

Published in final edited form as:

*J Affect Disord.* 2014 May ; 160: 68–73. doi:10.1016/j.jad.2013.12.020.

## Measuring Treatment Response in Psychotic Depression: The Psychotic Depression Assessment Scale (PDAS) Takes both Depressive and Psychotic Symptoms into Account

Søren D. Østergaard, M.D.<sup>1,2,3</sup>, Barnett S. Meyers, M.D.<sup>4</sup>, Alastair J. Flint, M.D.<sup>5,6</sup>, Benoit H. Mulsant, M.D.<sup>6,7,8</sup>, Ellen M. Whyte, M.D.<sup>8</sup>, Christine M. Ulbricht, M.P.H.<sup>9</sup>, Per Bech, M.D. DMSc<sup>10</sup>, and Anthony J. Rothschild, M.D.<sup>11</sup> on behalf of the STOP-PD Study Group

<sup>1</sup>Unit for Psychiatric Research, Aalborg Psychiatric Hospital, Aalborg University hospital, Aalborg, Denmark

<sup>2</sup>Institute of Clinical Medicine, Aarhus University Hospital, Aarhus, Denmark

<sup>3</sup>Depression Clinical and Research Program, Massachusetts General Hospital, Harvard Medical School, Boston, MA., USA

<sup>4</sup>Department of Psychiatry, Weill Medical College of Cornell University and New York Presbyterian Hospital, Westchester Division, White Plains, New York, USA

<sup>5</sup>Department of Psychiatry, University Health Network, Toronto, Ontario, Canada

<sup>6</sup>Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

<sup>7</sup>Centre for Addiction and Mental Health, Toronto, Ontario, Canada

<sup>8</sup>Western Psychiatric Institute and Clinic, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

<sup>9</sup>University of Massachusetts Medical School, Worcester, Massachusetts, USA

---

© 2013 Elsevier B.V. All rights reserved.

Corresponding author: Søren Dinesen Østergaard, Unit for Psychiatric Research, Aalborg Psychiatric Hospital, Aalborg University Hospital, Mølleparkvej 10, DK-9000 Aalborg, Denmark, sdo@rn.dk, Phone: + 45 61282753, Fax: +45 72137235.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

### **Conflicts of interest**

S.D. Østergaard has received speaking fees, consultant honoraria and travel support from Janssen-Cilag until April 2011. Furthermore, he has received travel support at one occasion in 2010 from Bristol Myers Squibb. A.J. Rothschild has received grant support from the National Institute of Mental Health (NIMH), Cyberonics, Takeda, and St. Jude Medical and has served as a consultant to Allergan, GlaxoSmithKline, Eli Lilly, Noven Pharmaceuticals, Pfizer, Shire Pharmaceuticals, and Sunovion. A.J. Flint has received grant support from the NIMH, the Canadian Institutes of Health Research, and Lundbeck and has received honoraria from Janssen-Ortho, Lundbeck Canada, and Pfizer Canada. B.H. Mulsant currently receives research support from the Canadian Institutes of Health Research (CIHR), the US National Institute of Health (NIH), Bristol-Myers Squibb (medications for a NIH-funded clinical trial), and Pfizer (medications for a NIH-funded clinical trial). He directly own stocks of General Electric (less than \$5,000). Within the past three years, he has also received some travel support from Roche. B. Meyers receives research support from the NIMH. He is receiving medication donated by Pfizer and Eli Lilly for his NIMH trial. During the last three years he has provided legal consultation to AstraZeneca and research consultation for Forest Laboratories. E.M. Whyte has received research support from the NIMH, the National Institute of Child Health and Human Development (NICHD), the Department of Defense (DOD) and through a Small Business Innovation Research (SBIR) grant from Fox Learning Systems / National Institute of Neurological Disorders and Stroke (NINDS). P. Bech and C. Ulbricht declare no conflicts of interest.

### **Contributors**

All authors contributed to the design of this reanalysis of the STOP-PD data. The article was drafted by S.D. Østergaard and was critically revised by the other authors. The final version of the article was approved by all authors.

<sup>10</sup>Psychiatric Research Unit, Psychiatric Center North Zealand, Copenhagen University Hospital, Hillerød, Denmark

<sup>11</sup>University of Massachusetts Medical School and University of Massachusetts Memorial Health Care, Worcester, Massachusetts USA

## Abstract

**Background**—There is no established psychometric instrument dedicated to the measurement of severity in psychotic depression (PD). The aim of this study was to investigate whether a new composite rating scale, the Psychotic Depression Assessment Scale (PDAS), covering both the psychotic and the depressive domains of PD, could detect differences in effect between two psychopharmacological treatment regimens.

**Methods**—We reanalyzed the data from the Study of Pharmacotherapy of Psychotic Depression (STOP-PD), which compared the effect of Olanzapine+Sertraline (n=129) versus Olanzapine+Placebo (n=130). The response to the two regimens was compared using both a mixed effects model and effect size statistics on the total scores of three rating scales: the 17-item Hamilton Depression Rating Scale (HAM-D<sub>17</sub>), its 6-item melancholia subscale (HAM-D<sub>6</sub>), and the 11-item PDAS consisting of the HAM-D<sub>6</sub> plus five items from the Brief Psychiatric Rating Scale covering psychotic symptoms.

