

Cross-cultural Translation and Validation of HRQOL Measures

HKSoQOL AGM, 17 July 2010

Cindy L.K. Lam, Wendy Wong, Elegance T.P. Lam

Family Medicine Unit,
The University of Hong Kong

cklam@hku.hk



Workshop Objectives

- Understand the concept of validity
- Understand the concept of cross-cultural equivalence
- Learn the skills of translation and cognitive debriefing
- Apply the standardized methods of cross-cultural adaptation of HRQOL measures



Workshop Programme

- Overview of cross-cultural validity & equivalence of HRQOL measures
- Exercises on translation & cognitive debriefing
- Break
- Discussion on exercise results
- Overview on psychometric testing of HRQOL measures

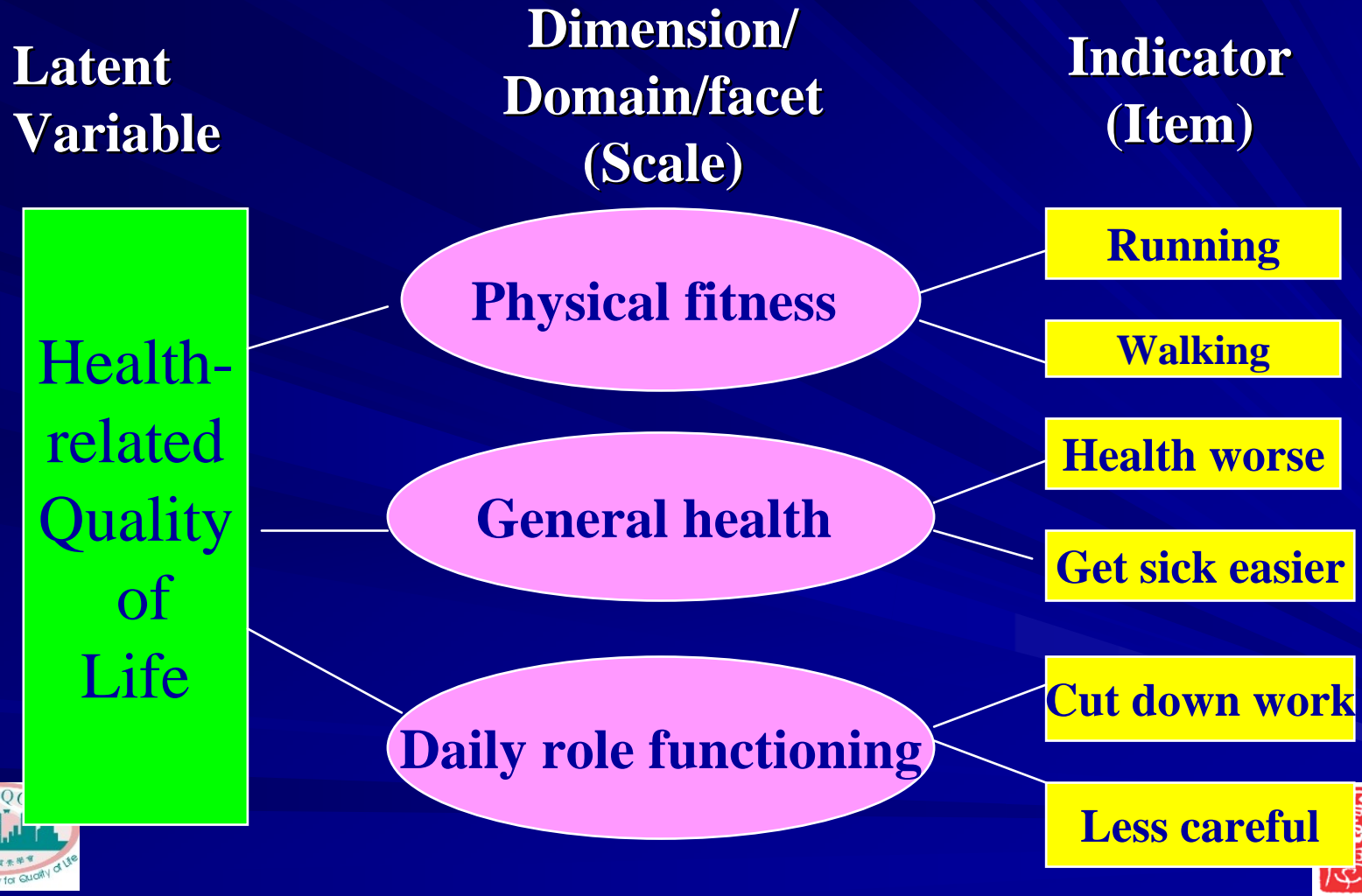


Cross-cultural HRQOL Measures

- Common definitions of concepts & indicators
- **A standard metric for easier interpretation**
- Continuing validation & development
- **Cross-cultural/global clinical trials**
- International comparative studies



Construction of a HRQOL Measure



Composition of a HRQOL Measure

- An observable indicator of HRQOL presented as an item
- Response choices for each item
 - Dichotomous/ Likert/ VA scales
- Collection of indicators of the same domain = a scale
- Scoring Method
 - Equal/ weighted summation
 - Norm-based scoring



Indicators of HRQOL

- **Effect indicators**
 - **Depend on and correlate with HRQOL**
- **Causal indicators**
 - **Affect HRQOL but variable correlation with HRQOL**
- **Representative and adequate**
- **Psychometric performance**
 - **Construct validity, reliability and sensitivity**



Cross-cultural HRQOL Measures

- Validity means it measures what it intends to measure
 - face, content, construct, criterion
- Equivalence means it gives a similar result for an equivalent situation in different cultures
 - conceptual , semantic, item, operational, measurement , functional



Sickness Impact Profile Items

- I don't write except to sign my name
