

Validating the Measurement of Real-World Functional Outcomes: Phase I Results of the VALERO Study

Philip D. Harvey, Ph.D.

Tenko Raykov, Ph.D.

Elizabeth W. Twamley, Ph.D.

Lea Vella, M.A., M.P.H.

Robert K. Heaton, Ph.D.

Thomas L. Patterson, Ph.D.

Objective: Treatment of cognitive impairment has been proposed as an intervention to reduce disability in people with schizophrenia. The Validation of Everyday Real-World Outcomes (VALERO) study was conducted to evaluate functional rating scales and to identify the rating scale or scales most robustly related to performance-based measures of cognition and everyday living skills.

Method: Adults with schizophrenia (N=198) were tested with the neurocognitive measures from the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery, the UCSD Performance-Based Skills Assessment–Brief Version, and the advanced finances subscale of the Everyday Functioning Battery. They and an informant (a friend, relative, clinician, or case manager) also reported their everyday functioning on six rating scales. Best judgment ratings were generated by an interviewer who administered the rating scales to patients and informants.

Results: Statistical analyses developed an ability latent trait that reflected scores on the three performance-based (i.e., ability) measures, and canonical correlation analysis related interviewer ratings to the latent trait. The overall fit of the model with all six rating scales was good. Individual rating scales that did not improve the fit of the model were systematically deleted, and a final model with two rating scales was fitted to the data. A regression analysis found that the Specific Levels of Functioning Assessment was a superior predictor of the three performance-based ability measures.

Conclusions: Systematic assessments of real-world functioning were related to performance on neurocognitive and functional capacity measures. Of the six rating scales evaluated in this study, the Specific Levels of Functioning Assessment was the best. Use of a single rating scale provides an efficient assessment of real-world functioning that accounts for considerable variance in performance-based scores.

(*Am J Psychiatry* 2011; 168:1195–1201)

Disability in multiple domains of everyday functioning is common in people with schizophrenia (1), encompassing social (2), vocational (3), and residential (4) domains. These impairments are present in the majority of patients and have been refractory to pharmacological interventions (5) and only partially responsive to rehabilitation (6). Some of the patient characteristics associated with these impairments include cognitive deficits (7), negative symptoms (8), and impairments in performance of social and daily living skills (i.e., “functional capacity” [3]). There are also environmental contributions to real-world functional deficits, which include disability compensation (associated with reduced vocational achievement [9]) and lack of social opportunities and community resources. Disability is thus a complex and multiply determined phenomenon that may have similarities across different severe mental illnesses (10).

There have been considerable efforts to treat the putative causes of these functional deficits, including pharmacological interventions aimed at cognitive enhancement and reduction of negative symptoms (5), cognitive reme-

diation interventions (11), supported employment (6), and various psychosocial interventions, including efforts to teach social and everyday living skills (12). These efforts have met with variable success, with combined interventions using cognitive remediation and psychosocial interventions yielding the most success (13). However, several studies have found that impairments on a variety of ability measures, including both neuropsychological tests and measures of functional capacity, were weakly related to ratings of real-world functioning (14–16). These findings raise the legitimate question of whether improving cognitive or functional abilities has the potential to exert a meaningful influence on real-world outcomes.

As we have observed before (17), many studies finding low correlations between ability or capacity measures and ratings of real-world outcomes used patient self-reports of real-world functioning. As shown in several systematic comparisons, patient self-reports of both ability (e.g., cognitive performance) and real-world functioning are unrelated to performance on performance-based ability measures and to informant ratings of ability and real-world

functioning. As a result of these findings, the Validation of Everyday Real-World Outcomes (VALERO) initiative was undertaken to evaluate functional rating scales and to identify the rating scale or scales most robustly related to performance-based measures of cognition and everyday living skills. The present study in the VALERO initiative was designed to directly compare a set of informant- or patient-rated instruments measuring real-world functional outcomes with cognitive and functional capacity assessments in people with schizophrenia.

