Cognition Working Group: Case Study

Tenth Annual
Patient-Reported Outcome Consortium Workshop

April 24 – 25, 2019 ■ Silver Spring, MD
Disclaimer

• The views and opinions expressed in the following slides are those of the individual presenters and should not be attributed to their respective organizations/companies, the U.S. Food and Drug Administration or the Critical Path Institute.

• These slides are the intellectual property of the individual presenters and are protected under the copyright laws of the United States of America and other countries. Used by permission. All rights reserved. All trademarks are the property of their respective owners.
Session Outline

• Cognition Working Group Overview
• Alzheimer’s Disease (AD): Defining the Disease and Clinical Staging
• University of California San Diego Performance-Based Skills Assessment (UPSA)
• Derivation and Validation of UPSA Short Forms for Persons with Mild Cognitive Impairment (MCI)
• Cognition Working Group: Next Steps
• Panel Discussion
• Question and Answer
Session Participants

Moderator
– Katy L. Benjamin, PhD - Director, HEOR, Patient Reported Outcomes, AbbVie Inc.

Presenters
– Richard Keefe, PhD – CEO, VeraSci, Inc.; Professor of Psychiatry, Psychology and Neurosciences, Duke University Medical Center
– Philip D. Harvey, PhD – Leonard M. Miller Professor, University of Miami Miller School of Medicine
– Katy L. Benjamin, PhD - Director, HEOR, Patient Reported Outcomes, AbbVie Inc.

Panelist
– Billy Dunn, MD – Director, Division of Neurology Products, OND, CDER, FDA
### 2010: Working group initiated

- **Goals:**
  - Develop PRO measure for mild cognitive impairment (MCI) due to AD
  - Capture patient’s perspective on functioning: instrumental activities of daily living (IADLs), interpersonal functioning

### 2010 – 2013: PRO instrument development

- Qualitative studies led to the development of the Interpersonal Function and Daily Activities Questionnaire (IFDAQ)
- Qualitative Research Summary Document submitted to FDA

### 2014: FDA concerns

- Low probability of successful PRO measure qualification
- Ability of patients with MCI due to AD to maintain sufficient cognitive insight to validly and reliably self-report

### 2015: Decision to pursue qualification of a performance outcome (PerfO) assessment tool

- **Goal:** To capture patient’s performance on aspects of day-to-day functioning (i.e., IADLs) expected to be impacted at predementia stages of AD
- Identified University of California San Diego Performance-based Skills Assessment (UPSA) as the candidate measure
  - Adaptable to MCI population
  - Short completion time
  - Substantial evidence of task reliability and validity in schizophrenia population
2018 National Institute on Aging—Alzheimer’s Association (NIA-AA) Research Framework*

• An aggregate of neuropathologic changes
• Defined \textit{in vivo} by biomarkers (amyloid \(\beta\) [A\(\beta\)] deposits and pathologic tau proteins) and by postmortem examination
• Not defined by clinical symptoms
• Defined by abnormal biomarkers in living people even if no cognitive or functional impairment is detected.

*Jack et al. NIA-AA research framework: toward a biological definition of Alzheimer’s disease. \textit{Alzheimer’s \\& Dementia} 2018;14:535-562
2018 NIA-AA Research Framework*

- For individuals biologically defined as having AD, the clinical staging along the AD continuum is as follows:
  - Stage 1 – asymptomatic/normal performance
  - Stage 2 – transitional; subtle decline in cognitive function
  - Stage 3 – mild cognitive impairment
  - Stage 4 – mild dementia
  - Stage 5 – moderate dementia
  - Stage 6 – severe dementia

University of California San Diego
Performance-Based Skills Assessment (UPSA)

Richard Keefe, PhD – CEO, VeraSci, Inc.; Professor of Psychiatry, Psychology and Neurosciences, Duke University Medical Center
• Performance-based measure of day-to-day living skills needed to function independently in a community setting

• Uses structured role-play scenarios that prompt participants to perform day-to-day tasks using props in controlled settings that mimic real world situations.