- I am not doing any of the maintenance or repair work in my home or yard

What do they intend to measure?

Are they valid in HK?



Validity of a HRQOL Measure

- **Face**
 - expert evaluation, essential but insufficient
- **Content**
 - Relevance, importance & representativeness
- **Construct (conceptual)**
 - Hypothesis on the construct
 - Scaling success & factor structure
 - Expected correlations with external criteria
- **Criterion**
 - gold standard or predictive of criterion outcome



**Sometimes in elections I vote for
people whom I know very little
about
(MMPI-2 L Scale Item)**

**I show less interest in other people'
s problems (SIP Item)**

What does each item mean in the US?

What does each item mean HK?

Conceptual Equivalence

- Does the concept (HRQOL) exist?
- Is the concept equally important?
- Does the concept have the same meaning?
- Does the concept have the same relevance?
- Does the concept have the same dimensions, domains & indicators?



Semantic Equivalence

- The translation has the same linguistic interpretation
 - literal
 - functional meaning
 - idiomatic
- Grammatically correct
- Easy to understand



Semantic Equivalence

- Were limited in the kind of work or other activities
工作或從事某些活動受到限制
- I am as healthy as anyone I know
您和所有您認識的人一樣健康
- Have you felt downhearted and blue?
您覺得心情不好，悶悶不樂？



Item Equivalence (content/ experiential equivalence)

- The items as indicators of the HRQOL domains are
 - **equally representative**
 - **equally important**
 - **equally relevant**
- **Qualitative cognitive debriefing**
- **Quantitative psychometric testing**



Operational Equivalence

- Same questionnaire format
 - paper and pen
 - computerized
 - web-based
 - VAS
- Same method of administration
 - self-completion
 - Interviewer-administration
 - Telephone administration



Measurement Equivalence

- Similar psychometric properties
 - **Reliability**
 - **Sensitivity**
- Response choices ordinality and intervals
- **Similar results for similar conditions**



