Journal of Affective Disorders 150 (2013) 384–388

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research report

Severity classification on the Hamilton depression rating scale

Mark Zimmerman *, Jennifer H. Martinez, Diane Young, Iwona Chelminski, Kristy Dalrymple

Department of Psychiatry and Human Behavior, Brown Medical School, Rhode Island Hospital, 146 West River Street, Providence, RI, United States

ARTICLE INFO

Article history:
Received 31 January 2013
Accepted 19 April 2013
Available online 4 June 2013

Keywords:
Depression
Severity
Hamilton depression rating scale

ABSTRACT

Background Symptom severity as a moderator of treatment response has been the subject of debate over the past 20 years. Each of the meta- and mega-analyses examining the treatment significance of depression severity used the Hamilton Depression Rating Scale (HAMD), wholly, or in part, to define severity, though the cutoff used to define severe depression varied. There is limited empirical research establishing cutoff scores for bands of severity on the HAMD. The goal of the study is to empirically establish cutoff scores on the HAMD in their allocation of patients to severity groups.
Methods Six hundred twenty-seven outpatients with current major depressive disorder were evaluated with a semi-structured diagnostic interview. Scores on the 17-item HAMD were derived from ratings according to the conversion method described by Endicott et al. (1981). The patients were also rated on the Clinical Global Index of Severity (CGI). Receiver operating curves were computed to identify the cutoff that optimally discriminated between patients with mild vs. moderate and moderate vs. severe depression.
Results HAMD scores were significantly lower in patients with mild depression than patients with moderate depression, and patients with moderate depression scored significantly lower than patients with severe depression. The cutoff score on the HAMD that maximized the sum of sensitivity and specificity was 17 for the comparison of mild vs. moderate depression and 24 for the comparison of moderate vs. severe depression.
Limitations The present study was conducted in a single outpatient practice in which the majority of patients were white, female, and had health insurance. Although the study was limited to a single site, a strength of the recruitment procedure was that the sample was not selected for participation in a treatment study, and exclusion and inclusion criteria did not reduce the representativeness of the patient groups. The analyses were based on HAMD scores extracted from ratings on the SADS. However, we used Endicott et al.'s (1981) empirically established formula for deriving a HAMD score from SADS ratings, and our results concurred with other small studies of the mean and median HAMD scores in severity groups.
Conclusions Based on this large study of psychiatric outpatients with major depressive disorder we recommend the following severity ranges for the HAMD: no depression (0–7); mild depression (8–16); moderate depression (17–23); and severe depression (≥24).

1. Introduction

The importance of severity in selecting and evaluating the efficacy of treatment for depressed patients has received increased attention recently. The Third Edition of the American Psychiatric Association's (APA's) guidelines for the treatment of major depressive disorder recommended both psychotherapy and pharmacotherapy as monotherapies for depression of mild and moderate severity, and pharmacotherapy (with or without psychotherapy) for severe depression (American Psychiatric Association, 2010). The National Institute for Health and Clinical Excellence (NICE) updated guidelines for the treatment and management of depression discouraged the use of antidepressant medication as the initial treatment option for mild depression, and recommended medication together with empirically supported psychotherapy for moderate and severe depression (National Collaborating Center for Mental Health, 2009). As reported by van der Lem et al. (2011), treatment guidelines in the Netherlands also recommended pharmacotherapy as the first treatment option for severely depressed patients, and either pharmacotherapy or psychotherapy for mildly and moderately depressed patients. While the recommendations in these guidelines are not entirely consistent, they are unanimous in recommending medication as the treatment of choice for severe depression.

* Corresponding author. Tel.: 1 401 444 7098.
E-mail address: mzimmerman@lifespan.org (M. Zimmerman).

0165-0327/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jad.2013.04.028

Symptom severity as a moderator of treatment response has been the subject of ongoing debate since the publication of the results from the National Institute of Mental Health Treatment of Depression Collaborative Research Program (TDCRP) suggesting that psychotherapy was not as effective as medication in the acute treatment of severe depression (Elkin et al., 1995, 1989). In 1999, DeRubeis et al. (1999) noted that treatment guidelines from the American Psychiatric Association (1993) and Agency for Health Care Policy and Research (1993) recommended antidepressant medication, and not psychotherapy, for severe depression, and these recommendations were largely based on the findings of the TDCRP. The importance of severity was recently renewed in another context—the range of effectiveness of antidepressant medication. Kirsch et al. (2008) conducted a meta-analysis of antidepressant treatment trials in the FDA data base and found that drug-placebo differences increased as baseline severity increased, and concluded that antidepressants were only minimally more effective than placebo for mildly and moderately depressed patients. This conclusion was reinforced by the results of a mega-analysis of 6 studies by Fournier et al. (2010), though a recent, larger, mega-analysis found that the efficacy of two antidepressants, fluoxetine and venlafaxine, over placebo was independent of severity (Gibbons et al., 2012). The severity of depression has also continued to retain importance in moderating the effect of psychotherapy, though a recent meta-analysis of psychotherapy studies found that greater symptom severity did not predict poorer response in controlled studies examining the moderating effect of severity (Driessen et al., 2010).

