Introduction to Assessing Health-Related Measures in Diverse Populations

Pre-conference Workshop 2
The International Society for Quality of Life Research

Center for Aging in Diverse Communities: A Resource Center for Minority Aging Research (RCMAR)
University of California San Francisco
October 19, 2005

Purpose of Workshop

◆ Present overview of measurement issues in health disparities research
◆ Inform health disparities researchers of importance of using measures that have been tested across diverse groups
◆ Highlight importance of conducting measurement studies in preliminary phases of research

Faculty

Anita L. Stewart, Ph.D.
Anna Nápoles-Springer, Ph.D.
Steven Gregorich, Ph.D.

Center for Aging in Diverse Communities

◆ A Resource Center for Minority Aging Research (RCMAR)
◆ Goal: to reduce health disparities between minority and non-minority older adults
  – Increasing number of researchers in this area
  – Improve recruitment/retention methods
  – Create culturally sensitive measures

Introducing Measurement Issues and Approaches

Anita L. Stewart, Ph.D.

Culturally Sensitive Research

Ongoing efforts to “mesh the process of inquiry with the cultural characteristics of the group being studied.”

Culturally Sensitive Research

- Consider cultural or group differences in all phases of research
  - Developing research questions
  - Concept development
  - Measurement
  - Recruitment
  - Methods of data collection
  - Interpretation of results

Today’s Presentation:

- Consider cultural or group differences in all phases of the research
  - Developing research questions
  - Concept development
  - Measurement
  - Recruitment
  - Methods of data collection
  - Interpretation of results

Distinctions: International versus U.S. Cross-Cultural Research

- ISOQOL – research across countries and languages
  - Approaches tend to assume non-equivalence of concepts and measures
- Health disparities and cross-cultural research in the U.S.
  - Cultural differences within the U.S.
  - Tend to assume equivalence of concepts and measures

Cross-cultural Research in the U.S.

- Currently, focus is on health disparities research
  - Why persons from ethnic minority groups have worse health
- Usually involves comparing scores on self-report measures across diverse groups in one study
  - Use of culture- or group-specific measures is impractical

Examples of Health Disparities Research Requiring Group Comparisons

- What is the nature of health and QOL differences among diverse populations?
- Does a treatment have the same health benefit in African Americans as in Whites?
- Does an intervention to improve cultural sensitivity of doctors reduce ethnic disparities in health outcomes?

The Measurement Problem

- Measurement goal - identify measures that can be used across all groups, and
  - are sensitive to diversity
  - have minimal bias between groups
- Most self-reported measures were developed and tested in mainstream, well-educated groups
  - little research on measurement characteristics in diverse groups
### Issues Concerning Group Comparisons

Differences in observed scores can be due to culturally- or group-mediated differences in true score (true differences)

--- OR ---

systematic differences in observed scores not attributable to true scores

» Observed scores not valid in one or more groups

### Bias - A Special Concern

- Measurement bias in any one group may make group comparisons invalid
- Bias can be due to group differences in:
  - the meaning of concepts or items
  - the extent to which measures represent a concept
  - cognitive processes of responding
  - use of response scales
  - appropriateness of data collection methods

### Typical Sequence of Developing New Self-Report Measures

1. Develop concept
2. Create item pool
3. Pretest/revise
4. Field survey
5. Psychometric analyses
6. Final measures

### Extra Steps in Sequence of Developing New Self-Report Measures for Diverse Groups

1. Develop concept
2. Create item pool
3. Pretest/revise
4. Field survey
5. Psychometric analyses
6. Final measures
7. Obtain perspectives of diverse groups
8. .. to reflect these perspectives

### Extra Steps in Sequence of Developing New Self-Report Measures for Diverse Groups

1. Develop concept
2. Create item pool
3. Pretest/revise
4. Field survey
5. Psychometric analyses
6. Final measures
7. Obtain perspectives of diverse groups
8. .. to reflect these perspectives
   in all diverse groups

---

Center for Aging in Diverse Communities. University of California San Francisco, 2005.
http://medicine.ucsf.edu/cadc/cores/measurement/index.html

