Assessment of Functional Outcomes

Dawn I. Velligan, Ph.D.

University of Texas, Health Science Center, San Antonio

Assessment of Functional Outcome

- What are the options for use in clinical trials?
- What are the psychometric properties of available assessments of functional capacity?
- How applicable are these assessments to different cultural contexts?

Bellack et al., 2006; Schiz Bul., 33: 805-882; McKibbon et al., 2004 Schiz Res 72: 53-67; Velligan et al., 2007 Psychiar Res,151: 55-66;

Why do we need functional outcome measures?

- For a medication to receive an indication for improving cognition, measures of cognition and functional outcome must improve
- Longer-term functional outcomes such as employment or marital status are multi-determined and unlikely to be improved during the course of a clinical trial
- Intermediate measures of functional outcome have face validity and are more likely to change during course of a clinical trial

Bellack et al., 2006; Schiz Bul., 33: 805-882; McKibbon et al., 2004 Schiz Res 72: 53-67; Velligan et al., 2007 Psychiar Res,151: 55-66;

Why do performance-based assessment?

- Self-report-most feasible, relies on insight and memory
- Informant report-provides information on behaviors not available to providers, behavioral samples not standard, availability may be an issue
- Clinician ratings- based on self/informant report with clinical judgment, validity may depend upon question format
- Naturalistic Observation-ecologically valid data; usually not feasible, inter-rater reliability must be addressed

Bellack et al., 2006 Schiz Bul, 33: 805-882; Velligan et al., 2007 Psychiar Res, 151: 55-66; McKibbon et al., 2004 Schiz Res 72: 53-67.

Validation of Intermediate Measures (VIM)

- Cognitive Assessment Interview (CAI)
- UCSD Performance-Based Skills Assessment (UPSA)
- Test of Adaptive Behavior in Schizophrenia (TABS)
- Independent Living Scales (ILS)
- Goal: to assess reliability, validity and utility of the intermediate measures

Cognitive Assessment Scale (CAI)

- 10 items assessing the domains of the MCCB.
  - Domains include: Speed of Processing, Attention/Vigilance, Working Memory, Verbal Learning and Memory, Visual Learning and Memory, Reasoning and Problem Solving, and Verbal Comprehension.
- Example Questions
  - "Do you have difficulty keeping figures in mind while paying bills?" "Do you have trouble learning or remembering instructions or other important information?" "Do you have trouble coming up with alternatives when your plans are disturbed?"

Bilder et al., 2008; Keefe et al., 2006; (Loeb, 1996)

The Independent Living Scales ILS

(Loeb, 1996)

- Memory/Orientation—e.g. person is asked to remember the name of a new doctor and the time of an appointment when asked later
- Managing Money—e.g., person is asked to make out a check/money order to a utility company;
- Managing Home and Transportation—e.g. person is asked how they would go about getting repairs made to their home
- Health and Safety—e.g. the person is asked what they would do if they cut their hand and it was bleeding badly
- Social Adjustment—e.g. person is asked to name two reasons why it is important to have relationships.

The UCSD Performance-based Skills Assessment (UPSA)

Patterson et al.2001

- Designed to assess an individual's ability to perform everyday tasks that are considered necessary for independent functioning in the community.
  - Finance, Communication, Organization and Planning, Transportation, Household Management

GRAND OPENING OF NEW WATER PARK! Wild Water is the state's newest and largest water theme park with…

Test of Adaptive Behavior in Schizophrenia (TABS)