In this study, six rating scales for assessing real-world functional outcomes, recommended as the best available by a RAND expert panel, were completed by patients with schizophrenia, by an informant (friend, relative, or high-contact clinician), and by a research examiner who conducted the interviews with the patient and informant. Interviewers' ratings reflected their best estimate of the patient's level of functioning. Examiners were instructed to base all judgments on what they thought was correct, including discounting any information they believed was inaccurate. These research examiner ratings were then related to the patient's performance on the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB; 18) and two functional capacity measures, the UCSD Performance-Based Skills Assessment–Brief Version (UPSA-B; 19) and the advanced finances subscale of the Everyday Functioning Battery (20). The UPSA-B is widely used in studies of functional capacity in schizophrenia, but the Everyday Functioning Battery is aimed at higher-functioning individuals and was included in order to avoid the possibility of ceiling effects in the assessment of higher-functioning patients. The goal of this study was to identify the rating scale or scales that measure real-world functioning that are most strongly correlated with patients' performance on measures of their ability: cognitive performance and functional capacity.

Method

Participants

The study participants were patients with schizophrenia who were receiving treatment at one of three different outpatient services, two in Atlanta and one in San Diego. In addition, informants were interviewed about the everyday functioning of each of the patients; informants were a friend or relative (80% of cases) or a high-contact clinician (case manager, psychiatrist, therapist, or residential facility manager; 20% of cases). In Atlanta, patients were recruited either at a psychiatric rehabilitation program (Skyland Trail) or from the general outpatient population of the Atlanta VA Medical Center. The Skyland Trail patients were receiving treatment because of functional disability, including impairments in both residential and vocational functioning, and these patients were recruited through their case managers. The Atlanta VA Medical Center patients were not selected for the presence of disability and were recruited through advertisements, by word of mouth, or from another research project. The San Diego patients were recruited from the UCSD Outpatient Psychiatric Services

clinic (a large public mental health clinic) and other local community clinics or by word of mouth.

All patients received a structured diagnostic interview by a trained interviewer. At the Atlanta sites, the Structured Clinical Interview for DSM-IV Axis I Disorders (21) was used, and at the San Diego site, the Mini International Neuropsychiatric Interview, 6th Edition (22), was used. All diagnoses were subjected to a consensus procedure at each site. Patients were excluded if they had a history of traumatic brain injury with unconsciousness >10 minutes, brain diseases such as a seizure disorder or a neurodegenerative condition, or the presence of another DSM-IV diagnosis that would exclude the diagnosis of schizophrenia. None of the patients were experiencing their first psychiatric admission. Comorbid substance use disorders were not an exclusion criterion, but patients who appeared intoxicated were rescheduled. Patients resided in a wide array of unsupported, supported, or supervised residential facilities, although inpatients were not recruited. Informants were not screened for psychopathology or substance abuse.

This study was approved by local institutional review boards in Atlanta and San Diego, and all participants provided written informed consent.

Procedure

All patients were examined with a performance-based assessment of neurocognitive abilities and functional capacity. They also provided self-reports of social, residential, and vocational functioning on six different functional outcomes scales administered either as an interview by a trained rater or in questionnaire format. Informants independently completed the same six scales, reporting on the functioning of the patients. The examiner who conducted the interviews with the patient and informant then generated ratings for all six rating scales, based on his or her impression of the "true" status of the patient. Rating scales were presented in a fixed, counterbalanced order across patients.

Performance-Based Assessment

Neurocognition. We examined cognitive performance with a modified version of the MCCB. For this study, we did not include the social cognition measure from the MCCB (the managing emotions subtest of the Mayer-Salovey-Caruso Emotional Intelligence Test) because there are several reasons to think that social cognition measures may have a different relationship with everyday outcomes compared to neurocognitive measures. This minor modification of the MCCB would make the results similar to previous work, such as our own studies (3, 10), that did not include social cognition measures. We calculated a composite score, an average of nine age-corrected T scores based on the MCCB normative program, as our critical dependent variable.

