Journal of Psychiatric Research 45 (2011) 213–219

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Validity of depression rating scales during pregnancy and the postpartum period: Impact of trimester and parity

Shuang Ji [a], Qi Long [a,*], D. Jeffrey Newport [b], Hyeji Na [b], Bettina Knight [b], Elizabeth B. Zach [b], Natalie J. Morris [b], Michael Kutner [a], Zachary N. Stowe [b,c]

[a] Department of Biostatistics and Bioinformatics, Emory University Rollins School of Public Health, 1518 Clifton Rd NE, Atlanta, GA 30322, USA
[b] Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 1365 Clifton Rd NE, Suite B6100, Atlanta, GA 30322, USA
[c] Department of Gynecology and Obstetrics, Emory University School of Medicine, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:
Received 15 March 2010
Received in revised form
12 May 2010
Accepted 17 May 2010

Keywords:
Depression
Pregnancy
Postpartum
Beck depression inventory
Hamilton rating scale for depression
Edinburgh postnatal depression scale

ABSTRACT

The objective of the current study was to delineate the optimal cutpoints for depression rating scales during pregnancy and the postpartum period and to assess the perinatal factors influencing these scores. Women participating in prospective investigations of maternal mental illness were enrolled prior to 28 weeks gestation and followed through 6 months postpartum. At each visit, subjects completed self-rated depression scales – Edinburgh Postnatal Depression Scale (EPDS) and Beck Depression Inventory (BDI) and clinician-rated scales – Hamilton Rating Scale for Depression ($HRSD_{17}$ and $HRSD_{21}$). These scores were compared to the SCID Mood Module for the presence of fulfilling diagnostic criteria for a major depressive episode (MDE) during 6 perinatal windows: preconception; first trimester; 2nd trimester; 3rd trimester; early postpartum; and later postpartum. Optimal cutpoints were determined by maximizing the sum of each scale's sensitivity and specificity. Stratified ROC analyses determined the impact of previous pregnancy and comparison of initial to follow-up visits. A total of 534 women encompassing 640 pregnancies and 4025 follow-up visits were included. ROC analysis demonstrated that all 4 scales were highly predictive of MDE. The AUCs ranged from 0.857 to 0.971 and were all highly significant ($p < .0001$). Optimal cutpoints were higher at initial visits and for multigravidas and demonstrated more variability for the self-rated scales. These data indicate that both clinician-rated and self-rated scales can be effective tools in identifying perinatal episodes of major depression. However, the results also suggest that prior childbirth experiences and the use of scales longitudinally across the perinatal period influence optimal cutpoints.

1. Introduction

Maternal depression during pregnancy and the postpartum period, i.e. perinatal depression, is a common problem that has been the focus of extensive investigation. Studies examining the prevalence of perinatal depression have demonstrated considerable variability that is a consequence, at least in part, of the assessment method used to identify the presence of depression, the timing of the assessment, and population characteristics (Gaynes et al., 2005). Authors of one review recommended that more precise determinants of the occurrence of perinatal depression are needed to estimate disease burden more accurately (Gaynes et al., 2005).

Depressive symptoms are common in pregnancy with most studies reporting rates comparable to non-gravid women (Cutrona, 1983; Kumar and Robson, 1984; Watson et al., 1984; Gotlib et al., 1989; O'Hara et al., 1991). A meta-analysis of depression during pregnancy (Bennett et al., 2004), utilizing data encompassing 19,284 gravidas from 21 studies in which depression was assessed by a structured clinical interview or self-rated scale such as the Beck Depression Inventory (BDI) (Beck et al., 1961), or the Edinburgh Postnatal Depression Scale (EPDS) (Cox et al., 1987), estimated the prevalence of depression as 7.4% in the first trimester, 12.8% in the second trimester, and 12.0% in the third trimester. However, the data were inadequate to render conclusions regarding comparative risk between trimesters. Furthermore, the authors reported that the BDI produced significantly higher prevalence estimates, whereas EPDS estimates were statistically equivalent to those of structured clinical interviews.

Depression during the postpartum period has also garnered considerable attention. An earlier meta-analysis by O'Hara and Swain (1996), encompassing 12,810 postpartum women from 59

* Corresponding author. Tel.: +1 404 712 9975; fax: +1 404 727 1370.
E-mail address: qlong@emory.edu (Q. Long).

0022-3956/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jpsychires.2010.05.017

studies utilizing a clinical interview or self-report scale, estimated the prevalence of depression in the postpartum period at 13.0%. Similar to the pregnancy data, self-report measures yielded higher estimates of postpartum depression than clinician-administered assessments. The postpartum timing of the assessment did not significantly affect the prevalence estimates in this meta-analysis. A review of the prevalence studies found that 7.1% may experience a major depressive episode (MDE) during the first 3 months postpartum (Gavin et al., 2005). Despite the historical assumption of increased vulnerability to depression in the postpartum period, the literature has not definitively demonstrated an increased risk (Gavin et al., 2005). In contrast, a recent large-scale epidemiological study provided evidence of increased risk for major depression in the postpartum period compared with non-pregnant/non-postpartum women (adjusted odds ratio: 1.52; 95% CI: 1.07–2.15) (Vesga-Lopez et al., 2008). Moreover, women are more likely to require psychiatric admission for depression during the postpartum period than outside the puerperium (Kendell et al., 1987; Munk-Olsen et al., 2006).

