Journal of Affective Disorders 174 (2015) 589–593

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research report

Test–retest reliability of a new questionnaire for the retrospective assessment of long-term lithium use in bipolar disorder

Sarah K. Tighe [a,*], Megan Ritchey [b], Barbara Schweizer [b], Fernando S. Goes [b], Dean MacKinnon [b], Francis Mondimore [b], J. Raymond DePaulo [b], Francis J. McMahon [c], Thomas G. Schulze [b,d,e], Peter P. Zandi [b,f], James B. Potash [a]

[a] University of Iowa Carver College of Medicine, Department of Psychiatry, Iowa City, IA, USA
[b] The Johns Hopkins University School of Medicine, Department of Psychiatry and Behavioral Sciences, Baltimore, MD, USA
[c] Human Genetics Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA
[d] Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Germany
[e] Institute of Psychiatric Phenomics and Genomics, Ludwig-Maximilians-University, Munich, Germany
[f] The Johns Hopkins University Bloomberg School of Public Health, Department of Mental Health, Baltimore, MD, USA

ARTICLE INFO

Article history:
Received 3 September 2014
Accepted 13 November 2014
Available online 24 November 2014

Keywords:
Bipolar disorder
Lithium
Retrospective assessment
Long-term therapy
Predictors of treatment response
The Alda Scale

ABSTRACT

Background: The identification of predictors of treatment response holds tremendous potential for the improvement of clinical outcomes in bipolar disorder (BP). The goal of this project is to evaluate the test–retest reliability of a new clinical tool, the Lithium Questionnaire (LQ), for the retrospective assessment of long-term lithium use in research participants with BP.

Methods: Twenty-nine individuals with BP-I ($n=27$), major depression ($n=1$), or schizoaffective disorder ($n=1$) were recruited for participation. The LQ was administered to all participants at two time-points, spaced 17 months apart on average, and used to determine each subject's score on the Retrospective Criteria of Long-Term Treatment Response in Research Subjects with Bipolar Disorder Scale, or the Alda Scale. Scores were confirmed through a best-estimate procedure, and test–retest reliability (intra-class correlation coefficient [ICC]) of the LQ was calculated.

Results: The correlation between the total Alda Scale scores at the two time-points was in the moderate range (ICC=0.60). Relevant clinical factors such as age or presence of Axis I psychiatric comorbidity did not influence the reliability.

Limitations: The validity of the LQ was not examined. Inclusion of two participants with non-BP diagnoses may have affected the LQ's reliability, but re-analysis of our data after exclusion of these participants did not influence the reliability. The absence of measures of mood and cognition at time of LQ may be a limitation of this work.

Conclusions: The LQ holds promise for the standardization of the retrospective assessment of long-term treatment in BP.

1. Introduction

Bipolar disorder (BP), a serious and chronic mood disorder, is associated with significant morbidity and mortality. BP affects up to 3% of the general population (Merikangas et al., 2007) and is characterized by recurrent episodes of depression and mania. Placing individuals at significant risk for self-inflicted harm, BP increases the risk of suicide ~20-fold compared to the general population (Osby et al., 2001). Due to its severe and recurring pathologic mood fluctuations, BP is disruptive to functioning at home, interpersonally, and in the workplace. In light of this, the World Health Organization deems BP to be the sixth-leading cause of disability (Lopez and Murray, 1998).

Mood stabilizers, such as lithium and valproic acid, are the cornerstone of BP pharmacotherapy. The efficacy of lithium for maintenance therapy in BP is well established, and thus, lithium is widely regarded as the first-line agent for BP long-term therapy. Converging evidence from naturalistic studies and clinical trials suggest that roughly 30% of BP patients have a full therapeutic response to long-term lithium treatment (Garnham et al., 2007; Rybakowski et al., 2001). In support of this, a meta-analysis of five randomized, placebo-controlled clinical trials of long-term lithium

* Correspondence to: University of Iowa Carver College of Medicine, Department of Psychiatry, W278 GH, 200 Hawkins Drive, Iowa City, IA 52242.
Tel.: +1 319 384 7136; fax: +1 319 353 8300.
E-mail address: sarah-k-tighe@uiowa.edu (S.K. Tighe).

http://dx.doi.org/10.1016/j.jad.2014.11.021
0165-0327/

use in BP demonstrated that lithium reduces BP recurrences by 35% with greater prophylaxis against mania than depression (Geddes et al., 2004). However, the corollary of these findings is that a substantial fraction of individuals with BP are partially or completely non-responsive to maintenance lithium therapy.

Because of the substantial rates of treatment non-response in BP and the potential harm that can befall poorly controlled patients, considerable interest exists in identifying predictors of a positive response. Some clinical predictors are known (Tighe et al., 2011), but their predictive power is modest. Biomarkers may be a more powerful approach to determining which patients are most likely to have a positive response to lithium. Such interest extends to other medical disciplines and has given rise to the field known as personalized medicine.