• The subscales of functioning assessed by the UPSA include recreational planning, finance, communication, transportation, household chores, and medication management.

• Administered and scored by a trained technician.
• Reduced functional capacity due to cognitive impairment
• The total score, and most of the subscale scores, have consistently differentiated schizophrenia-spectrum patients from healthy comparison subjects, indicating reduced functional capacity in schizophrenia (Heinrichs et al., 2006; Patterson et al., 2001).

• Used in several other psychiatric populations:
  • major depressive disorder (MDD) (McIntosh et al., 2011; Harvey et al., 2017)
  • bipolar disorder (Allen et al., 2015; Henry et al., 2013; Mausbach et al., 2010; McIntosh et al., 2011)
  • posttraumatic stress disorder (PTSD) (Kaye et al., 2014)
  • personality disorders (McClure et al., 2013)
  • substance abuse (Henry et al., 2010).
  • mild cognitive impairment (Goldberg et al., 2010; Gomar et al., 2011) and Alzheimer's disease (Goldberg et al., 2010; Gomar et al., 2011)
Demonstrated Validity of the UPSA

- UPSA demonstrates strong external (aka ecological) validity across many studies of schizophrenia and other serious mental illnesses.

- For example, the UPSA has been shown to predict real-world functioning such as failures to achieve milestones in vocational, residential, and social subscales, or ratings of functioning provided by high quality informants (Leifker et al., 2010; Mausbach et al., 2011, 2010).

- Further, the correlation between UPSA performance and cognitive tests has been consistent and substantial, typically $r = .60$ or greater (Allen et al., 2015, Bowie et al., 2006, 2010; Green et al., 2008; Harvey et al., 2009; Keefe et al., 2011).
Nine (!) distinct versions of the UPSA have been used in published studies

The original measure, the UPSA-1 (Patterson et al., 2001), assesses 5 functional subscales.

- **Communication Skills:** subjects are asked to perform tasks with a telephone (dial an emergency number, call information to get a specific phone number) and to read a letter confirming a medical appointment and demonstrate how they would prepare for the appointment.

- **Household Skills:** subjects are provided with a recipe for rice pudding, asked to check an array of items in a mock kitchen pantry, and prepare a shopping list.

- **Transportation:** subjects use a bus schedule to plan a trip to a specific location.

- **Financial Skills:** measures subjects’ ability to make change and fill out a check to pay a utility bill.

- **Comprehension/Planning:** subjects read stories that depict outings to the beach and the zoo, and then answer questions to evaluate their comprehension.
## Subscales Included in UPSA Versions

<table>
<thead>
<tr>
<th>Description of Instrument</th>
<th>SKILLS ASSESSED</th>
<th>UPSA-1</th>
<th>UPSA-2</th>
<th>UPSA SHORT FORM</th>
<th>UPSA-“LONG FORM”</th>
<th>UPSA-2-VIM</th>
<th>UPSA-B International</th>
<th>UPSA-2ER International</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONCEPT COVERAGE BASED ON IADLS (REPORTED IN THE LITERATURE)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication Skills</td>
<td>“Dialing telephone,” “Reschedule appt.,” and “Memory questions”</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Household Skills</td>
<td>“Write shopping list” and for UPSA-2ER includes meal preparation</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Transportation</td>
<td>“Read bus map” or “Read bus map &amp; schedule”</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Medication Management (MMAA)</td>
<td>“Develop and follow medication regimen”</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial Skills</td>
<td>“Counting change,” and either “Writing check” or “Understanding bill”</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>COVERAGE OF CONCEPTS NOT DEFINED AS IADLS IN THE LITERATURE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension/Planning</td>
<td>“Planning an outing”</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
</tbody>
</table>
UPSA-2ER Kit Materials

Kit Contents

1. UPSA-2ER Administration Manual
2. Big button telephone
3. Currency
4. Set of 3 Bus Maps
5. Set of 4 pill bottles and pill organizer
6. Timer

Additionally, a set of pantry items is needed for the Household Skills subscale.
General Psychometric Characteristics

• **Test-retest reliability**: Test-retest reliability is generally quite strong. Published studies and findings from our cohort show ICCs > 0.70.