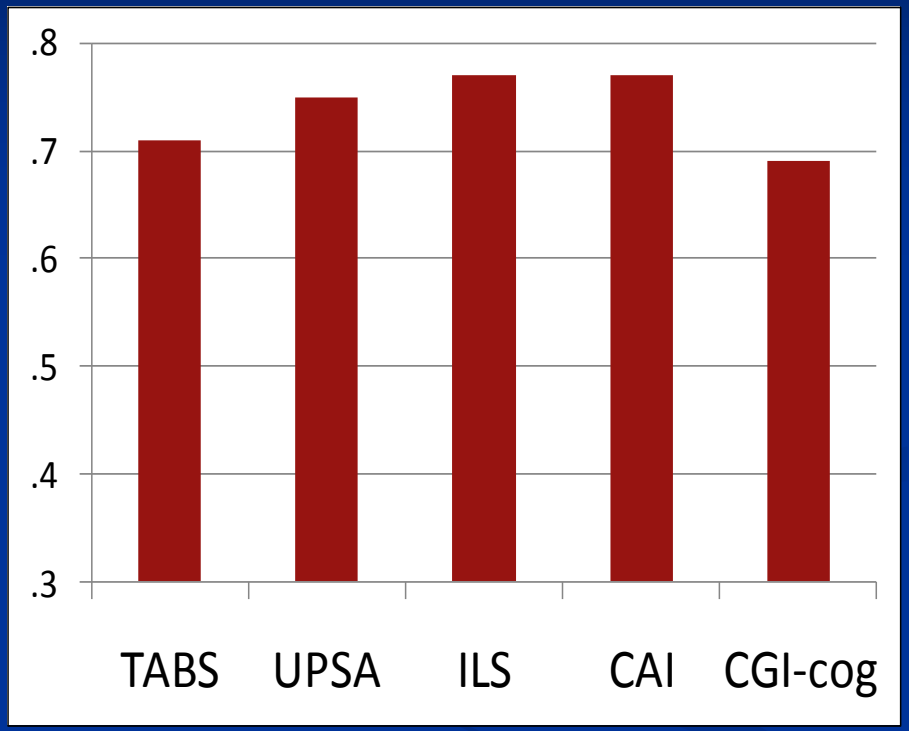
- Designed to assess underlying abilities needed to complete goal-directed adaptive behavior such as initiation, planning and sequencing, and problem identification.
- Subtests include: Medication Management, Work and Productivity, Empty Bathroom, Grocery Store, Clothes Closet, Social Skills

(Velligan et al., 2007)

VIM Study Results
Test-retest reliability

N=166 stable patients at BL,
147 at 4 weeks

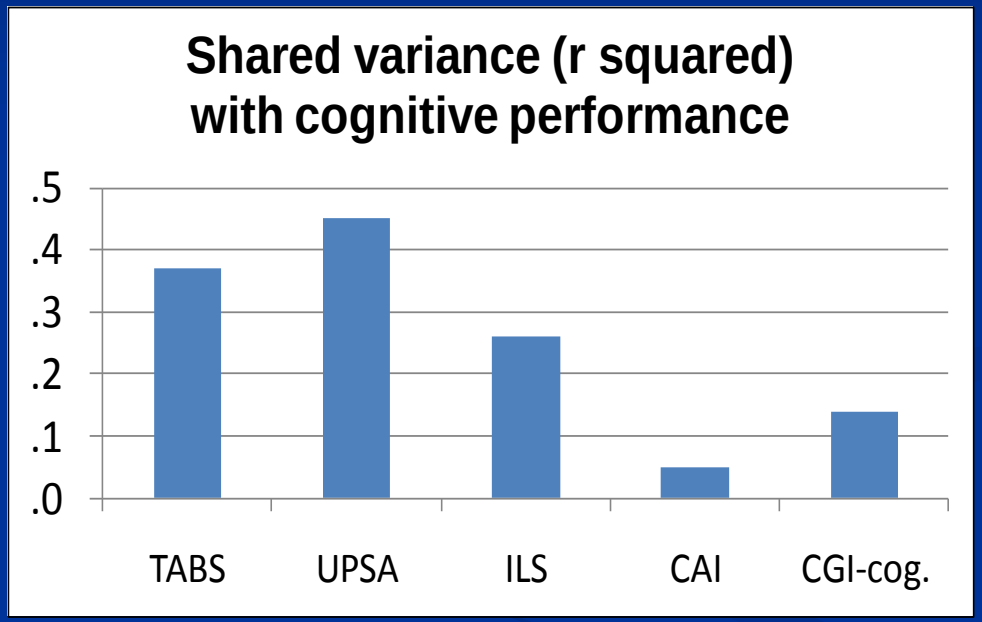
	r	ICC	CI for r
TABS	.71	.69	.62-.78
UPSA	.75	.74	.67-.81
ILS	.77	.76	.69-.83
CAI*	.77	.76	.69-.83



