PMH46

Qualitative Development and Cognitive Evaluation of the Symptoms of Major Depressive Disorder Scale (SMDDS): A Consortium-Developed Patient-Reported Outcome Measure

McCarrier KM[1], Martin ML[1], Bushnell DM[1], Abraham L[2], Bush EN[3], Houle CR[4], Mathews M[5], Carpenter LL[6], Blum SI[5] on behalf of the PRO Consortium Depression Working Group

[1]Health Research Associates, Inc., Mountlake Terrace, WA, USA; [2]Pfizer Ltd, Tadworth, Surrey, UK; [3]Eli Lilly and Co., Indianapolis, IN, USA; [4]AbbVie Inc., North Chicago, IL, USA; [5]Forest Research Institute, Jersey City, NJ, USA; [6]Butler Hospital/Brown University, Providence, RI, USA

INTRODUCTION:

- Major Depressive Disorder (MDD) is a severe mental health disorder affecting 16.9% of the U.S. adult population, nearly 340 million people worldwide, and is a leading cause of disability, with disproportionate impact on women.[1]
- Because depression is primarily experienced subjectively, and the severity of MDD symptoms is directly related to the degree of impairment that patients experience, the assessment of depressive symptoms is an essential endpoint for clinical studies, particularly where the use of clinical indicators will be limited.
- By exploring the patient experience with MDD through qualitative interviews, it is possible to better understand and document the specific depression-related concepts that are relevant to the patient as well as understand the patient's assessment of improvement in his or her condition.[2,3]
  - Ultimately, a well-developed instrument that has firmly established content validity (supported by qualitative data from patients) will be expected to demonstrate greater sensitivity in clinical studies of treatment benefit.
- Prior to conducting the qualitative interviews a systematic review was conducted to evaluate existing depression instruments[4] as well as previously published qualitative data[5]
  - The systematic review of qualitative data helped to inform on the development of the interview guides.
  - The systematic review of existing instruments helped to assess their content coverage, measurement properties and determine the extent to which existing measures were developed with direct input from patients. Suitable existing instruments could provide the basis for qualification or modification

OBJECTIVES:

- Complete qualitative concept elicitation and cognitive interviews with subjects diagnosed with MDD to support preliminary development of a patient-reported outcome (PRO) measure to assess treatment benefit in MDD clinical trials.

METHODS:

Study Population

- Recruitment was designed to enroll a diverse sample of patients similar to those who would be using the PRO instrument in future clinical trials of MDD treatments.
  - No formal recruitment quotas were employed, each site targeted recruitment of a mix of patients with varying severity of MDD and MDD-treatment histories, as well as broad representation across demographic characteristics such as age, sex, race/ethnicity, marital status, and educational attainment and employment status.
  - Subjects were recruited from 6 U.S. clinical sites (CT, FL, IL, NY, OK, WA)
- The eligibility criteria for the targeted interview population were designed to reflect common entry criteria for clinical trials in major depression:
  - Inclusion Criteria: Male and Female subjects between the ages of 18 to 65, inclusive, who met DSM-IV-TR criteria[6] for MDD; and were being treated on an outpatient basis; and had experienced a major depressive episode within the previous 6 months; and had a Hamilton Rating Scale for Depression (HAM-D) score of > 18 at the time of screening
  - Exclusion Criteria: Current or past history of a personality disorder, schizophrenia or other psychotic disorder, obsessive compulsive disorder, or post-traumatic stress disorder; significant risk of suicide; positive urine drug screen or recent clinically significant alcohol abuse or drug use.

Concept Elicitation (CE) Interviews

- Semi-structured qualitative interviews[2] were conducted by trained research staff with a representative sample of adult MDD patients in the US who recently experienced a major depressive event.
  - Interviews followed a pre-approved interview guide and used open-ended questions and day-reconstruction exercises to elicit spontaneous reports of symptom/impact concepts.
  - Subsequent probing was used to assess concepts not spontaneously reported by subjects.
  - Subjects were asked to rate the severity and level of bother or difficulty for reported symptoms and impacts
  - To guide item development, subjects were asked about appropriateness of measuring the severity, frequency, or duration of each concept

Content Analysis

- All interview sessions were audio recorded and transcribed.
- The concept elicitation interview transcripts were coded and analyzed by trained qualitative coders using Atlas.ti, and were summarized by like-content using an iterative coding framework.
  - Coded concepts were grouped by similarity of content and analyzed to identify the most relevant expressions and most common language used by patients.
- A Saturation Grid was used to track symptoms and impacts expressed during the interviews and assess saturation of concept.
  - Transcripts were ordered chronologically in groups of 8 transcripts. Codes from each group were compared with previous groups to determine whether any new additional unique concepts emerged.

