

# The Validation and Utilization of PROMs and PREMs for Health Services and Clinical Practice

---

Rick Sawatzky, PhD, RN

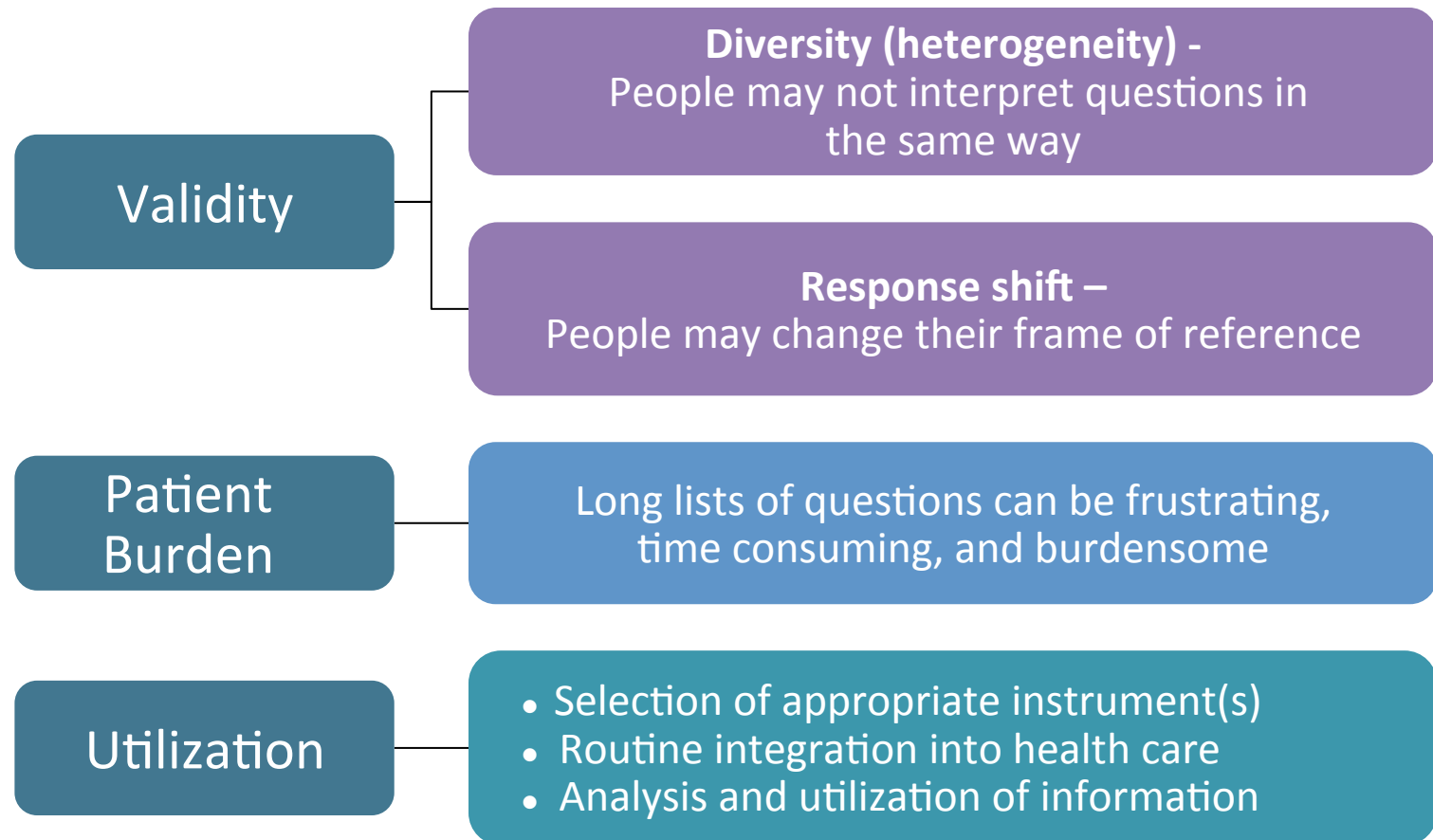
Canada Research Chair in Patient-Reported Outcomes

# An “explanatory” perspective of measurement validation

“an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment”

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.

# Challenges and opportunities in patient-reported measurement validation



# Purposes for patient-reported measurement

## Health professionals

- At point of care, to inform treatment decisions, monitor patients' conditions, promote patient-clinician communication, reveal health and quality of life concerns that may otherwise have not been noticed
- Quality improvement and service recovery

## Health researchers

- Examine the effectiveness of treatments and the impact of healthcare interventions
- Better understand the impacts of treatments and services on people's health from their point of view

## Health service decision makers

- Evaluate the efficacy, effectiveness and cost-effectiveness of healthcare services and programs

## Health care recipients

- Monitor symptoms and concerns and communicate with health care professionals

# Validation

---

diversity and response shift

# The Draper-Lindley-de Finetti (DLD) framework of measurement validation

Measurement items	
People	<i>EXCHANGEABLE</i>
	<i>NOT EXCHANGEABLE</i>
	<i>EXCHANGEABLE</i>
	<i>NOT EXCHANGEABLE</i>
<i>EXCHANGEABLE</i>	General Measurement Inference
	Specific Sampling Inference
<i>NOT EXCHANGEABLE</i>	Specific Domain Inference
	Initial Calibrative Inference

Adapted from: Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26: Psychometrics, pp. 45-79). Amsterdam: Elsevier Science.

# Validation of PROMs

## **Population heterogeneity (Diversity)**

- Differences in how people interpret and respond to questions
- **Threatens the comparability of scores across individuals or groups**

## **Response shift**

- An individual's frame of reference may change in response to a health event or intervention
- **Threatens the comparability of scores over time**

# Diversity in the population



A conventional assumption underlying PROMs is that individuals interpret and respond to questions about their health in the same way, such that scores are equivalently applicable to all people in the population.



# The challenge of heterogeneity

Is it reasonable to believe that people from different backgrounds and with different life experiences interpret and respond to questions about their health and quality of life in the same way?