Measurement Equivalence – Response Choices

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

Excellent	1
Very good	2
Good	3
Fair	4
Poor	5

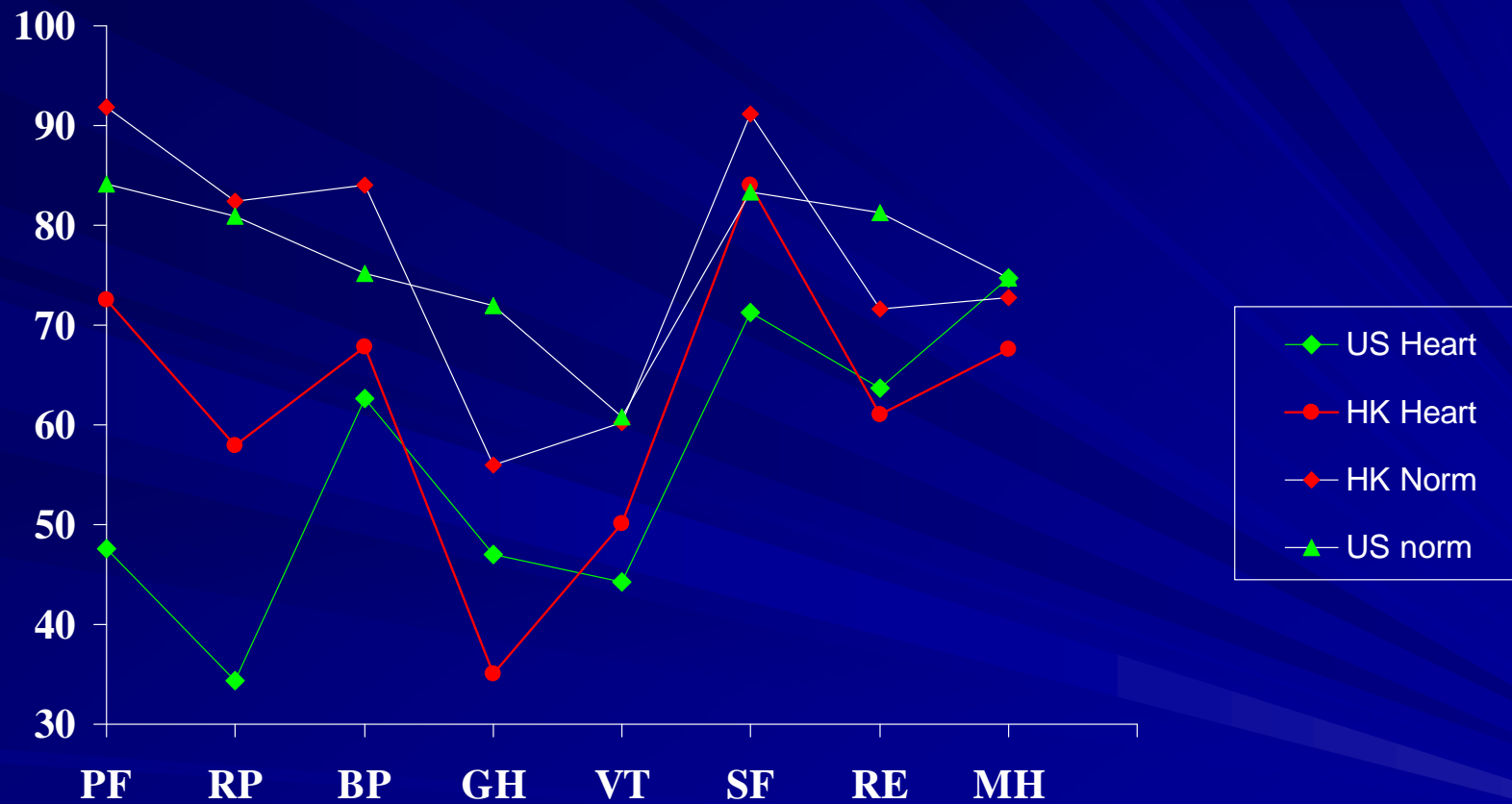
- Does ‘excellent’ have the same standard in the US and HK?



How much is ‘good’ better than ‘fair’?