**Results**—According to both statistical approaches, the PDAS, the HAM-D<sub>17</sub> and the HAM-D<sub>6</sub> were all able to detect significant differences in treatment effect between Olanzapine+Sertraline and Olanzapine+Placebo (Olanzapine+Sertraline being superior). Notably, 45% of the trial participants were at least “probable psychotic” at their last assessment in the trial.

**Limitations**—The STOP-PD was not designed specifically to answer the research questions of the present study.

**Conclusions**—The Psychotic Depression Assessment Scale (PDAS) is a sensitive measure of treatment response in PD. The fact that 45% of the patients still experienced psychotic symptoms at their last trial assessment emphasizes the need to include items pertaining to psychotic symptoms in rating scales for PD.

## Keywords

Affective Disorders; Psychotic; Antidepressive Agents; Antipsychotic Agents; Psychiatric Status Rating Scales

## Introduction

Major depressive disorder (MDD) with psychotic features or “psychotic depression” (PD) is a severe and debilitating condition, which needs intensive treatment (Østergaard, et al, 2012a; Østergaard, et al, 2013; Rothschild, 2009). According to most established guidelines, PD should be treated either with electroconvulsive therapy (ECT) or a combination of an antidepressant and an antipsychotic (Farahani and Correll, 2012; Leadholm, et al, 2013). These recommendations are based on clinical trials indicating superior effect of these treatment modalities (Birkenhager, et al, 2003; Loo, et al, 2011; Meyers, et al, 2009; Petrides, et al, 2001; Rothschild, et al, 2004; Spiker, et al, 1985; Wijkstra, et al, 2010). All of these studies have used various versions of the Hamilton Depression Rating Scale (HAM-D) as part of the definition of the primary outcome. However, until recently (Østergaard, et al, 2013), none of the HAM-D scales had ever been subjected to clinical and psychometric validation in relation to PD. Furthermore, the HAM-D only covers a fraction of the psychotic symptoms in PD as only delusions and hallucinations with very specific themes

can be rated (Østergaard, et al, 2013). Therefore, the HAM-D scales may have limited content validity for assessing the severity and improvement of PD.

In a recent analysis we have tested the clinical and psychometric validity of the 11-item Psychotic Depression Assessment Scale (PDAS) covering both depressive and psychotic symptoms among patients with PD (Østergaard, et al, 2013). This scale consists of the 6-item melancholia subscale (HAM-D<sub>6</sub>) (Bech, et al, 1975), derived from the 17-item Hamilton depression rating scale (HAM-D<sub>17</sub>) (Hamilton, 1967), plus five items from the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 1962). The 11 items are: *depressed mood, guilt feelings, work and activities, psychomotor retardation, psychic anxiety and somatic symptoms (general)* from the HAM-D<sub>17</sub> and *hallucinatory behavior, unusual thought content, suspiciousness, emotional withdrawal and blunted affect* from the BPRS. In our analysis the PDAS demonstrated clinical validity, unidimensionality, and responsiveness and therefore seems to offer a promising alternative to pure depression scales in clinical trials of PD (Østergaard, et al, 2013). In contrast, the same analysis showed that the HAM-D<sub>17</sub> was not a unidimensional measure of PD, i.e. the sum of the individual item scores (the total score) is not a psychometrically valid measure for the severity of PD.

In order to further investigate the PDAS, we compared its performance to that of the HAM-D<sub>17</sub> and the HAM-D<sub>6</sub>, using data from the Study of Pharmacotherapy of Psychotic Depression (STOP-PD), which tested the effect of Olanzapine+Sertraline versus Olanzapine+Placebo among patients with PD (Meyers, et al, 2009). More specifically, we addressed the three following research questions:

1. Are the PDAS, the HAM-D<sub>17</sub> and the HAM-D<sub>6</sub> sensitive to difference in the effects of Olanzapine+Sertraline versus Olanzapine+Placebo on the severity of psychotic depression?
2. Is the measured response to the treatment regimens employed in STOP-PD captured similarly across the PDAS, the HAM-D<sub>17</sub>, and the HAM-D<sub>6</sub>?
3. What proportion of subjects in STOP-PD trial was still psychotic at the end of their participation in the trial?

## Methods

### Patient data

This analysis was based on data from the Study of the Pharmacotherapy of Psychotic Depression (Clinical Trial Registration: NCT00056472). As reported in detail elsewhere, STOP-PD is a twelve-week, randomized controlled trial (RCT) comparing the remission rates among PD patients treated with either Olanzapine+Sertraline or Olanzapine+Placebo (Meyers, et al, 2009). A total of 259 patients who met DSM-IV-TR criteria for MDD with psychotic features (American Psychiatric Association, 1994) and presented with a minimum total score of 21 on the GRID-HAMD (a modified version of the HAM-D<sub>17</sub>) (Williams, et al, 2008) participated in the study. The inclusion also required presence of a delusion, rated as 2 on at least one of the conviction items of the Delusional Assessment Scale (Meyers, et al, 2006) and a severity score of 3 on the delusion item of the Schedule of Affective Disorders and Schizophrenia (SADS) (Spitzer and Endicott, 1979). The 259 participants in the trial were recruited at four psychiatric facilities in Canada and the United States. The institutional review boards at each of the participating institutions and a data safety monitoring board at the National Institute of Mental Health approved study consent forms. Informed consent was obtained from all subjects, either directly or through approved surrogate consent procedures. The investigators were allowed to withdraw subjects who demonstrated clinically significant worsening at any time during the trial or who met criteria

