Introduction & Session overview

**Topic**
Developing PROMs to have more impact

**Panelists**

1. **Konrad Pesudovs, PhD** - Foundation Chair of Optometry and Vision Science at Flinders University, Adelaide, Australia
2. **Matthias Rose, MD, PhD** - Head of the Psychosomatic Medicine Centre at Medical Clinic of the Charité, Universitätsmedizin Berlin, Germany
3. **James Willig, MD, MSPH** - Associate Professor at the Department of Medicine, Division of Infectious Diseases at University of Alabama at Birmingham, Alabama, USA
4. **Helen Burstin, MD, MPH, FACP** - Chief Scientific Officer of The National Quality Forum, Washington DC, USA

**Agenda**
- Introduction (5 minutes)
- Panel presentations (~10-12 minutes each)
- Discussion, Q&A (35 minutes)
- Summary and closing (2 minutes)
Psychometrics concerns itself with turning qualitative measures into quantitative data

“I am in pain”

Intelligence

Ability to ambulate

Instrument

1, 2, 3, 4, …
Psychometrics concerns itself with turning qualitative measures into quantitative data

Construct, domain, trait, patient-reported outcome

“I am in pain”

Intelligence

Ability to ambulate

PROM

Instrument

Score

1, 2, 3, 4, …
Psychometrics concerns itself with turning qualitative measures into quantitative data.

Construct, domain, trait, patient-reported outcome

“I am in pain”

Intelligence

Ability to ambulate

PROM

Instrument

Score

1, 2, 3, 4, ...

PRO-PM - PRO-based Performance Measure
Important to differentiate between ordinal and interval scales, health outcomes are mostly measured on ordinal scales

<table>
<thead>
<tr>
<th>Data</th>
<th>Scale</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes can be rank-ordered (sorted), but distances between attributes have no meaning</td>
<td>Ordinal</td>
<td>Level of education, IQ, ability to ambulate, health outcomes, ...</td>
</tr>
<tr>
<td>Distance between attributes have meaning (averages can be computed), but ratio’s have no meaning</td>
<td>Interval</td>
<td>Temperature (°C), date</td>
</tr>
<tr>
<td>A meaningful absolute zero exists, a meaningful ratio can therefore be calculated</td>
<td>Ratio</td>
<td>Weight, mass, length, temperature (K)</td>
</tr>
</tbody>
</table>

“There is no rational basis for adding up a set of ordinal Likert scores, unless they have been shown to scale”*

*Source: Matthew Hankins, Senior lecturer in Public Health at University of Southampton, [2011 presentation](#)
**Item response theory** is the field of work concerning the translation of ordinal health outcomes scores to interval scores.

<table>
<thead>
<tr>
<th>Data</th>
<th>Scale</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes can be rank-ordered (sorted), but <strong>distances between attributes have no meaning</strong></td>
<td><strong>Ordinal</strong></td>
<td>Level of education, IQ, ability to ambulate, health outcomes, …</td>
</tr>
<tr>
<td><strong>Distance between attributes have meaning</strong> (averages can be computed), but ratio’s have no meaning</td>
<td><strong>Interval</strong></td>
<td>Temperature (°C), date</td>
</tr>
<tr>
<td>A meaningful absolute zero exists, a meaningful ratio can therefore be calculated</td>
<td><strong>Ratio</strong></td>
<td>Weight, mass, length, temperature (K)</td>
</tr>
</tbody>
</table>
Determining a questionnaire’s psychometric properties is an important step in its development.

Rasch analysis is one of many possible Item Response Theory models...

...which results in the “anchoring” of questions to a scale.

How much do you like...

