Chain Reactions: Information Technology and Biomedical Discovery

Russ B. Altman, MD, PhD

Departments of Bioengineering, Genetics, Medicine & (by courtesy) Computer Science

Stanford University

Biomedical Informatics

The use of advanced IT to organize & understand biomedical data:

- DNA sequence (genome)
- Molecular (binding & catalysis)
- Cellular (gene expression)
- Organ (cellular communication)
- Organism (Electronic Medical Records)
- Populations (public health, social networks)

Informatics exists because of IT explosion

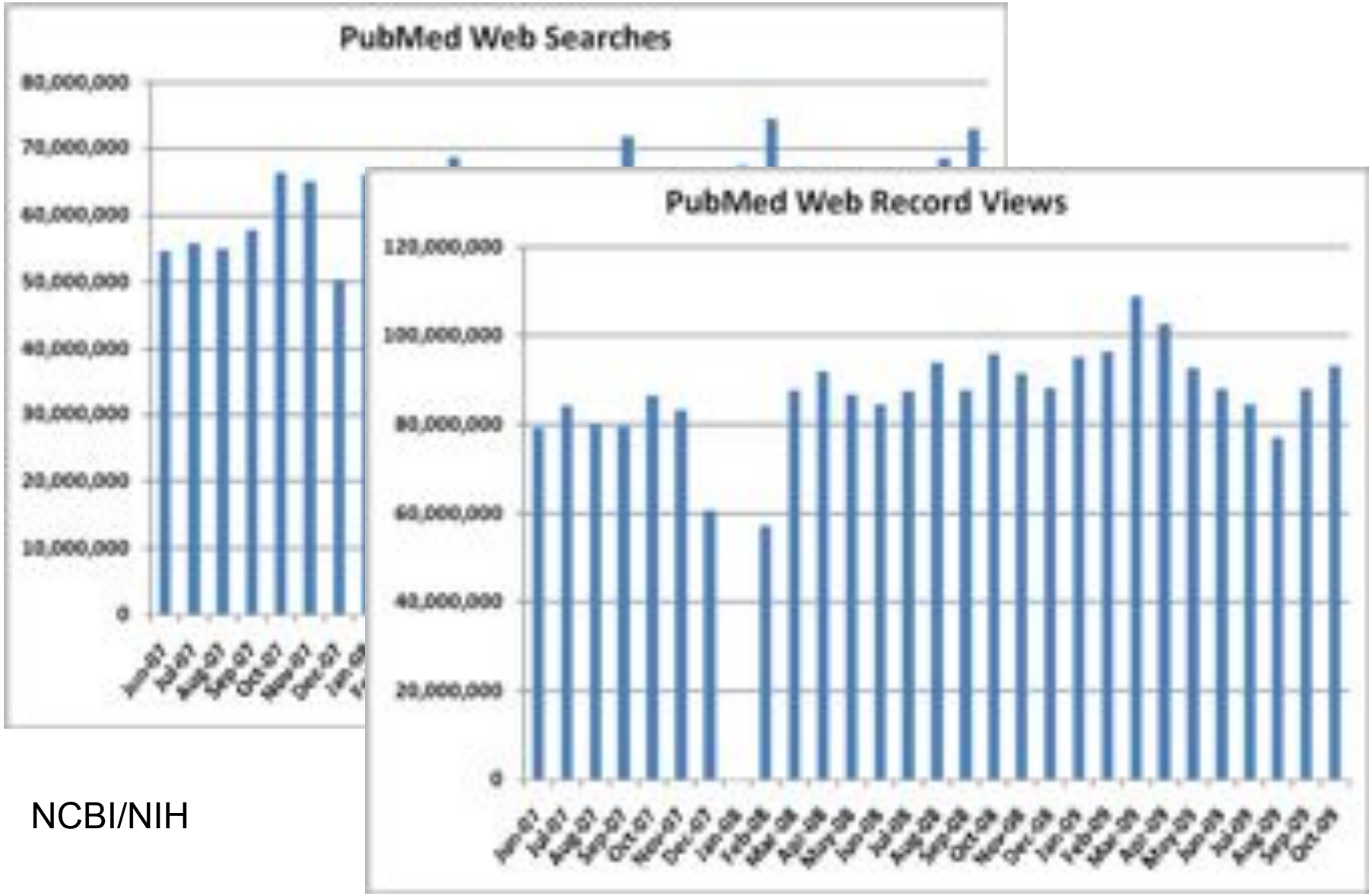
- NIH Institutes support
- My course introduced 1994: "Representations & Algorithms for Molecular Biology"
- DOE leadership in human genome
- NSF role in biological database activities, supercomputer centers
- Industry (Microsoft, IBM, Oracle & others)