for insufficient clinical improvement after five weeks of randomized treatment. Insufficient improvement was operationalized as having both a Clinical Global Impressions - Improvement (CGI-I) score of 3 or more (minimal improvement, no change, or worsening) and a Clinical Global Impressions – Severity (CGI-S) score of 4 or more (moderately or more severely ill) (Guy, 1976). The primary outcome in STOP-PD was remission, which was defined as a HAM-D<sub>17</sub> score  $\leq 10$  at two consecutive ratings and the absence of delusions, defined as a score of 1 on the SADS delusion item. The 18-item Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 1962) was used as secondary outcome measure.

### Effects size statistics

For each of the three rating scales, the HAM-D<sub>17</sub>, the HAM-D<sub>6</sub> and the PDAS, we employed effect size statistics to compare the response to Olanzapine+Sertraline versus Olanzapine+Placebo. This statistical method provides a measure for the difference in response between two treatments assessed by the same rating scale, in relation to the standard deviation on the rating scale total scores (Cohen, 1969). Therefore, the resulting effect sizes are dimensionless and independent of the raw scores, which allows for comparison of results obtained by different scales (Bech, 2012; Cohen, 1969). The effect sizes were calculated as follows: Effect size = change in rating scale total score (Olanzapine +Sertraline) - change in rating scale total score (Olanzapine+Placebo) / standard deviation on the change in rating scale total score for the entire sample. In order to match the score range on the HAM-D (0–4), the BPRS scores (1–7) on the five items included in the PDAS were converted to a score between 0 and 4 according to this formula: (BPRS score – 1)  $\times$  2/3 (Østergaard, et al, 2013). The “last observation carried forward” (LOCF) approach was used throughout. In accordance with the recommendations by the Federal Drug Administration (FDA), we applied the “70% rule” dictating that effect sizes are only valid when at least 70% of the patients are retained in each treatment group (Angst, et al, 1989; Bech, 2001). Positive effect sizes, which did not include 0 in the 95% confidence interval, were considered to document a statistically significant superior effect of Olanzapine +Sertraline over Olanzapine+Placebo. The analysis was conducted by means of the SAS statistical package (version 9.00, 2002).

### Mixed effect statistics

We also compared the response (defined as a decrease of 50% in ratingscale total score) to Olanzapine+Sertraline vs Olanzapine+Placebo on the HAM-D<sub>17</sub>, the HAM-D<sub>6</sub> and the PDAS respectively. Response was chosen as outcome for this analysis because it allows a direct comparison of the three rating scales due to its definition as a proportional change in the total score.

The analysis of response to treatment was carried out by means of an intention-to-treat, mixedeffects logistic regression (Hedeker and Gibbons, 1994) with a random intercept that included treatment and time (after baseline) as fixed effects and a treatment $\times$ time interaction effect. The analysis was performed with time expressed as both weeks and the square root of week (  $\sqrt{\text{week}}$ ) respectively. The two different time-units were employed since the logistic regression analysis depends on a linear relationship between time and outcome and we could not know a priori whether week or  $\sqrt{\text{week}}$  would yield the best linear correlation with the response rates. The degree of linear dependence between the two time units and response rates was determined by the Pearson productmoment correlation and expressed as the Pearson’s  $r$  coefficient. Potential differences in response rates between the two treatment regimens were assessed by testing for statistical significance of the treatment $\times$ time interaction at the .05 level. In order to match the score range on the HAM-D (0–4), the BPRS scores (1–7) on the five items included in the PDAS were converted to a score

between 0 and 4 according to this formula:  $(\text{BPRS score} - 1) \times 2/3$  (Ostergaard, et al, 2013). The analysis was conducted using the SAS statistical package (version 9.00, 2002).

### Psychosis-status at the end of trial participation

The status regarding psychotic symptoms was evaluated based on the scores on the SADS items representing delusions (SADS1) and hallucinations (SADS2) at the end of trial participation (LOCF). Psychotic symptoms were considered to be present if either the SADS1 score was  $\geq 2$  (*delusion(s) suspected or likely*) or if the SADS2 score was  $\geq 2$  (*probable hallucination(s)*). Accordingly, psychotic symptoms were considered to be absent when both the SADS1 and SADS2 scores were 1 (*delusions are absent & no hallucinations*).

## Results

The mean HAM-D<sub>17</sub> total score at baseline was 29.8 (SD=5.3), 36% of the subjects were males and the mean age was 58.8 years (SD=17.7). The full demographic and clinical description of the study groups can be found elsewhere (Meyers, et al, 2009). The overall conclusion of the primary publication of the STOP-PD data was that the combination of Olanzapine+Sertraline was associated with significantly higher remission rates than Olanzapine+Placebo among the patients with PD (Meyers, et al, 2009).