• **Practice effect**: Small practice effects were found across studies, with the magnitude being somewhat larger in our cohort (ds > .30) than in published studies (ds < .23).

• **Floor-ceiling effects**: Not substantial for schizophrenia patients.
Distribution Data From Four Schizophrenia Studies Using the UPSA-2-VIM

- Skewness Statistic = -0.819
- At Baseline, two subjects had a score of 100 (highest possible score; both subjects from the control group) and 18 subjects had scores approaching ceiling (i.e., scores on the UPSA-2-VIM Composite ≥ 90) – two subjects from the schizophrenia group and 16 subjects from the control group.)
Consensus Development Meetings for Derivation of the UPSA-MCI

VeraSci coordinated two meetings in early 2018 with Advisory Panel members, the Cognition Working Group members, and C-Path staff to address key questions necessary for the derivation of the **UPSA for Mild Cognitive Impairment (UPSA-MCI)**.

**Advisory Panel Members**

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution and Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terry E. Goldberg, PhD</td>
<td>Professor of Medical Psychology specializing in the study of aging and dementia, Columbia University</td>
</tr>
<tr>
<td>Philip D. Harvey, PhD</td>
<td>Professor of Psychiatry and Behavioral Sciences and Chief of the Division of Psychology, University of Miami Miller School of Medicine</td>
</tr>
<tr>
<td>Thomas L. Patterson, PhD</td>
<td>Professor of Psychiatry and developer of the UCSD Performance-Based Skills Assessment (UPSA), University of California San Diego</td>
</tr>
<tr>
<td>Pierre M. Tariot, MD</td>
<td>Director, Banner Alzheimer’s Institute; Research Professor, University of Arizona College of Medicine</td>
</tr>
<tr>
<td>Kathleen Welsh-Bohmer, PhD</td>
<td>Professor of Psychiatry and Neurology, Duke University; Director of Alzheimer’s Disease Clinical Trials, Duke Clinical Research Institute; VP of Neurodegenerative Disorders, VeraSci</td>
</tr>
</tbody>
</table>
Key Questions for the Consensus Meetings

1. Which subscales are the most important to include?
2. Should alternate forms be used?
3. Should informant report be used?
4. Should the tasks within the UPSA be updated to be more reflective of contemporary life?
5. What versions of the UPSA have been used in MCI to date? What subscales were included? Did some subscales provide more valuable information than others?
Which Subscales are Important to Include?

• **Communication difficulty** is an important component of the MCI profile that differentiates individuals with MCI from the cognitively healthy older adult population and may therefore be a key target for intervention efforts designed to improve multiple subscales of well-being in individuals with MCI (Johnson et al., 2014).

• Individuals with MCI demonstrate difficulty with tasks associated with **day-to-day planning and organization**. This difficulty is related to the deficits in memory as well as abilities to perform complex tasks.

• All Advisory Panel members agreed that **handling financial information** is important and relevant in an MCI population, which demonstrates performance deficits in measures of financial capacity assessed in the clinic as well as impairment in more complex real-world activities (Marson et al., 2000).
<table>
<thead>
<tr>
<th>UPSA-2-VIM subscales</th>
<th>MCCB subscales</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speed of Processing</td>
</tr>
<tr>
<td>Financial Skills</td>
<td>0.49</td>
</tr>
<tr>
<td>Communication Skills</td>
<td>0.52</td>
</tr>
<tr>
<td>Comprehension/Planning</td>
<td>0.50</td>
</tr>
</tbody>
</table>

MCCB = MATRICS Consensus Cognitive Battery
Which Subscales are Important to Include?