Item Generation

- An item-generation meeting was held by the development team, where concepts identified from published literature[5], existing instruments[4], and the qualitative data from the CE interviews were reviewed as the basis for selection of concepts for inclusion in PRO measurement.
  - This initial evaluation process resulted in the selection of candidate symptom concepts to be targeted for PRO measurement.
  - During subsequent review by the team, these targeted concepts were further reduced by removing synonymous or problematic concepts, and a draft version of the questionnaire was prepared for evaluation in cognitive interviews and a translatability assessment.[7]

Cognitive Interviews

- Cognitive interviews were conducted to evaluate concept relevance, understandability, and structure of the draft items, and to facilitate further instrument refinement.[3]
- Three separate waves of interviews with patients with MDD were conducted
  - Following each wave, the development team considered the findings and used the information to modify the draft instrument
- A semi-structured cognitive interview guide was designed to capture the subject's comprehension of items and ability to complete the draft PRO instrument.
  - Updated versions of the interview guide were created for each of the three interview waves
  - Questions in the interview process asked about: the comprehension and relevance of the individual items; the fit of the response scales; the appropriateness of the recall period and item wording; and any lack of clarity of items, terminology, instructions, or sentence structure.
- During the cognitive interviews, the draft instrument items were completed and evaluated by patients with MDD, recruited through the same process and eligibility criteria as used previously for the CE interviews.
- Cognitive interview transcripts were summarized in cognitive report tables for analysis.
- In parallel with the cognitive interview process, a translatability assessment (TA) was conducted on the draft instrument to assess the potential for difficulty in translating the items to maintain content equivalency[7]
  - The findings from the TA process were used to make revisions to select PRO items prior to the closure of the cognitive interview process.

RESULTS:

Table 1: ***Demographic Characteristics***

Characteristic	Concept Elicitation N=40	Cognitive Interviews N=15
Age in years: mean (SD); [range]	46.2 (11.8); [21-63]	44.6 (13.4); [18-59]
Gender: Female: n (%)	27 (67.5%)	9 (60.0%)
Racial and Ethnic Group: n (%)		
White (Non-Hispanic):	18 (45.0%)	11 (73.3%)
White (Hispanic):	9 (22.5%)	1 (6.7%)
White (Ethnicity not reported):	1 (2.5%)	---
Black/African American:	9 (22.5%)	2 (13.3%)
Other:	3 (7.5%)	1 (6.7%)
Highest Level of Education Completed: n (%)		
High School	9 (22.5%)	7 (46.7%)
Some College	17 (42.5%)	5 (33.3%)
Bachelor's Degree	7 (17.5%)	---
Graduate or Professional School	7 (17.5%)	2 (13.3%)
Clinical Characteristics		
Years since diagnosis with MDD: mean (SD); [range]	7.8 (8.7); [0-40]	12.3 (12.0); [0.9-42.8]
Years since onset of most recent MDE: mean (SD); [range]	1.0 (1.8); [0-8]	1.9 (1.5); [0.5-4.8]
HAM-D Total Score at Screening: mean (SD); [range]	24.4 (4.3); [19-39]	24.4 (5.3); [19-36]

Concept Elicitation

- A total of 40 subjects participated in the CE interviews. They were an average of 46.2 years old, were 67.5% female, 45.0% white (non-Hispanic), and had an average HAM-D total score of 24.4 at enrollment (Table 1).
- Analysis of the transcripts resulted in 3022 symptom expressions and 830 impact expressions
- Expressions were coded and grouped into 105 concepts (91 symptom and 14 impact) in 15 hypothesized domains (11 symptom and 4 impact)
- Saturation of concept (the point at which no new concepts were elicited) was achieved after the fourth of five transcript groups (eight transcripts per group) (Table 2).
- Inter-rater reliability was assessed in five transcript pairs, and was observed to be high with 85.4 to 92.1% agreement between raters for the identification of symptom concepts being expressed in the transcripts, and 97.5 to 99.1% agreement between raters on code assignment for identified concepts.

Item Generation

- The item generation evaluation process resulted in the selection of candidate symptom concepts to be targeted for PRO measurement.
  - Predominance of symptom mentions as well as whether such mentions were spontaneous or probed and the relative severity and bother of the symptoms/impacts provided a context for evaluating individual concepts
  - The development team agreed to focus on symptoms and not disease impacts for the measure
- Because no existing PRO comprehensively assessed the selected concepts[4], the development team decided to develop a new measure, rather than attempting to either qualify or modify an existing measure
- During subsequent review by the development team, the targeted concepts were further reduced by removing synonymous or problematic concepts, and a 36-item draft questionnaire was prepared for evaluation in cognitive interviews and the translatability assessment[7]

