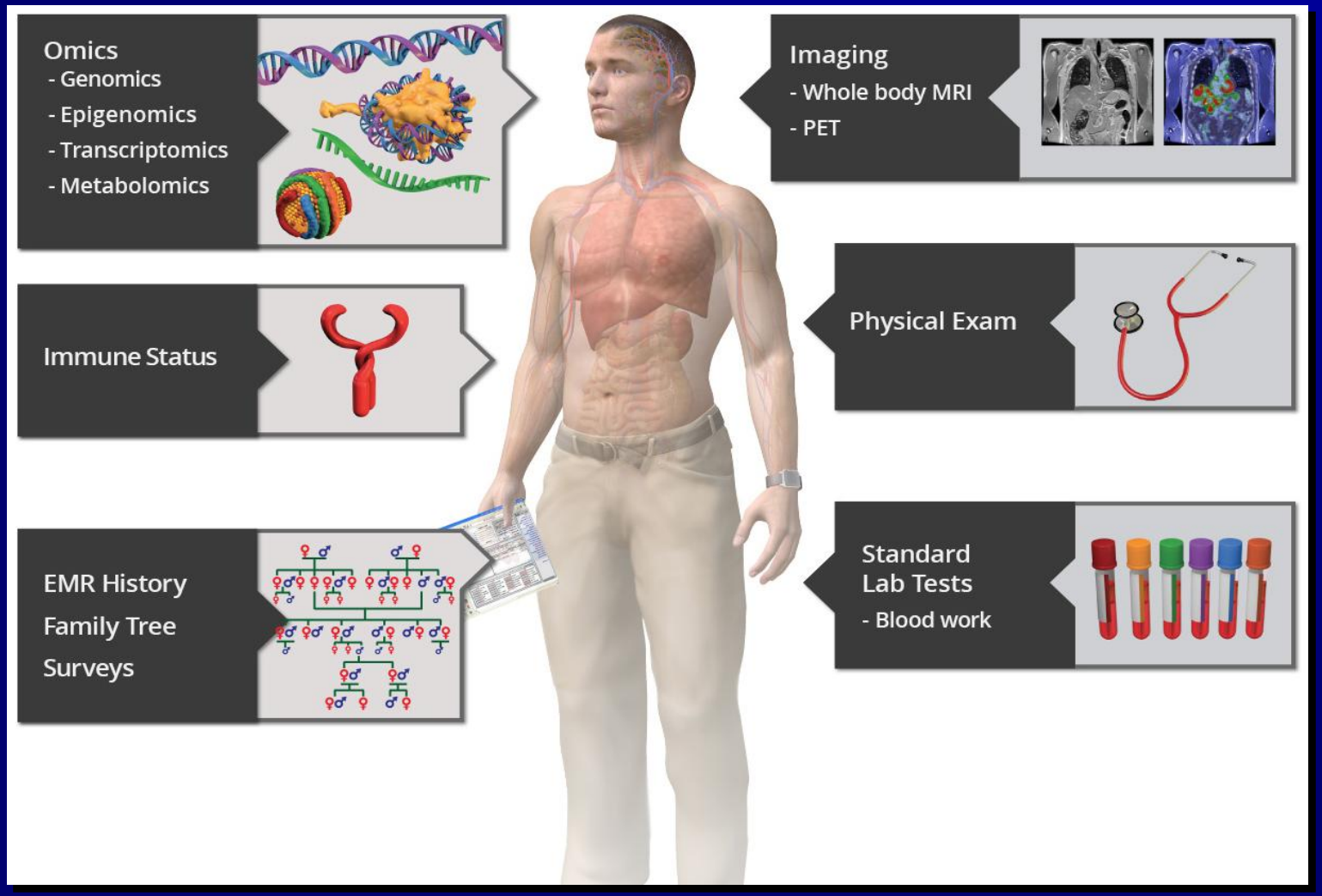


Modified from GD Lewis, R Gerszten et. al. JACC 2008;52;118

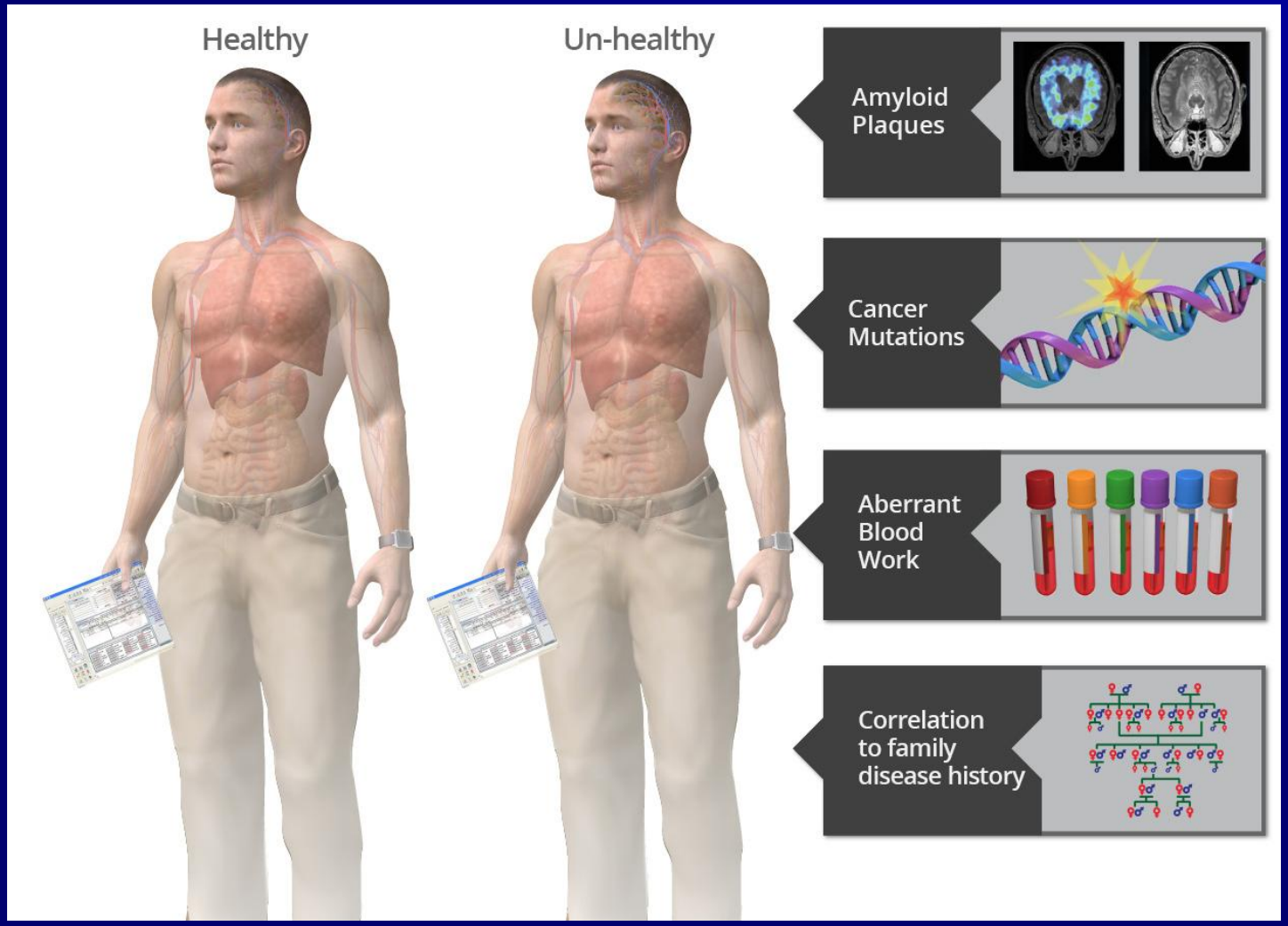
Big Challenges in Biomedicine

- **Lack of significant information over the time dimension**
  - Measurements to assess biology and human health are made periodically in visits to healthcare or for research
- **Missing systems biology**
  - When developing concepts of human biology or drug development we make limited measurements focused on specific mechanisms—we're looking "under the lamppost"
- **Missing the opportunity to measure the interactions of biology, sociology, environment and decision-making that could enable optimization of individualized and population health**
  - Although we know that health and disease are the product of the interactions of genes, multiple derivative biological systems, environment, social context and personal decisions, we tend to look at one part of the time