Bioinformatics current challenges

Key network & information technologies today

- Entire biomedical literature available for search every day (20 x 10^6 abstracts)
- Large stores of DNA data, 3D structure data, gene expression data freely available
- High precision physical simulation on clusters, GPUs to get to millisecond timescale
- Imaging data: increasing resolution and dynamic range
- Crowd sourced health data: from medical records to public health to search engines and social networks.

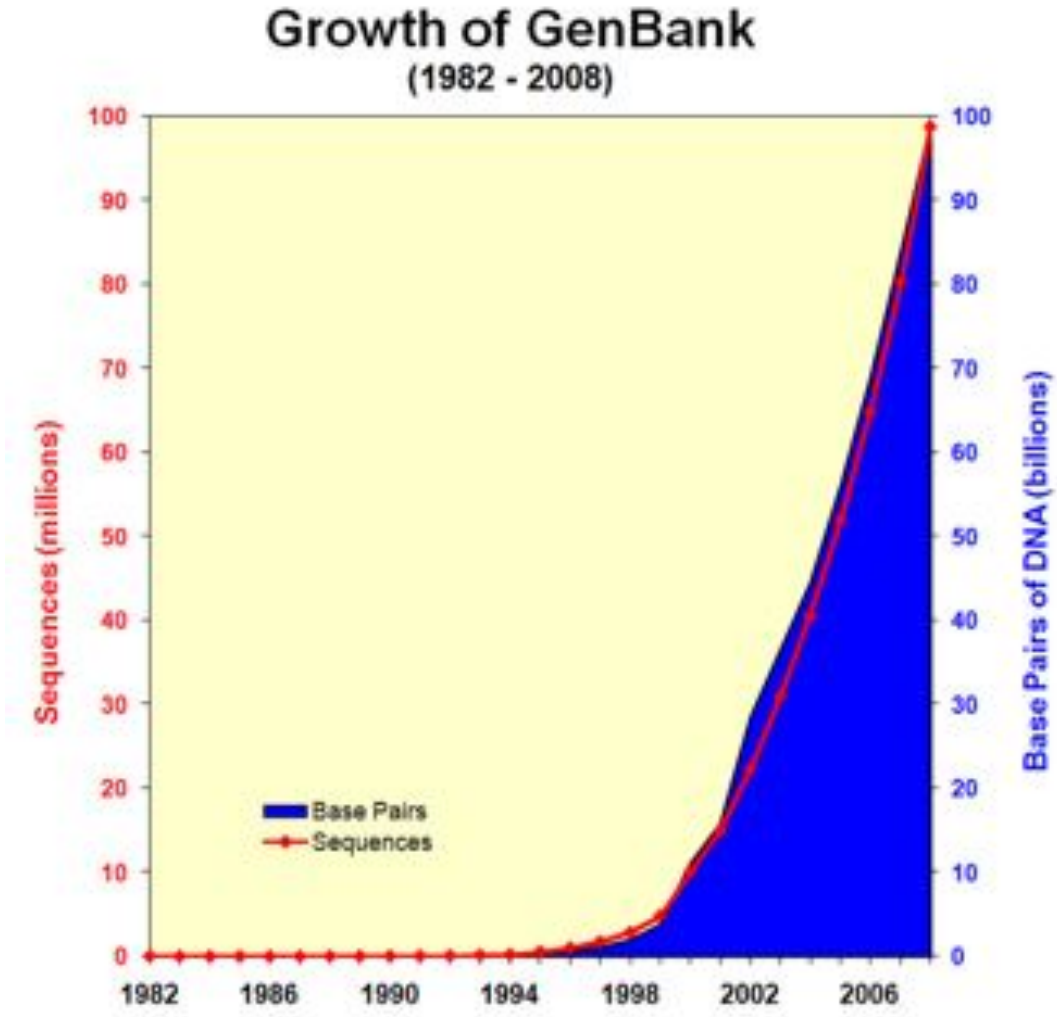
Routine access to complete biomedical literature