People may respond to QOL and PROM questions in systematically unique ways because of:

- Cultural, developmental, or personality differences
- Contextual factors or life circumstances
- Different health experiences or events

In this situation, the PROMs will produce biased scores that are not comparable across different individuals or groups

# Examining the implications of heterogeneity

Qual Life Res  
DOI 10.1007/s11136-011-9976-6

---

## Latent variable mixture models: a promising approach for the validation of patient reported outcomes

Richard Sawatzky · Pamela A. Ratner ·  
Jacek A. Kopec · Bruno D. Zumbo

Accepted: 7 July 2011  
© Springer Science+Business Media B.V. 2011

### Abstract

**Purpose** A fundamental assumption of patient-reported outcomes (PRO) measurement is that all individuals interpret questions about their health status in a consistent manner, such that a measurement model can be constructed that is equivalently applicable to all people in the target

measurement model, (b) implications of sample heterogeneity with respect to model-predicted scores (theta), and (c) sources of sample heterogeneity. An example is provided using the 10 items of the Short-Form Health Status (SF-36<sup>®</sup>) physical functioning subscale with data from the Canadian Community Health Survey (2003) ( $N = 7,030$

Sawatzky, R., Ratner, P. A., Kopec, J. A., & Zumbo, B. D. (2011). Latent variable mixture models: A promising approach for the validation of patient reported outcomes. *Quality of Life Research*. doi: 10.1007/s11136-011-9976-6

# SF-36 physical function

Physical function items	
SFRC_03	Vigorous activities
SFRC_04	Moderate activities
SFRC_05	Lifting or carrying groceries
SFRC_06	Climbing several flights of stairs
SFRC_07	Climbing one flight of stairs
SFRC_08	Bending, kneeling, or stooping
SFRC_09	Walking more than one kilometer
SFRC_10	Walking several blocks
SFRC_11R	Walking one block
SFRC_12R	Bathing and dressing

Response options
0. No limitations
1. Limited a little
2. Limited a lot

Conventional scoring method*
1. Add all items
2. - 20 (reverses the scale)
3. x 5 (scaled from 0 – 100)

# Data from the Canadian Community Health Survey (cycle 2.1) (2003)

<b>Description</b>	Repeated cross-sectional national survey of health status, healthcare utilization & health determinants
<b>Target</b>	Canadians 12 years and older
<b>Coverage</b>	~ 98% of people in 10 provinces and 71% - 97% of people in 3 northern territories.
<b>Data collection</b>	Computer assisted telephone interviewing Core and optional content (e.g., SF-36)
<b>Sampling methods</b>	Clustered stratified sampling represents 121 health regions in Canada

# The validation of PROMs in heterogeneous populations

- The cumulative probability of an item response at or above category  $j$  within a latent class can be computed as follows:

$$P_{ijk}(Y \geq j | \theta, C = k) = \frac{\exp(-\tau_{ijk} + \lambda_{ik}\theta)}{1 + \exp(-\tau_{ijk} + \lambda_{ik}\theta)}$$

- Each class has a unique set of parameters that are estimated simultaneously in the latent variable mixture model:

$$f(x) = \sum_{k=1}^K \pi_k f_k(x)$$

, where  $f$  is the mixture of the class-specific distributions, and  $\pi_k$  is the mixing proportion.

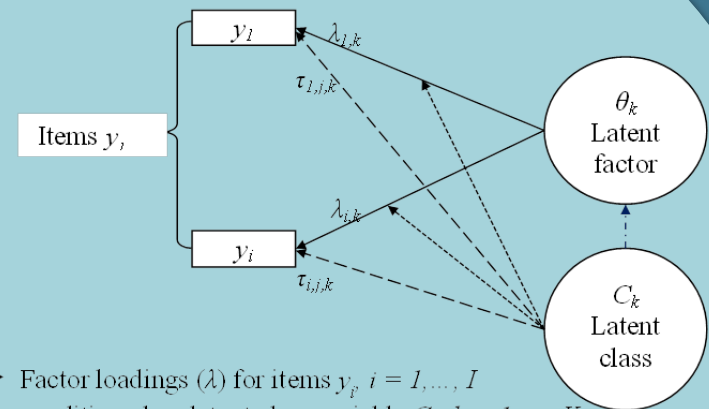
- The cumulative probability of an item response at or above category  $j$  within a latent class can be computed as follows:

$$P_{ijk}(Y \geq j | \theta, C = k) = \frac{\exp(-\tau_{ijk} + \lambda_{ik}\theta)}{1 + \exp(-\tau_{ijk} + \lambda_{ik}\theta)}$$

- The cumulative probability of an item response at or above category  $j$  within a heterogeneous population is obtained by:

$$P_{ij}(Y \geq j | \theta) = \sum_{k=1}^K (X_k * P_{ijk}(Y \geq j | \theta))$$

, where  $X_k$  is the posterior probability of an individual being in class  $k$ .



- > Factor loadings ( $\lambda$ ) for items  $y_i$ ,  $i = 1, \dots, I$  conditioned on latent class variable  $C_k$ ,  $k = 1, \dots, K$ .
- > Thresholds ( $\tau$ ) for  $j - 1$  response categories per item conditioned on latent class variable  $C_k$ ,  $k = 1, \dots, K$ .
- > Variance of the latent factor ( $\theta$ ) conditioned on latent class variable  $C_k$ ,  $k = 1, \dots, K$ .

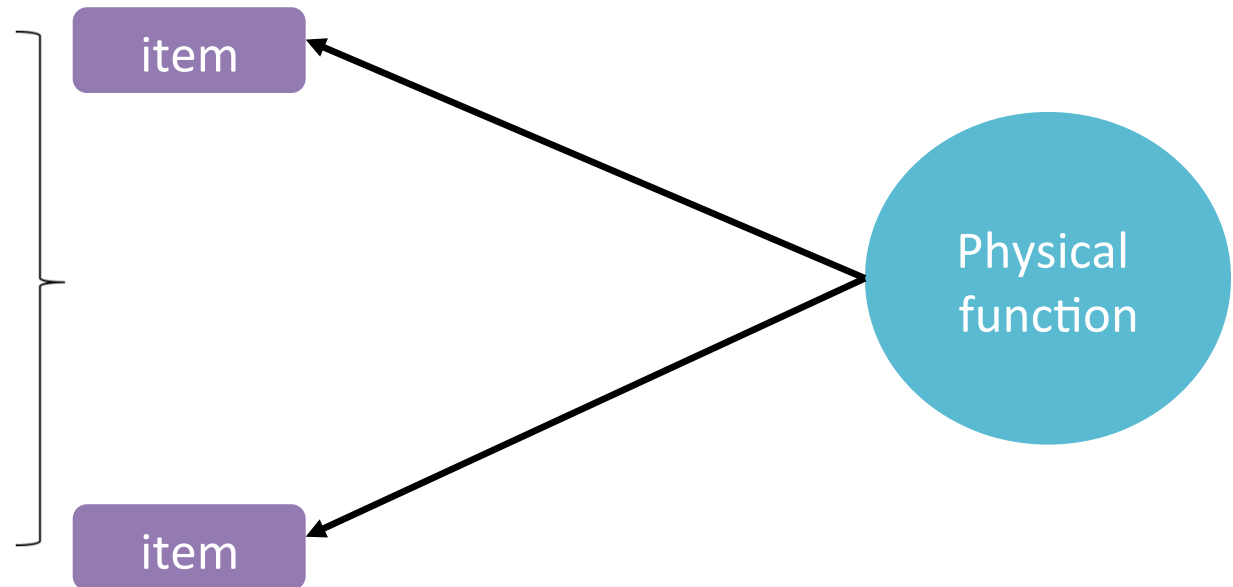
# Conventional PROM model

Does your health limit you in any of the following activities:

- Walking one block

- Moderate activities,  
such as .....