Characterize Human Health



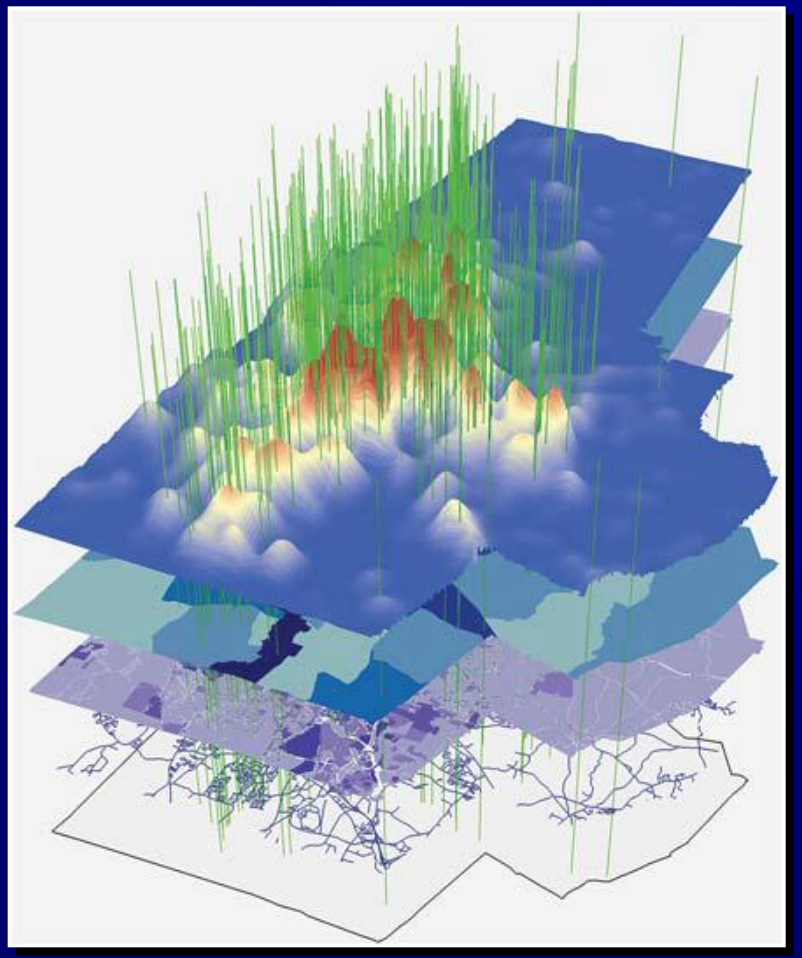
Characterize Transition to Disease

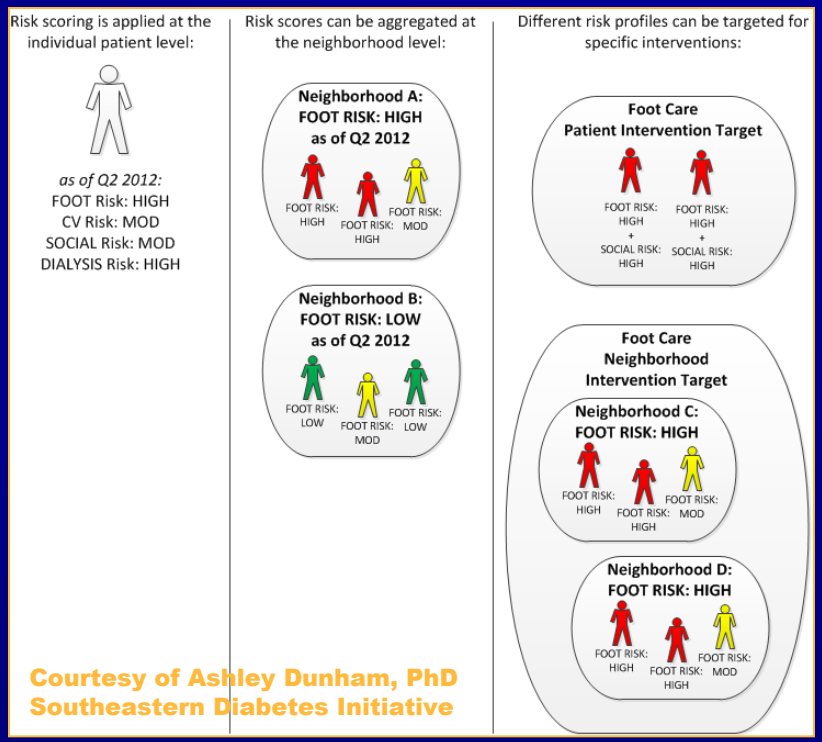


Envisioning a Healthcare Ecosystem of Big Data
Ultimate Goals

- Integrate personal (clinical and biological) and external information to enable individuals, neighborhoods and populations to:
  - optimize health
  - prevent disease
  - monitor treatment
  - enable people to be as functional as possible
- Provide physicians and healthcare systems with continuously updated estimates of individual risk and the health and health behaviors of neighborhoods and populations
  - enable directed education, prevention and treatment programs

Envisioning a Healthcare Ecosystem of Big Data





Courtesy of Ashley Dunham, PhD
Southeastern Diabetes Initiative

Top layer: concentrations of diabetes patients

Next layer down: percentage single female head of household

Purple layer: another indicator of economic status.

Bottom layer maps the county boundary and streets

Vertical green spines: longitude and latitude coordinates where diabetes patients live and locations of key social or commercial institutions, that can be used to link all of these disparate data sets together based on shared geography.

Miranda ML, et al Health Affairs 2013;32:608-1615

Envisioning a Healthcare Ecosystem of Big Data
Ultimate Goals

- Use a more profound understanding of health and disease to inform development of new therapeutics and diagnostics
  - Enable “precision cardiovascular medicine”
  - Leverage an infrastructure of well characterized individuals for future studies
  - Leverage multiple data sources in population selection, post-marketing surveillance of new drugs and devices
- Provide these opportunities at a very low cost per individual at a large-scale