### Effects size statistics

The effect-size statistics for the comparison of response to Olanzapine+Sertraline versus Olanzapine+Placebo assessed by the HAM-D<sub>17</sub>, the HAM-D<sub>6</sub> and the PDAS are listed in table 1.

For all three rating scales, the effect sizes were in favor of the Olanzapine+Sertraline combination throughout the 12-weeks and the significance level was reached at week 5 (HAM-D<sub>17</sub>: 0.31, 95%CI: 0.06–0.55. HAM-D<sub>6</sub>: 0.30, 95%CI: 0.06–0.55. PDAS: 0.30, 95%CI: 0.05–0.54) when there were still more than 70% of the subjects retained in each treatment group (Olanzapine+Placebo: 71.5%, Olanzapine+Sertraline: 77.5%). There was relatively little difference in the effect sizes obtained by the HAM-D<sub>17</sub>, the HAM-D<sub>6</sub> and the PDAS throughout the study.

### Mixed effect statistics

The results from the mixed-effects analysis are illustrated in figure 1.

As evident from the scatterplots, the proportion of subjects responding to treatment was higher in the Olanzapine+Sertraline group than in the Olanzapine+Placebo group according to all three rating scales. The treatment×time interaction effect was significant for all three rating scales when time was expressed as week, but only for HAM-D<sub>17</sub> and HAM-D<sub>6</sub> when time was expressed as weeks (PDAS: P=0.057).

### Psychosis-status at the end of trial participation

According to the SADS criteria outlined in the method section, 45% of the trial participants still suffered from at least probable psychotic symptoms at the end of their participation in STOP-PD. Of the 142 individuals who completed all 12 weeks of the study, 20% presented psychotic symptoms at their last assessment: 16% of the 61 subjects receiving Olanzapine +Placebo and 22% of the 81 subjects receiving Olanzapine+Sertraline. The proportion of patients with persisting psychotic symptoms at week 12 was not significantly different between groups (Fisher's exact test, p=0.523).

## Discussion

In this analysis of the data from STOP-PD, the effect of two different psychopharmacological regimens in patients with psychotic depression were compared by means of three different rating scales and two different statistical approaches. The answers to the main research questions were as follows:

### 1. Are the PDAS, the HAM-D<sub>17</sub> and HAM-D<sub>6</sub> sensitive to differences in the effects of Olanzapine+Sertraline versus Olanzapine+Placebo on the severity of psychotic depression?

Yes, all three scales appear to be able to distinguish between the effects of different treatments in PD. According to the effect-size statistics and all three rating scales, the effect of the Olanzapine+Sertraline combination was significantly superior to Olanzapine+Placebo already from week 5. Due to the considerable attrition in the STOP-PD (only 55% of participants completed the 12-week trial), the effect-sizes from week 6 and beyond are difficult to interpret. This is mainly due to the fact, that attrition was not distributed evenly between the two treatment groups, as reflected by the number of trial completers in the Olanzapine+Placebo (61 (47%)) and the Olanzapine+Sertraline (81 (63%)) group respectively. The results of the logistic regression analysis based on the mixed-effects model pointed in the same direction as the effect size statistics. The treatment×time interaction effect was significant for five of six tests (figure 1), again indicating that the PDAS, the HAM-D<sub>17</sub> and the HAM-D<sub>6</sub> are sensitive to relatively small differences in treatment effects.

### 2. Is the measured response to the treatment regimens employed in STOP-PD captured similarly across the PDAS, the HAM-D<sub>17</sub>, and the HAM-D<sub>6</sub>?

The effect sizes obtained by the three rating scales were very similar. However, towards the end of the trial the effect sizes for the PDAS were slightly lower than those for the HAM-D<sub>6</sub> and the HAM-D<sub>17</sub>. A similar pattern was seen when response (50% reduction in rating scale total score) was used as outcome (Figure 1). This is also reflected by the fact that the treatment×time interaction effect in the mixed-effects logistic regression analysis was only statistically significant in one of two analyses when the PDAS was used as outcome measure. When examining the response curves in detail, it is clear that the discrepancy in the results obtained by the pure depression scales (HAM-D<sub>17</sub> and HAM-D<sub>6</sub>) versus the composite psychosis-depression scale (PDAS) is mainly explained by the difference in their evaluation of response in the Olanzapine+Placebo group, which was highest according to the PDAS. Consequently, it appears that the HAM-D<sub>17</sub> and HAM-D<sub>6</sub>, covering mainly depressive symptoms, favor the addition of an antidepressant relatively more than does the PDAS. To test this hypothesis we reanalyzed the data using the five BPRS psychosis items (BPRS<sub>5</sub>) from the PDAS as outcome measure. Both the effect size and the mixed effects statistical approaches were used. Indeed, the effect of Olanzapine+Sertraline was not significantly different from that of Olanzapine+Placebo on the BPRS<sub>5</sub> (effect-size at week 5: 0.22 (-0.02–0.46). P-value for treatment×week interaction: 0.24. P-value for treatment× week interaction: 0.17). Consequently, it appears that the effect of adding Sertraline to Olanzapine, was indeed driven by an antidepressant effect. This also indicates that using a pure depression scale as outcome measure in PD may result in an overestimation of the clinical benefit of antidepressants, because the psychotic symptoms, which respond less to antidepressants, are not taken properly into account. A related reason for preferring a scale covering both depressive and psychotic symptoms is that it will be more likely to detect potential worsening in psychotic symptoms during treatment, which has indeed been reported previously for the selective serotonin reuptake inhibitors in PD (Bourgeois, et al, 1998; Capaldi and Carr, 2010; Narayan, et al, 1995; Popli, et al, 1997). This clinical phenomenon is unlikely to be captured in full by the HAM-D<sub>17</sub> and HAM-D<sub>6</sub>.