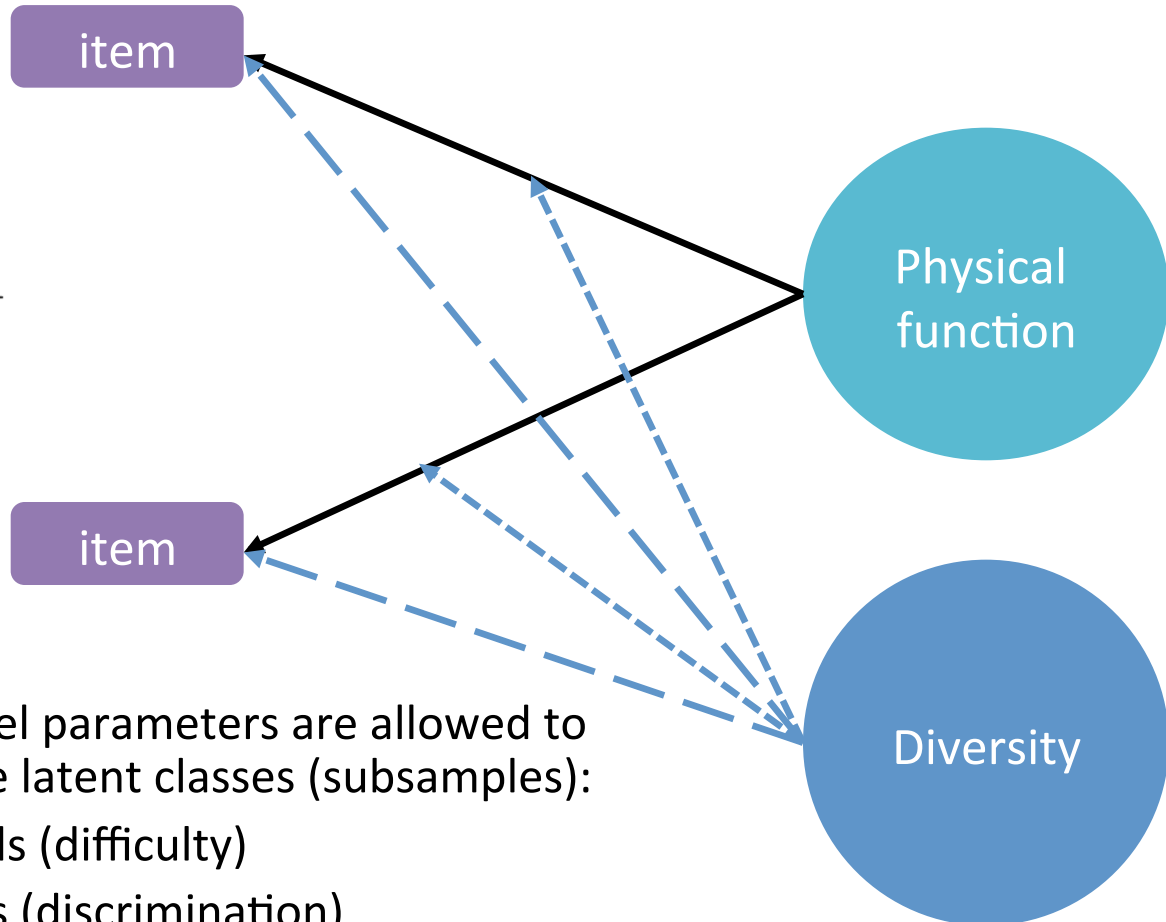
- Vigorous activities,  
such as .....



# PROM model that accommodates diversity

Does your health limit you in any of the following activities:

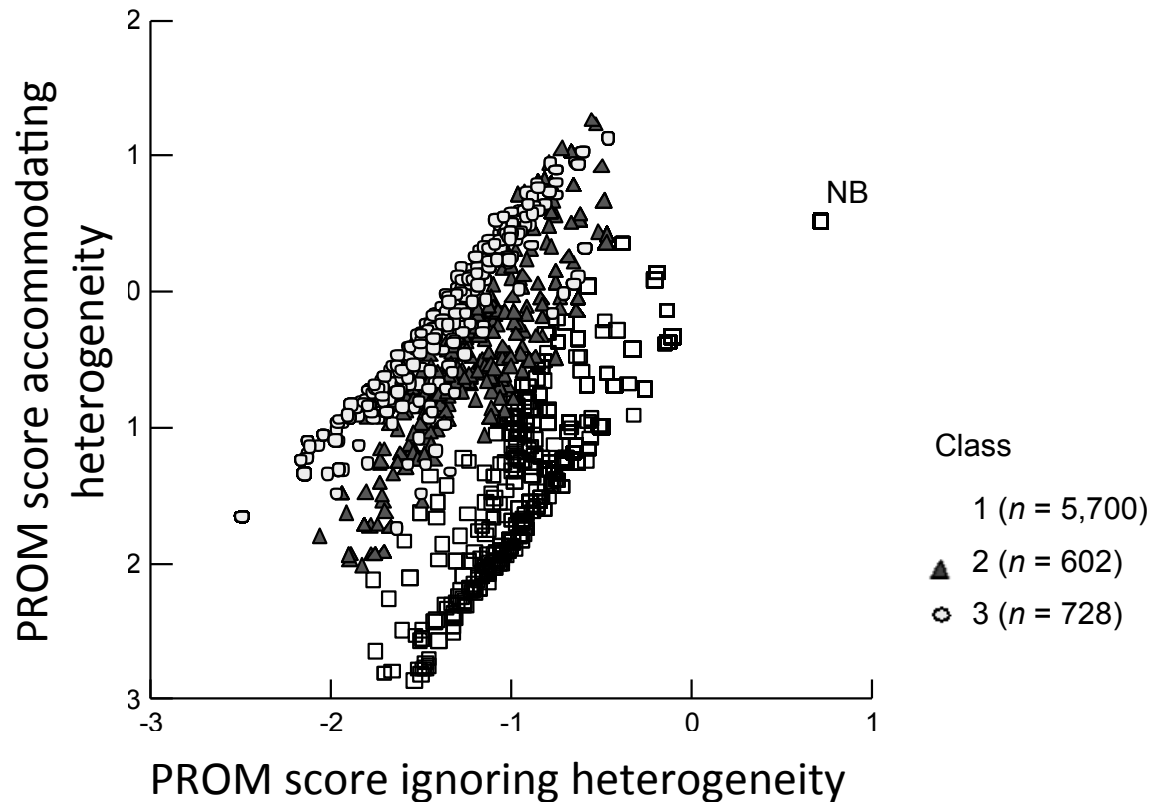
- Walking one block
- Moderate activities, such as .....
- Vigorous activities, such as .....



