Diagnostic Accuracy and Uptake of the PHQ-9 Depression Scale in Pregnancy and the Postpartum Period

Larry Wang1, Kurt Kroenke2

1Indiana University School of Medicine, 2Indiana University School of Medicine, Division of General Internal Medicine and Regenstrief Institute

Purpose: Perinatal depression (PND) is a prevalent and disabling problem both during pregnancy and the postpartum period globally. While the legacy screening measure has been the Edinburgh Postnatal Depression Scale (EPDS), the PHQ-9 is the most widely used depression measure in clinical practice and is considered a viable alternative in many PND guidelines. This systematic review examines the criterion and convergent validity of the PHQ-9 as a screener for PND.

Methods: The following databases were searched in June 2020: MEDLINE, Embase, and PsychInfo. Search terms were the PHQ-9 AND perinatal AND depression. Variants of each were also included as an OR statement such as “patient health questionnaire” as a variant of the PHQ-9. Only studies published in English and after 2001 were accepted as the PHQ-9 was not available as a screening test before then.

Results: A total of 35 articles were eligible for criterion (n=10) or convergent validity (n=25). Studies with criterion validity had sensitivity and specificity ranging from 65 to 95 and 62 to 89, respectively. For the 7 articles using the standard PHQ-9 cut point ≥ 10, median sensitivity and specificity were 0.82 and 0.84, respectively. The median AUC was 0.88. Of the 9 correlation values shared between the PHQ-9 and EPDS, values range from 0.36 to 0.77 with a median of 0.60. The median Cohen’s kappa value of 0.535 in four studies showed good agreement.

Conclusion: Of the criterion validity articles, the sensitivity, specificity, and AUC indicate the PHQ-9 may be used to screen for PND. The convergent validity comparison articles between the PHQ-9 and EPDS also support use of the PHQ-9 in perinatal populations.