Numerous scales have been developed for identifying postpartum depression or risk factors for the development of postpartum depression (Beck, 1995; Fergerson et al., 2002; Morris-Rush et al., 2003; Perfetti et al., 2004; Austin et al., 2005). The EPDS has emerged as a well-validated and widely-utilized instrument for postpartum depression screening and detection. Conversely, validated tools to assess depression during pregnancy are lacking (Gaynes et al., 2005). By default, the EPDS, developed for postpartum use, has been increasingly used to identify depression during pregnancy (Adouard et al., 2005; Thoppil et al., 2005; Felice et al., 2006) and to screen for those at risk for developing depression during pregnancy (Evans et al., 2001; Rubertsson et al., 2005; Gordon et al., 2006). Beyond this *ad hoc* use of the EPDS, no scale exists to identify major depressive disorder during pregnancy. Moreover, only one screening test, an unvalidated scale consisting of only two items, has been developed specifically for depression in pregnancy (Campagne, 2004). Our group in collaboration (Altshuler et al., 2008), recently completed an individual item analysis of the 28 item Hamilton Rating Scale for Depression (HRSD) compared to SCID Mood Module to identify the items most predictive of an accurate identification of an episode of major depression across all trimesters of pregnancy. The seven items most predictive of the presence of depression were tested as a screening tool for depression during pregnancy (Altshuler et al., 2008).

The urgent need to identify reliable instruments for detecting perinatal depression is underscored by: 1) numerous reports of adverse obstetrical, neonatal, and developmental outcomes in association with maternal stress, depressive symptoms, and episodes of major depression during the perinatal period (Paton et al., 1977; Zuckerman and Bresnahan, 1991; Steer et al., 1992; Hedegaard et al., 1993; Pritchard and Teo, 1994; Orr and Miller, 1995; Chung et al., 2001; Andersson et al., 2004; Mancuso et al., 2004; Dayan et al., 2006; Diego et al., 2006; Neggers et al., 2006); 2) accurate diagnosis of an MDE during the peripartum is complicated by the fact that purportedly normal perinatal symptoms (e.g., fatigue, sleep disturbance, appetite and weight changes, diminished libido) potentially overlap with the neurovegetative symptoms comprising part of the diagnostic criteria for major depression; 3) lower estimates of maternal mental illness during pregnancy may be in part secondary to limited recognition (Vesga-Lopez et al., 2008); and 4) validated assessment tools are a requisite step in the design and completion of much needed controlled treatment studies during the perinatal period.

The overall aim of the current study was to provide clinicians and researchers alike with information regarding the sensitivity and specificity of commonly used depression rating scales during pregnancy and the postpartum period. The specific objectives of the study were: 1) to identify optimal cutpoints (maximizing the summation of sensitivity and specificity) for commonly used depression rating scales during each trimester of pregnancy and the postpartum period; 2) to determine whether previous pregnancy and childbirth experience influences the performance of the rating instruments; and 3) to determine whether repeated administration of a depression rating scale over the course of pregnancy and the postpartum period is associated with learning effects that alter the optimal cutpoints for the rating scales. With respect to these objectives, our *a priori* hypotheses were: 1) that the performance of the scales including optimal cutpoints would be altered during pregnancy, particularly during the third trimester when many of the physical symptoms of pregnancy most closely mirror the neurovegetative symptoms of depression; 2) that multigravid women (having previously experienced the physical sequelae of gestation) would be more likely to report physical symptoms of depression on a depression rating scale than primigravid women producing higher cutpoints on the scales during pregnancy; and 3) that optimal cutpoint scores would be impacted by repeated administration of both clinician-administered and self-rated depression scales.

2. Methods

2.1. Subjects

The study was conducted at the Women's Mental Health Program (WMHP) at the Emory University School of Medicine. Women with a lifetime history of mental illness participating in one of two prospective longitudinal perinatal investigations of the pharmacokinetics of psychotropic medications and/or maternal stress (P50 MH 68036; P50 MH 77928) were screened for inclusion in the current analysis. The schedule and methods for assessing maternal depression were identical in the two studies. Participants were enrolled no later than week 28 of gestation and evaluated at 4–6 week intervals across pregnancy and through 26 weeks postpartum. At each visit, subjects completed the self-rated BDI and EPDS. In addition, a research interviewer masked to treatment status administered the Structured Interview Guide (Williams, 1988) for the Hamilton Rating Scale for Depression (Hamilton, 1960) to obtain 17-item ($HRSD_{17}$) and 21-item ($HRSD_{21}$) scores and the Mood Module of the *Structured Clinical Interview for DSM-IV Axis I Disorders* (First et al., 2002). To ensure consistent administration of the clinician-rated instruments, research interviewers were trained to use a "rate as you see" approach when scoring items, eschewing any subjective judgment as to whether symptoms were due to depression or pregnancy/postpartum. Quarterly inter-rater reliability assessments were conducted throughout the course of both investigations to ensure maintenance of *kappa* statistics $\geq$0.8 on all clinician-administered instruments. All scales were coded with a HIPAA compliant identifier and entered into a centralized database. Subjects were included in the current analysis if they had two or more perinatal visits during which the *SCID* Mood Module and one or more of the depression rating scales were completed. The investigation was carried out in accordance with the latest version of the Declaration of Helsinki. The study was reviewed and approved by the Emory University Institutional Review Board. Informed consent of the participants was obtained after the nature of the procedures had been fully explained.

2.2. Data analysis

Each visit was assigned to one of 6 distinct perinatal epochs including: 1) preconception; 2) 1st trimester (0–12 weeks gestation); 3) 2nd trimester (13–24 weeks gestation); 4) 3rd trimester (25