* inter-rater reliability for CAI; ICC = .73
VIM study used same rater 88% of the time

VIM Study Results
Correlation with Cognitive Performance

	Pearson r	r^2
TABS	.61	.37
UPSA	.67	.45
ILS	.51	.26
CAI	.23	.05
CGI-cog.	.38	.14



UPSA > ILS > CGI-cog.
CAI
TABS > CGI-cog, CAI

VIM Study Results
Additional Scientific Criteria
Utility as a repeated measure / Correlation with functioning

	Scores at Floor/ CeilingT1	Scores at Floor/ CeilingT2	T1-T2 effect size
TABS	0/0	0/0	.24
UPSA	0/0	0/0	.18
ILS	4/0	2/0	.15
CAI	1/0	1/0	.10
CGI-cog.	6/0	5/0	.03

	Corr. w/ comm. function
TABS	.23
UPSA	.25
ILS	.30
CAI	.27

Practicality and Tolerability/Correlation with Community Function

	Practicality	Tolerability *	Duration (min) **
TABS	5.2	5.9	33
UPSA	6.0	5.6	27
ILS	4.6	5.4	46
CAI	4.1	6.0	25

	Corr. w/ comm. function
TABS	.23
UPSA	.25
ILS	.30
CAI	.27

Practicality and Tolerability on 1-7 scale where 7 = best

Tolerability: *CAI, TABS > UPSA > ILS

Administration Time: **ILS > TABS > UPSA > CAI

VIM Study Results
Conclusions

1. VIM Committee followed a clearly defined process for evaluation of study results for 5 measures: TABS, ILS, UPSA, CAI, CGI-cognition

2. For the full measures, the UPSA was the leading measure because:
   - good test-retest reliability
   - excellent shared variance w/ cognitive performance
   - good utility as repeated measure; no floor / ceiling effects
   - reasonable tolerability and practicality

3. For short forms, the TABS and UPSA were the leading measures because:
   - well-defined short forms
   - moderate shared variance w/ cognitive performance
   - acceptable utility as repeated measure
   - but: lower test-retest reliability, requiring larger samples

Cultural Adaptation of Intermediate Measures CIM-Study

- Goal was to have 31 sites participate; each with a bilingual PI and at least one bilingual research assistant.
- More sites (n=5) were recruited from less-westernized countries; Fewer (n=2) from more westernized countries
- Specific countries were chosen based upon languages for MCCB translation; Germany Russia, India, China, Spain, Argentina, and Mexico. U.S. sites were included for baseline comparison
- The goal was 62+ expert raters and we obtained data from 55

Example C-CARS Item

Rate the extent to which the section/item would work in in your country with typical patients

1	2	3	4	5	6	7
Doesn't Work At all						Works Extremely Well

Comment (If you gave a rating of 4 or lower) :

C-CARS scores for overall level of cultural adaptability

	FULL SCALES					BRIEF SCALES	
	CAI	TABS	ILS	UPSA	Mean	TABS	UPSA
US	6.03	6.07	6.12	6.00	6.06	5.80	6.10
GERMANY	6.92	6.29	5.45	5.95	6.15	6.38	6.25
ARGENTINA	6.60	6.04	5.95	5.30	5.97	5.63	5.81
SPAIN	6.46	5.88	5.45	5.45	5.81	4.88	5.13
RUSSIA	6.39	6.01	5.55	5.22	5.79	5.59	5.10
MEXICO	6.31	5.39	5.70	***4.84****	5.56	4.89	***4.61****
CHINA	5.81	5.71	***5.17****	***4.63****	5.33	5.50	***4.50****
INDIA	6.10	***4.60****	***4.98****	***3.98****	4.91	***4.13****	***3.19****
Mean	6.33	5.75	5.54	5.17	5.65	5.35	5.09