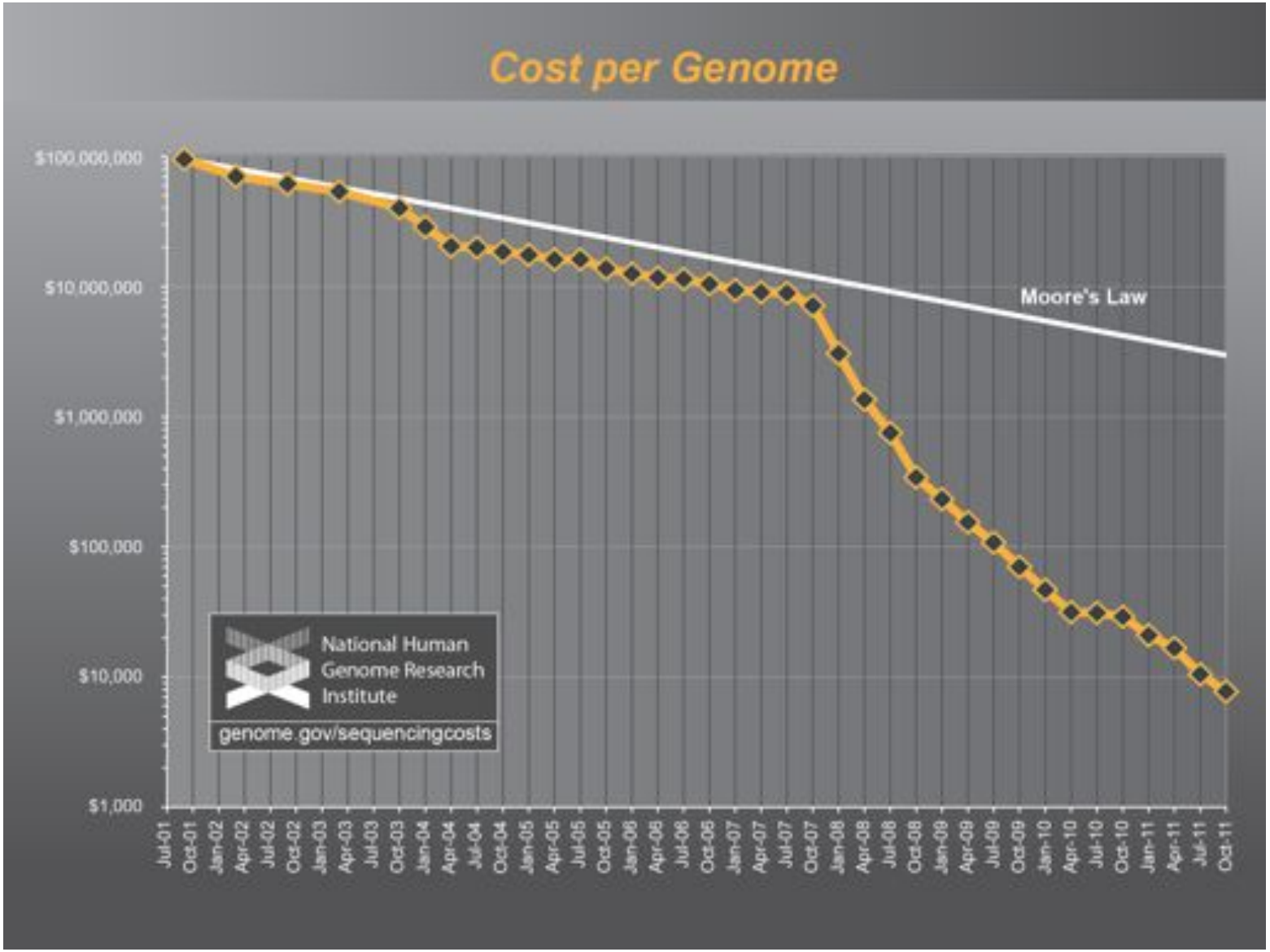
NCBI/NIH

Freely available large stores of DNA, protein, metabolite data

National Center for Biotechnology Information (NCBI) at National Library of Medicine (NLM), NIH



