Journal of Psychiatric Research 61 (2015) 40–45

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Rating depression over brief time intervals with the Hamilton Depression Rating Scale: Standard vs. abbreviated scales

David A. Luckenbaugh, Rezvan Ameli, Nancy E. Brutsche, Carlos A. Zarate Jr. *

Experimental Therapeutics and Pathophysiology Branch, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, and Department of Health and Human Services, Bethesda, MD, USA

ARTICLE INFO

Article history:
Received 31 July 2014
Received in revised form
24 November 2014
Accepted 18 December 2014

Keywords:
Antidepressant
Bipolar disorder
Depression
HDRS
Ketamine
Ratings

ABSTRACT

Although antidepressant trials typically use weekly ratings to examine changes in symptoms over six to 12 weeks, antidepressant treatments may improve symptoms more quickly. Thus, rating scales must be adapted to capture changes over shorter intervals. We examined the use of the 17-item Hamilton Depression Rating Scale (HDRS) to evaluate more rapid changes. Data were examined from 58 patients with major depressive disorder or bipolar disorder enrolled in double-blind, placebo-controlled, cross-over studies who received a single infusion of ketamine (0.5 mg/kg) or placebo over 40 min then crossed over to the other condition. HDRS subscales, a single HDRS Depressed mood item, and a visual analogue scale were used at baseline, after a brief interval (230 min), and one week post-infusion. Effect sizes for the ketamine-placebo difference were moderate ($d > 0.50$), but one and two-item HDRS subscales had the smallest effects. Response rates on active drug were lowest for the complete HDRS (43%); the remaining scales had higher response rates to active drug, but the shortest subscales had higher response rates to placebo. Correlations between the changes from baseline to 230 min post-ketamine across scores were similar for most subscales ($r = 0.82–0.97$), but correlations using the single items were lower ($r < 0.74$). Overall, effect sizes for drug-placebo differences and correlations between changes were lower for one- and two-item measures. Response rates were lower with the full HDRS scale. The data suggest that, to best identify rapid antidepressant effects, a scale should have more than two items, but fewer items than a full scale.

Published by Elsevier Ltd.

1. Introduction

For decades, clinical trials in depression—major depressive disorder (MDD) and bipolar disorder (BD)—have typically used weekly ratings to examine changes after treatment. Some recent experimental treatments, however, appear to exert antidepressant effects within hours or days, underscoring the need for rating instruments capable of detecting much more rapid antidepressant effects (Machado-Vieira et al., 2008). Current common scales used to assess depressive symptoms, such as the Hamilton Depression Rating Scale (HDRS) (Hamilton, 1960) and the Montgomery-Asberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979) were designed to obtain measurements at weekly intervals. In addition, these scales measure certain symptoms that cannot be evaluated over a short time frame (e.g., changes in sleep or weight).

Various research groups have suggested different ways to use the original HDRS items to assess depressive symptoms. In a series of clinical trials, Entsuah et al. (2002) and Santen et al. (2008) examined the individual items of the HDRS to determine those that were most sensitive to changes in depressive symptoms. Santor et al. (2008) took an item response theory approach in an attempt to differentiate patients with higher and lower degrees of symptom severity.

Because these approaches led to different subscales of the HDRS, additional studies examined how well various subscales performed compared to total HDRS score. While some studies showed that shorter subscales improved the rate of response to the outcome measure (e.g. Bech et al., 2010; Faries et al., 2000; Mallinckrodt et al., 2011; Revicki et al., 2010; Santen et al., 2009; Silverstone et al., 2002), others found no noticeable difference (e.g.

* Corresponding author. National Institutes of Health, National Institute of Mental Health, Experimental Therapeutics & Pathophysiology Branch, Building 10, Clinical Research Center (CRC), 10 Center Dr., Room 7-5342, Bethesda, MD 20892, USA. Tel.: +1 301 451 0861; fax: +1 301 480 8792.

E-mail addresses: dave.luckenbaugh@nih.gov (D.A. Luckenbaugh), rezvan.ameli@nih.gov (R. Ameli), nancy.brutsche@nih.gov (N.E. Brutsche), zaratec@mail.nih.gov, carlos.zarate@nih.gov (C.A. Zarate).

http://dx.doi.org/10.1016/j.jpsychires.2014.12.015
0022-3956/Published by Elsevier Ltd.

Ballesteros et al., 2007; McIntyre et al., 2005; Revicki et al., 2010; Ruhe et al., 2005). Boessen et al. (2013) pointed out that some of the differences across studies were likely due to the type of studies used to evaluate the scales. For instance, they noted that total HDRS score appeared to work better when evaluating tricyclic antidepressants (TCAs), but HDRS subscales often worked better with selective serotonin reuptake inhibitors (SSRIs). Santen et al. (2009) found that a specific subscale of the HDRS worked best for both of these drug classes.