Table 2: Saturation Grid of Disease-Relevant Concepts

Domain	# of New Concepts Identified Per Domain — Transcript Group (8 transcripts/group)				
	1	2	3	4	5
Physical Symptoms	10	2	---	---	---
Energy	6	---	1	---	---
Motivation	8	---	---	---	---
Emotions/Mood	10	2	---	---	---
Negative Affect	6	---	---	---	---
Cognition	11	---	---	1	---
Sleep Disturbances	5	1	---	---	---
Sense of Self	5	---	---	---	---
Self-Harm/Suicide	3	---	---	1	---
Eating Behaviors	6	---	---	---	---
Anxiety	6	---	---	---	---
Social/Relationship	5	---	---	---	---
Aspects of Burden	3	---	1	---	---
Difficulty with Activities	7	---	---	---	---
Coping Strategies	5	---	---	---	---
Total new concepts per transcript group (n/105)	96 (91.4%)	5 (4.8%)	2 (1.9%)	2 (1.9%)	0 (0.0%)

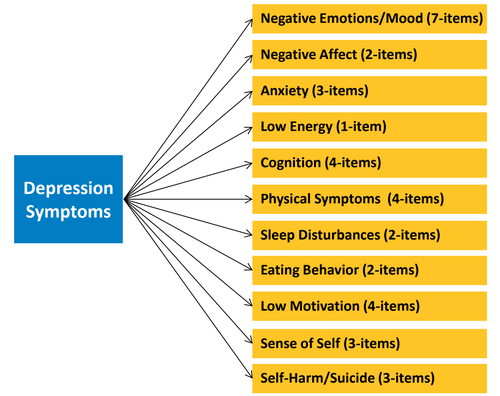
Cognitive Interviews (CIs)

- A total of 15 subjects participated in three waves of CIs. The subjects were an average of 44.6 years old, were 60.0% female, 73.3% white (non-Hispanic), and had an average HAM-D total score of 24.4 at enrollment (Table 1).
- Over the three waves, one item was removed and four others were substantially modified based on cognitive interview findings and recommendations from a formal translatability assessment.[7]
- Other minor instrument formatting and wording modifications were made based on the results of a formal migratability assessment for electronic PRO data collection platforms (ePRO).

Symptoms of Major Depressive Disorder Scale (SMDDS)

- The newly-created scale, the Symptoms of Major Depressive Disorder Scale (SMDDS), is a 35- item instrument that measures each concept using a 5-point verbal rating scale and a 7-day retrospective recall period for each of the items.
  - Items in the SMDDS are hypothesized to be organized into 11 sub-domains (Figure 1)
  - Sixteen of the items focus on the intensity of symptoms with a rating scale from "not at all" to "extremely."
  - Nineteen items focus on frequency or the amount of time a symptom was experienced and employ a rating scale from "never" to "always."

Figure 1: ***Proposed Conceptual Framework***



CONCLUSIONS:

- The SMDDS is a 35-item PRO measure intended for use as an endpoint in MDD clinical trials to support medical product labeling
- The SMDDS was developed in accordance with the FDAs PRO Guidance and best practices.
  - Qualitative interviews have provided evidence for content validity.
  - Cognitive interviews provided evidence that the instructions, items and response options are comprehensible, easy to complete and address key symptoms of MDD that are relevant to patients with the condition.
- Future quantitative studies will confirm the measurement properties of the SMDDS and support FDA qualification.

FINANCIAL DISCLOSURES:

- Funding for this research was provided by the following PRO Consortium member firms: AbbVie Inc; Bristol-Myers Squibb; Eli Lilly and Company; Forest Laboratories; Janssen; Pfizer; Shire, Sunovion Pharmaceuticals and Takeda Pharmaceuticals.
- Critical Path Institute's PRO Consortium is supported by grant No. U01FD003865 from the United States Food and Drug Administration and by Science Foundation Arizona under Grant No. SRG 0335-08.

Presented at:
International Society for Pharmacoeconomics and Outcomes Research
18th Annual International Meeting
May 18-22, 2013
New Orleans, LA, USA

REFERENCES:

1. Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, Wang P, Wells KB, Zaslavsky AM. Prevalence and Treatment of Mental Disorders, 1990 to 2003. N Engl J Med, 2005, 352: 2515-2523.
2. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molson E, Ring L. Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1—Eliciting Concepts for a New PRO Instrument. Value Health 2011; 14:967-977
3. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molson E, Ring L: Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2—Assessing Respondent Understanding. Value Health 2011; 14:978-988
4. Blum SI, Bush EN, Bushnell DM. Systematic Review of Patient-Reported Outcomes Measures Used to Assess Symptoms Associated with Major Depressive Disorder. J Ment Health Policy Econ, 2012, 15:S2-S3
5. Ball S, Dedios C, Abraham L. The Patient's Perspective on Major Depressive Disorder: What Do We Know? J Ment Health Policy Econ, 2012, 15:S1
6. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th Edition, text revised (DSM-IV-TR). 2000 Washington, DC
7. Martin ML, McCarrier KP, Basse S, Ramasamy A, Mathews M, Bush EN, Greco N. Using Translatability Assessment to Refine a Patient-Reported Outcome (PRO) Measure During the Development Process. ISPOR 18th Annual International Meeting, PMH48

PRO CONSORTIUM CRITICAL PATH INSTITUTE