The measurement model parameters are allowed to vary across two or more latent classes (subsamples):

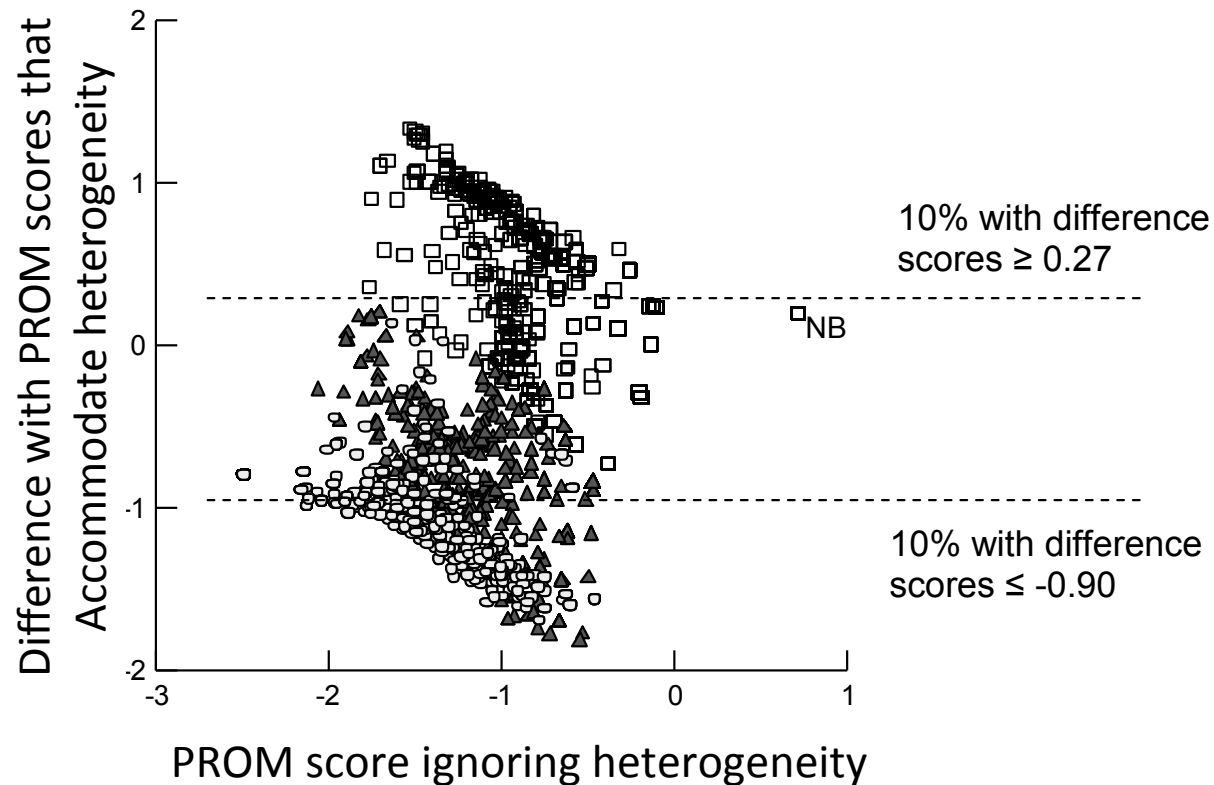
- — — — — item thresholds (difficulty)
- - - - - factor loadings (discrimination)

# Implications of ignoring heterogeneity on item response theory predicted scores





# Implications of ignoring heterogeneity on item response theory predicted scores



# What we have learned to date

## The challenge of diversity in the population

People may not interpret and respond to questions about their health and quality of life in the same way.

Differences among people that may explain such inconsistencies include:

- Differences in health experiences
- Differences in age
- Cultural differences
- Gender differences

## Application to PRO measurement

Accuracy in PRO measurement is improved when we use approaches that accommodate for differences in how people interpret and respond to PRO questions

# What is response shift?

Schwartz and Sprangers defined response shift as “a change in the meaning of one’s self-evaluation of a target construct as a result of change in”:

recalibration

- internal standards of measurement

reprioritization

- values (i.e. the importance of component domains constituting the target construct)

reconceptualization

- definition of the target construct

# Theoretical model of response shift

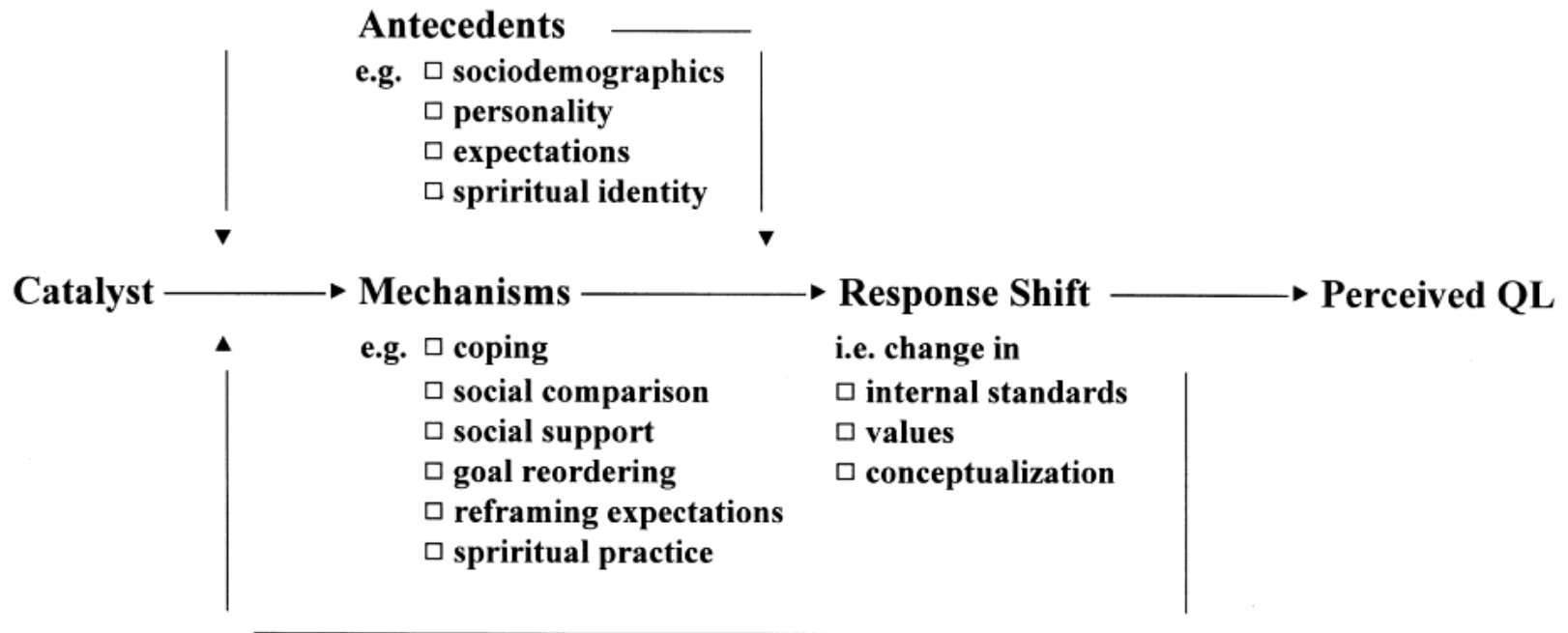


Fig. 1. A theoretical model of response shift and quality of life (QOL).

Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine*, 48(11), 1507-1515.

# Why care about response shift?

- From a validation point of view, it is important to distinguish “true change” from RS change
  - Ignoring RS could lead to measurement bias:
    - Decreased sensitivity to detect change over time
    - Detecting change over time that does not exist
- Contributes to understanding regarding the meaning of scores
  - Unexpected health outcomes
- May want to promote response shift
  - Palliative care
  - Rehabilitation
  - Self-management
  - Other non-curative interventions

# Patient burden

---

need for the efficient collection of PRO data

# Patient Burden

Long lists of questions can be frustrating, time consuming, and burdensome

## Methods for addressing patient burden

- Appropriate reading difficulty and mode of administration
- Use of short forms
- Computerized adaptive tests (CATs)



# Computer adaptive tests

## **Advantages of CATs:**

Only questions that are most likely to be meaningful and relevant to an individual's condition are administered, based on their responses to prior questions.

- Improved efficiency
- Reduced response burden
- Uniquely targeted to the individual's conditions





# Utilization

---

clinical practice & health-service decision making

# Use of PROMs and PREMs in clinical practice

---

An electronic system to  
support the routine integration  
of quality of life assessments  
in clinical practice



# Benefits of e-QOL assessment instruments

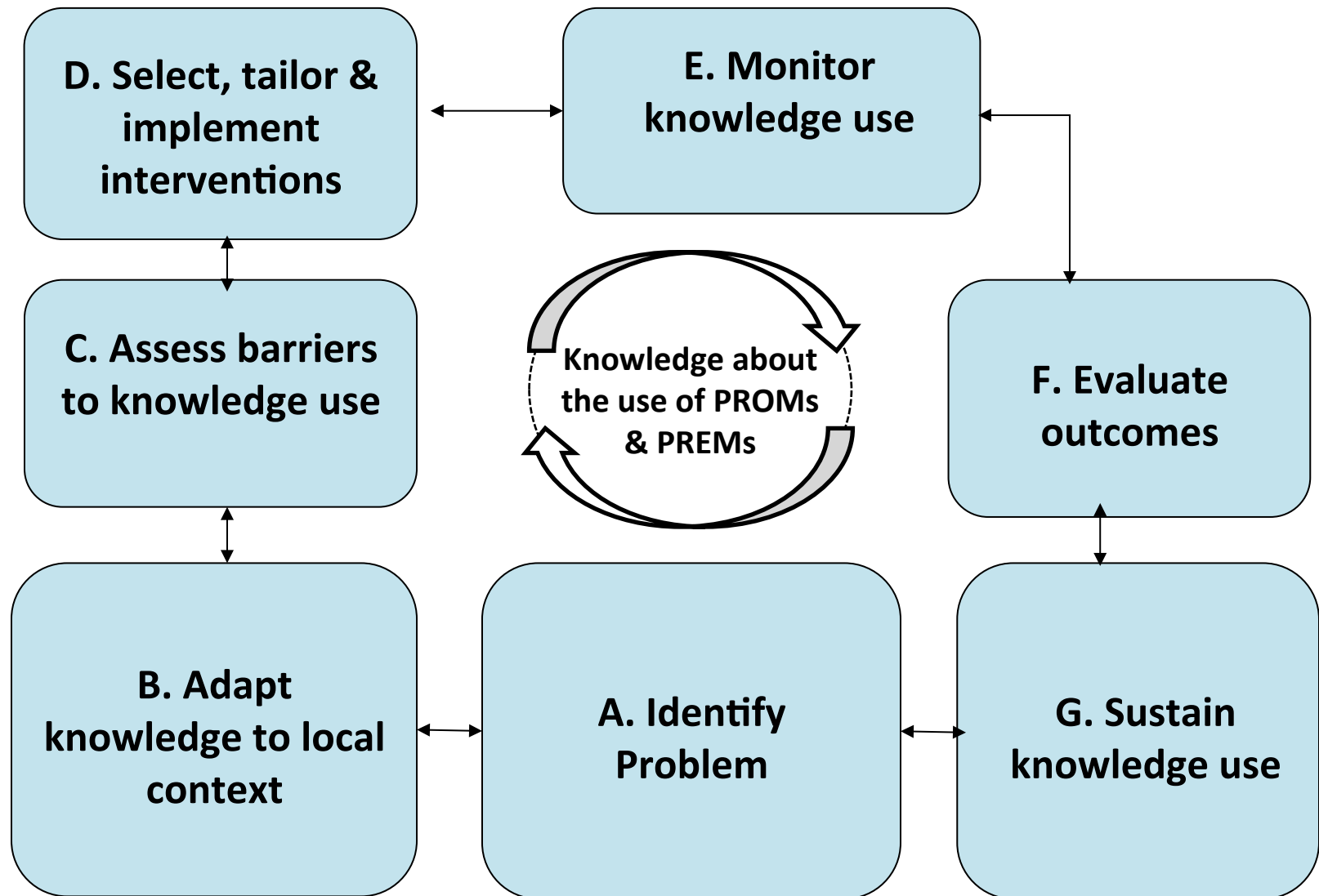
- Reduced patient burden
  - Ease of administration
  - Computer adaptive testing
- Reduced clinician burden
  - Information is automatically analyzed
  - No additional forms to complete
- Enhanced visualization and monitoring of patient concerns through ongoing and immediate feedback
- PROM & PREM information become part of administrative data for program evaluation, cost-effectiveness analysis, resource allocation

# Researching the QPSS

**How can we best design and integrate electronic quality of life assessments with best-practice recommendations as practice support tools and integrate them into routine acute care for older adults who have advancing chronic life-limiting illnesses and their family caregivers?**



# QPSS design and implementation



# PROMs and PREMs for seriously ill older adults: Results to date

## **Literature review**

- Nearly 200 PROMs and PREMs have been used in this population.
- Our synthesis focuses on categorizing these instruments and providing recommendations for making informed decisions about the selection and utilization of PROMs and PREMs for seriously ill older adults.

## **Focus groups with clinicians**

- Instruments must measure symptoms, physical function and emotional, psychological, and existential concerns.
- Concern regarding potential response burden.
- Importance of distinguishing “screening” and “assessment”.
- Potential for linkage with clinical-reported measurements.

