Postpartum depression screening: a Comment on Leung et al.

Leung et al.\(^1\) reported that women screened with the Edinburgh Postnatal Depression Scale (EPDS) 2 months post partum were significantly less likely to score $\geq 10$ on the EPDS 6 months later than control group women. Women screened with the EPDS were referred for depression treatment if they had EPDS scores $\geq 10$, reported suicidal ideation, or were assessed as ‘probably’ depressed based on a separate clinical assessment, described as ‘observing participants’ expression and behavior, enquiring about feelings, appetite, sleep pattern, childcare and suicidal ideas’ (p. 294). Control group women were similarly referred for treatment if they were evaluated as probably depressed via the same clinical assessment. Several reasons, however, suggest that the results reported by Leung et al.\(^1\) should be viewed cautiously.

First, whereas screening is intended to select patients for more comprehensive assessment,\(^2,3\) in this study, all patients in both groups received a clinical assessment. No women identified as possibly having depression in either group were further evaluated to determine whether or not they had depression and whether depression treatment was indicated. Rather, women were only evaluated to determine the format of treatment they would receive. Assuming a 12% rate of postnatal depression\(^1\) and EPDS $\geq 10$ sensitivity and specificity of 92 and 77%, respectively,\(^4\) just over one-third of treated women likely had depression.

Despite this, the standardized mean difference (SMD) effect size for EPDS scores at 6 months was 0.34 (calculated from their Table 2), even though only 24% of screening group patients received treatment (55/231) and even though 11 patients in the control group were treated. Assuming no outcome differences between treated patients in the screening and control groups and non-treated patients in the two groups, this is roughly equivalent to SMD = 1.81 for the 44 additional patients treated in the screened group—many times larger than results from even well-controlled depression treatment trials. The SMD from 30 collaborative depression care intervention trials, for example, was 0.25.\(^5\) Ultra-large treatment effects from relatively small numbers of treated patients, as in Leung et al., often fail to replicate.\(^6\)

Finally, in their 2005 trial registration (NCT00251342), Leung et al. declared two primary outcome measures, the EPDS and the General Health Questionnaire-12 (GHQ-12) (http://clinicaltrials.gov/ct2/show/NCT00251342). In their article, however, they stated that there was only one primary outcome, EPDS scores (statistically significant) by which to judge screening effectiveness. They listed GHQ-12 scores (not statistically significant) as secondary. Clinical trial registration requirements were implemented to improve research transparency, including reducing the likelihood that null or equivocal trials are presented as positive in the research literature.\(^7,8\) Changing the status of outcome variables from primary to secondary based on trial results misleads research users about the trial design, and, generally, raises concerns about the fidelity of the trial’s reporting. In the case of the trial by Leung et al., based on its registered design, it was an equivocal, not a positive trial.

Post partum depression is an important problem, and screening may be a solution. This trial, however, did not establish whether or not this is the case.

Acknowledgements

Dr. Thombs is supported by a New Investigator Award from the Canadian Institutes of Health Research (CIHR) and an Établissement de Jeunes Chercheurs award from the Fonds de la Recherche en Santé Québec.

References


