European Psychiatry 53 (2018) 17–22

Contents lists available at ScienceDirect

European Psychiatry

journal homepage: http://www.europsy-journal.com

Patient centric measures for a patient centric era: Agreement and convergent between ratings on The Patient Global Impression of Improvement (PGI-I) scale and the Clinical Global Impressions – Improvement (CGI-S) scale in bipolar and major depressive disorder

M. Mohebbi[a,b,*], S. Dodd[b,d,e], O.M. Dean[b,c,d], M. Berk[b,c,d,e]

[a] Deakin University, Faculty of Health, Biostatistics Unit, Geelong, Victoria, Australia
[b] Deakin University, School of Medicine, IMPACT Strategic Research Centre, Barwon Health, Geelong, Victoria, Australia
[c] Florey Institute for Neuroscience and Mental Health, University of Melbourne, Kenneth Myer Building, Parkville, Australia
[d] University of Melbourne, Department of Psychiatry, Level 1 North, Main Block, Royal Melbourne Hospital, Parkville, Australia
[e] Orygen, The Centre of Excellence in Youth Mental Health, The Department of Psychiatry and the Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Australia

ARTICLE INFO

Article history:
Received 19 December 2017
Received in revised form 15 May 2018
Accepted 15 May 2018
Available online xxx

Keywords:
Patient global impression scale of improvement (PGI-I)
Clinician global impression scale of improvement (CGI-I)
Convergent validity
Agreement
Depression
Bipolar disorder
Psychology
randomized control trial

ABSTRACT

Background: Concordant with an increased emphasis on consumer engagement, the Patient Global Impression Scale of Improvement (PGI-I) is commonly used as an outcome measure in studies evaluating the efficacy of treatments in medical and psychiatric conditions with subjective symptom domains. The current study evaluated the agreement between PGI-I and Clinician Global Impression Scale of Improvement (CGI-I) ratings and convergent validity of PGI-I among individuals with bipolar or major depressive disorders.

Method: Data were derived from three double-blind, placebo-controlled, multicentre studies conducted from 2007 to 2015 among adult individuals (N = 472). Clinicians were asked to rate participants symptoms using the CGI-I as well as severity of depression by the Montgomery-Åsberg Depression (MADRS), quality of life (Q-LES-Q), social and occupational functioning (SOFAS), and functional impairment (LIFE–RIFT). Participants were asked to assess their symptom improvement with the PGI-I. Bland-Altman agreement plots and Intra-class correlations were used to evaluate agreement, and Spearman correlation coefficients were implemented to examine convergent validity. Sub-group analyses for disorder type (bipolar and major depression) were performed.

Results: There was high agreement between the PGI-I and CGI-I ratings across follow-up time points (weeks 2, 4, 6, 8, 12, 16, 20, 24, and 28). Similar results were observed in male only and female only data and after adjustment for age and gender. Both PGI-I and CGI-I ratings were robustly positively correlated with MADRS, and LIFE-RIFT and negatively correlated with SOFAS and Q-LES-Q, supporting the convergent validity of the PGI-I. Sub-group analyses for bipolar and major depressive disorder showed similar findings.

Conclusion: Our findings support the utility of the PGI-I as a participant rated measure of global improvement among individuals with bipolar or major depressive disorders.

1. Introduction

There is increasing emphasis on objective assessment of patient centric outcomes that span function, symptoms and quality of life. Validated measures of patient-reported outcomes with standardized questionnaires are thus critical to clinical and research outcome assessment [1]. When patients seek treatment, a determination of severity before treatment and improvement after treatment is mostly based on their subjective reporting of symptoms, and are linked to objective measures in those disorders

Abbreviations: CGI-I, clinician global impression scale of improvement; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition; ICH-GCP, International Conference of Harmonisation for Good Clinical Practice Guidelines; ICC, Intraclass Correlation Coefficient; MDD, major depressive disorder; MADRS, Montgomery-Asberg Depression; NAC, *N*-acetylcysteine; PGI-I, patient global impression scale of improvement; Q-LES-Q, Quality of Life Enjoyment and Satisfaction Questionnaire–Short Form; LIFE–RIFT, Range of Impaired Functioning Tool; SOFAS, Social and Occupational Functioning Scale.

* Corresponding author at: Deakin University, Faculty of Health, Biostatistics Unit, Geelong, Victoria, Australia.
E-mail address: m.mohebbi@deakin.edu.au (M. Mohebbi).

http://dx.doi.org/10.1016/j.eurpsy.2018.05.006
0924-9338/

where they are salient. Little is known about how the patient's subjective and the clinician's objective rating of disorder improvement or treatment effect are aligned.

The Patient's Global Impressions of Improvement (PGI-I) scale has been included in several studies conducted worldwide to assess patients' overall perception of their condition by a simple and easy-to-use validated questionnaire [2]. The PGI-I is a 1-item questionnaires that ask an individual patient to rate the perceived change in his/her condition in response to therapy at endpoint. It is derived from the Clinical Global Impressions – Improvement scale (CGI-I) which was first developed for use in psychopharmacology trials as part of the NIMH collaborative study of schizophrenia [2]. Since then, it has been used as a standard primary outcome measure in studies investigating the efficacy of pharmacological treatments for psychiatric and medical conditions where subjective symptoms predominate, including pain, fatigue and mood [3–6] as well as secondary outcomes and responder analysis in many more studies; for example [7–9]. The CGI-I address the patient's improvement from baseline rated by the clinician. Both PGI-I and CGI-I show a bipolar scaling from 1 (very much improved) to 7 (very much worse). These types of measures have been validated in clinical studies of patients with stress incontinence [10], urogenital prolapse [11], fibromyalgia [12] major depressive disorder [13] and stress urinary incontinence [14].