<table>
<thead>
<tr>
<th>Question</th>
<th>High ‘difficulty’</th>
<th>Low ‘difficulty’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Q9</td>
<td>Q10</td>
</tr>
<tr>
<td>Q2</td>
<td>Q6</td>
<td>Q7</td>
</tr>
<tr>
<td>Q3</td>
<td>Q5</td>
<td>Q8</td>
</tr>
<tr>
<td>Q4</td>
<td>Q11</td>
<td>Q12</td>
</tr>
<tr>
<td>Q5</td>
<td>Q3</td>
<td>Q13</td>
</tr>
<tr>
<td>Q6</td>
<td>Q2</td>
<td>Q1</td>
</tr>
<tr>
<td>Q7</td>
<td>Q10</td>
<td>Q11</td>
</tr>
<tr>
<td>Q8</td>
<td>Q7</td>
<td>Q4</td>
</tr>
<tr>
<td>Q9</td>
<td>Q12</td>
<td>Q2</td>
</tr>
<tr>
<td>Q10</td>
<td>Q11</td>
<td>Q13</td>
</tr>
<tr>
<td>Q11</td>
<td>Q6</td>
<td>Q5</td>
</tr>
<tr>
<td>Q12</td>
<td>Q9</td>
<td>Q1</td>
</tr>
<tr>
<td>Q13</td>
<td>Q4</td>
<td>Q8</td>
</tr>
</tbody>
</table>

Copyright © 2015 by the International Consortium for Health Outcomes Measurement. All rights reserved.
PANEL PRESENTATIONS
Developing PROMs to have more impact

Konrad Pesudovs, PhD

Item Banking: A Generational Change in Patient-Reported Outcome Measurement

Konrad Pesudovs"
PROM Measurement

Level of Visual Ability

Most visually disabled person

Less demanding activity

Watching TV

Public transport

Day driving

Equal intervals

Reading

Night driving

More demanding activity

Level of Difficulty

How visually demanding is the activity?

Tall

Short

How impaired is the person?

Least visually disabled person
Item Banking

Figure 2.

Actual person/item map for all the VRAL items drawn from all 19 instruments. Persons and items are calibrated along the same linear scale with the ruler from Fig. 1 included to illustrate the parallel between the two figures. Each item is identified with the acronym of the instrument’s title, and its item number. The distribution of persons and items mismatches slightly indicating that more “difficult” items need to be added.
Computer adaptive testing
The Eye-tem Bank Project

- To develop the most comprehensive and technologically advanced patient-reported outcome measurement system for eye disease—“the Eye-tem Bank”
- To develop, validate, and implement an item banking and CAT system to assess ophthalmic quality of life (QOL) across all eye diseases

**Eye-tem Bank modules**
- Glaucoma
- Diabetic retinopathy
- Age-related macular degeneration
- Retinal detachment
- Other vitreo-retinal
- Cataract and Corneal opacities
- Corneal diseases
- Refractive error
- Strabismus and Amblyopia
- Uveitis spectrum of diseases
- Ocular inflammation other than uveitis
- Ocular-surface and Lacrimal
- Neuro-ophthalmic

**Eye-tem Bank QOL domains**
- Activity limitation
- Mobility
- Ocular comfort symptoms
- Visual symptoms
- General symptoms
- Emotional well-being
- Health concerns
- Convenience
- Social participation and well-being
- Economic (work and finance)
The Eye-tem Bank

<table>
<thead>
<tr>
<th>Domains</th>
<th>Glaucoma</th>
<th></th>
<th>DR</th>
<th></th>
<th>AMD</th>
<th></th>
<th>Common items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Unique</td>
<td>Total</td>
<td>Unique</td>
<td>Total</td>
<td>Unique</td>
<td></td>
</tr>
<tr>
<td>Activity limitation</td>
<td>88</td>
<td>14</td>
<td>120</td>
<td>46</td>
<td>114</td>
<td>29</td>
<td>65</td>
</tr>
<tr>
<td>Mobility</td>
<td>20</td>
<td>3</td>
<td>19</td>
<td>2</td>
<td>23</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Ocular comfort symptoms</td>
<td>23</td>
<td>15</td>
<td>10</td>
<td>2</td>
<td>25</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Visual symptoms</td>
<td>18</td>
<td>6</td>
<td>18</td>
<td>6</td>
<td>14</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>General symptoms</td>
<td>16</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Convenience</td>
<td>40</td>
<td>22</td>
<td>30</td>
<td>12</td>
<td>34</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Emotional well-being</td>
<td>54</td>
<td>16</td>
<td>48</td>
<td>10</td>
<td>72</td>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td>Health concerns</td>
<td>45</td>
<td>19</td>
<td>36</td>
<td>10</td>
<td>39</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>Social</td>
<td>23</td>
<td>8</td>
<td>21</td>
<td>6</td>
<td>21</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Economic</td>
<td>22</td>
<td>12</td>
<td>12</td>
<td>2</td>
<td>17</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>349</strong></td>
<td><strong>131</strong></td>
<td><strong>314</strong></td>
<td><strong>96</strong></td>
<td><strong>359</strong></td>
<td><strong>112</strong></td>
<td><strong>169</strong></td>
</tr>
</tbody>
</table>
Developing PROMs to have more impact