Functional capacity. We administered two different performance-based functional capacity measures. Participants' functional abilities were assessed using the UPSA-B, which is a measure of functional capacity in which patients are asked to perform everyday tasks related to communication and finances. For the communication subtest, participants role play exercises using an unplugged telephone (e.g., making an emergency call; dialing a number from memory; calling to reschedule a doctor's appointment). For the finances subtest, participants count change, read a utility bill, and write and record a check for paying the bill. The UPSA-B takes 10–15 minutes, and raw scores are converted into a total score ranging from 0 to 100, with higher scores indicating better functional capacity. We also administered the advanced finances subscale of the Everyday Functioning Battery, which was designed to examine financial management in higher-functioning individuals. The advanced finances test requires the individual to prepare bank deposits, write checks to pay bills, maintain a

checkbook balance, and organize payments such that a prespecified amount of money is left available at the end of the task. This instrument was selected because it measures abilities considered important for independent living, and at the time the study was planned, we were concerned that younger individuals with schizophrenia might evidence ceiling effects on the UPSA-B. Possible scores on the advanced finances subtest range from 0 to 13.

Real-world functional outcomes. As we previously reported (17), the initial phase of the VALERO study included a RAND panel that selected six functional outcome scales from a much larger group of candidate scales as most suitable for current use at the time of the panel: the Heinrichs-Carpenter Quality of Life Scale (23), the Specific Levels of Functioning Assessment (24), the Social Behavior Schedule (25), the Social Functioning Scale (26), the Life Skills Profile (27), and the Independent Living Skills Survey (28).

There are several important features of these functional scales. Two of them were pure social functioning scales (the Social Behavior Schedule and the Social Functioning Scale), while two others examined only community functioning (the Life Skills Profile and the Independent Living Skills Survey). The remaining two (the Quality of Life Scale and the Specific Levels of Functioning Assessment) were hybrid scales examining social, residential, and vocational outcomes. Of the six scales, two were administered as self-report questionnaires (the Independent Living Skills Survey and the Specific Levels of Functioning Assessment) and the others as interviews using the standard instructions for the scale. Although all of these scales have multiple subscales, for the purposes of the present study, we examined only total scores. If these scales were used as outcome measures in a clinical trial, a single predefined primary outcome measure would be selected, and we wished to make this information available as straightforwardly as possible.

Some of these instruments were modified by deletion of some subscales following the suggestions of the RAND panel. For instance, the social acceptability and personal care subscales were omitted from the Specific Levels of Functioning Assessment total score, and the intrapsychic foundations subscale of the Quality of Life Scale was not included in the analyses because it measures deficit (i.e., negative) symptoms. While negative symptoms are known to affect functional outcomes (10), we were interested in examining the association between performance-based measures and functioning. Similar to our decision to exclude social cognition from the neurocognitive predictor set, we wanted to exclude negative symptoms as an outcome measure.

Data Analysis

The primary goal of the data analysis was to find the real-world functional outcome scale (or scales) that was most strongly related to the three performance-based indices of functional ability. The analysis was conducted using robust maximum-likelihood parameter estimation, which included all available data and did not assume multivariate normality for observed measures (29). With this analytic approach, the model is fitted to all available data from all study subjects, whether or not they have missing data on any of the variables in any model.

To accomplish these aims, a structural equation model was developed and fitted to the available data from the three sites with the latent variable modeling software program Mplus (29). A single latent trait reflecting the shared variance of the three performance-based “ability” variables was developed using hierarchical linear modeling. This single trait was then statistically related to examiner-generated total scores on all six of the real-world functional outcome scales. This overall model with a single ability latent trait and six scales as predictors was tested for its goodness of fit with standard indices, including the chi-square test, degrees of freedom, associated *p* value, and root mean square error of ap-

proximation (RMSEA). As discussed in the literature on structural equation modeling and confirmatory factor analysis (30), smaller scores on both chi-square and RMSEA are indicators of desirable fit (useful approximation to the analyzed data). Shared variance statistics are calculated with an R^2 statistic for the shared multiple correlation between real-world functioning measures and the performance-based ability latent trait.

After the fitting of this model, the real-world outcome variables were considered for deletion from the model in a sequential order based on the lowest correlations for their loadings on the ability latent trait. After deletion of the real-world functional scale with the smallest correlation, the overall model fit was recalculated, and if there was room for improvement in the fit, the next scale with a low correlation was deleted. Because the dimension of real-world functioning needed to be defined by at least two rating scales, four of the six scales could be considered for deletion, and the final two could be compared using regression analysis for their relative ability to predict the ability latent trait. Additional information about the statistical techniques we used is available in the data supplement that accompanies the online edition of this article.