Measurement Equivalence



SF-36 Scale Scores

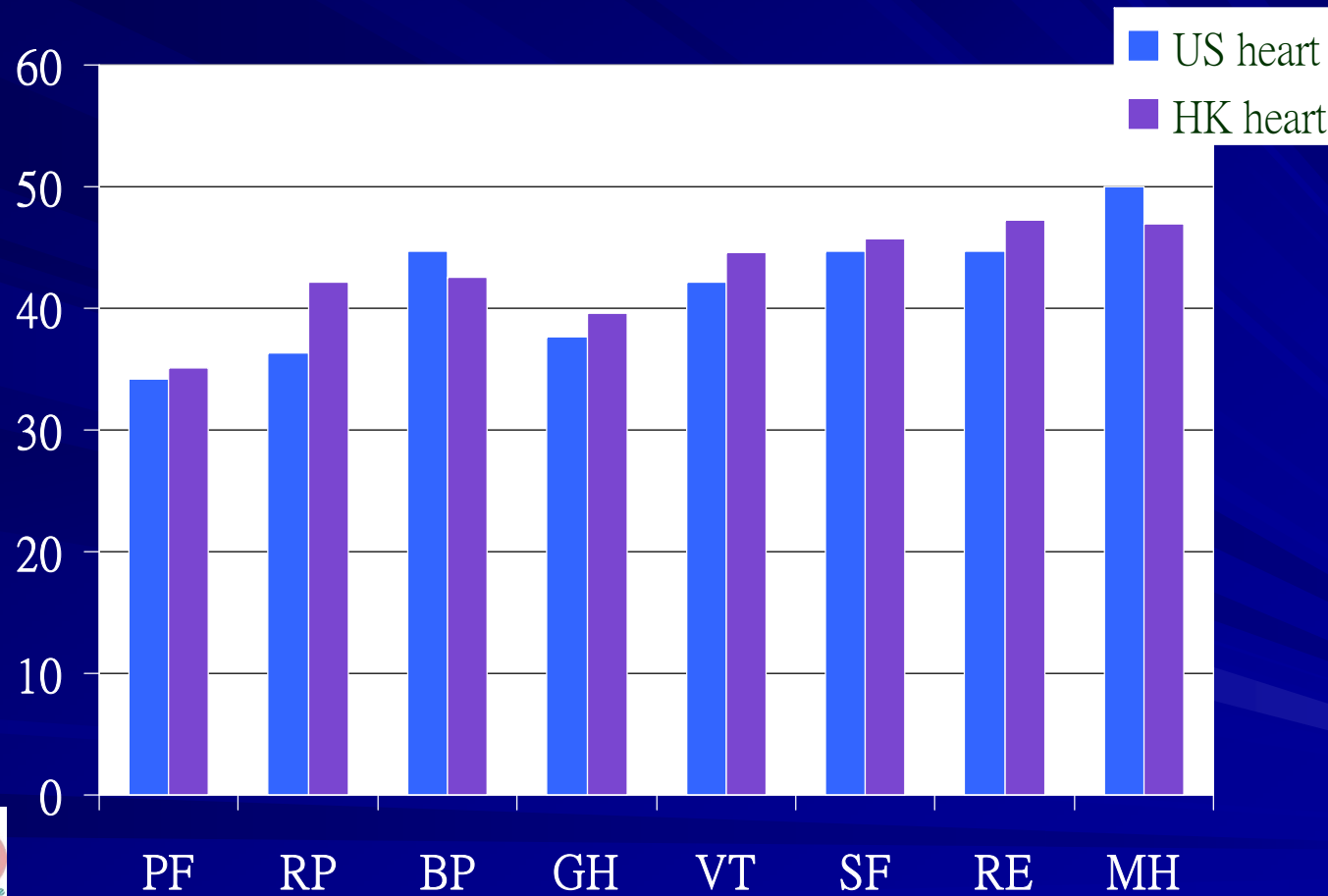


Norm-based Scoring

- **Z score transformation**
(Observed score – Population mean)
/ Population SD x 10 + 50
- **Population Mean = 50**
- **Population SD = 10**