To determine whether the results obtained with the BPRS<sub>5</sub> were psychometrically valid, we tested the unidimensionality of this subscale. We have recently established that the HAM-D<sub>6</sub> and the PDAS, but not the HAM-D<sub>17</sub>, were unidimensional measures in PD (Østergaard, et al, 2013). The analysis of unidimensionality was performed according to the method described by Mokken (Mokken, 1971), a non-parametric item response theory model based on Loevinger's coefficient of homogeneity (Loevinger, 1957). The Loevinger coefficient of homogeneity expresses whether the items of a scale are ordered according to their relationship to the severity of the syndrome they define, such that scorings on lower prevalence items presuppose scorings on higher prevalence items. When this is the case, the item scores are additive and the total score is a valid (unidimensional) measure for the severity of the syndrome (Loevinger, 1957). As suggested by Mokken, a Loevinger coefficient 0.40 is considered a demonstration of unidimensionality (Mokken, 1971). Our analysis was based on BPRS<sub>5</sub> ratings from week 4 in order to have sufficient variance in the total scores (Østergaard, et al, 2013). The resulting Loevinger coefficient for the BPRS<sub>5</sub> was 0.44, i.e. demonstrating the unidimensionality of the scale. Consequently, we have established that the PDAS is a unidimensional composite rating scale for PD (Østergaard, et al, 2013), consisting of two unidimensional subscales, the HAM-D<sub>6</sub> evaluating depressive symptoms, and the BPRS<sub>5</sub> focusing on psychotic symptoms.

### **3. What proportion of the subjects in the STOP-PD trial were still psychotic at the end of their participation in the trial?**

A considerable proportion (45%) of the patients included in the STOP-PD still experienced at least probable psychotic symptoms at the end of their participation in the trial. This finding is partly explained by the fact that the study design led the investigators to withdraw a number of patients due to lack of clinical improvement or direct worsening of their psychotic symptoms. However, even among the patients who completed the entire 12-week trial (n=142), probable psychotic symptoms remained prevalent (20%). From a psychometric/clinical point of view, this finding underscores the importance of including items that assess psychotic symptoms in rating scales for PD in order to ensure proper content validity.

There are some limitations to this analysis, which should be taken into account. First and foremost, the STOP-PD was not designed specifically to develop and validate the PDAS. The implications of this limitation are discussed elsewhere (Østergaard, et al, 2013). Furthermore, the results are not representative for all patients with DSM-IV psychotic depression since the presence of at least one delusion was one of the inclusion criteria of STOP-PD. In contrast, patients can fulfill the DSM-IV (American Psychiatric Association, 1994), DSM-5 (American Psychiatric Association, 2013) and ICD-10 (World Health Organization, 1993) criteria for PD without being delusional if they suffer from hallucinations in addition to major depression. Therefore, the results of the present study are only valid for PD patients with “delusional depression” and a severity of illness that resulted in contact with an academic psychiatric service in the USA or Canada.

In conclusion, we have demonstrated that the Psychotic Depression Assessment Scale (PDAS) is sensitive to relatively small differences in treatment effects in psychotic depression. The fact that almost half of the patients in STOP-PD still experienced at least probable psychotic symptoms at the end of their participation in the trial also emphasizes the need to include items pertaining to psychotic symptoms in rating scales for PD (Østergaard, et al, 2012b; Østergaard, et al, 2013). Furthermore, it seems that depression scales such as the HAM-D<sub>17</sub> and HAM-D<sub>6</sub> may overestimate the clinical benefit of antidepressants in PD, because the psychotic symptoms, which respond less to antidepressants, are not measured properly. Finally, it makes more sense from a mere semantic/intuitive perspective that the

measurement of severity and treatment response in *psychotic depression* should take both *psychotic* and *depressive* symptoms into account. The PDAS meets this criterion and through our studies we have established that this unidimensional rating scale of PD consists of two unidimensional subscales, the HAM-D<sub>6</sub> (depression) (Ostergaard, et al, 2013) and the BPRS<sub>5</sub> (psychosis). Consequently, when using the PDAS as outcome measure in PD, the overall treatment effect (PDAS) can be stratified on antidepressant (HAM-D<sub>6</sub>) and antipsychotic (BPRS<sub>5</sub>) effects respectively.

Despite the fact that the PDAS seems to perform well in the evaluation of PD, it will require further study before it can be more widely implemented. Currently, a validation study of the PDAS among patients with psychotic depression according to ICD-10 criteria (World Health Organization, 1993) is being carried out in Denmark ([ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01518049) Identifier: NCT01518049).

## Acknowledgments

The authors are grateful to the STOP-PD participants.

