Journal of Affective Disorders 196 (2016) 252–258

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

Differentiating between Transient and Enduring distress on the Edinburgh Depression Scale within screening contexts

Stephen Matthey [a,b,c,*,1]

[a] South Western Sydney Local Health District, Sydney, Australia
[b] University of Sydney, School of Psychology, Sydney, Australia
[c] UNSW, School of Psychiatry, Sydney, Australia

ARTICLE INFO

Article history:
Received 5 November 2015
Received in revised form
24 December 2015
Accepted 3 February 2016
Available online 18 February 2016

Keywords:
Perinatal screening
postnatal depression
Distress duration
Diagnosis

ABSTRACT

Background: Research has shown that women screened as being 'possibly depressed' on the Edinburgh Depression Scale consist of two distinct duration types: those with Transient distress, and those with Enduring distress. This paper reports on the exploration of antenatal data to ascertain if information from the initial EDS screening can help determine which women may have Transient, and which Enduring, distress after just a few weeks.

Methods: Data from three antenatal studies were explored, where the EDS had been given twice within a psychosocial screening setting. Repeat testing of the EDS, together with a diagnostic interview, was conducted 2–5 weeks later.

Results: Women with Enduring distress (those scoring high on both occasions) were significantly more likely to meet criteria for a depressive disorder than those with Transient distress. They also scored significantly and clinically meaningfully higher on their initial EDS, though no cut-off score was optimal in discriminating between the two duration categories. Differentiation could also not be made from the endorsement of the self-harm question, but was best when women were asked to predict how they would be feeling, and why.

Limitations: The data come from three studies just with English-speaking women with slightly different methodologies, producing information on a fairly small number of women with Transient (n=12–29) and Enduring (n=14–25) distress. In addition the EDS re-test interval of between 2 and 5 weeks was quite wide.

Conclusions: Clinical implications are that women who score high initially on the EDS are most likely to continue to score high (have Enduring distress) if they themselves think this will be the case, or if they only give wishful thinking as the reason as to why they think they will feel better. Research studies should also therefore analyse their data taking into account this duration category.

The Edinburgh Postnatal Depression Scale (EPDS, or EDS outside of the postnatal period: Cox et al., 1987; Cox et al., 1996) is arguably the most widely used self-report measure in clinical services and research studies for screening for, or detecting, possible depression in women in the perinatal period. It is also increasingly being used to screen for mood difficulties in fathers (eg., Massoudi et al., 2013; Matthey et al., 2001; Ramchandani et al., 2008). Validation studies have shown a variety of optimal cut-off scores to screen for diagnostic major or minor depression (Kozinszky and Dudas, 2015), though in English-speaking women the scores of 10 or more, or 13 or more, appear to be the most oft-used, both antenatally and postnatally, (eg., Bell et al., in press; Clark et al., 2015; Woolhouse et al., 2015). This is despite the lower value not necessarily being supported by the scientific evidence for the antenatal period (Matthey et al., 2006). The scale has also been shown to have some capacity to assist in the detection of women with anxiety (e.g., Matthey et al., 2013; Phillips et al., 2009; Swalm et al., 2010). While this has led some investigators to suggest that the anxiety subscale on the EDS (EDS-3A in English-speaking women: Matthey, 2008) should also therefore be scored separately from the total score to screen for possible anxiety disorders (eg., Phillips et al., 2009; Tuohy and McVey, 2008; though note the contrary view expressed by Reichenheim et al., 2011), this procedure has rarely been reported within clinical settings or research studies.

* Corresponding author at: Sydney South West Local Health District, Liverpool Hospital, Elizabeth Drive, Liverpool, Sydney 2170, NSW, Australia
E-mail addresses: stephen.matthey@sswahs.nsw.gov.au, stephenresearch@hotmail.com

[1] Note: Neither the author, or his institutions, have paid for this ms to be published.

http://dx.doi.org/10.1016/j.jad.2016.02.004
0165-0327/

The fact that the EPDS may also detect women with high levels of anxiety, rather than just depressive mood, has led to some researchers using the term 'distress', rather than 'depression' to describe high EDS scores (eg., Gawlik et al., 2013). This term will thus be used in this paper.

Thus there has been a plethora of research using this instrument, investigating routine assessment, treatment outcomes, and epidemiological issues, including determining which risk factors are associated with developing postnatal depression (eg., Dennis and Ross, 2006; Sutter-Dallay et al., 2004), and the probable rates of antenatal or postnatal depression or distress (e.g. Grote et al., 2010; Woolhouse et al., 2015). Nearly all such studies have used a single administration of the EDS to report on their findings, with the inherent untested assumption being that all women scoring high are similar with respect to their mood persistence – that is, all such women (and men) with a 'high' EPDS score will be experiencing an ongoing clinically significant mood difficulty, and hence they are 'possibly depressed'. This is also reflected in the methodology of studies determining the optimum cut-off score on the EDS, or commenting upon its relationship to diagnostic status, with a single administration of the EPDS being compared to diagnostic status, either determined at the same time or within a few weeks (eg., Boyce et al., 1993; Clark et al., 2015; Lee et al., 2003; Edmondson et al., 2010; Ghubash et al., 1997; Massoudi et al., 2013).