Estimates from mixed effects regression. Asterisks indicate significantly poorer adaptability than in the US value (p<0.05, ** P<0.01 Holm-Bonferroni adjusted). Root mean squared standard error=0.26 (range: 0.17-0.36). Overall scale means: CAI>TABS>ILS>UPSA, all pair-wise differences significant at Holm-Bonferroni p=0.001 except TABS vs. ILS, p=0.017.*

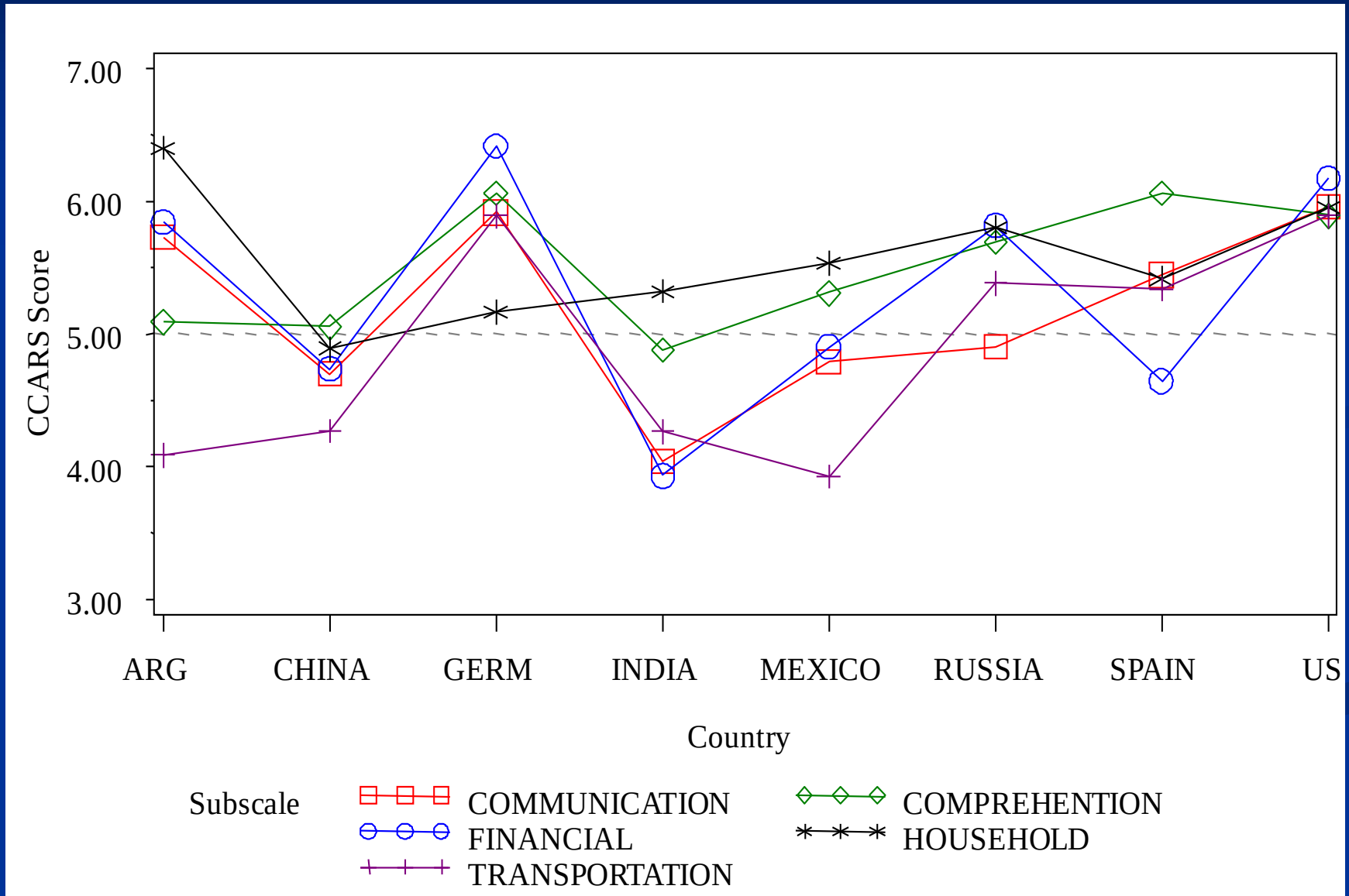
CAI>ILS=TABS>UPSA; Biggest problems in adaptation for lower SES, and rural residents