Matthias Rose, MD, PhD
Developing PROMs to have more impact

“Harmonisation”

ICHOM
London 2016

Matthias Rose

Department of Psychosomatic Medicine
Center for Internal Medicine and Dermatology
Charité Universitätsmedizin Berlin, Germany

Department of Quantitative Health Sciences
University of Massachusetts, USA
What is wrong with today’s instruments?
What is Wrong with Today‘s Measurement?

- **not precise**
  measurement precision does not meet required standards for individual clinical decision making

- **restricted measurement range**
  clinical instruments do not provide information for the general population

- **not comparable**
  same constructs are often measured with different instruments without being comparable
## Psychometric and Utility Related Concepts

<table>
<thead>
<tr>
<th>Concepts and Characteristics</th>
<th>SIP</th>
<th>HIE</th>
<th>NHP</th>
<th>QLI</th>
<th>COOP</th>
<th>QLQ</th>
<th>MOS FWBP</th>
<th>MOS SF-36</th>
<th>QWB</th>
<th>EURO-QOL</th>
<th>HUI</th>
<th>SF-6D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONCEPTS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Social functioning</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Role functioning</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Psychological distress</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Health perceptions (general)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Pain (bodily)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Energy/fatigue</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Sleep</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Cognitive functioning</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Quality of life</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Reported health transition</td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Financial Impact</td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

**SIP** = Sickness Impact Profile (1976)
**HIE** = Health Insurance Experiment surveys (1979)
**NHP** = Nottingham Health Profile (1980)
**QLI** = Quality of Life Index (1981)
**COOP** = Dartmouth Function Charts (1987)
**QLQ-C30** = EORTC

**MOS FWBP** = MOS Functioning and Well-Being Profile (1992)
**MOS SF-36** = MOS 36-Item Short-Form Health Survey (1992)
**QWB** = Quality of Well-Being Scale (1973)
**EUROQOL** = European Quality of Life Index (1990)
**HUI** = Health Utility Index (1996)
**SF-6D** = SF-36 Utility Index (Brazier, 2002)

adapted from Ware, Medical Care 1995
What is Wrong with Today’s Measurement?

- **not precise**
  
  measurement precision does not meet required standards for individual clinical decision making

- **restricted measurement range**
  
  clinical instruments do not provide information for the general population

- **not comparable**
  
  same constructs are often measured with different instruments without being comparable

- **not anchored**
  
  score values have no intuitive meaning
Anchoring and Standardization

<table>
<thead>
<tr>
<th>Fahrenheit</th>
<th>Reaumur</th>
<th>Celsius</th>
<th>Kelvin</th>
<th>Rankine</th>
</tr>
</thead>
<tbody>
<tr>
<td>1724</td>
<td>1731</td>
<td>1742</td>
<td>1848</td>
<td>1859</td>
</tr>
</tbody>
</table>

- Body temperature: ≈100°F
- Water freezes: 0°C (32°F)
- Water boils: 100°C (212°F)
- Absolute zero: 0 K (−273°C, −459°F)

Example: temperature
Can IRT item banks also solve those problems?
Item Bank - Depression