Critical to the discovery of biomarkers predictive of treatment response is the classification of BP phenotypes according to treatment response. The gold standard approach toward this end is the prospective, blinded ascertainment of treatment response during a clinical trial of lithium monotherapy (Lopez de Lara et al., 2010; Manchia et al., 2013). However, this approach is not always feasible under real-world conditions, and it may fail to provide a sample size that yields adequate power for the identification of biomarkers in disciplines such as pharmacogenetics or neuroimaging. An alternative method is the retrospective evaluation of treatment response.

The aim of this project is the evaluation of the reliability of a new clinical interview, the Lithium Questionnaire (LQ, Supplement 1), for the retrospective assessment of lithium response in individuals with BP. The Consortium on Lithium Genetics (ConLiGen, www.ConLiGen.org, Schulze et al., 2010), a large, international group which collectively aims to define the primary phenotype and genotype of a good lithium response, has adopted a particular rating scale, the Retrospective Criteria of Long-Term Treatment Response in Research Subjects with Bipolar Disorder Scale (the Alda Scale; Grof et al., 2002), for retrospective studies of lithium response. In an effort to standardize the use of the Alda Scale across academic centers, we developed the LQ, a semi-structured clinical interview, which aims to obtain the clinical information necessary for the completion of the scale. Here we present the results from the test–retest reliability assessment of the LQ. We hypothesized that the reliability of the LQ at the two different timepoints would be negatively impacted by the following factors: greater duration between the initial and follow-up interviews; interviews performed by research-clinicians with lower levels of training; the presence of Axis I psychiatric comorbidity; and increasing age of the participant.

2. Methods

2.1. Ascertainment

One hundred and five participants with a mood disorder diagnosis were contacted as part of a larger ConLiGen initiative. These individuals were recruited from previous genetic studies conducted by the Johns Hopkins Mood Disorders Center. Of the 105 participants, 29 agreed to participate in the present reliability study.

Subjects were contacted by phone and screened for eligibility. After the study was explained in detail, subjects who agreed to participate provided written consent, and a clinical interview with a study clinician was scheduled. All study procedures were reviewed and approved by the Institutional Review Board at the Johns Hopkins University School of Medicine (Baltimore, MD, USA).

2.2. Assessment

All subjects received a diagnostic evaluation in person or by telephone by an advanced research-clinician using the Diagnostic Interview for Genetic Studies (DIGS; Nurnberger et al., 1994), a semi-structured interview that assesses the severity and lifetime course of depressive, manic, and hypomanic symptoms, as well as several other categories of psychopathology required to make axis I and II diagnoses. Mood disorder diagnoses were made according to a best-estimate procedure using Diagnostic Statistical Manual (DSM)–III-R or DSM-IV criteria. Diagnoses incorporated all available clinical information including results from the DIGS, screening data, and medical records.

Each participant completed a telephone interview with an expert research clinician trained in mood disorders using the new, semi-structured clinical tool, the LQ, which is described in detail below. The level of training of the interviewers ranged from master's level psychologist (LN) to psychiatrist-in-training (SKT) to psychiatrist with mood disorder subspecialty expertise (FSG, DM, FM, TGS). After each interview, the evaluating clinician used the LQ to score each subject's response to long-term lithium treatment according to the Alda Scale. This score was then confirmed through a best-estimate review by a different research clinician (SKT, LN, FSG, DM, FM, JBP), who was blind to the Alda Scale results from the interviewer. Any disagreement in the scores between the two clinicians was adjudicated through discussion or review by a third clinician.

To determine the test–retest reliability of the LQ, subjects were contacted for a second interview with the questionnaire. After a thorough explanation of this second phase of the study, all subjects who agreed to participate provided written consent, and the second interview occurred 17 months on average (standard deviation [SD] 3.6 months) after the initial interview. All follow-up interviews were performed by an M.D. level research clinician (SKT, TGS) blind to the findings of the initial interview. The identical method as described above was used to obtain a best-estimated second score.

2.3. Assessment of treatment response according to the Alda Scale

The retrospective assessment of long-term lithium treatment was determined according to the previously validated Alda Scale, which is described in detail elsewhere (Grof et al., 2002). In brief, the total score is a function of the magnitude of clinical improvement during treatment (A score) and factors that reduce one's confidence that the observed clinical improvement is a result of treatment (B score). With scores ranging from zero to ten, the A score considers the impact of treatment on the frequency and severity of mood episodes. The B score is the sum of five potential confounders whose values range from zero to two: the number of episodes before or off the treatment (B1); the frequency of episodes before or off the treatment (B2); the duration of the treatment (B3); compliance during period (s) of stability (B4); and the use of additional medication during the period of stability (B5). The total score is calculated by subtracting the total B score from the A score.

2.4. Lithium questionnaire

With the goal of standardizing the acquisition of clinical data necessary for the completion of the Alda Scale, the LQ was developed by research clinicians from the Johns Hopkins Mood Disorders Center. Three sources were taken into consideration during the development of the questionnaire: the lithium response rating form developed by Dr. John Kelsoe of the University of California San Diego; a phone-based assessment of retrospective lithium response used by researchers at the Genetic Basis of Mood and Anxiety Disorders Section of the National Institute of Mental Health (NIMH) in Bethesda, Maryland; and the collective expertise of the Johns Hopkins Mood Disorders Center.

The LQ consists of six different parts (Parts 1–6) and is informed by all available clinical information including the actual