DNA sequencing outpacing Moore's law



National Human Genome Research Institute (NHGRI), NIH

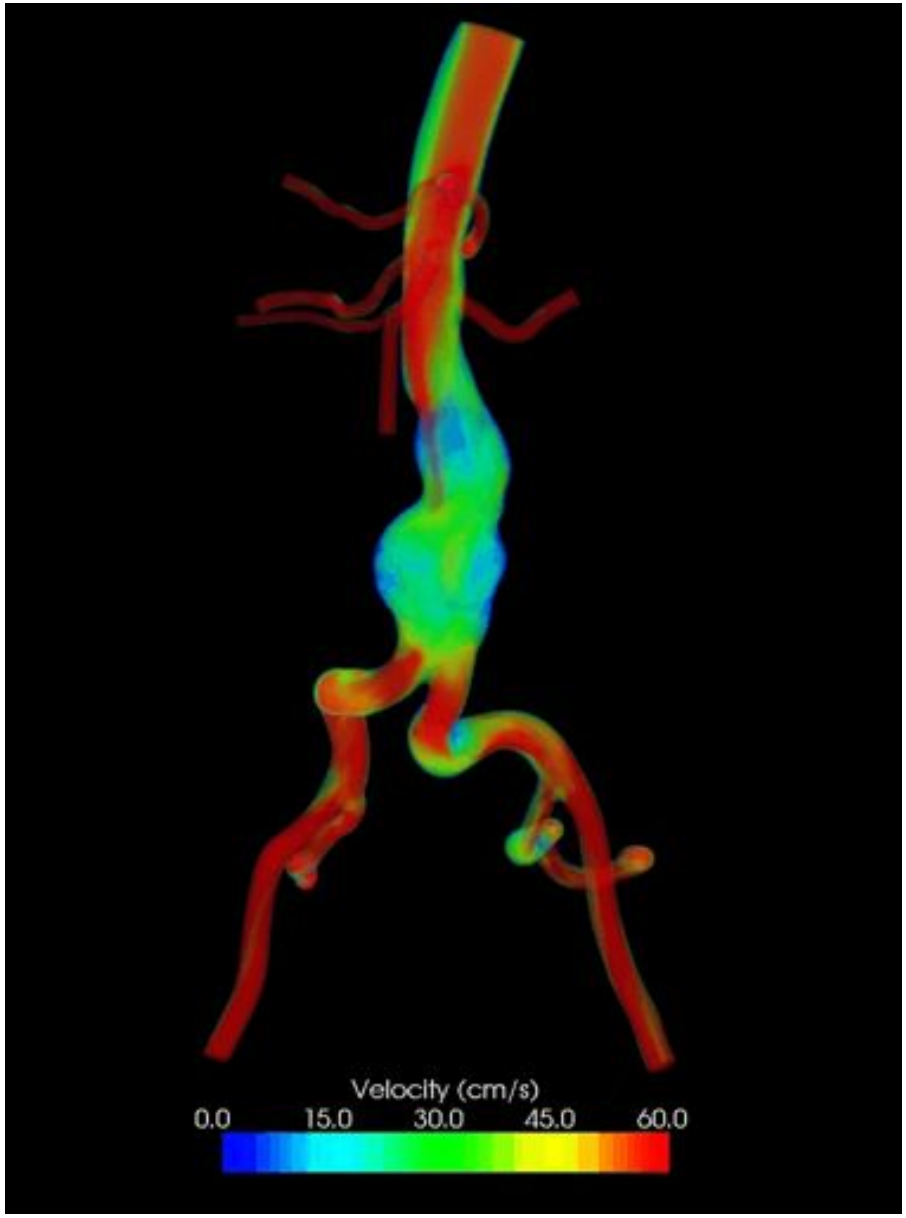
High precision simulations of 3D dynamics of molecules

Clusters, Screensavers, GPUs…getting to biological timescales.

Folding of Villin protein.

Vijay Pande,
Stanford Dept of Chemistry

Biomedical image data moving through hospitals every day



3D reconstruction of abdominal aortic aneurysm with simulation of arterial blood flow.