C-CARS Report Card for Intermediate Measures

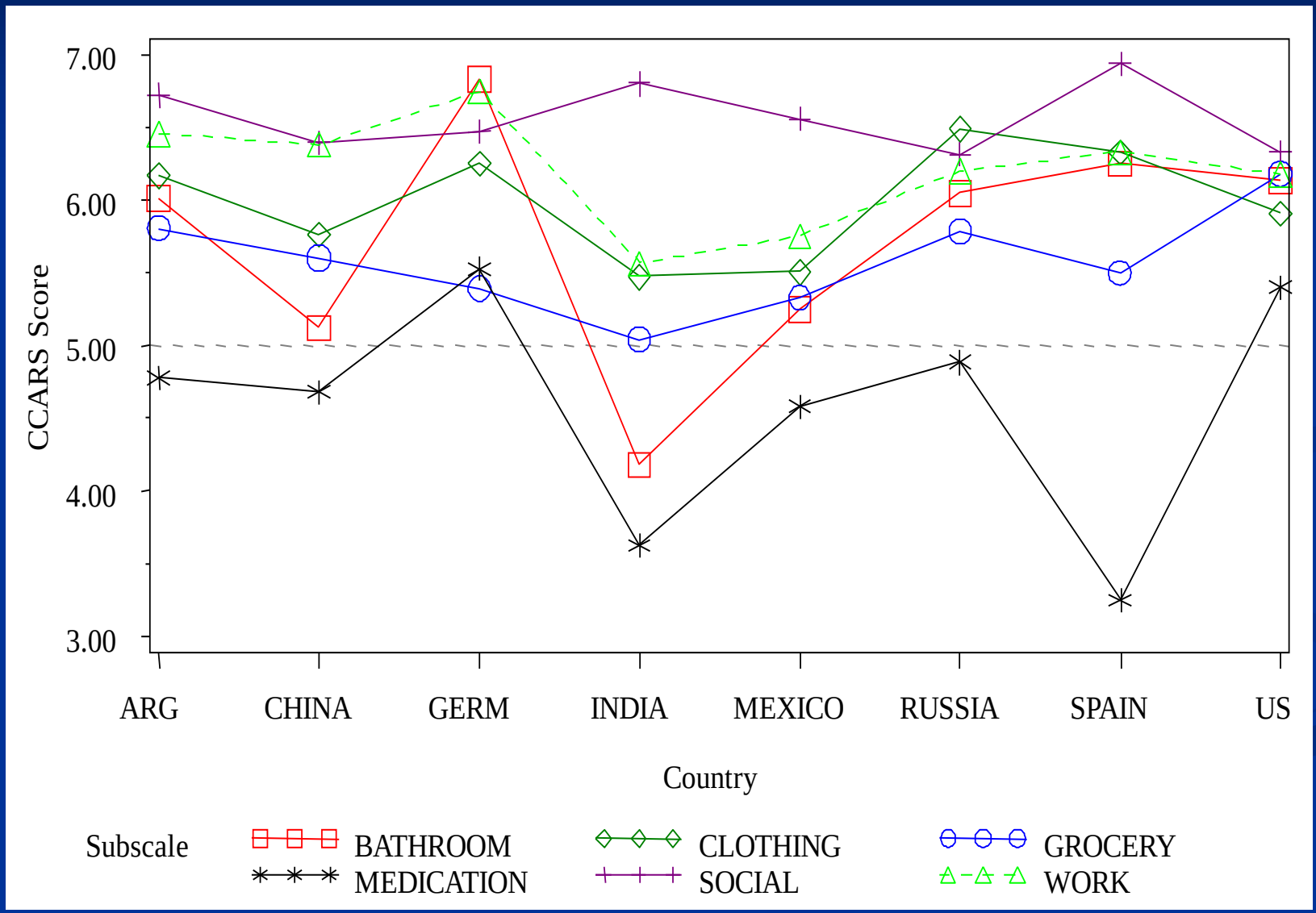
	Full Scales				Brief Scales	
	CAI	TABS	ILS	UPSA	TABS	UPSA
US	A	A	A	A	A	A
GERMANY	A	A	A	A	A	A
ARGENTINA	A	A	A	C	B	A
SPAIN	A	A	A	**F**	**F**	**F**
RUSSIA	A	A	A	D	A	B
MEXICO	A	D	B	**F**	C	**F**
CHINA	B	B	C	**F**	C	**F**
INDIA	A	**F**	**F**	**F**	**F**	**F**
GPA	**3.9**	***3.0***	**3.1**	**1.4**	**2.4**	**1.9**