This article aims to evaluate the agreement between patient- and clinician-rated global impression of improvement (PGI-I, CGI-I) scales. We also examine convergent validity of PGI-I compared with CGI-I correlation other clinician-assigned ratings of disease severity, functioning and quality of life. Data were derived from a three double-blind, placebo-controlled, multicentre, randomized controlled trials in adult outpatients with bipolar depression and major depressive disorder.

2. Methods

2.1. Study design and participants

This was a secondary analysis of data from 3 clinical trials. Details of the study designs and populations have previously been published [15–18]. Study 1 was a randomized, double-blind, placebo-controlled, parallel-design study to evaluate the efficacy of 2 g/day *N*-acetylcysteine (NAC) as adjunct maintenance treatment for bipolar disorder [15,16]. Participants (n = 149) had a Montgomery Asberg Depression Rating Score of (MADRS) $\geq$12 at trial entry and, after eight weeks of open-label NAC treatment, were randomized to adjunctive (in addition to treatment as usual) NAC or placebo for a further 24 weeks. Study participants were men and women residing in Australia and Brazil (www.anzctr.org.au: ACTRN12607000074493).

Study 2 was a randomized, double-blind, placebo-controlled, parallel-design study to evaluate the efficacy of 1 g/day NAC for major depressive disorder (MDD) in addition to existing treatments [17]. Participants (N = 252) had MADRS $\geq$18 at the time of entry with a current episode of MDD diagnosed according to DSM-IV-TR criteria. Participants were treated with NAC or placebo in addition to treatment as usual for 12 weeks and were followed to 16 weeks. Study participants were men and women residing in Australia (www.anzctr.org.au: ACTRN12607000134426).

Study 3 was a randomized, double-blind, placebo-controlled, parallel-design study to evaluate the efficacy of 200 mg/day of adjunctive minocycline or placebo for major depressive disorder (MDD) in addition to existing treatments [18]. Participants (N = 71) had MADRS $\geq$25 at the time of entry and met criteria for unipolar depression, based on Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition (DSM-IV) criteria. Participants were randomized to NAC or placebo (parallel groups) over 12 weeks of treatment and were followed to week 16. Study participants were men and women residing in Australia and Thailand (www.anzctr.org.au: ACTRN12612000283875).

2.2. Instruments

2.2.1. Patient Global Impression of Improvement (PGI-I) and Clinician Global Impression of Improvement (CGI-I)

Patient global impression of improvement scale (PGI-I) is a single-item global rating of change scale that ask an individual patient to rate the severity of a specific condition at baseline and or to rate at endpoints the perceived change in his/her condition in response to therapy. There are seven possible responses (scored 1–7): very much better, much better, a little better, no change, a little worse, much worse, and very much worse. The clinical global impression of improvement scale (CGI-I) is the clinician rated single-item scale that uses the same seven-point response criteria as the PGI-I [2] (see Appendix A in Supplementary material).

2.2.2. Depression severity

Severity of depressive symptomatology across studies time points were measured using the Montgomery–Åsberg Depression Rating Scale (MADRS) [19].

2.2.3. Quality of life

Quality of Life Enjoyment and Satisfaction Questionnaire–Short Form (Q-LES-Q) [20] was used for measuring quality of life.

2.2.4. Functional impairment

Functional impairment was measured using the Range of Impaired Functioning Tool (LIFE–RIFT) [21].

2.2.5. Social and occupational functioning

The Social and Occupational Functioning Scale (SOFAS) [22] was used to measure functioning over the duration of the study.

2.3. Ethics

All trials were conducted according to the Declaration of Helsinki 1964 as revised in 2008, the requirements of the Australian National Statement on Ethical Conduct in Human Research, the federal patient privacy (HIPAA) law and the International Conference of Harmonisation for Good Clinical Practice Guidelines (ICH-GCP) and were approved by institutional review boards at all sites.

2.4. Statistical analysis

Weighted agreement [23] was reported as a descriptive measure. The weights were given by $1 - \frac{|i-j|}{(k-1)}$, where i and j index the rows of columns of the ratings for CGI-I and PGI-I, $|i - j|$ indicate absolute difference and k is the maximum number of possible ratings. A weight of 1 indicates that an observation should count as perfect agreement and a weight of, say, 0.66 means that CGI-I and PGI-I are in two-thirds agreement (which happens if CGI-I and PGI-I are "two apart"). The agreements between clinician and patient ratings were assessed using Intraclass Correlation Coefficient (ICC) and its 95% confidence interval (CI) by implementing two-way random-effects model [24]. According to Fleiss [23], ICC values lower than 0.40 can be interpreted as poor, between 0.41 and 0.75 as fair, and above 0.75 as excellent agreement. The Bland-Altman plot [25] was used to visually inspect agreement. This analysis involved plotting the difference between CGI-I and PGI-I measurements against the average of the two measurements $\pm$ 1.96 times its SD known as the 95% limits of agreement.