### **Role of the funding source**

The STOP-PD received funding by the US Public Health Service - MH 62446, MH 62518, MH 62565, and MH 62624; National Center for Research Resources - M01-RR024153, RR000056, CTSC UL1RR024996; the National Institute of Mental Health (NIMH) - MH069430, MH067710, and P30 MH068368; The NIMH supported the STOP-PD and participated in its implementation through the UO1 mechanism. They did not participate in the collection, analysis, or interpretation of study data or in the preparation, review, or approval of this manuscript. A data safety monitoring board at the NIMH provided data and safety monitoring during the STOP-PD. Eli Lilly donated olanzapine and Pfizer donated sertraline and placebo/sertraline for the STOP-PD. Neither Eli Lilly nor Pfizer participated in the design, implementation, collection, analysis, or interpretation of data in the STOP-PD or in the preparation, review, or approval of this manuscript

## REFERENCES

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th Edition. Washington, DC: 1994.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th Edition. Washington, DC: 2013.
- Angst J, Bech P, Boyer P, Bruinvels R, Engel R, Helmchen H, Hippus H, Lingjaerde O, Racagni G, Saletu B, Sedvall G, Silverstone JT, Stefanis CN, Stoll K, Woggon B. Consensus conference on the methodology of clinical trials of antidepressants, Zurich, march 1988: Report of the Consensus Committee. *Pharmacopsychiatry*. 1989; 22:3–7.
- Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG. Quantitative rating of depressive states. *Acta Psychiatr. Scand*. 1975; 51:161–170. [PubMed: 1136841]
- Bech P. Meta-analysis of placebo-controlled trials with mirtazapine using the core items of the Hamilton Depression Scale as evidence of a pure antidepressive effect in the short-term treatment of major depression. *Int. J. Neuropsychopharmacol*. 2001; 4:337–345. [PubMed: 11806859]
- Bech, P. *Clinical Psychometrics*. Oxford, UK: Wiley-Blackwell; 2012.
- Birkenhager TK, Pluijms EM, Lucius SA. ECT response in delusional versus nondelusional depressed inpatients. *J. Affect. Disord*. 2003; 74:191–195. [PubMed: 12706521]
- Bourgeois JA, Thomas D, Johansen T, Walker DM. Visual hallucinations associated with fluoxetine and sertraline. *J. Clin. Psychopharmacol*. 1998; 18:482–483. [PubMed: 9864082]
- Capaldi VF 2nd, Carr RB. Citalopram-induced hallucinations and delusions in a young adult. *Gen. Hosp. Psychiatry*. 2010; 32:648.e1–648.e3. [PubMed: 21112465]
- Cohen, J. *Statistical Power Analysis for the Behavioural Sciences*. Second Ed. New York: Lawrence Erlbaum; 1969.
- Farahani A, Correll CU. Are antipsychotics or antidepressants needed for psychotic depression? A systematic review and meta-analysis of trials comparing antidepressant or antipsychotic



monotherapy with combination treatment. *J. Clin. Psychiatry.* 2012; 73:486–496. [PubMed: 22579147]

Guy, W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD, USA: US Department of Health, Education and Welfare pub no (AMD) 76-338, NIMH; 1976. Clinical Global Impressions Scale.

Hamilton M. Development of a rating scale for primary depressive illness. *Br. J. Soc. Clin. Psychol.* 1967; 6:278–296. [PubMed: 6080235]

Hedeker D, Gibbons RD. A random-effects ordinal regression model for multilevel analysis. *Biometrics.* 1994; 50:933–944. [PubMed: 7787006]

Leadholm AK, Rothschild AJ, Nolen WA, Bech P, Munk-Jorgensen P, Ostergaard SD. The treatment of psychotic depression: is there consensus among guidelines and psychiatrists? *J. Affect. Disord.* 2013; 145:214–220. [PubMed: 23021823]

Loevinger J. Objective tests as instruments of psychological theory. *Biol. Psychiatry.* 1957; 3:635–694.

Loo CK, Mahon M, Katalinic N, Lyndon B, Hadzi-Pavlovic D. Predictors of response to ultrabrief right unilateral electroconvulsive therapy. *J. Affect. Disord.* 2011; 130:192–197. [PubMed: 20961620]

Meyers BS, English J, Gabriele M, Peasley-Miklus C, Heo M, Flint AJ, Mulsant BH, Rothschild AJ. STOP-PD Study Group. A delusion assessment scale for psychotic major depression: Reliability, validity, and utility. *Biol. Psychiatry.* 2006; 60:1336–1342. [PubMed: 17046724]

Meyers BS, Flint AJ, Rothschild AJ, Mulsant BH, Whyte EM, Peasley-Miklus C, Papademetriou E, Leon AC, Heo M. STOP-PD Group. A double-blind randomized controlled trial of olanzapine plus sertraline vs olanzapine plus placebo for psychotic depression: the study of pharmacotherapy of psychotic depression (STOP-PD). *Arch. Gen. Psychiatry.* 2009; 66:838–847. [PubMed: 19652123]

Mokken, RJ. *Theory and Practice of Scale Analysis.* Berlin, Germany: Mouton; 1971.

