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?

Clinical rheumatology

Cohort studies

Clinical Trial

Aging and Frailty

Gouty Arthritis

Teaching and mentoring

PROMIS work

Epidemiology and Psychometric work.
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

U.S. Department of Health and Human Services

Supported by the

National Institutes of Health

National Institute of Arthritis and Musculoskeletal and Skin Diseases
Advancing Patient-Centered Outcomes

PROMIS: A Single Source of PROs

Clinical Practice

Clinical research

Surveys (CDC)

Clinic

Hospital

NIH

Industry

FDA
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)

Mental
- Anxiety 29
- Depression 28

Physical
- Fatigue 95
- Pain Interference 41
- Sleep Disturbance 27
- Physical Function 86

Social
- Satisfaction with Roles 14
# PROMIS® Pediatric Banks: v1.0

<table>
<thead>
<tr>
<th>Domains</th>
<th>Items in Bank</th>
<th>Items in Short Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Distress – Anger</td>
<td>n/a</td>
<td>6</td>
</tr>
<tr>
<td>Emotional Distress – Anxiety</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>Emotional Distress – Depression</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Fatigue</td>
<td>23</td>
<td>10</td>
</tr>
<tr>
<td>Pain – Impact</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Physical Function-Mobility</td>
<td>125</td>
<td>10</td>
</tr>
<tr>
<td>Physical Function-Upper Extremity</td>
<td>29</td>
<td>8</td>
</tr>
<tr>
<td>Peer Relationships</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>Asthma</td>
<td>17</td>
<td>8</td>
</tr>
</tbody>
</table>
>30

Domains

And growing
>400 items

And growing
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.

<table>
<thead>
<tr>
<th>Your Score</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>48</td>
</tr>
<tr>
<td>Anxiety</td>
<td>64</td>
</tr>
<tr>
<td>Depression</td>
<td>62</td>
</tr>
<tr>
<td>Fatigue</td>
<td>34</td>
</tr>
<tr>
<td>Pain Behavior</td>
<td>43</td>
</tr>
<tr>
<td>Pain Impact</td>
<td>56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Your Score</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Function</td>
<td>38</td>
</tr>
<tr>
<td>Social Activity</td>
<td>48</td>
</tr>
<tr>
<td>Social Role</td>
<td>45</td>
</tr>
</tbody>
</table>
We have not reached the fullpotential
Challenge 2

How to Analyze and Integrate the data?
>30 Domains
>400 items
Repeated over Time
Magnitude and complexity

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

- Replication: DNA duplicates
- Transcription: RNA synthesis
- Translation: Protein synthesis

DNA → RNA → Protein
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
What is a DNA Microarray? (cont.)
Clustering of example

Campbell & Heyer, 2003
Real DNA Microarray

Campbell & Heyer, 2003
Hierarchical Clustering

Eisen et al., 1998
Challenge 3

Melding patient outcomes and genomic data