In part this assumption might seem to be supported by the moderate-high test–retest reliability coefficient of the EPDS. Thus Bergink et al. (2011) reported test–retest coefficients across the three trimesters of between .55 and .63, which they considered to be "high" (p. 388), and Kirpinar et al. (2010) reported a correlation between 7 days and 6 weeks postpartum of .60, interpreting this to indicate that the scale had good psychometric properties and showed stability of mood during this period.

However, within a screening context, where the focus is on 'high scorers', this test–retest reliability coefficient can be very misleading. If most low scoring women continue to score low over the brief time interval (of a few weeks), then, given that they are the majority of the participants (usually 70–90% score 'low'), a moderate-high retest reliability will be obtained regardless of whether or not the 'high scorers' remain stable or not.

Evidence has in fact shown that repeating the administration of the EPDS within a short time-period (eg., 2–4 weeks), when no specialist intervention has occurred, reveals that there are two distinct categories of women initially scoring 'high' on the scale: those with Transient distress, and those with Enduring distress. The former are women who initially score in the distressed range on the EDS (i.e., above the validated cut-off score), but when retested after just a few weeks no longer score in this distressed range, while the latter are those high-scoring women who continue to score in the distressed range a few weeks later.

This phenomenon has been found in studies conducted postnatally (Ballestrem et al., 2005; Morrell et al., 2009; Wickberg and Hwang, 1996), during pregnancy (Matthey and Ross-Hamid, 2012), and within a clinical setting (Harvey and Pun, 2007). These studies have shown that around half of the women scoring high initially on the EPDS show Transient distress when the scale is re-administered just a few weeks later and no specialist intervention has occurred.

This finding appears to be a valid representation of the women's emotional experience, and is not due to measurement error. Thus, in the Harvey and Pun (2007) study, women themselves reported that they no longer needed the clinical appointment for their mood as they were feeling substantially better. In the Matthey and Ross-Hamid (2012) study, the size of EDS score reduction across the two periods for nearly all of the Transient distress women was more than the reliable change index for the EDS (3–4 points: see Measures: EPDS section below), indicating that the change in score was unlikely to be a reflection of measurement error. Also in this antenatal study the reasons given by these women for their 'improved' mood had good face validity. They included having had re-assuring test results in the intervening period; their morning sickness having subsided since the first assessment; and having adjusted to the idea of being pregnant over this short time span.

Also importantly is that the antenatal study (Matthey and Ross-Hamid, 2012) found that almost all (~95%) of women scoring low on the EDS at their first administration (ie., below the cut-off score) continued to score low just a few weeks later. And this therefore produced a high test–retest correlation coefficient in this sample of .73 (not previously reported), despite around half of the high scorers no longer scoring high just a few weeks later.

What, then, are the clinical and research implications of the evidence that there are in fact two categories of women who score in the distressed range at the first administration of the EDS?

Clinically it would strongly support the need to repeat the administration of the EPDS for any woman who initially scores high within a screening procedure. This practice was indeed recommended by the first author of the EPDS (Cox and Holden, 2003), though, as stated, it is rarely done. Referral to a specialist team based simply upon an initial high EDS score will result in many women being referred who will have substantially improved in their mood prior to their appointment, thus resulting in an inefficient use of the clinical service. As Mann and Gilbody (2011) state, in discussing accuracy in psychosocial screening, "(if there is a large number of false positives, they will require) unnecessary follow-up, a situation which (will) infer burden and negative impact to local health service delivery" (p. 393).

Such inappropriate referrals may also create undue anxiety in women, and their partners, who are led to believe, by such a referral, that there is something 'wrong' with them, or that they have a possible 'mental illness' (Thombs and Stewart, 2014).

Within a screening context it would thus be extremely useful to know if there was some way of distinguishing which women were likely to experience Transient distress, and which Enduring distress, at the time of the first EDS administration. While repeat testing of high-scoring women is now being recommended by some services to elucidate this (eg., NSW Department of Health, 2009), doing so poses significant practical problems. For example, it may be difficult to do such repeat testing where women have just a single contact with the hospital's antenatal clinic, and then have their care managed outside of the hospital system. Or if the next hospital appointment for the woman is not for many weeks, trying to make phone contact with her to assess her ongoing mood will often entail significant resources (i.e., multiple phone calls to make contact at a convenient time to re-administer the EDS) that thus diminish the service's capacity to provide good clinical care to others.

Within research contexts the reporting of rates of high scorers as indicating 'probable rates of depression', as is currently done based upon the single administration, will be incorrect if in fact the rates of caseness in Transient and Enduring distress women are substantially different. This then could have implications for further in-depth analyses investigating the nature of pnd (eg., risk-factor analyses, structural equation modelling) as not all women with the single high EDS score should be deemed to 'have the condition' which such analyses assume. And naturally if prevention or treatment programs are studied that mix the two categories of women, the findings will be less robust than if only women with Enduring distress were included as participants.

Data on this topic will thus be investigated to answer the following research questions:

i) Are the rates of 'caseness', as defined by DSM diagnostic