Representative Samples

Patients with Depression
7 Clinical Sites / 12 Health Centers
total n > 33,000

Questionnaire A

Questionnaire B

11 established instruments
Standardized Metric

<table>
<thead>
<tr>
<th>Beck Depression Inventory</th>
<th>Hospital Anxiety Depression Scale</th>
<th>Well-Being Index</th>
<th>Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>12</td>
<td>20</td>
<td>75</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>40</td>
<td>65</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>50</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>60</td>
<td>55</td>
</tr>
</tbody>
</table>

- Cut-Off Depression
- Mean Representative Sample
Standardized Metric

Temperature

- 36° Celsius
- 37° Celsius
- 38° Celsius
- 38.5° Celsius
- 39° Celsius
- 39.5° Celsius

- Cut-Off Fever
- Mean Representative Sample

Mercury (1714)
Quartz (1965)
Infrared (1992)

Celsius (1742)
Comparability

**HbA1c**

- Week 1: 8.6%
- Week 6: 6.5%

**Depression**

- Week 1: 10%
- Week 3: 12%
- Week 6: 8%

**Standard tool (HADS)**

- Week 1: 95%
- Week 3: 17%
- Week 6: 9%

**Cut off Depression**

- HADS: 8
- BDI: 4

---

Common Metrics
Comparing Scores from Different Patient Reported Outcomes using Item Response Theory

Background
In the field of measurement of health there is a plethora of instruments and questionnaires, so called Patient-Reported Outcomes. For example, it has been estimated that about 100 instruments are designed to measure (aspects of) depression or depressive severity, such as
- PHQ-9
- BDI I and II
- CES-D
- and many more.

This also applies to various other constructs such as Anxiety, Fatigue or Quality of Life, and: the more important a construct is considered, the more instruments have been developed.

These instruments differ in various ways, for example their underlying philosophies, their emphasis on different aspects of the construct, and their validation. Data obtained with different measures is often hardly to compare, making

Metrics
Please send any questions and feedback:
info@common_metrics.org

Contact
Please let us know, what you think about this site and feel free to send us your questions. We are grateful for any feedback and suggestions and provide support happily!
- info@common_metrics.de
- +49 30 450 529 104

Links
We would like to draw your attention to some other websites that might be interesting for you.

PROMIS
This page was built using bootstrap. The conversion app was made using R, shiny, ggplot2, mint and psjr.

Fisher et al. 2014
Are patient-reported health parameters disease specific?
Charaterizing the phenotype of multiple sclerosis–associated depression in comparison with idiopathic major depression


Unconstrained Model:

\[ \chi^2_{278} = 1.628, \ p < 0.01, \ CFI = 0.867, \ RMSEA = 0.053 \]

Constrained Model:

\[ \chi^2_{269} = 1.615, \ p < 0.01, \ CFI = 0.862, \ RMSEA = 0.053 \]

Model Comparison:

\[ \chi^2_{20} = 27.451, \ p = 0.123 \]
Health Parameters

Nephrology

- kidney failure

Hematology

- plasmacytoma

\( \gamma \)-globulin

- creatinine

\( \text{CH}_3 \)

σύμπτωμα (symptoma) = it falls together with something

but is not unique to it

disease specific measures = compilation of health parameters

fatigue
For which constructs do harmonized measurements make sense?
Suitable Constructs

Main health constructs: interval scales, e.g.:

- Depression
- Anxiety
- Fatigue
- Pain
- Shortness of breath
- Physical functioning
- Cognitive functioning
- Sleep functioning
- Social role functioning

Composite scores, preference instruments, treatment satisfaction, e.g.:

- Asthma control Questionnaire
- Anti-Clot Treatment Scale
- Alzheimer’s Disease Care Giver Performance Questionnaire
- Ascites Impact Measure
- Comprehensive Assessment of Satisfaction with Care
- Diabetes Foot Ulcer Scale
Conclusion
Conclusion

I. Item banks allow to move away from an instrument to a construct defined measurement system

II. Item banks provide a common metric for existing tools measuring the same construct

III. Item banks make sense for key health indicators, and will allow to treat subjective health assessments similar to other clinical markers
Developing PROMs to have more impact

James Willig, MD, MSPH
PROMs in Clinical Care Settings

James H. Willig, MD, MSPH
University of Alabama at Birmingham
One person monitoring completion real time