Based on the psychometric properties, face validity and impairments in MCI, the advisory panel members and Cognition Working Group agreed that the following UPSA subscales should be included in the UPSA-MCI:

• Financial Skills
• Communication Skills
• Comprehension/Planning
Should Alternate Forms Be Used?

• In earlier work during the development of the UPSA completed by Dr. Patterson, the use of alternate forms was considered.

• Variability greater than practice/learning effects.

• The added variability between the versions was viewed as a larger risk than the potential benefit of reducing the learning effects found when using one version in an MCI population.

• The Advisory Panel and members of the consortium determined that alternate versions would not be used for this population as the practice effects were minimal and the ceiling effects were nonexistent.
Should Informant Report be Used?

• Informant reports in an MCI population can create psychometric challenges.
• Data generated from informants are often skewed because once a caregiver discovers that the person with MCI has cognitive impairment, the caregiver may prohibit the patient from engaging in the activities that the patient would normally do.
• The informant-based assessment may not be an objective indicator for how well a person is able to perform real-world tasks.
• Performance-based measures are designed to be administered without the use of an informant.
• UPSA was correlated with informant questionnaires at approximately 0.6, suggesting that these measures are complementary but with significant non-overlapping variance.
• The Advisory Panel and members of the working group determined that the UPSA-MCI would not include information collected from an informant, and would solely be a performance-based measure.
Should the UPSA Tasks Be Modernized?

• Existing versions of the UPSA include tasks such as writing a check to pay a utility bill, calling directory assistance to obtain a telephone number, and recalling a telephone number from memory.

• With advances in technology, these tasks are being replaced in the real world with activities like online banking, internet search engines to find telephone numbers, and smart phone storage processes to store telephone numbers for future use.

• Should the tasks included be updated to reflect these aspects of contemporary life?

• Small study by Dr. Harvey of patients with MCI and healthy controls over the age of 75 suggested that most have minimal exposure to ATM use and never using online banking.

• While updated items for the UPSA-MCI may need to be employed for the next generation of older individuals, the collective opinion of the expert panel and members of the consortium was that updating the tasks at this time would not be the most appropriate course of action, and could be counter-productive.
Subscales of the UPSA-MCI

Financial Skills
• Count money
• Make change
• Demonstrate understanding of information included in a sample utility bill
Subscales of the UPSA-MCI

Communication Skills

• Telephone: Demonstrate how to use a telephone by role-playing a number of scenarios (e.g., what number to dial in case of emergency and dialing a number from memory)

• Medical Appointment Letter: Demonstrate how to schedule a medical appointment based on information from a letter, then role-playing a number of scenarios (e.g., rescheduling the appointment, describing the letter’s instructions for preparing for the appointment and the items to bring)
Subscales of the UPSA-MCI

Comprehension/Planning

• Demonstrate comprehension of a newspaper article about the opening of a new city park (comprehension)

• List seven items necessary to bring or wear in order to spend the day at the park (planning)
## UPSA-MCI Summary Scoring Worksheet

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Skills</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>Communication Skills</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>Comprehension/Planning</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
</tbody>
</table>

**UPSA-MCI Total Score**
(Range = 0-60) ________
Importance of Training

Adequate training procedures are key to ensuring consistency of administration in any paper-based assessment

• During training, raters:
  o are expected to administer the assessment to an UPSA certifier in the same way that he/she would administer it to a subject in a clinical trial
  o prompt when necessary

• Training helps to ensure that raters:
  o administer the assessment according to the administration manual
  o know how/when to provide prompting and corrective feedback when appropriate
  o understand the importance of using solid judgement when scoring certain items for a performance-based assessment (i.e., scoring variations in a company name for Financial Skills)
Up Next! Dr. Phil Harvey

• What versions of the UPSA have been used in MCI to date? What subscales were included?
• Did some subscales provide more valuable information than others?
Derivation and Validation of UPSA Short Forms for Persons with MCI