# PROMs and PREMs for the seriously ill

## Preliminary selection

### **PREMs for seriously ill older adults**

- Canadian Health Care Evaluation Project (CANHELP) Lite
  - Individualized Patient Questionnaire
  - Individualized Family Caregiver Questionnaire

### **PROMs for seriously ill older adults**

#### **For patients**

- Edmonton Symptom Assessment System – Revised ESAS-R
- McGill Quality of Life Questionnaire (MQOL)

#### **For family caregivers**

- Quality of Life in Life-Threatening Illness (QOLLI-F)
- Carer Support and Needs Assessment Tool (CSNAT)



# Feedback system

## **Reporting of assessment results**

- Graphical displays that present changes in identified concerns over time
- Ranking of areas of most important areas of concerns or needs
- Assessment results must be accessible in “real time” at point of care
- Importance of producing printable reports that can be used in rounds and filed in paper charts

## **Linkage with current practices**

- Integration of prompts for potential interventions to address identified concerns or needs
- Tracking of interventions that have been applied



# Use of PROMs in health services administration

The slide features a blue header with two logos: the 'CENTRE for Clinical Epidemiology and Evaluation' on the left and the 'UBC CENTRE FOR HEALTH SERVICES AND POLICY RESEARCH' on the right. The main title, 'Patient-reported outcome measurement (PROMs) and integrated primary and community care (IPCC)', is centered in a large blue box. Below the title, the text 'Workshop: January 14<sup>th</sup>, 2013' is displayed. The footer contains four logos: 'CIHR IRSC' with a green leaf icon, a text-based acknowledgment of support from the 'Canadian Institutes of Health Research', the 'CHÉOS' logo (Centre for Health Evaluation & Outcome Sciences), and the 'Vancouver Coastal Health Research Institute' logo with the tagline 'Healthier lives through discovery' and the UBC crest.

CENTRE for  
Clinical Epidemiology  
and Evaluation

UBC CENTRE FOR  
HEALTH SERVICES AND  
POLICY RESEARCH

Patient-reported outcome measurement  
(PROMs) and integrated primary and  
community care (IPCC)

Workshop: January 14<sup>th</sup>, 2013

CIHR IRSC

This study was supported by the  
Canadian Institutes of Health  
Research.

CHÉOS  
Centre for Health Evaluation  
& Outcome Sciences

Vancouver  
Coastal Health  
Research Institute  
Healthier lives through discovery

UBC

Bryan (PI), J. Davis, S. Lewis, K. McGrail, M. McGregor, R. Sawatzky (co-investigator), & M. Dawes. In collaboration with BC Ministry of Health, Michael Smith Foundation for Health Research, and the Canadian Institutes of Health Research.

# Project objective

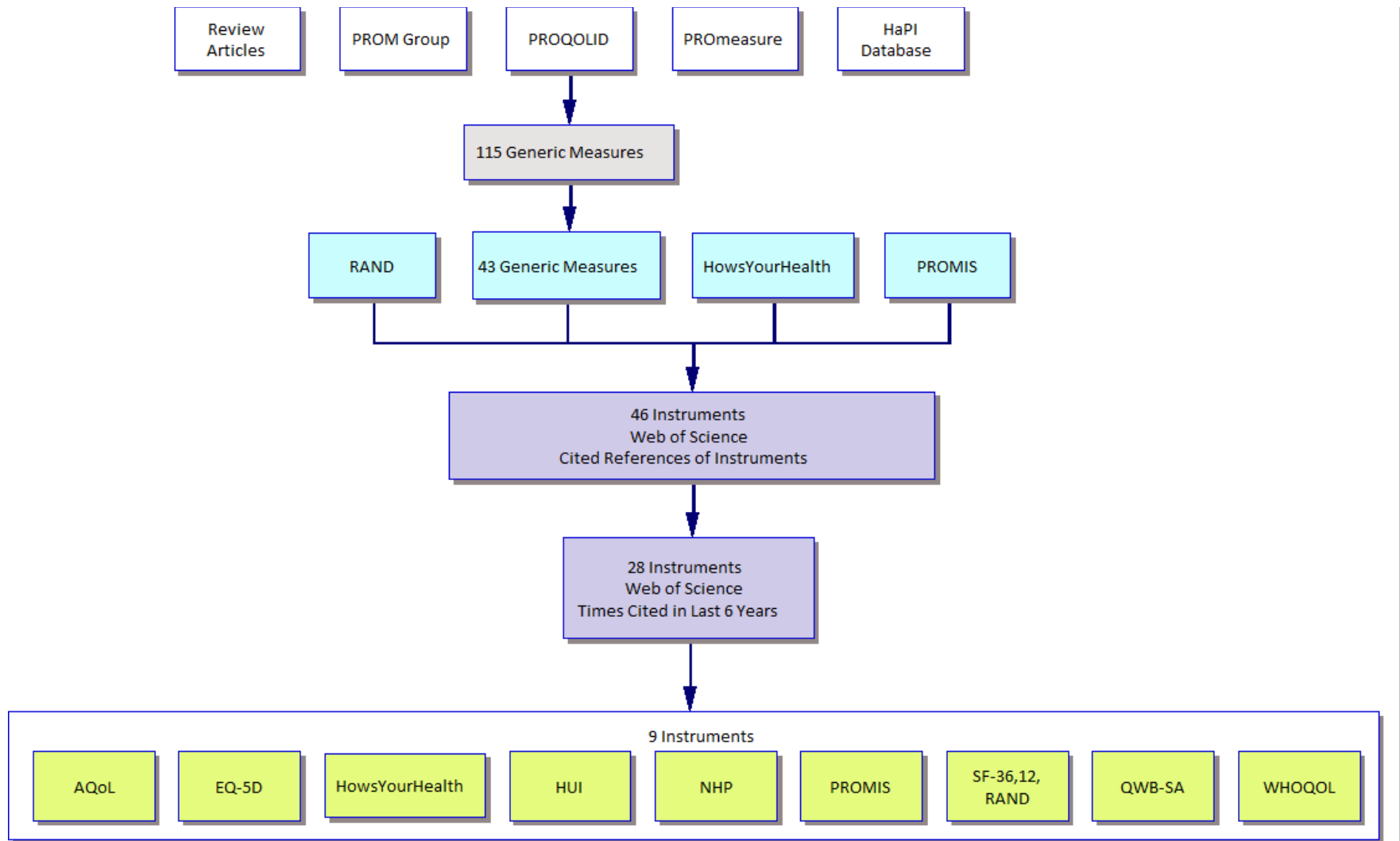
- “To explore the utility of existing generic instruments for the measurement of patient reported outcomes in obtaining reliable, valid and useful information from patients in assessing the impact of primary and community health care reform initiatives in Canada.”