3
Extra Steps in Sequence of Developing New Self-Report Measures for Diverse Groups

- Develop concept
- Create item pool
- Pretest/revise
- Field survey
- Psychometric analyses
- Final measures

Obtain perspectives of diverse groups
.. to reflect these perspectives
.. in all diverse groups

Obtain perspectives of diverse groups
.. to reflect these perspectives
.. in all diverse groups

Measurement studies across groups
Psychometric analyses
Final measures

A Framework: Conceptual and Psychometric Adequacy and Equivalence

Conceptual
- Concept meaningful within one group
- Concept equivalent across groups

Psychometric
- Psychometric properties meet minimal standards within one group
- Psychometric properties invariant (equivalent) across groups

Left Side of Matrix: Issues in a Single New “Diverse” Group

Conceptual
- Concept meaningful within one group
- Concept equivalent across groups

Psychometric
- Psychometric properties meet minimal standards within one group
- Psychometric properties invariant (equivalent) across groups

Right Side of Matrix: Issues in Studies of Group Comparisons

Conceptual
- Concept meaningful within one group
- Concept equivalent across groups

Psychometric
- Psychometric properties meet minimal standards within one group
- Psychometric properties invariant (equivalent) across groups
## Conceptual Adequacy in One Group

<table>
<thead>
<tr>
<th>Conceptual Adequacy in 1 Group</th>
<th>Conceptual Adequacy in 1 Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concept meaningful within one group</td>
<td>Concept equivalent across groups</td>
</tr>
<tr>
<td>Adequacy</td>
<td>Equivalence</td>
</tr>
<tr>
<td>Psychometric properties meet minimal standards within one group</td>
<td>Psychometric properties invariant (equivalent) across groups</td>
</tr>
</tbody>
</table>

## Conceptual Adequacy in One Group

- Is concept relevant, meaningful, and acceptable in that group?
- Traditional research
  - Conceptual adequacy = simply defining a concept
  - Mainstream population “assumed”
- Minority and cross cultural research
  - Mainstream concepts may be inadequate
  - Concept should correspond to how a particular group thinks about it

## Example of Inadequate Concept

- **Patient satisfaction** typically conceptualized in mainstream populations in terms of, e.g.,
  - access, technical care, communication, continuity, interpersonal style
- In minority and low income groups, additional relevant domains include, e.g.,
  - discrimination by health professionals
  - sensitivity to language barriers

## Conceptual Equivalence Across Groups

<table>
<thead>
<tr>
<th>Conceptual Equivalence Across Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concept meaningful within one group</td>
</tr>
<tr>
<td>Adequacy</td>
</tr>
<tr>
<td>Psychometric properties meet minimal standards within one group</td>
</tr>
</tbody>
</table>

## Example of Nonequivalent Concept: Decision Making

- Mainstream concept of decision making in health care settings focuses on patient autonomy
  - shared decision-making with provider
- In traditional Latino or Chinese cultures, family participation in decision-making is the norm, especially for serious medical concerns
  - individual autonomy is less relevant
Qualitative Approaches to Address Conceptual Equivalence in Diverse Groups

- Literature reviews
  - ethnographic and anthropological
- In-depth interviews and focus groups
  - discuss concepts, obtain their views
- Expert consultation from diverse groups
  - review concept definitions
  - rate relevance of items

Example: Subjective Test of Conceptual Equivalence of Spanish FACT-G

- Bilingual/bicultural expert panel reviewed all 28 items
  - One item had low cultural relevance to quality of life
  - One concept was missing – spirituality
- Developed new spirituality scale (FACT-SP) with input from cancer patients, psychotherapists, and religious experts
  - Sample item “I worry about dying”

Cella D et al. Med Care 1998: 36;1407

Generic/Universal vs Group-Specific (Etic versus Emic)

- Concepts unlikely to be defined exactly the same way across diverse ethnic groups
- Generic/universal (etic)
  - features of a concept that are appropriate across groups
- Group-Specific (emic)
  - idiosyncratic portions of a concept

Etic versus Emic (cont.)