<table>
<thead>
<tr>
<th>MRN</th>
<th>First Name</th>
<th>Last Name</th>
<th>Birthdate</th>
<th>Earliest Answer</th>
<th>Recent Answer</th>
<th>Progress</th>
<th>Patient Page</th>
<th>Screening Report</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-11 09:20:16</td>
<td>2014-09-11 10:02:44</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-11 07:58:03</td>
<td>2014-09-11 08:07:11</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-11 06:48:49</td>
<td>2014-09-11 07:20:45</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 14:13:18</td>
<td>2014-09-10 14:20:07</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 11:51:36</td>
<td>2014-09-10 12:06:46</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 11:36:27</td>
<td>2014-09-10 11:36:27</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 09:53:51</td>
<td>2014-09-10 10:15:20</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 09:17:43</td>
<td>2014-09-10 10:01:44</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 09:01:23</td>
<td>2014-09-10 10:27:02</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 08:35:39</td>
<td>2014-09-10 09:02:09</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 08:11:21</td>
<td>2014-09-10 08:34:30</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2014-09-10 07:49:06</td>
<td>2014-09-10 08:04:09</td>
<td></td>
<td>Go</td>
<td>Go</td>
</tr>
</tbody>
</table>
## Patient-Based Measures Provider Feedback

**Printed 9/11/2014**

<table>
<thead>
<tr>
<th>NAME</th>
<th>MRN</th>
<th>DOB</th>
<th>Depression (PHQ-9)</th>
<th>Suicide Ideation (PHQ-9)</th>
<th>Tobacco Use</th>
<th>Alcohol Score</th>
<th>MINI Score</th>
<th>Substance Use (Past 3 months)</th>
<th>Antiretroviral Adherence (Past 4 weeks)</th>
<th>High-risk Behavior (Past 6 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>9/11/2014</td>
<td>0</td>
<td>0</td>
<td>Previously smoked</td>
<td>1</td>
<td>N/A</td>
<td>Marijuana</td>
<td>Excellent</td>
<td>Anal sex condom use</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3/6/2014</td>
<td>0</td>
<td>0</td>
<td>Currently</td>
<td>1</td>
<td>N/A</td>
<td>Marijuana</td>
<td>Excellent</td>
<td>All the time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9/5/2013</td>
<td>0</td>
<td>0</td>
<td>Currently</td>
<td>3</td>
<td>N/A</td>
<td>Marijuana</td>
<td>Excellent</td>
<td>Never had vaginal sex</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Less than half a pack a day</td>
<td>0</td>
<td>N/A</td>
<td>Less than half a pack a day</td>
<td>Currently</td>
<td>Never used</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Not at all</td>
<td>0</td>
<td>N/A</td>
<td>Not at all</td>
<td>Currently</td>
<td>Within the last week</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Not at all</td>
<td>0</td>
<td>N/A</td>
<td>Not at-risk (AUDIT-C)</td>
<td>Not at-risk (AUDIT-C)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>Not at all</td>
<td>0</td>
<td>N/A</td>
<td>Not at-risk (AUDIT-C)</td>
<td>Not at-risk (AUDIT-C)</td>
<td></td>
</tr>
</tbody>
</table>
Routine, Self-Administered, Touch-Screen, Computer-Based Suicidal Ideation Assessment Linked to Automated Response Team Notification in an HIV Primary Care Setting

Sarah T. Lawrence,¹ a James H. Willig,¹ a Heidi M. Crane,⁵ Jiatao Ye,² Inmaculada Aban,³ William Lober,⁴ Christa R. Nevin,¹ D. Scott Batey,¹ Michael J. Mugavero,¹ Cheryl McCullumsmith,³ Charles Wright,¹ Mari Kitahata,⁵ James L. Raper,¹ Micheal S. Saag,¹ and Joseph E. Schumacher¹

¹Department of Internal Medicine, Division of Infectious Diseases, ²Department of Biostatistics, School of Public Health, ³Department of Psychiatry, and ⁴Department of Internal Medicine, Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham; and ⁵Department of Medicine and ⁶School of Nursing, University of Washington, Seattle
1917 Palliative Care Clinic