Note: Letter grades assigned conventionally (90-100=A, 80-89=B, 70-79=C, 60-69=D, <60%=F). GPA based on A=4, B=3, C=2, D=1.

Are there individual subtests that do not work well in some countries?-UPSA



Are there individual subtests that do not work well in some countries?-TABS



UPSA-B International Use

- The UPSA-B was used internationally for the first time in a trial sponsored by Dainippon Sumitomo Pharma (DSP)
- It was implemented in the following countries and languages:
  - Russia
  - Colombia
  - Romania
  - Ukraine
  - India
    - Gujarati
    - Hindi
    - Kannada
    - Marathi
    - Tamil
    - Telugu

Challenges of International Translation and Implementation

- Adaptation difficulties
  - For almost every adaptation, changes to items and/or stimuli were necessary for cultural appropriateness
    - Alter the assumption of face validity
    - Challenge assumption of ecological validity
  - Even after translation finalized, sites and raters were not satisfied with the translation
    - Regional differences in everyday activities seem to be greater in non-US countries

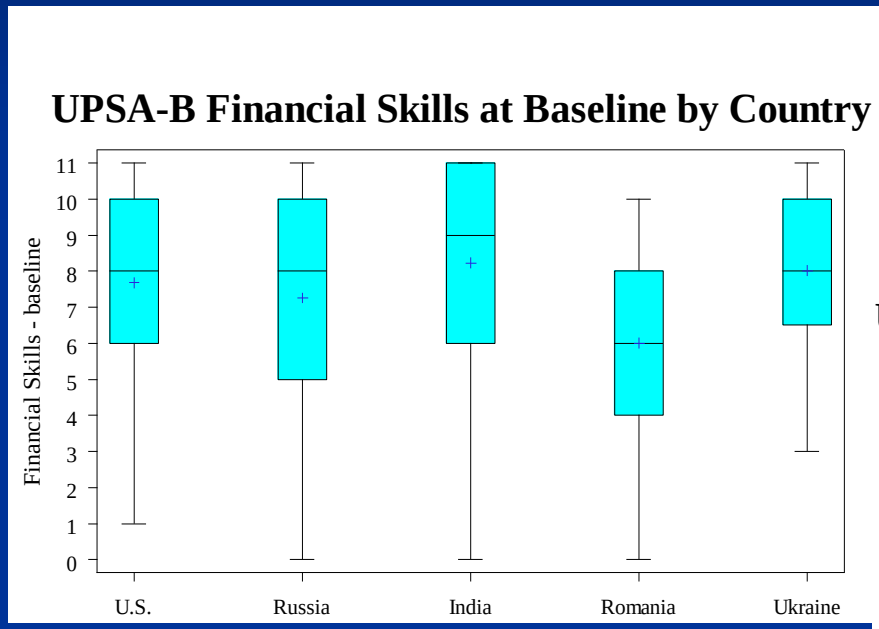
Challenges of International Translation and Implementation