Philip D. Harvey, PhD
Leonard M. Miller Professor
University of Miami Miller School of Medicine
Pharvey@med.Miami.edu
Disclosures

• Dr. Harvey has grants from:
  – Takeda and the Stanley Medical Research Foundation

• Served as a consultant (past year) for:
  – Alkermes, Boehringer Ingelheim, Jazz Pharma, Lundbeck, Otsuka America, Roche, Sanofi (Regeneron), Sunovion (Sumitomo Dainippon), Takeda, and Teva Pharma
Goals of the Presentation

• Uses of the UPSA in the Alzheimer’s disease spectrum
• Selection factors for critical subscales
• Psychometric characteristics
• Longitudinal stability vs. sensitivity to change
First use of the UPSA in Dementia Spectrum

- Comparison of clinically-defined Alzheimer’s disease (AD), mild-cognitive impairment (MCI), and healthy controls (HC) cross-sectionally on an UPSA short form
- UPSA subscales
  - Communication Skills
  - Financial Skills
  - Comprehension/Planning
  - Transportation

Validators

• Neuropsychological (NP) performance
  – Speed; Episodic memory; Semantic fluency; and Clock drawing

• Everyday Functioning
  – Alzheimer’s Disease Cooperative Study-Activities of Daily Living (ADCS-ADL)
Participant Populations

- AD MMSE 13-23
  - N=22
- MCI MMSE 24+
  - N=26
- HC MMSE 24+, no other impairments
  - N=50

MMSE = Mini-Mental State Examination
UPSA and ADCS-ADL Scores

Comparative Functional Performance

A = A > B

A > B > C

HC, MCI, AD

UPSA, HC, MCI, AD

A > B > C
Correlations of NP Performance and Functioning

• After adjusting for age and education, Trail Making Part A, Logical Memory Immediate, and Semantic Fluency accounted for 51% of the variance in the UPSA.
• Thus, the UPSA is sensitive to impairments across groups and correlated with NP deficits.
Selection of UPSA Subscales

• Re-analysis of Goldberg et al.
• Used stepwise regression to enter UPSA subscales into a regression predicting total UPSA scores
• Communication Skills and Comprehension/Planning account for 90% of the total UPSA score.

Separation of Groups with UPSA Short Form

Note: All groups differ from each other. Correlation of long form and short form p<.001 in all three groups. Overall shared variance: .89
Unanswered Questions

• Sensitivity to decline
• Test-retest stability in non-declining groups
• Importance of financial skills in the prediction of MCI
  – Many studies suggest financial changes are early and prominent
Current Study with 3-Subscale UPSA
(Financial Skills, Communication Skills, and Comprehension/Planning)

• HC
  – N=43; MMSE Mean=29
• MCI
  – N=20; MMSE Mean=26
• AD
  – N=26; MMSE Mean=20

Goldberg TE, Harvey PD, Devanand DP, Keefe RS, Gomar JJ, Submitted
Methods

- 3-Subscale UPSA
- Assessments at baseline, 6 weeks, and 12 months
- Analyses
  - whole groups
  - comparison of participants with 12-month MMSE decline, defined as any change in MMSE from BL to month 12
Decline from BL to month 12 significant for MCI and AD; 3-group x 3 time interaction p=.11
Practice Effects

Baseline Vs. Week 6
- HS
- MCI
- AD

Baseline Vs. Year 1
- HS
- MCI
- AD
UPSA Declines in Participants with Declining MMSE
Other Psychometrics

• For the total group test retest reliability was $r=.90$.