Each of these meta- and mega-analyses used the Hamilton Depression Rating Scale (HAMD) (1960), wholly, or in part, to define severity, though the cutoff used to define severe depression varied. DeRubeis et al. (1999) conducted a mega-analysis of 4 studies comparing cognitive-behavioral therapy and medication. Following the precedent of the TDRCP, DeRubeis et al. defined severe depression as a cutoff of 20 or more on the 17-item HAMD. Likewise, the recent mega-analysis of placebo-controlled trials of fluoxetine and venlafaxine cited the TDRCP in using a cutoff >20 to define severe depression. Of note, no empirical justification was given in the TDRCP for using this threshold to define severe depression (Elkin et al., 1989). In fact, Elkin et al.'s (1989) did not refer to the patients scoring above 20 on the HAMD in absolute terms (i.e., having severe depression), but instead referred to them in relative terms (i.e., having more severe depression than the patients scoring 20 and below). In Kirsch et al. (2008) meta-analysis the authors noted that the mean baseline HAMD scores of the antidepressant efficacy trials were in the very severe range (i.e., ≥23 based on the APA's Handbook of Psychiatric Measures (Rush et al., 2008) for all but 2 of the 35 studies included in the analysis. Prior to the report by Kirsch and colleagues, Khan et al. (2002) examined the FDA data base, and also found that drug-placebo differences increased with increasing mean baseline HAMD scores. Khan et al. (2002) divided the studies into 3 groups based on pretreatment HAMD scores (≤24, 25–27, ≥28) without indicating the basis for using these cutoff scores to define the groups. Fournier et al. (2010) used the thresholds recommended in the APA's Handbook of Psychiatric Measures (Rush et al., 2008) to define grades of severity on the HAMD (mild to moderate≤18; severe 19 to 22; very severe ≥23).

In contrast to all of these studies, and the APA guidelines, most pharmacotherapy studies have used a cutoff of 25 on the 17-item HAMD to define severe depression (Dunner et al., 2005; Kasper, 1997; Montgomery et al., 2003; Schmitt et al., 2009; Shelton et al., 2007; Versiani et al., 2005), and this cutoff has been recommended by several experts to define severe depression (Hirschfeld, 1999; Montgomery and Lecrubier, 1999; Schatzberg, 1999).

Fundamental to all of these studies on the treatment implications of severity is the validity of the cutoffs on the HAMD to define the severity categories. In none of the discussion sections of these reports were questions raised about the cutoffs used to define the grades of severity. The APA's Handbook of Psychiatric Rating Scales (Rush et al., 2008) cited two studies in support of the cutoff scores to identify severity subtypes. One was a study examining the validity of deriving a HAMD equivalent score on the Schedule for Affective Disorders and Schizophrenia (Endicott et al., 1981). In fact, this study did not attempt to determine the cutoff scores on the HAMD indicating grades of severity. Rather, when examining the agreement between the extracted and original HAMD in classifying patients into severity categories, the authors used a cutoff of 25 to indicate severe depression (and a cutoff of 18 to distinguish mild and moderate depression). The second study cited as evidence for using a cutoff of 23 to indicate severe depression examined the association between HAMD scores and global ratings of severity in 59 depressed inpatients (Kearns et al., 1982). The authors did not derive (or recommend) cutoff scores corresponding to severity levels. In Figure 2 of their paper, the authors graphed the mean HAMD for patients rated at different levels of severity. Visual inspection of this figure suggests that very severe depression corresponded to a mean HAMD score of approximately 29 and severe depression corresponded to a mean HAMD score of 21. If these groups were combined, the mean HAMD for the severe category would be approximately 25. Thus, it is unclear why a cutoff of 19 was recommended in the APA Handbook to identify severe depression. We are aware of only 2 other small studies comparing HAMD scores to clinical global severity ratings. Knesevich et al. (1977) evaluated a sample of 26 outpatients, 9 of whom were rated in the severe range. Visual inspection of the figure plotting the distribution of scores suggests that the median score for these patients was 24. Muller et al. (2003) evaluated 85 depressed inpatients, 26 of whom were rated severe. The results of a receiver operating curve analysis to determine the optimal cutoff score on the HAMD to indicate severe depression found that a cutoff of 25 provided the best balance of sensitivity and specificity.

There is thus a limited amount empirical research establishing cutoff scores for bands of severity on the HAMD. Because of the significance accorded severity by treatment guidelines it is important to empirically establish cutoff scores on the HAMD in the allocation of patients to severity groups. Accordingly, in the present report from the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project we compared HAMD scores to clinician global ratings of severity in a large sample of depressed outpatients.

2. Methods

The Rhode Island MIDAS project represents an integration of research methodology into a community-based outpatient practice affiliated with an academic medical center (Zimmerman, 2003). A comprehensive diagnostic evaluation is conducted upon presentation for treatment. This private practice group predominantly treats individuals with medical insurance (including Medicare but not Medicaid) on a fee-for-service basis, and it is distinct from the hospital's outpatient residency training clinic that predominantly serves lower income, uninsured, and medical assistance patients. Data on referral source was recorded for the last 1600 patients enrolled in the study. Patients were most frequently referred from primary care physicians (30.0%), psychotherapists (16.1%), and family members or friends (18.8%). The Rhode Island Hospital institutional review committee approved the research protocol, and all patients provided informed, written consent.

The patients were interviewed by a diagnostic rater who administered a modified version of the Structured Clinical