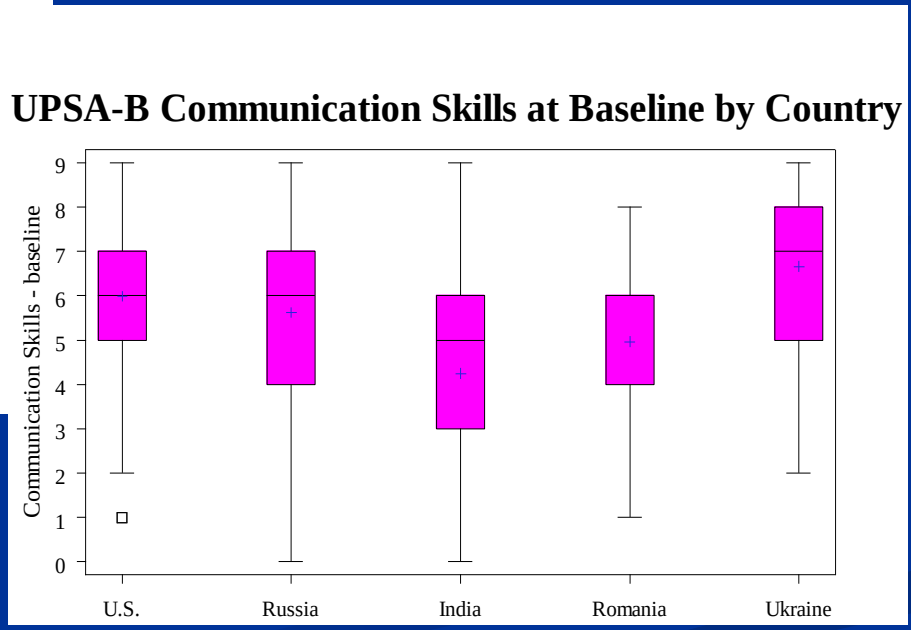
- Irrelevant activities outside US
  - Many countries do not use checks
  - Several countries don't send utility bills to customers and many natives have never read one
    - In Mexico, residents get "payment due" slips
    - In Romania, payment is automatically drafted from bank account
    - In India, utilities are sometimes obtained illegally, therefore no payment is made
  - Some countries don't have a universal emergency number
    - In some Indian regions, calling one's mother is appropriate

Challenges of International Translation and Implementation

- In Mexico, directory assistance is only for businesses
- Rescheduling a doctor's appointment is unheard of in some regions of India
- Insurance cards are not issued in Israel
- Patients sometimes reluctant to perform unfamiliar tasks
  - Raters reported that, for some patients, even if they are able to deduce what should be done, they are reluctant to do so if it's not something they do in their everyday lives
  - CRAs reported that some raters appear not to encourage patients to try since they feel like those activities are not relevant to the patient's everyday life

UPSA-B subscale scores by country





U.S n=149
Russia n=87
India n=97
Romania n=49
Ukraine n=76

Summary

- VIM results indicate: UPSA; UPSA-Brief, Brief- TABS have the best psychometric properties for full and brief measures respectively

- CIM results indicate: The CAI was rated most adaptable to other cultural contexts. TABS and ILS were rated highest for performance-based measures. Subscale analysis indicated that there were at least some subtests on all scales that were rated poorly. China, India, and Mexico are countries where the scales were rated lowest on adaptability. Across countries, low SES and rural groups pose the biggest problems in adaptation.

- DSP results indicate: Use of UPSA-B in ex-US countries presented challenges due to activities being irrelevant, regional differences, and lack of satisfaction with final translations/adaptations

Conclusions

- China, India, and Mexico are countries where the scales were rated lowest on adaptability.
- These countries may need specialized adaptations for use in clinical trials
- Trials could be planned in countries where adaptation is easier
- Psychometric properties of a hybrid measure using the UPSA comprehension and household management and the Work TABS will be examined in India