# Project components

- A comprehensive long-list of all generic PROMs
- A shorter-list to include:
  - Patient self-report, truly generic, true assessment of HRQL, developed for adult population
- A short-list, reduced on the basis of citations
- Characteristics of short-listed PROMs
- Review of each instrument's 'performance':
  - Psychometrics
  - Decision-making attributes
- Additional information:
  - Norms, value sets, examples of use of in primary and community care context, other jurisdictions
- Stakeholder engagement and recommendations

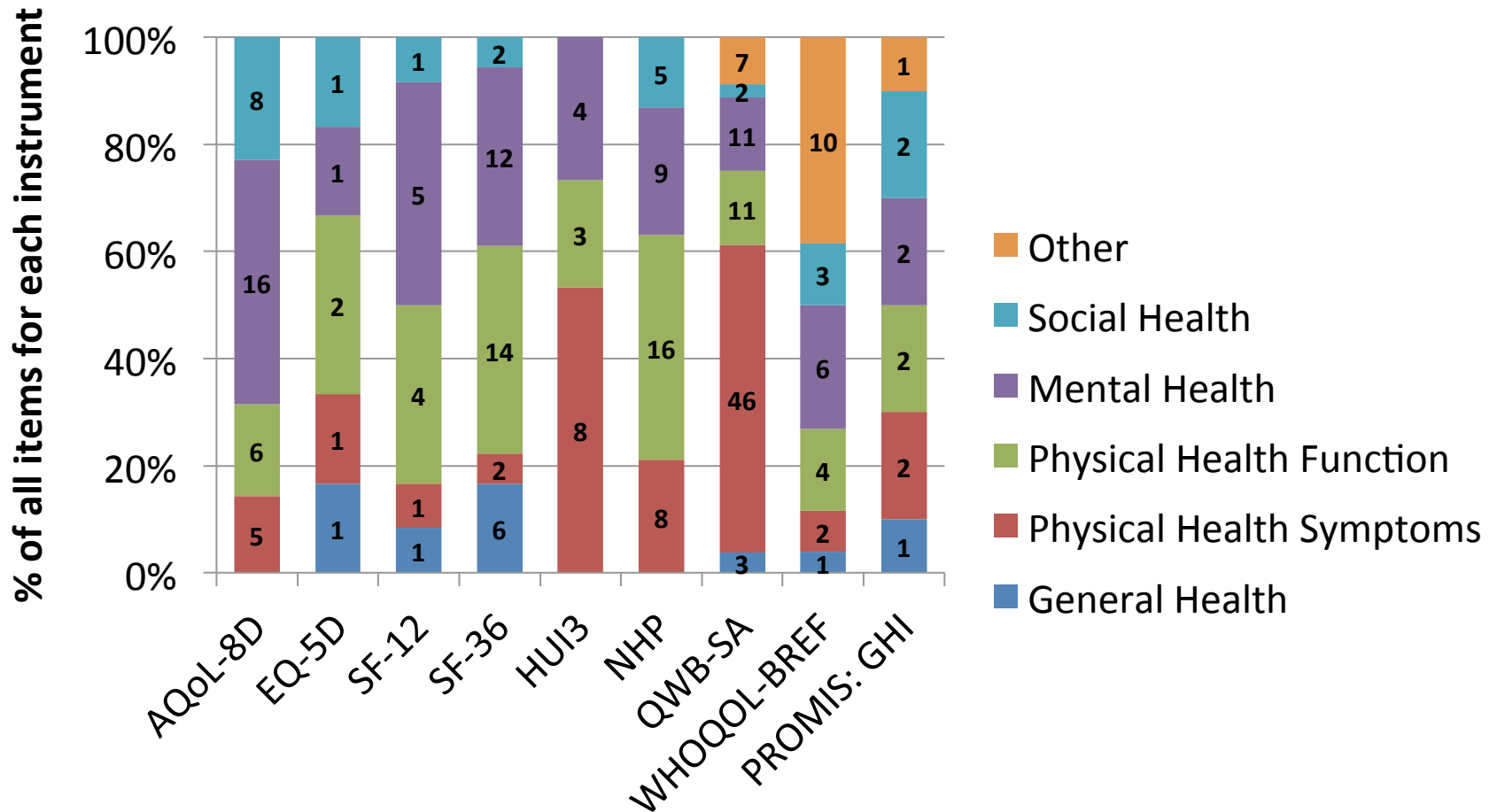
# PROMs search



# Framework for the Review

- **Purpose:**
  - Examine the evidence pertaining to the measurement properties (psychometrics) of the candidate PROM instruments
- **Scope:**
  - Review Articles on Psychometric properties
  - General Population Context
- **Comparative analysis:**
  - Use of COSMIN: Consensus-based Standards for the selection of health measurement instruments ([www.cosmin.nl](http://www.cosmin.nl))
  - Comparison of PROM characteristics relevant to decision making
- **Examples of use in a primary and community care contexts**

# Domain coverage of selected instruments\*



\*Refers to the representation of domains in the pool of items. Note some instrument do not provide summary scores for individual domains.

# Other instrument features

Instrument	Utility Scores?	Canadian Utility Scores?	Population Norms?	Canadian Norms?	Minimally Important Differences Published?
AQoL	Yes	No	Yes	No	Yes
EQ-5D	Yes	Yes	Yes	Yes	Yes
SF-36	Yes	No	Yes	Yes	Yes
HUI3	Yes	Yes	Yes	Yes	Yes
NHP	No	-	Yes	No	No
QWB	Yes	No	Yes	No	Yes
WHOQoL-BREF	No	-	Yes	No	No
PROMIS/GHS	No <sup>1</sup>	-	Yes	No	No

1. Equations for converting to EQ-5D scores have been published

# Summary of Evidence about Psychometric Characteristics

	AQoL	EQ-5D	SF-36	HUI	NHP	QWB	WHOQoL	PROMIS
<b>Internal consistency</b>		n/a		?		n/a		
<b>Reliability</b>	?					?	?	?
<b>Content validity</b>							?	
<b>Construct validity</b>	?	?		?		?		
<b>Cross-cultural validity</b>	?	?		?		?		?
<b>Criterion validity</b>	?		?	?	?	?	?	?
<b>Responsiveness</b>							?	

= negative evidence     
 = positive evidence     
 = conflicting evidence     
 ? = Unknown/Not Reported



# Key Strengths and Weaknesses

Instrument	Strengths	Weaknesses
AQoL	Discriminates between groups with clinical variations in health.	Smaller evidence base.
EQ-5D	Discriminates between groups with clinical variations in health.	Not as comprehensive. Not sensitive to small changes, limited responsiveness in healthy populations.
SF-36	Top instrument in most psychometric categories. Widely used, multiple cultural contexts, and many versions available.	
HUI	Can distinguish between groups with clinical variations in health, and widespread use in a variety of cultural contexts.	Lacking in mental health. Less reliability. Less responsive in populations of fairly good health.
NHP	More responsive than SF-36 in populations with poor health. Widespread use in a variety of cultures.	Not ideal for use in general population, or outside of populations with major health issues.
QWB	Good for capturing change in primarily healthy populations.	Lacking on mental health, may overweight minor conditions.
WHOQoL	Very strong cross-cultural validity. Correlated with groups with clinical variations in health.	Smaller evidence base.
PROMIS GHS	Good internal consistency, responsiveness and correlation with other instruments.	Smaller evidence base.