Narayan M, Meckler L, Nelson JC. Fluoxetine-induced delusions in psychotic depression. *J. Clin. Psychiatry.* 1995; 56:329. [PubMed: 7615486]

Ostergaard SD, Leadholm AK, Rothschild AJ. Persistent delusional theme over 13 episodes of psychotic depression. *Acta Neuropsychiatr.* 2013 <http://dx.doi.org/10.1017/neu.2013.33>.

Ostergaard SD, Rothschild AJ, Uggerby P, Munk-Jorgensen P, Bech P, Mors O. Considerations on the ICD-11 Classification of Psychotic Depression. *Psychother. Psychosom.* 2012a; 81:135–144. [PubMed: 22398817]

Ostergaard SD, Bille J, Soltoft-Jensen H, Lauge N, Bech P. The validity of the severity-psychosis hypothesis in depression. *J. Affect. Disord.* 2012b; 140:48–56. [PubMed: 22381953]

Ostergaard SD, Meyers BS, Flint AJ, Mulsant BH, Whyte EM, Ulbricht CM, Bech P, Rothschild AJ. on behalf of the STOP-PD Study Group. Measuring psychotic depression. *Acta Psychiatr. Scand.* 2013

Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychol. Rep.* 1962; 10:799–812.

Petrides G, Fink M, Husain MM, Knapp RG, Rush AJ, Mueller M, Rummans TA, O'Connor KM, Rasmussen KG Jr, Bernstein HJ, Biggs M, Bailine SH, Kellner CH. ECT remission rates in psychotic versus nonpsychotic depressed patients: a report from CORE. *J. ECT.* 2001; 17:244–253. [PubMed: 11731725]

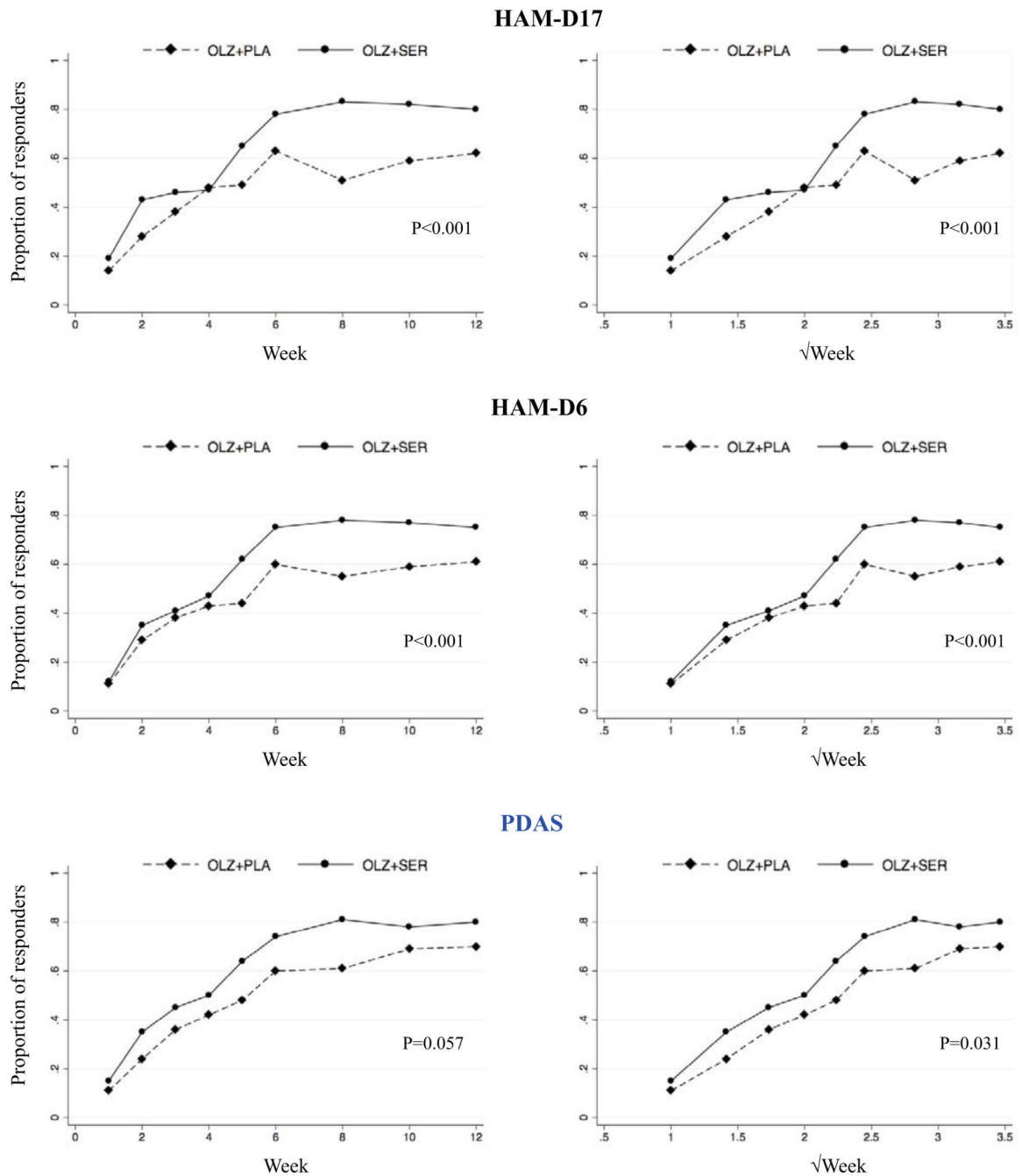
Popli AP, Fuller MA, Jaskiw GE. Sertraline and psychotic symptoms: a case series. *Ann. Clin. Psychiatry.* 1997; 9:15–17. [PubMed: 9167832]