Norm-based SF-36 Scores of Patients with Heart Diseases



Cross-cultural Translation & Adaptation of HRQOL Measures

- Exploration of conceptual equivalence
- **Linguistically valid & equivalent translation**
- **Assessment of content validity**
- Pilot testing on psychometrics, & operational & measurement equivalence
- Clinical testing for sensitivity & responsiveness
- Population study for norm-based interpretation



Linguistic Translation/ Adaptation

Wild et al. ISPOR Guidelines Value Health 2005

- **Double forward translation**
- **Expert reconciliation**
- **Single/ double backward translation**
- **Review by original developer / expert**
- **Revision**
- **Cognitive debriefing with 5-8 subjects**
- **Revision and evaluation by experts**



Exercises in Groups of 4

Iterative Translations – 30 minutes

1. Two persons translate 2 items from English to Chinese independently
2. Reconciliation of the forward Chinese translation
3. Swap with the other 2 persons to back translate the items to English
4. Panel of 4 reconcile on the Chinese translations of the 4 items



Cognitive Debriefing – 15 minutes

Reconciled Translation of the item	Do you understand this item? (If no, please explain the difficulty) 你明白此的意思嗎？(若是不明白，請指出難以明白的地方)	Would you suggest any rewording of this item ? 你建議改寫此項目嗎？	Can you please explain what it means in your own words? 你能用你自己的字句解釋這項目嗎？.



Final Revision – 15 min

- 1. Is the item well understood?**
- 2. Is the item interpreted with an equivalent meaning as original?**
- 3. Revision of non-equivalent item**
- 4. Reword the item to improve understanding & grammar**
- 5. Write the your final translation for presentation**



Thurstone Scaling Exercise 1

- Each line is a 100mm ruler from 0 on the left to 100 on the right.
- Each person mark an 'X' on the line to represent the relative position of the response option independently



Thurstone Scaling Exercise 2

- Each calculate the score for each response option = (distance of X from the left end in mm/ 100) x score range + 1
- Calculate the mean score of the response value
- Rank order the response options



Cognitive Debriefing of HK CLDQ

Scale	Item no.	Original item	Problems	Action
FA	11	感到精力下降 (decreased energy)	Misinterpreted the meaning as physical strength only	No change
FA	13	昏昏欲睡 (drowsy)	Had difficulty to differentiate sleepy and drowsy	No Change but add explanatory notes
EF	19	心情起伏不定 (mood strings)	Did not interpret the term clearly	No Change
WO	28	擔心自己的病情日後也不會好轉 (worried about never feeling any better)	Did not include the meaning of 'worries about never feeling any better'	Revised to 因為自己的病情日後也不會好轉而擔心 (worry that your health condition will not improve)



Mean Values of Response Choices

Response Choice	SF-36	Keller et al*	HK
All the time	6.0	6.0	6.0
Most of the time	5.0	5.0	5.3
A good bit of the time	4.0	4.5	4.6
Some of the time	3.0	3.1	3.2
A little of the time	2.0	2.1	2.3
None of the time	1.0	1.0	1.0

* Data from 17 populations, Keller et al. *J Clin Epidemiol* 1998; 51: 933-944



Content Validity Index of HK CLDQ

	Physicians (n=13)			CHB patients (n=23)		
	Hepatologists (n=6)	Non-hepatologists (n=7)	Total (n=13)	Uncomplicated (n=14)	Complicated (n=9)	Total (n=23)
Clarity	1.00	0.71-1.00	0.85-1.00	0.53-1.00	0.88-1.00	0.70-1.00
Relevance	0.17-1.00	0.57-1.00	0.46-1.00	0.07-1.00	0.13-0.88	0.17-0.96
AS	0.5-1	0.71-1	0.69-1	0.6-0.73	0.5-0.75	0.65-0.65
FA	0.33-1	0.57-1	0.54-1	0.67-1	0.5-0.88	0.61-0.96
SS	0.33-0.67	0.57-0.86	0.46-0.77	0.07-0.73	0.38-0.63	0.17-0.7
AC	0.33-0.33	0.57-0.86	0.46-0.62	0.53-0.6	0.63-0.63	0.57-0.61
EF	0.17-0.33	0.71-1	0.46-0.69	0.4-0.6	0.13-0.5	0.35-0.57
WO	0.5-1	0.86-1	0.77-1	0.67-0.8	0.63-0.88	0.7-0.83