# Preferred PROMs for BC Integrated Primary and Continuing Care

## Short Form Health Survey instruments

**Your Health and Well-Being**

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities. *Thank you for completing this survey!*

For each of the following questions, please mark an ☒ in the one box that best describes your answer.

1. In general, would you say your health is:

Excellent	Very good	Good	Fair	Poor
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Compared to one year ago, how would you rate your health in general now?

Much better now than one year ago	Somewhat better now than one year ago	About the same as one year ago	Somewhat worse now than one year ago	Much worse now than one year ago
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

SF-36® Health Survey © 1988, 2002 by Medical Outcomes Trust and QualityMetric Incorporated. All Rights Reserved. SF-36® is a registered trademark of Medical Outcomes Trust. (11-5) Standard, US Version 1.0

## PROMIS Global Health Scale

PROMIS v.1.0.1.1 Global

**Global Health Scale**

Please respond to each item by marking one box per row.

	Excellent	Very good	Good	Fair	Poor
1. In general, would you say your health is: .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. In general, would you say your quality of life is: .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. In general, how would you rate your physical health? .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. In general, how would you rate your mental health, including your mood and your ability to think? .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. In general, how would you rate your satisfaction with your social activities and relationships? .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. In general, please rate how well you carry out your usual social activities and roles. (This includes activities at home, at work and in your community, and responsibilities as a parent, child, spouse, employee, friend, etc.) .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair? .....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

© 2008-2017 PROMIS Health Organization and PROMIS Cooperative Group Page 1 of 2

## EQ-5D

By placing a check-mark in one box in each group below, please indicate which statements best describe your own state of health today.

**Mobility**

I have no problems in walking about ☐

I have some problems in walking about ☐

I am confined to bed ☐

**Self-Care**

I have no problems with self-care ☐

I have some problems washing or dressing myself ☐

I am unable to wash or dress myself ☐

**Usual Activities (e.g. work, study, housework, family or leisure activities)**

I have no problems with performing my usual activities ☐

I have some problems with performing my usual activities ☐

I am unable to perform my usual activities ☐

**Pain/Discomfort**

I have no pain or discomfort ☐

I have moderate pain or discomfort ☐

I have extreme pain or discomfort ☐

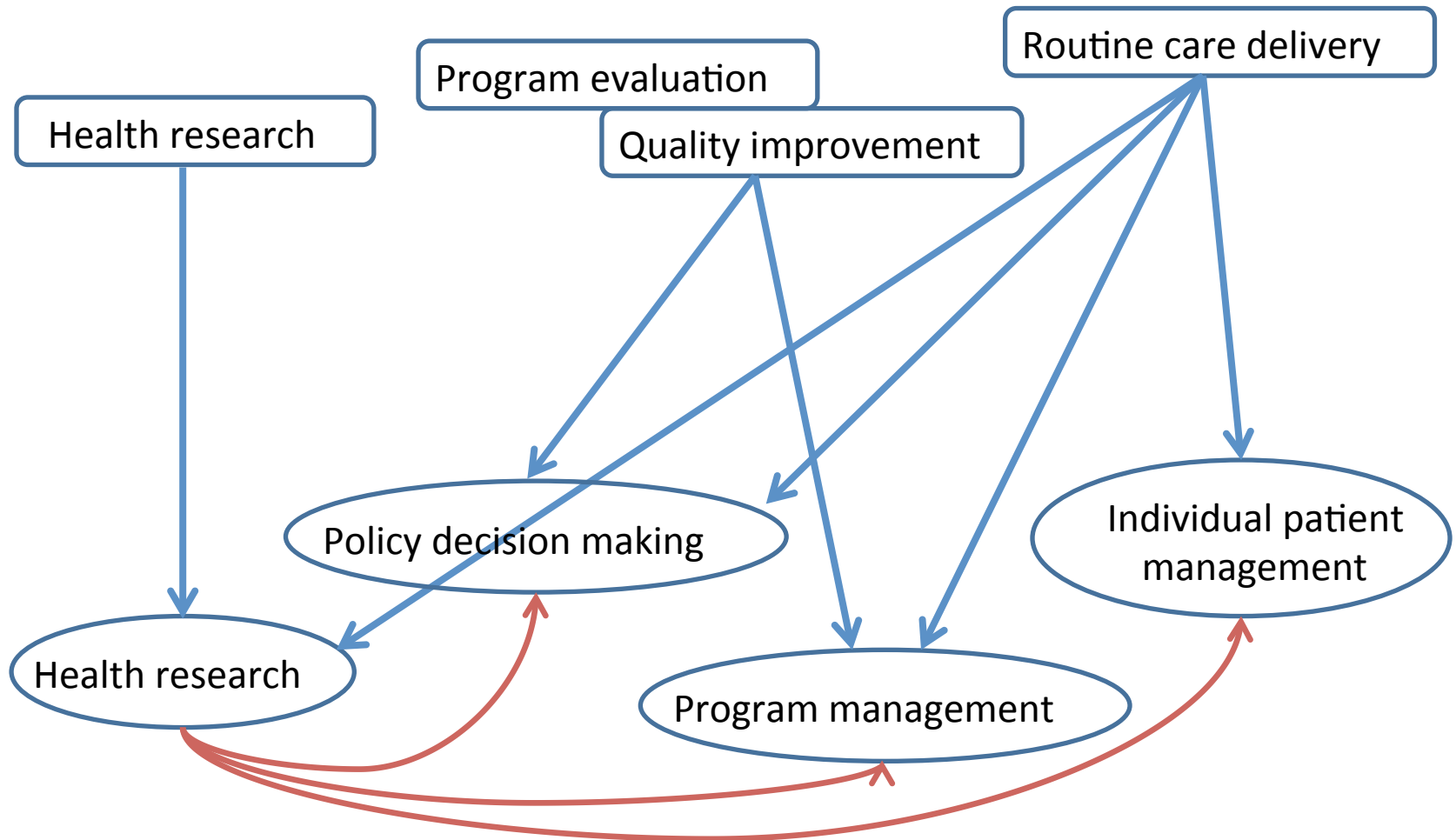
**Anxiety/Depression**

I am not anxious or depressed ☐

I am moderately anxious or depressed ☐

I am extremely anxious or depressed ☐

## Settings for PROM data collection



## Contexts for use of PROMs data



## **Richard (Rick) Sawatzky PhD**

*Canada Research Chair in Patient-Reported Outcomes*

Associate Professor | School of Nursing

Trinity Western University | t: 604.513.2121 (3274)

Langley, BC, V2Y 1Y1, Canada

[twu.ca/nursing](http://twu.ca/nursing) | [www.ricksawatzky.com](http://www.ricksawatzky.com)

Research Scientist | Centre for Health Evaluation and Outcome Sciences

St. Paul's Hospital, Vancouver, BC

<http://www.cheos.ubc.ca/>