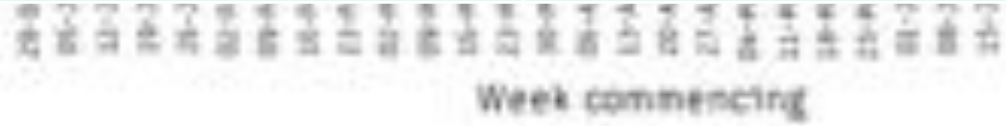
Charles Taylor
Dept of Bioengineering
Stanford University
&
HeartFlow Inc.

Crowd sourcing for detecting disease, drug interactions, patient support


FIGURE 3

Influenza-like illness (ILI) from GP sentinel surveillance,

Drugs in search with…	All words associated with hyperglycemia		More specific words associated with
Paxil ONLY	11.3%	(1.4x)	
Pravachol ONLY	7.8%	(1x)	
Paxil & Pravachol	25.7%	(3.3x)	


Week commencing

Two data trends

Basic biological measurements of genome and biomarkers (metabolites, proteins) are getting smaller, cheaper and more accurate.

Healthcare system is increasingly instrumented with electronic medical records, capture of individual patient physiology

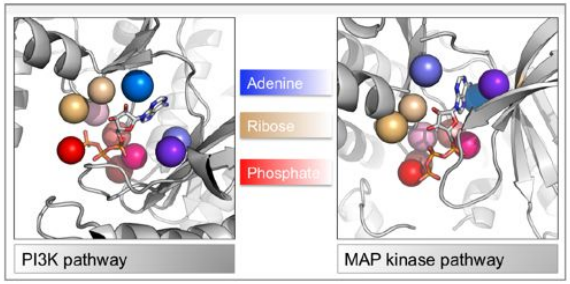
Emerging challenges

Measure, store and integrate big data ranging from the molecular and organismal level.

Discover patterns that allow us to improve diagnosis, prognosis and treatment = personalized medicine.

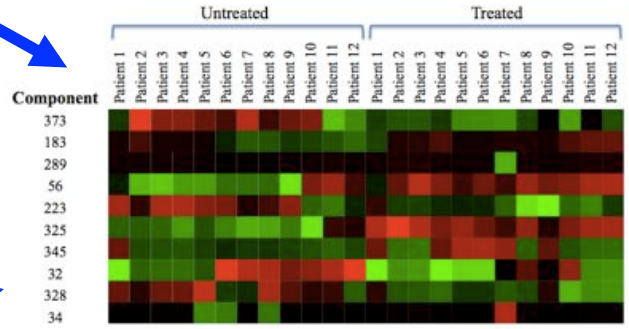
Thus, the emerging network of data….