Symptoms
- Pain
- Anxiety
- QOL
- Depression

Improved Pain

PRO Implementation
- Selection of instruments
- Sequence
- **Internal logic**
- Frequency
- Alerts

Notification Triggered
- Clinical: SI, IPV (starts clinic response protocol)
- Study enrollment
<table>
<thead>
<tr>
<th>Settings</th>
<th>Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Services(^1)</td>
<td>2,610</td>
</tr>
<tr>
<td>Palliative(^1)</td>
<td>1,208</td>
</tr>
<tr>
<td>Signs &amp; Symptoms(^1)</td>
<td>52,249</td>
</tr>
<tr>
<td>Viral Hepatitis(^1)</td>
<td>1,213</td>
</tr>
<tr>
<td>Supportive Care(^2)</td>
<td>1,376</td>
</tr>
<tr>
<td>Breast Health(^2)</td>
<td>106</td>
</tr>
<tr>
<td>iEngage(^3) &amp; BA2C(^3)</td>
<td>471</td>
</tr>
<tr>
<td>Pediatric Neurosurgery</td>
<td>65</td>
</tr>
</tbody>
</table>

(1) 1917 Clinic; (2) TKC; (3) Research
(2) As of 3/31/2016
Developing PROMs to have more impact

Helen Burstin, MD, MPH, FACP
Developing PROMs to Have More Impact

Helen Burstin, MD, MPH, FACP
Chief Scientific Officer, NQF

ICHOM 2016 Annual Conference

May 2016
What is NQF?

NQF is an independent, nonprofit, membership organization that brings together all stakeholders working to improve health and healthcare through quality measurement.
U.S. Policy Environment: From Volume to Value

All Medicare Fee-For-Service (FFS) payments

- 2011: 68% FFS linked to quality, ~20% Alternative payment models
- 2014: >80% FFS linked to quality, 30% Alternative payment models
- 2016: 85% FFS linked to quality, 15% Alternative payment models
- 2018: 90% FFS linked to quality, 10% Alternative payment models
Measuring Value

Outcomes
- Defined by patient
- Measured for patient’s condition over entire episode of care

Cost
- Measured for patient’s condition over entire episode of care

Value for Patients over their condition = Health Outcomes
Cost of delivering outcomes

VALUE FOR PATIENTS
Patient Focused Episodes

- Functional Status
- Quality of Life
- Shared decision-making
- Clinical outcomes & PROs
- Costs
Selection and Use of PROs in Measurement

- Guiding principles for PRO selection: psychometric soundness, person-centered, meaningful, actionable, implementable

- Challenges to use PROs for accountability and performance improvement:
  - Frequently used in research, but not in clinical use
  - Aggregation of patient-reported information to measure provider performance challenging
  - Proprietary tools
The Pathway from PROs to PRO-PMs

**PRO**
- Identify the quality performance issue (include broad input)
- Identify outcomes meaningful to target population and amenable to change
- Determine whether (PRO) is the best way to assess the outcome of interest

**PROM**
- Identify existing PROMs for measuring the outcome in the target population
- Select PROM suitable for use in performance measurement (e.g., reliable, valid, feasible)
- Use the PROM in real world with the intended target population and setting

**PRO-PM**
- Specify the outcome performance measure (e.g., average change, percentage improved or meeting a benchmark)
- Test PRO-PM for reliability, validity, and threats to validity (e.g., measure exclusions; missing data; poor response rate; risk adjustment; discrimination of performance; equivalence of results across PROMs)
The Pathway from PROs to PRO-PMs

Symptom: Depression

PHQ-9 © standardized tool to assess depression

Percentage of depressed patients with remission or progress toward remission in PHQ-9 scores at 6 months and at 12 months
NQF Measure Incubator: Getting to Quality Measures that Matter

Improved Patient Care and Outcomes
Helen Burstin, MD, MPH, FACP
hburstin@qualityforum.org
@HelenBurstin
DISCUSSION AND Q&A
A warm thank you to our panelists and to all of you for your participation!