Within the scientific community, considerable interest exists in developing or adapting instruments capable of evaluating improvement over much shorter periods of time. Several groups have pointed out that antidepressants appear to improve symptoms long before the commonly reported value of three to six weeks (e.g. Papakostas et al., 2006; Posternak et al., 2005; Stassen and Angst, 1988) and, notably, some current trials examine symptoms within the course of a single day (e.g. Zarate et al., 2006). Because some depression severity scales include items that cannot change over short intervals (e.g., insomnia), the need exists to understand which scales, if any, will be sensitive to changes within these very brief time frames but nevertheless remain relevant over longer periods.

The present study used the HDRS to illustrate change in depressive symptoms over the course of a brief treatment trial with the N-methyl-D-aspartate (NMDA) antagonist ketamine. This scale was chosen as representative of typical clinical trial scales because 1) it is widely used; 2) it was the primary outcome measure for our initial ketamine study; and 3) we had the most available data with it. The standard HDRS scale and HDRS subscales were used to determine better approaches for handling data in studies involving rapid changes in depressive symptoms. The analysis examined depressive symptoms assessed at baseline with changes assessed at 230 min post-ketamine infusion to examine extremely short time frames; the symptoms were also assessed at seven days post-ketamine infusion in order to reflect a more standard antidepressant time point.

2. Material and methods

Fifty-eight treatment-resistant inpatients with either BD (n = 36) or MDD (n = 22)—as assessed via the Structured Clinical Interview for the DSM IV-R, clinical interviews, and patient history—were recruited to participate in double-blind, placebo-controlled, crossover studies of ketamine to reduce depressive symptomatology. All studies were approved by the Combined Neuroscience IRB at the NIH. Patients provided informed consent prior to participation. The methodology and results of these studies have been published elsewhere (DiazGranados et al., 2010; Zarate et al., 2006, 2012), but additional patients recruited in the process of generating those manuscripts are included here. Briefly, patients were randomized to receive a single infusion of either placebo or 0.5 mg/kg of ketamine hydrochloride and then crossed over to the other condition after a week for MDD patients and two weeks for BD patients. MDD patients were medication free for at least two weeks prior to the study and BD patients were on stable doses of either lithium or valproic acid. Trained clinicians rated patients on the HDRS, a visual analogue scale (VAS) for “depressed, sad, blue”, the MADRS, the Beck Depression Inventory (BDI) (Beck and Beamesderfer, 1974), and several other psychiatric rating scales at 60 min prior to infusion, then at 40, 80, 120, and 230 min post-infusion, and finally on days 1, 2, 3, and 7 post-infusion. Ratings were made relative to the most recent time period assessed.

The 17-item HDRS was examined as a whole and in several subscales designed to evaluate depressive symptoms in a clinical trial. Outcome measures included: 1) the total HDRS; 2) seven subscales of the HDRS drawn from the extant literature (see Table 1) (Bech et al., 1981; Evans et al., 2004; Gibbons et al., 1993; Maier et al., 1985; McIntyre et al., 2002; Santen et al., 2008; Silverstone et al., 2002); 3) a shortened version of the HDRS that eliminated items that would not change over brief time intervals (e.g., early, middle, and late insomnia and weight change) (Leibenluft et al., 1993); 4) the Depressed mood item from the HDRS, and 5) the VAS rating. Table 1 shows the items used for each of the subscales. The insight item was not included in the analyses because it was a constant zero throughout the dataset presented here; this is likely due to patients accepting that they have an illness in order to participate in research.

2.1. Statistics

To understand the coherence (internal consistency reliability) of the full scale HDRS as well as that of its subscales, Cronbach's alpha was calculated with data from the 230-min time point following ketamine infusion because that was the primary point of interest. Baseline measures were not used due to the restricted variance likely for a clinical trial where participants would have to meet a certain severity criterion for entry into the study. Once the study began, subjects could have any score within the range of the scale in question, which would provide a more realistic assessment of the relationships among scale items.

To understand the degree of overlap among scales (criterion related validity), Pearson correlations were used to examine the inter-relationships among scores at 230 min and percent change in

Table 1
Items included in HDRS subscales.

Item	Description	Silverstone	Maier	Bech	Evans	Santen	McIntyre	Gibbons	Shortened
1	Depressed mood	•	•	•	•	•	•	•	•
2	Guilt		•	•	•	•	•	•	•
3	Suicide					•	•	•	•
4	Early insomnia								
5	Middle insomnia								
6	Late insomnia								
7	Work & activities		•	•	•	•	•	•	•
8	Motor retardation		•	•		•			•
9	Agitation		•					•	•
10	Anxiety: psychic	•	•	•	•	•	•	•	•
11	Anxiety: somatic				•		•	•	•
12	Somatic symptoms: G.I.								•
13	Somatic symptoms: general		•	•	•	•	•		•
14	Genital symptoms							•	•
15	Hypochondriasis								•
16	Weight loss								
17	Lack of insight								