Results

The demographic and clinical characteristics of the patients are summarized in Table 1. There were differences between the sites in age, education level, race, ethnicity, and type of informant. All of the informants at the Atlanta VA Medical Center were friends or relatives. There were also differences in residential status, with the majority of the San Diego and Atlanta VA patients living in the community, whereas the Skyland Trail patients were more residentially disabled. There were no differences in employment status across the sites, and total scores on the Positive and Negative Syndrome Scale were essentially identical across sites. The scores on the three performance-based measures were similar across the sites as well, although the Atlanta VA patients had somewhat higher UPSA-B scores. Thus, the demographic differences between the samples and sites did not correspond to major differences in the performance-based measures.

Table 2 presents the overall fit of the complete baseline model with the ability latent trait based on three performance-based indicators and examiner-generated total scores for the six real-world functioning scales. All three performance-based variables were significantly related to the ability latent trait, and only one of the six rating scales was significantly related to the ability latent trait—the Specific Levels of Functioning Assessment in the baseline analyses. Nonetheless, the overall fit of the model still would be considered acceptable, given the small RMSEA and a *p* value close to nonsignificant. However, this model clearly had room for improvement, in that most of the rating scales were not independently related to the performance-based measures.

Next, we sequentially omitted real-world functioning scales, as presented in Figure 1. The sequence was determined by rank-ordering the nonsignificant *p* values for the factor loadings between the real-world rating measure and the ability latent trait. Using these criteria, the order of de-

TABLE 1. Demographic and Clinical Characteristics and Performance-Based Scores for Participants at the Three Study Sites

Variable	Site						Analysis		
	Skyland Trail (N=55)		Atlanta VA Medical Center (N=40)		UCSD (N=100)				
	N	%	N	%	N	%	χ^2	df	p
Male	38	69	30	75	66	66	1.08	1	0.582
Race							41.23	2	<0.001
Caucasian	43	78	9	23	54	54			
African American	10	18	30	77	34	34			
Other	2	4	0	0	12	12			
Hispanic ethnicity	2	4	0	0	21	21	15.57	1	<0.001
Informant type							24.31	1	<0.001
Clinician	22	40	0	0	17	17			
Friend or relative	33	60	40	100	83	83			
Residential status							78.99	3	<0.001
Independent and financially responsible	7	13	24	60	57	57			
Independent, not financially responsible	11	20	5	13	17	17			
Unsupervised residential facility	4	7	5	13	23	23			
Supervised residential facility	33	60	6	15	3	3			
Employment status							10.12	3	0.120
Employed part-time	10	18	3	8	7	7			
Employed full-time	3	5	2	5	1	1			
Unemployed	40	73	35	88	89	89			
Retired	2	4	0	0	3	3			
	Mean	SD	Mean	SD	Mean	SD	F	df	p
Age	35.78	14.13	47.3	8.58	47.25	8.89	23.27	2, 192	<0.001
Education (years)	14.13	2.81	13.07	1.64	12.32	2.34	9.95	2, 178	<0.001
Baseline scores									
Positive and Negative Syndrome Scale (range: 30–210)	62.11	13.55	64.55	15.86	62.21	14.47	0.43	2, 192	0.653
UCSD Performance-Based Skills Assessment–Brief Version (range: 0–100)	75.47	14.38	81.65	13.19	75.32	12.02	3.73	2, 192	0.026
Everyday Functioning Battery (range: 0–13)	9.51	3.36	9.06	3.52	8.18	4.02	1.93	2, 168	0.148
MATRICS Consensus Cognitive Battery, Modified Composite (average T score)	39.01	7.87	37.75	6.3	37.34	6.64	1.03	2, 191	0.358

letion was Social Behavior Schedule, Independent Living Skills Survey, Quality of Life Scale, and Social Functioning Scale. In each of the resulting models, there was still room for improvement in the fit of the model, although some began to show evidence of fitting the data acceptably (e.g., the model in which Social Behavior Schedule and Independent Living Skills Survey were deleted). When four of the six scales had been deleted, the model judged as best fitting was revealed, which suggested that the Life Skills Profile and the Specific Levels of Functioning Assessment were the two best scales in combination for the prediction of the ability latent trait.