- Goal in health disparities research
  - identify generic/universal portion of a concept (could be entire concept) that can be applied across all groups
- For within-group analyses or studies
  - the culture-specific portion is also relevant

Psychometric Adequacy in One Group

<table>
<thead>
<tr>
<th>Conceptual</th>
<th>Psychometric properties meet minimal standards within one group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concept meaningful within one group</td>
<td>Psychometric properties invariant (equivalent) across groups</td>
</tr>
</tbody>
</table>

Psychometric Adequacy in any Group

- Minimal standards:
  - Sufficient variability
  - Minimal missing data
  - Adequate reliability/reproducibility
  - Evidence of construct validity
  - Evidence of responsiveness to change
- Basic classical test theory approach
Psychometric Equivalence

Conceptual
Concept meaningful within one group

Adequacy in 1 Group
Psychometric properties meet minimal standards within one group

Psychometric

Psychometric

Psychometric properties invariant (equivalent) across groups

Equivalence Across Groups

Concept equivalent across groups

Measurement Invariance Studies Across Groups

◆ Confirmatory factor analysis (CFA)
  – goal is to identify a factor structure that is equivalent across all groups

◆ Item response theory (IRT)
  – goal is to determine whether any individual items are biased with respect to the underlying trait or latent construct in a particular “group”

What if Measures Are Not Equivalent Across Groups?

◆ If measure can be modified/adapted
  – Can adapt for use in a new group

◆ If measure cannot be modified
  – Limited to concepts/items of original measure

When a Measure Can be Modified

◆ Adapt original measure to be appropriate for mainstream and new group
  – Pretest/revise
  – Once psychometric testing shows equivalence
    – Adapted measure can be used for group comparisons

Example: Measure Can be Modified

◆ GHAA Consumer Satisfaction Survey
  – Adapted to be appropriate for African American patients

  – Focus groups conducted to obtain perspectives of African Americans

  – New domains added (e.g., discrimination/stereotyping)

  – New items added to existing domains


When a Measure Cannot be Modified

◆ All translations must reflect original English
  – No modifications can be made to original
  – Translation attempts to match meaning of the original as closely as possible

◆ Assumes perspectives of diverse group are similar to mainstream (universality)
  – “Cultural hegemony” (Guyatt)
  – “Middle-class ethnocentrism” (Rogler)
When Measures Are Not Equivalent in a Specific Study

- Need guidelines for how to handle data when substantial non-comparability is found in a study
  - Drop bad or biased items from scores
  - Compare results with and without biased items
  - Analyze study by stratifying diverse groups
- The current challenge for measurement

International Versus U.S. Approach

- Assess Conceptual Equivalence (Qualitative)
- Assess Psychometric Equivalence (Quantitative)

Typical International Approach

Begin here (assumes conceptual differences across countries)

Assess Conceptual Equivalence (Qualitative)
Assess Psychometric Equivalence (Quantitative)

- If new domains or definitions are found, can revise and add items
- Translated “adapted” version is the goal
- May achieve conceptual equivalence before testing psychometric adequacy

Typical U.S. Approach When No Translation is Done

Assess Conceptual Equivalence (Qualitative)
Assess Psychometric Equivalence (Quantitative)

Most studies begin here (assumes universality of constructs)

Typical U.S. Approach When No Translation is Done

Assess Conceptual Equivalence (Qualitative)
Assess Psychometric Equivalence (Quantitative)

Most studies begin here (assumes universality of constructs)

No Guidelines!
If refine items based on qualitative studies, no longer have comparable instrument

Typical U.S. Approach When No Translation is Done

Assess Conceptual Equivalence (Qualitative)

Most studies begin here (assumes universality of constructs)

- Proceed with analysis. May miss important domains and definitions
- If problems

If equiv.

Proceed with analysis. May miss important domains and definitions
<table>
<thead>
<tr>
<th>Conclusions</th>
</tr>
</thead>
</table>
| - Measurement in health disparities research is a relatively new field  
  - Few guidelines  
| - Encourage first steps  
  - Test and report adequacy and equivalence  
| - As evidence grows, concepts and measures that work better across diverse groups will be identified  