High-throughput Health Status Measurement using CAT in the Era of Personal Genomics: Opportunities and Challenges

> Eswar Krishnan MD M.Phil Stanford University, CA

What Am I?



Asilomar October 2011



Disclosures*

Abbott Amgen Ardea* **Centocor Ortho-Biotech** Takeda* **URL** Pharma* **IMS Health** Metabolex* Savient* Pfizer* Merck* **Proctor and Gamble Bristol Myers Squibb**

***NONE RELEVANT FOR THIS TALK**

Definitions

- Throughput-
 - Average rate of data collection delivery
 - Number of hits per experiment
- High throughput in medical research thousands of data points collected over hours
- Personalized medicine

– Medicine customized to individual patients

PROMIS **Over \$70 million investment**



J.S. Department of Health

Supported by the





Musculoskeletal and Skin Diseases



PROs are different from educational status

- Patient reported outcome- Information communicated directly by the patient

 Often has independent clinical value or meaning
- Motivation dependent
- No 'cheating'
- No right or wrong answer
- Cannot put constraints
- Group data vs. individual patient data
- Repeated measures
 - Irregular time intervals

Older days

- Simple questionnaire
 - Pain questionnaire
 - Depression Questionnaire
 - Fatigue questionnaire



PROMIS[®] Profile Short Forms (29-43-57 items)



PROMIS® Pediatric Banks: v1.0

Domains	ltems in Bank	ltems in Short Form
Emotional Distress – Anger	n/a	6
Emotional Distress – Anxiety	15	8
Emotional Distress – Depression	14	8
Fatigue	23	10
Pain – Impact	13	8
Physical Function-Mobility	125	10
Physical Function-Upper Extremity	29	8
Peer Relationships	15	8
Asthma	17	8



And growing



And growing

Challenge 1

administration

Good CAT

NOT SO GOOD CAT



Your scores for the CATs you completed are shown below.

The diamond + is placed where we think your score lies. This diamond is placed on your T-Score, which is a standardized score that is based on an average score of 50, based on responses to the same questions in the United States general population. The T-score also has a standard deviation of 10 points, so a score of 40 or 60 represents a score that is one standard deviation away from the average score of the general US population.

The Standard Error (SE) is a statistical measure of variance and represents the possible range of your score. The lines on either side of the diamond in your profile report show the possible range of your actual score around this estimated score. It is very likely that your score is in the range of these lines.



We have not reached the fullpotential



Challenge 2

How to Analyze and Integrate the data ?





Repeated over Time

Magnitude and complexity

ARAMIS example:

- 30 measurements over 15 years;
- >121,000 observations
- 12 PRO variables; 120 other variables
- Non-independent



Data

Me

Lets talk about genomics

Central dogma

The expression of the genetic information stored in the DNA molecule occurs in two stages:

- –(i) transcription, during which DNA is transcribed into mRNA;
- (ii) translation, during which mRNA is translated to produce a protein.

Central dogma



The Central Dogma of Molecular Biology

What is a DNA Microarray?

- Also known as DNA Chip
- Allows simultaneous measurement of the level of transcription for every gene in a genome (gene expression)
- Transcription?
 - Process of copying of DNA into messenger RNA (mRNA)
 - Environment dependant!
- Microarray detects mRNA, or rather the more stable cDNA



Fields, S. (Science, 2001)

What is a DNA Microarray? (cont.)



Clustering of example



fold	fold repressed			fold induced		
				2	10-	

Campbell & Heyer, 2003

Real DNA Microarray



Campbell & Heyer, 2003

Hierarchical Clustering



Eisen et al., 1998

Challenge 3

Melding patient Outcomes and genomic data