The Life Skills Profile factor coefficient was not statistically significant in the last and best-fitting model, which suggests that it did not add any information to that provided by the Specific Levels of Functioning Assessment. For this reason, a simultaneous regression analysis was performed, with these two examiner-rated total scores entered as predictors of the ability latent trait. The analysis found that the Specific Levels of Functioning Assessment total score was significantly related to the ability latent trait ($t=3.09$, $df=192$, $p=0.002$), while the Life Skills Profile total score was not. In a forced-entry regression analysis,

when the Specific Levels of Functioning Assessment was entered first, the results were significant ($t=4.52$, $df=193$, $p<0.001$), while the Life Skills Profile total score did not enter. When the Life Skills Profile total score was forced into the regression analysis first, it still did not enter the equation, while the Specific Levels of Functioning Assessment still contributed variance to the ability latent trait above and beyond the nonsignificant contribution of the Life Skills Profile ($t=4.21$, $df=192$, $p<0.001$). When we examined the total variance accounted for between the ability latent trait and the Specific Levels of Functioning Assessment total scores, we found that the Specific Levels of Functioning Assessment ratings accounted for 24% of the variance in the ability latent trait. Thus, reducing the number of real-world functional rating scales from six to one leads to a reduction in reliable variance of 17% but allows for the collection of functional information with a single questionnaire.

Discussion

The results of this study indicate that real-world functional outcomes in people with schizophrenia, rated with

TABLE 2. Factor Loadings on the Ability Latent Trait in the Full Baseline Model, Based on Three Performance-Based Indicators and Examiner-Generated Total Scores for Six Real-World Functioning Scales^a

Variable	Analysis		
	Standardized factor loading	t	p
Performance-based ability measures			
UCSD Performance-Based Skills Assessment–Brief Version	0.33	3.13	0.005
Everyday Functioning Battery	0.55	4.14	0.001
MATRICES Consensus Cognitive Battery, modified composite	0.57	4.26	0.001
	Nonstandardized factor loading	t	p
Ratings of real-world functioning			
Life Skills Profile	4.459	3.116	0.152
Social Behavior Schedule	1.761	0.617	0.537
Social Functioning Scale	1.451	0.304	0.192
Independent Living Skills Survey	1.683	0.758	0.448
Quality of Life Scale	1.206	1.271	0.204
Specific Levels of Functioning Assessment	7.093	2.435	0.015

^a Overall model fit: $\chi^2=78.10$, $df=56$, $p=0.027$; root mean square error of approximation=0.078, 90% confidence interval=0.027–0.117. Total variance shared between the ability latent trait and all six real-world measures: $R^2=0.41$.

an array of preselected rating scales and using information from the patient and an informant and condensed into a judgment rating by the examiner, are globally related to performance-based assessments of ability. This is an important finding, because several studies (reviewed above) reported modest to negligible correlations between cognitive and functional abilities and real-world outcomes, raising questions about the well-accepted relationship between cognitive impairment and disability in people with schizophrenia. Our data suggest that this relationship is actually as strong as suggested by Green et al. (7) but also indicate that both the rating scale and the methods for rating real-world functioning may influence the strength of the association.

Our results also indicate that many rating scales that address real-world functioning, even when using multiple sources of information and a systematic approach to ratings in optimal research conditions, are not strongly related to the indices of functional abilities that are employed as state-of-the-art outcome measures in treatment studies. Finally, it is possible to tentatively endorse a functional outcome scale—the Specific Levels of Functioning Assessment—that measures social, vocational, and everyday living outcomes and is related to performance on measures of everyday functional ability at a level that provides substantial information about everyday functioning.