Rothschild AJ, Williamson DJ, Tohen MF, Schatzberg A, Andersen SW, Van Campen LE, Sanger TM, Tollefson GD. A double-blind, randomized study of olanzapine and olanzapine/fluoxetine combination for major depression with psychotic features. *J. Clin. Psychopharmacol.* 2004; 24:365–373. [PubMed: 15232326]

Rothschild, AJ. *Clinical Manual for Diagnosis and Treatment of Psychotic Depression.* Washington, DC, USA: American Psychiatric Publishing, Inc.; 2009.

Spiker DG, Weiss JC, Dealy RS, Griffin SJ, Hanin I, Neil JF, Perel JM, Rossi AJ, Soloff PH. The pharmacological treatment of delusional depression. *Am. J. Psychiatry.* 1985; 142:430–436. [PubMed: 3883815]

- Spitzer, R.; Endicott, J. Schedule for Affective Disorders and Schizophrenia. 3rd ed. Biometrics Research Dept, New York State Psychiatric Institute; 1979.
- Wijkstra J, Burger H, van den Broek WW, Birkenhager TK, Janzing JG, Boks MP, Bruijn JA, van der Loos ML, Breteler LM, Ramaekers GM, Verkes RJ, Nolen WA. Treatment of unipolar psychotic depression: a randomized, double-blind study comparing imipramine, venlafaxine, and venlafaxine plus quetiapine. *Acta Psychiatr. Scand.* 2010; 121:190–200. [PubMed: 19694628]
- Williams JB, Kobak KA, Bech P, Engelhardt N, Evans K, Lipsitz J, Olin J, Pearson J, Kalali A. The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. *Int. Clin. Psychopharmacol.* 2008; 23:120–129. [PubMed: 18408526]
- World Health Organization. Diagnostic criteria for research. Geneva: WHO; 1993. The ICD-10 Classification of Mental and Behavioural Disorders.



**Figure 1. Proportion of patients responding to Olanzapine+Placebo or Olanzapin+Sertraline according to the HAM-D<sub>17</sub>, the HAM-D<sub>6</sub> and the PDAS respectively**  
 Potential differences in response (defined as a decrease of 50% in rating-scale total score) to Olanzapine+Placebo versus Olanzapin+Sertraline was tested using an intention-to-treat, mixed-effects logistic regression with a random intercept that included treatment and time as fixed effects and a treatment×time interaction effect. In the three figures to the left, time is expressed as weeks and in the three figures to the right, time is expressed as the square root of weeks (  $\sqrt{\text{week}}$  ). The P-values reflect the significance of the treatment×time interaction effect in the logistic regression.

**Table 1**

Effect-size statistics for the comparison of Olanzapine+Sertraline versus Olanzapine+Placebo

	<b>HAM-D<sub>17</sub></b>	<b>HAM-D<sub>6</sub></b>	<b>PDAS</b>	<b>% retained</b>
Week 1	0.23 (-0.01 – 0.48)	0.18 (-0.07 – 0.42)	0.19 (-0.05 – 0.43)	96.5
Week 2	0.28 (0.04 – 0.53)	0.22 (-0.02 – 0.47)	0.19 (-0.06 – 0.43)	91.5
Week 3	0.27 (0.03 – 0.52)	0.21 (-0.03 – 0.46)	0.25 (0.01 – 0.50)	85.7
Week 4	0.22 (-0.03 – 0.46)	0.23 (-0.01 – 0.47)	0.23 (-0.02 – 0.47)	81.5
Week 5	0.31 (0.06 – 0.55)	0.30 (0.06 – 0.55)	0.30 (0.05 – 0.54)	74.5
Week 6	0.31 (0.07 – 0.56)	0.35 (0.10 – 0.60)	0.33 (0.09 – 0.58)	67.2
Week 8	0.48 (0.24 – 0.73)	0.52 (0.28 – 0.77)	0.46 (0.22 – 0.71)	62.2
Week 10	0.46 (0.21 – 0.71)	0.48 (0.24 – 0.73)	0.42 (0.17 – 0.67)	58.3
Week 12	0.51 (0.26 – 0.76)	0.49 (0.25 – 0.74)	0.45 (0.21 – 0.70)	54.8

The effect sizes for the comparison of Olanzapine+Sertraline (n=129) versus Olanzapine+Placebo (n=130) for the HAM-D<sub>17</sub>, the HAM-D<sub>6</sub> and the PDAS respectively. The effect sizes are calculated as follows: Effect size = (change in rating scale total score (Olanzapine+Sertraline) - (change in rating scale total score (Olanzapine+Placebo) / standard deviation on the change in rating scale total score for the entire sample). The "Last observation carried forward" approach was used. The column "% retained" lists the proportion of patients still active in the study at each assessment.