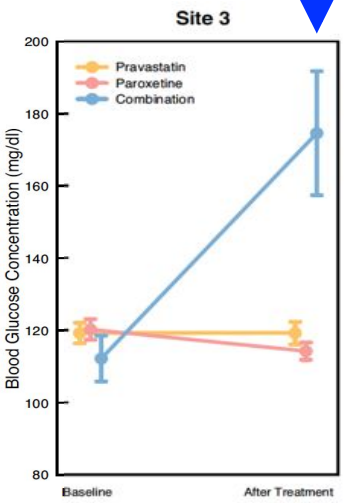
Target structure & dynamics



Drug recognition & binding

Cellular response & pathways

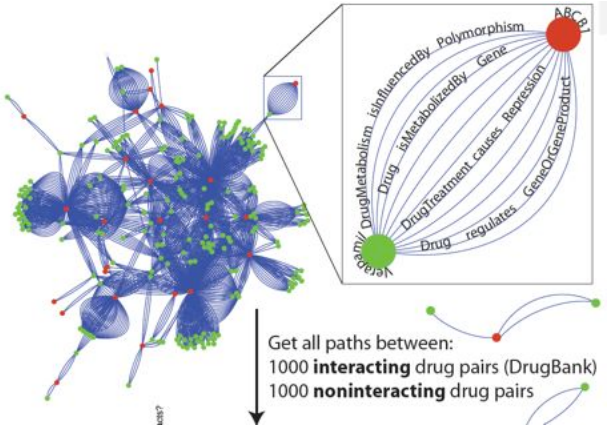




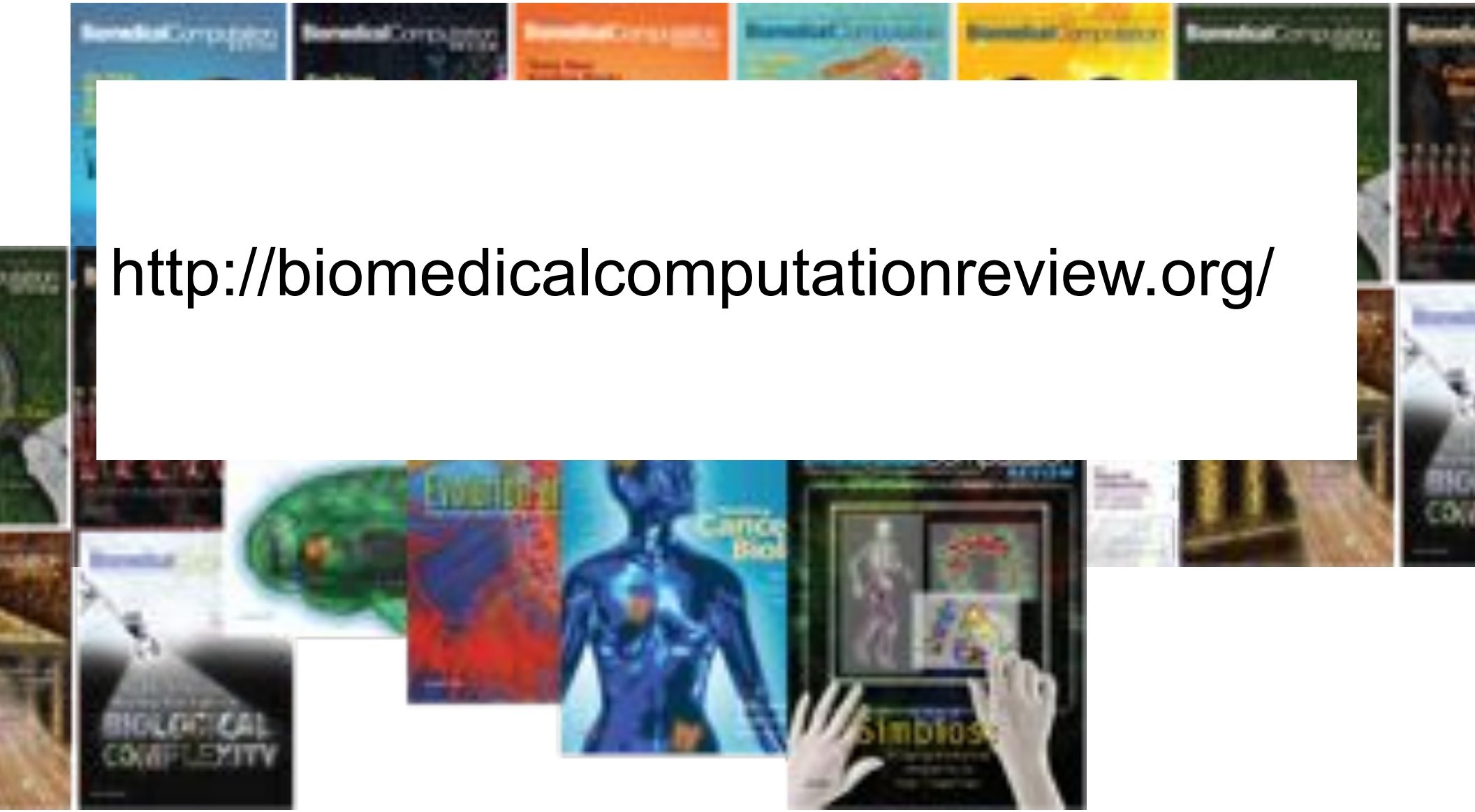
Clinical response EMR datamining

Clinical & Population effects reporting/sensing

Text mining of gene, drug, phenotype associations



Biomedical Computation Review



Thanks!

russ.altman@stanford.edu