In future studies using these performance-based measures as outcome measures, it appears that the Specific Levels of Functioning Assessment could be a suitable baseline measure to index ability-relevant real-world functioning or as an outcome measure in a long-term study, as well as for use in clinical assessment. Most real-world functional outcome scales seem to be largely redundant with each other when utilized simultaneously, and our results suggest that none of them has a stronger relationship to functional abilities than the Specific Levels of Functioning Assessment. However, this finding does not prove that other rating scales would not be suitable

FIGURE 1. Model Fittings During Sequential Deletion of Real-World Functioning Scales^a**Original model (baseline model, Table 2)**

$\chi^2=78.100$, $df=56$, $p=0.027$; RMSEA=0.078, 90% CI=0.027–0.117

Social Behavior Schedule deleted

$\chi^2=65.764$, $df=48$, $p=0.046$; RMSEA=0.075, 90% CI=0.012–0.118

Social Behavior Schedule and Independent Living Skills Survey deleted

$\chi^2=53.944$, $df=40$, $p=0.069$; RMSEA=0.073, 90% CI=0–0.120

Social Behavior Schedule, Independent Living Skills Survey, and Quality of Life Scale deleted

$\chi^2=49.275$, $df=32$, $p=0.026$; RMSEA=0.091, 90% CI=0.032–0.139

Social Behavior Schedule, Independent Living Skills Survey, Quality of Life Scale, and Social Functioning Scale deleted^b

$\chi^2=32.059$, $df=24$, $p=0.126$; RMSEA=0.072, 90% CI=0.000–0.131

^a RMSEA=root mean square error of approximation.

^b Best-fitting model: smallest chi-square value, largest p value, and smallest RMSEA; the remaining scales were the Life Skills Profile and the Specific Levels of Functioning Assessment.

for use with similar rating methods. In future analyses, we intend to examine whether the subscales of these instruments are useful in augmenting the Specific Levels of Functioning Assessment for assessment of ability-relevant real-world functioning and whether any of the informant ratings could be substituted for the comprehensive interviewer judgments.

A number of limitations must be considered when evaluating these results. As noted, it is possible that items or subscales of the scales we examined are more closely related to ability measures than the total scores. Also, our analyses did not address whether self-reports or informant reports of functioning might be superior to interviewer judgments and whether either could reliably be used on their own with the scales used in this study; this question

will be addressed in detail in subsequent analyses. Most of the available real-world functional status scales that we did not examine here were excluded either in the scale screening stage or by the RAND panel in the early phases of this study because not enough data were available to evaluate their suitability. Some scales have been developed since that time, such as the Schizophrenia Outcomes Functioning Interview (31), or have had several new studies published, such as the Personal and Social Performance Scale (32). Finally, since we did not use the MCCB social cognition measure, we could not compute the composite score with the MCCB normative sample scoring program. This program compares composite scores of individuals with schizophrenia with those of healthy individuals, producing composite overall scores in impaired populations that are lower than the average of the individual item scores. For example, in a recently published clinical trial (33), the screening composite T score on the MCCB was 24.7, but the average of the nine tests administered was 34.2. This score is similar to the average score of 38.6 on the same tests in our sample, indicating that this is not a relatively unimpaired sample.

As previous studies by our team and others have shown, other factors may contribute to the prediction of real-world outcomes, and some of them have occasionally been found to have a greater impact on real-world functioning than ability (e.g., disability compensation, depression, and negative symptoms). However, even if other factors also predict outcome, treatments aimed at disability reduction try to improve ability, not to change environmental, cultural, or emotional variables that influence disability. Later phases of the VALERO study will examine the usefulness of different informants in generating real-world functional ratings on the Specific Levels of Functioning Assessment, and later analyses of the present data set will address issues of site and informant differences in the validity of real-world functional ratings. The analyses reported here were aimed at identifying the most broadly useful real-world outcome rating scale, when administered to the patient and an informant and followed by a best-estimate rating by the examiner. Such a rating procedure has been successfully implemented in previous treatment studies (34), suggesting that it is a feasible procedure.

Received Dec. 14, 2010; revisions received Feb. 4 and Feb. 22, 2011; accepted March 3, 2011 (doi: 10.1176/appi.ajp.2011.10121723). From the Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami; Research Service, Miami VA Medical Center; the Department of Counseling, Educational Psychology, and Special Education, Michigan State University, East Lansing; the Department of Psychiatry, University of California San Diego; and the Joint Doctoral Program in Clinical Psychology, San Diego State University/University of California San Diego. Address correspondence and reprint requests to Dr. Harvey, Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, 1120 NW 14th St., Suite 1450, Miami, FL 33136; philipdharvey1@cs.com (e-mail).