AS, Abdominal Symptom; FA, Fatigue; SS, Systemic Symptoms; AC, Activity; EF, Emotional Function; WO, worry



Content Validity of HK CLDQ

HK CLDQ Items	CVI on Relevance among Physicians			CVI on Relevance among Patients		
	Hepatologist (n=6)	Non- hepatologist (n=7)	Total (n=13)	Uncomplicated CHB (n=14)	Complicated CHB (n=9)	Total (n=23)
EF						
10 感到焦慮 (Anxiety)	0.17	0.86	0.54	0.4	0.25	0.35
12 覺得不開心 (Unhappiness)	0.17	0.86	0.54	0.4	0.38	0.39
16 在晚上難以入睡 (Difficulty in sleeping at night)	0.17	0.71	0.46	0.47	0.5	0.48
20 不能一覺睡到天光 (Difficulty in falling asleep at night)	0.17	1	0.62	0.53	0.38	0.48
26 難以集中精神 (Problems with concentration)	0.33	1	0.69	0.47	0.13	0.35
SS						
6 因氣促而影響你的日常活 動 (Shortness of breath)	0.33	0.71	0.54	0.27	0.38	0.3
21 肌肉有多常會抽筋 (Muscle cramps)	0.67	0.57	0.62	0.33	0.5	0.39
23 感到口乾 (Dry mouth)	0.5	0.57	0.54	0.33	0.38	0.35
27 因為痕癢而感到困擾 (Itching)	0.67	0.86	0.77	0.07	0.38	0.17



Pilot Psychometric Test

- Operational equivalence, feasibility & acceptability
- Item validity & equivalence
 - **Factor structure**
 - **Scaling success**
 - **Concurrent (convergent) construct validity**
- Measurement properties & equivalence
 - **Ceiling/ floor effects**
 - **Reliability**
 - **Sensitivity**



Factor Structure of the HK CLDQ

Item	Factor 1 EF	Factor 2 WD	Factor 3 SS+FA	Factor 4 AS	Factor 5 AC	Factor 6 SL
<i>Abdominal Symptom (AS)</i>						
1: Abdominal bloating	0.18	0.27	0.16	0.70	0.14	0.15
5: Abdominal pain	0.20	0.06	0.13	0.81	0.08	0.14
17: Abdominal discomfort	0.17	0.23	0.17	0.82	0.20	0.08
<i>Fatigue (FA)</i>						
2: Tiredness or fatigue	0.33	0.03	0.51	0.31	0.45	0.22
4: Feel sleepy during the day	0.28	-0.05	0.47	0.14	0.44	0.14
8: Decreased strength	0.37	0.14	0.37	0.24	0.60	0.24
11: Decreased energy	0.56	0.06	0.37	0.11	0.51	0.12
13: Drowsiness	0.49	0.09	0.47	0.15	0.36	-0.17
<i>Systemic Symptoms (SS)</i>						
3: Bodily pain	0.49	0.00	0.18	0.39	0.14	0.22
6: Shortness of breath	0.22	0.22	0.59	0.39	0.18	0.03
21: Muscle cramps	0.09	0.08	0.76	0.17	0.10	0.14
23: Dry mouth	0.24	0.42	0.48	0.16	0.08	-0.02
27: Itching	0.10	0.27	0.65	-0.01	0.02	0.28
<i>Activity (AC)</i>						
7: Not able to eat as much as you would like	0.05	0.29	0.04	0.24	0.81	0.12
9: Trouble in lifting or carrying heavy objects	0.41	0.18	0.37	-0.05	0.35	0.17
14: Bothered by a limitation of the diet	0.11	0.43	0.09	0.10	0.67	0.12
<i>Emotional Function (EF)</i>						
10: Anxiety	0.67	0.32	0.15	0.39	0.14	0.04
12: Unhappiness	0.77	0.25	0.15	0.20	0.13	0.01
15: Irritability	0.78	0.27	0.15	0.17	0.06	0.13
16: Difficulty in sleeping at night	0.27	0.17	0.17	0.18	0.14	0.79
19: Mood swings	0.78	0.24	0.10	0.19	0.20	0.21
20: Difficulty falling asleep at night	0.16	0.14	0.25	0.22	0.25	0.71
24: Depression	0.80	0.31	0.15	0.08	0.03	0.22
26: Problems on concentration	0.52	0.28	0.29	0.18	0.37	0.27
13: Worries about the impact of the liver disease	0.40	0.47	0.03	0.50	0.25	0.03
18: Worries about what symptoms will develop into major problem	0.29	0.80	0.11	0.16	0.04	0.03
25: Worries about what the condition is getting worse	0.29	0.84	0.13	0.17	0.17	0.03
28: Worries about never feeling any better	0.27	0.77	0.13	0.18	0.26	0.24
29: Availability of a liver for transplant	0.16	0.64	0.26	0.14	0.30	0.10