• Ceiling and floor effects:
  – Ceiling
    • Baseline HC:1, MCI 1, AD 1
    • Month 12 HC:0, MCI 0, AD 0
  – Floor
    • Baseline HC:1, MCI 1, AD 2
    • Month 12 HC:1, MCI 0, AD 0
Conclusions

• The 3-subscale UPSA is stable in HC, sensitive to decline in MCI and AD
• Generally sensitive to decline in identified declining cases
• Practice effects seen at 6 weeks, absent at third assessment at 12 months
• Test-retest reliability substantial
Cognition Working Group: Next Steps

Katy L. Benjamin, PhD - Director, HEOR, Patient Reported Outcomes, AbbVie Inc.
Next Steps

• Submissions to FDA
• Pilot study to include qualitative and quantitative components
Submissions to FDA

• Initial Briefing Package
  • Summarizes the work done to date:
    • To identify the UPSA subscales most relevant to Stage 2/3 AD
    • To document the changes made to create the UPSA-MCI

• Qualification Plan
  • Proposed study protocol and analysis plan to generate evidence of validity and other measurement properties to support qualification of the UPSA-MCI
  • Proposed study: pilot study to include qualitative and quantitative components
Pilot Study: Qualitative Phase (1 of 2)

• Modified UPSA for use in Stage 2 – 3 AD patients
  • Minor modifications to the instructions and tasks
  • Now called UPSA-MCI

• Collect additional qualitative evidence to confirm that the UPSA-MCI is relevant and applicable to the target population
  • Study sample will need to include a sufficient number of the following subgroups to confirm that the tasks are understood and can be performed by participants with and without AD
    • Asymptomatic Stage 1 (biomarker positive)
    • Stage 2
    • Stage 3
    • Age-matched asymptomatic healthy controls (biomarker negative)

• Anticipated to be small sample size (n=40) from two to three sites
Pilot Study: Qualitative Phase (2 of 2)

• Study design: pilot-test of the UPSA-MCI followed by cognitive interviews with study participants
  • Cognitive interview study to:
    • Confirm that tasks are related to everyday life and clinically meaningful to participants
    • Explore if change in these tasks reflects difficulty with actual task performance in everyday life
    • Conduct preliminary exploration of clinically meaningful change with participants

• The tester/administrator will be debriefed
  • Provide his/her perception of how well the patient understood the tasks and performance instructions
  • Feedback on experience administering and scoring the measure
    • Adequacy of training, clarity of instructions for administering task, duration of the tasks
    • Clarity of criteria for scoring
Quantitative evidence in a larger study sample with the same four subgroups (n=TBD) needed to evaluate cross-sectional measurement properties:

- Construct validity
  - Convergent validity
  - Known-groups validity
- Test-retest reliability (intra-rater): half the sample tested by same tester
- Inter-rater reliability
  - Evaluate within test-retest design by half the sample tested by a different person at retest timepoint
- Internal consistency reliability
- Scoring
Study Design Challenges

• Eligibility requires biological confirmation of AD and clinical staging
• Will need to find sites that collect this information routinely so it is not something that needs to be performed for the observational study
• Sample size for quantitative phase: based on number of testers or number of participants/tester?
• Need to determine appropriate measures for convergent validity assessment
  • Cognitive test (e.g., Symbol Digit Modalities Test)
  • Informant assessment (e.g., Alzheimer’s Disease Cooperative Study-Activities of Daily Living)
    • Clinician assessment (e.g., Mini-Mental State Examination)
• Time interval between test and retest to avoid practice effect for participants (e.g., is one month appropriate?)
Panel Discussion and Q&A

Moderator
- Katy L. Benjamin, PhD - Director, HEOR, Patient Reported Outcomes, AbbVie Inc.

Presenters
- Richard Keefe, PhD – CEO, VeraSci, Inc.; Professor of Psychiatry, Psychology and Neurosciences, Duke University Medical Center
- Philip D. Harvey, PhD – Leonard M. Miller Professor, University of Miami Miller School of Medicine
- Katy L. Benjamin, PhD - Director, HEOR, Patient Reported Outcomes, AbbVie Inc.

Panelist
- Billy Dunn, MD – Director, Division of Neurology Products, OND, CDER, FDA