Dr. Harvey has received consulting fees from Abbott Labs, Bristol-Myers Squibb, Cypress Bioscience, En Vivo, Genentech, Merck, Shire,

Sunovion, and Teva. The other authors report no financial relationships with commercial interests.

Supported by NIMH grants MH078775 to Dr. Harvey and MH078737 to Dr. Patterson.

References

1. Murray CJ, Lopez AD: Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997; 349:1436–1442
2. Wiersma D, Wanderling J, Dragomirecka E, Ganey K, Harrison G, An Der Heiden W, Nienhuis FJ, Walsh D: Social disability in schizophrenia: its development and prediction over 15 years in incidence cohorts in six European centres. *Psychol Med* 2000; 30:1155–1167
3. Bowie CR, Leung WW, Reichenberg A, McClure MM, Patterson TL, Heaton RK, Harvey PD: Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. *Biol Psychiatry* 2008; 63:505–511
4. Twamley EW, Doshi RR, Nayak GV, Palmer BW, Golshan S, Heaton RK, Patterson TL, Jeste DV: Generalized cognitive impairments, ability to perform everyday tasks, and level of independence in community living situations of older patients with psychosis. *Am J Psychiatry* 2002; 159:2013–2020
5. Harvey PD: Pharmacological cognitive enhancement in schizophrenia. *Neuropsychol Rev* 2009; 19:324–335
6. Cook JA, Leff HS, Blyler CR, Gold PB, Goldberg RW, Mueser KT, Toprac MG, McFarlane WR, Shafer MS, Blankertz LE, Dudek K, Razzano LA, Grey DD, Burke-Miller J: Results of a multisite randomized trial of supported employment interventions for individuals with severe mental illness. *Arch Gen Psychiatry* 2005; 62:505–512
7. Green MF, Kern RS, Braff DL, Mintz J: Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the "right stuff"? *Schizophr Bull* 2000; 26:119–136
8. Ventura J, Hellemann GS, Thames AD, Koellner V, Nuechterlein KH: Symptoms as mediators of the relationship between neurocognition and functional outcome in schizophrenia: a meta-analysis. *Schizophr Res* 2009; 113:189–199
9. Rosenheck R, Leslie D, Keefe R, McEvoy J, Swartz M, Perkins D, Stroup S, Hsiao JK, Lieberman J, CATIE Study Investigators Group: Barriers to employment for people with schizophrenia. *Am J Psychiatry* 2006; 163:411–417
10. Bowie CR, Depp C, McGrath JA, Wolyniec P, Mausbach BT, Thornquist MH, Luke J, Patterson TL, Harvey PD, Pulver AE: Prediction of real-world functional disability in chronic mental disorders: a comparison of schizophrenia and bipolar disorder. *Am J Psychiatry* 2010; 167:1116–1124
11. McGurk SR, Twamley EW, Sitzler DI, McHugo GJ, Mueser KT: A meta-analysis of cognitive remediation in schizophrenia. *Am J Psychiatry* 2007; 164:1791–1802
12. Patterson TL, Mausbach BT, McKibbin C, Jeste DV: Functional adaptation skills training (FAST): a randomized trial of psychosocial intervention for middle-aged and older patients with chronic psychotic disorders. *Schizophr Res* 2006; 86:291–299
13. McGurk SR, Mueser KT, Feldman K, Wolfe R, Pascaris A: Cognitive training for supported employment: 2–3 year outcomes of a randomized controlled trial. *Am J Psychiatry* 2007; 164:437–441
14. Green MF, Nuechterlein KH, Kern RS, Baade LE, Fenton WS, Gold JM, Keefe RSE, Mesholam-Gately R, Seidman LJ, Stover E, Marder SR: Functional co-primary measures for clinical trials in schizophrenia: results from the MATRICS psychometric and standardization study. *Am J Psychiatry* 2008; 165:221–228
15. McKibbin C, Patterson TL, Jeste DV: Assessing disability in older patients with schizophrenia: results from the WHODAS-II. *J Nerv Ment Dis* 2004; 192:405–413