Move to AC

New factor

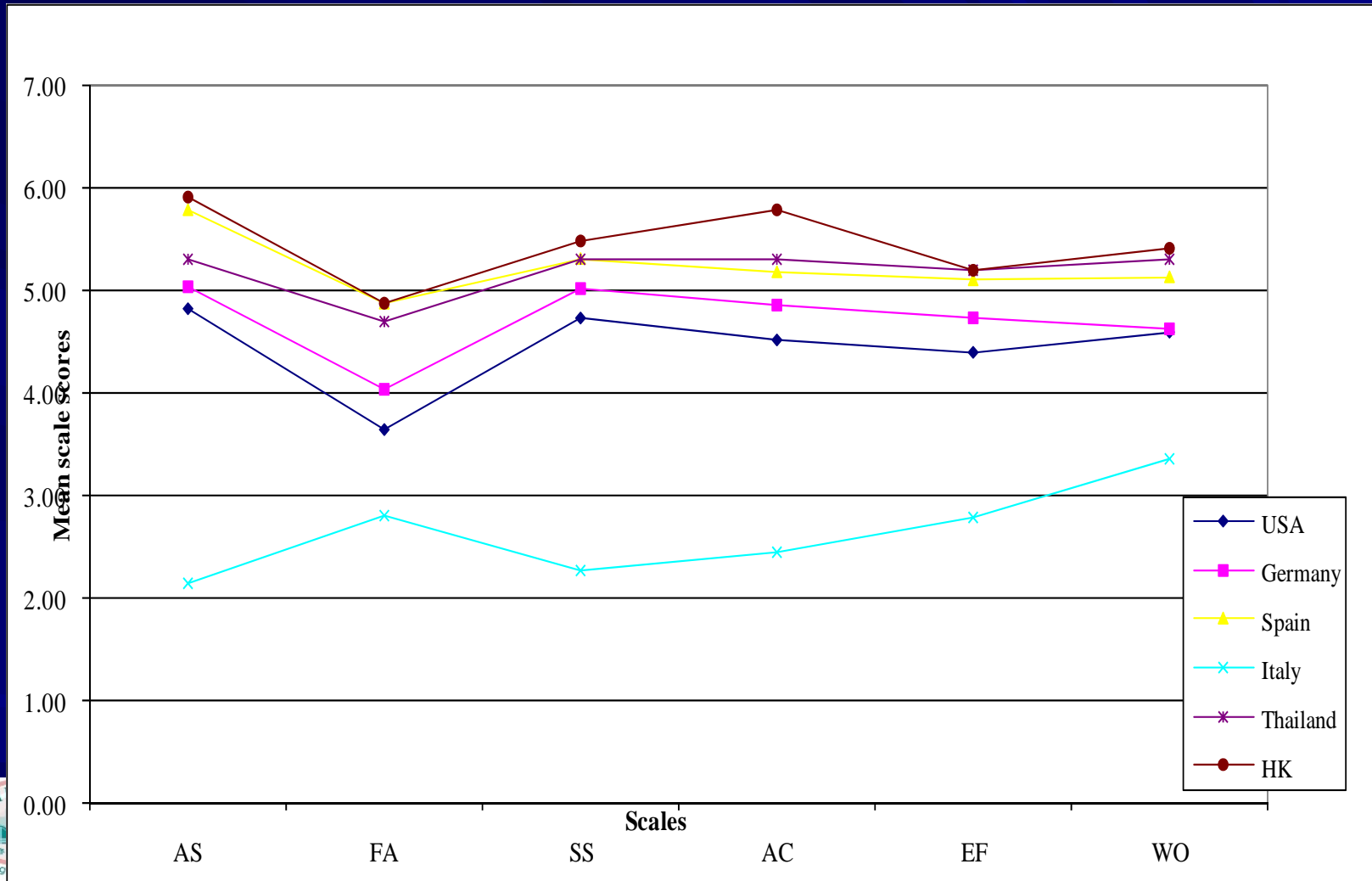


Psychometrics Properties of HK CLDQ

Subscales	Item-Scale correlations	Scaling success	Correlations corresponding SF-36v2 scales	Cronbach's alpha	ICC
AS	0.57-0.73	100	NA	0.84	0.58
FA	0.66-0.79	100	0.79 (VT)	0.88	0.82
SS	0.44-0.57	64	0.62 (BP)	0.74	0.86
AC	0.42-0.65	73	0.46-0.67 (PF, RP, RE)	0.72	0.66
EF	0.52-0.79	100	0.78 (MH)	0.90	0.86
WO	0.62-0.85	100	0.50-0.61 (MH & RE)	0.90	0.89
Overall	N.A.	NA	0.70-0.75 (PCS & MCS)	0.90	0.85



Measurement Equivalence of CLDQ



Sensitivity of the HK CLDQ

	Uncomplicated CHB (n=72)		Complicated CHB (n=78)		ES†
	Mean	(SD)	Mean	(SD)	
CLDQ					
AS	6.3	(1.0)	5.6	(1.5)	0.5‡
FA	5.2	(1.1)	4.6	(1.5)	0.4‡
SS	5.8	(1.0)	5.2	(1.2)	0.5‡
AC	6.2	(1.1)	5.4	(1.6)	0.6‡
EF	5.5	(1.0)	5.0	(1.4)	0.4‡
WO	5.9	(1.3)	5.0	(1.7)	0.6‡
Overall	5.8	(0.8)	5.1	(1.2)	0.6‡
SF36-v2 (HK norm)					
PCS (50)	46.4§	(9.9)	39.5§	(12.2)	0.6‡
MCS (50)	50.1	(10.4)	47.8	(13.9)	0.2



†Effect size was calculated as the difference between uncomplicated and complicated mean domain score, divided by the overall SD.

‡Significant difference between uncomplicated and complicated CHB patients by independent sample t-test ($p < 0.05$).

§Significant difference between CHB groups and HK norm by independent sample t-test ($p < 0.05$).



Conclusions

- **An indicator may not be valid in a different culture**
- **An indicator may not measure the same concept in different cultures**
- **An equivalent translation is the first essential step**
- **Cognitive debriefing is becoming a standard**
- **Pilot psychometric testing should be done before application**



Further Reading

- **Wild D et al. Principles and good practice for the translation and cultural adaptation process for PRO measures: report of the ISPOR Task Force for translation and cultural adaptation. Value in Health 2005; 8:94-104.**
- **Herdman H et al. A model of equivalence in cultural adaptation of HRQOL instruments: the universalist approach. QoL Res 1998; 7:323-335.**
- **Guillemin et al. Cross-cultural adaptation of HRQOL measures. J Clin Epidemiol 1993; 46:1417-1432.**