16. Bowie CR, Twamley EW, Anderson H, Halpern B, Patterson T, Harvey PD: Self-assessment of functional status in schizophrenia. *J Psychiatr Res* 2007; 41:1012–1018
17. Leifker FR, Patterson TL, Heaton RK, Harvey PD: Validating measures of real-world outcome: the results of the VALERO expert survey and RAND panel. *Schizophr Bull* 2011; 37:334–343
18. Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, Essock S, Fenton WS, Frese FJ 3rd, Gold JM, Goldberg T, Heaton RK, Keefe RS, Kraemer H, Mesholam-Gately R, Seidman LJ, Stover E, Weinberger DR, Young AS, Zalcman S, Marder SR: The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry* 2008; 165:203–213
19. Mausbach BT, Harvey PD, Goldman SR, Jeste DV, Patterson TL: Development of a brief scale of everyday functioning in persons with serious mental illness. *Schizophr Bull* 2007; 33:1364–1372
20. Heaton RK, Marcotte TD, Mindt MR, Sadek J, Moore DJ, Bentley H, McCutchan JA, Reicks C, Grant I; HNRC Group: The impact of HIV-associated neuropsychological impairment on everyday functioning. *J Int Neuropsychol Soc* 2004; 10:317–331
21. First MB, Spitzer RL, Gibbon M, Williams JBW: Structured Clinical Interview for DSM-IV Axis I Disorders (SCID). New York, New York State Psychiatric Institute, Biometrics Research, 1995
22. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998; 59(suppl 20):22–33
23. Heinrichs DW, Hanlon TE, Carpenter WTJ: The Quality of Life Scale: an instrument for rating the schizophrenia deficit syndrome. *Schizophr Bull* 1984; 10:388–396
24. Schneider LC, Struening EL: SLOF: a behavioral rating scale for assessing the mentally ill. *Soc Work Res Abstr* 1983; 19:9–21
25. Wykes T, Stuart E: The measurement of social behaviour in psychiatric patients: an assessment of the reliability and validity of the SBS schedule. *Br J Psychiatry* 1986; 148:1–11
26. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S: The Social Functioning Scale: the development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *Br J Psychiatry* 1990; 157:853–859
27. Rosen A, Hadzi-Pavlovic D, Parker G: The Life Skills Profile: a measure assessing function and disability in schizophrenia. *Schizophr Bull* 1989; 15:325–337
28. Wallace CJ, Liberman RP, Tauber R, Wallace J: The Independent Living Skills Survey: a comprehensive measure of the community functioning of severely and persistently mentally ill individuals. *Schizophr Bull* 2000; 26:631–658
29. Muthén LK, Muthén BO: *Mplus User's Guide*. Los Angeles, Calif, Muthén & Muthén, 2010
30. Browne MW, Cudeck R: Alternative ways of assessing model fit, in *Testing Structural Equation Models*. Edited by Bollen KA, Long JS. Beverly Hills, Calif, Sage Publications, 1993, pp 136–162
31. Kleinman L, Lieberman J, Dube S, Mohs R, Zhao Y, Kinon B, Carpenter W, Harvey PD, Green MF, Keefe RS, Frank L, Bowman L, Revicki DA: Development and psychometric performance of the Schizophrenia Objective Functioning Instrument: an interviewer administered measure of function. *Schizophr Res* 2009; 107:275–285
32. Patrick DL, Burns T, Morosini P, Rothman M, Gagnon DD, Wild D, Adriaenssen I: Reliability, validity, and ability to detect change of the clinician-rated Personal and Social Performance Scale in patients with acute symptoms of schizophrenia. *Curr Med Res Opin* 2009; 25:325–338
33. Keefe RSE, Fox KH, Harvey PD, Cucchiaro J, Siu C, Loebel A: Characteristics of the MATRICS Consensus Cognitive Battery in a 29-site antipsychotic schizophrenia clinical trial. *Schizophr Res* 2011; 125:161–168
34. Harvey PD, Ogasa M, Cucchiaro J, Loebel A, Keefe RSE: Performance and interview-based assessments of cognitive change in a randomized, double-blind comparison of lurasidone vs ziprasidone. *Schizophr Res* 2011; 127:188–194