Journal of Affective Disorders 228 (2018) 49–59

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

Developing and validating a perinatal depression screening tool in Kenya blending Western criteria with local idioms: A mixed methods study

Eric P. Green[a,*], Hawa Tuli[a], Edith Kwobah[b], D. Menya[c], Irene Chesire[c], Christina Schmidt[a]

[a] Duke University, United States
[b] Moi Teaching and Referral Hospital, Kenya
[c] Moi University, Kenya

ABSTRACT

Background: Routine screening for perinatal depression is not common in most primary health care settings. The U.S. Preventive Services Task Force only recently updated their recommendation on depression screening to specifically recommend screening during the pre- and postpartum periods. While practitioners in high-income countries can respond to this new recommendation by implementing one of several existing depression screening tools developed in Western contexts, such as the Edinburgh Postnatal Depression Scale (EPDS) or the Patient Health Questionnaire-9 (PHQ-9), these tools lack strong evidence of cross-cultural equivalence, validity for case finding, and precision in measuring response to treatment in developing countries. Thus, there is a critical need to develop and validate new screening tools for perinatal depression that can be used by lay health workers, primary health care personnel, and patients.

Methods: Working in rural Kenya, we used free listing, card sorting, and item analysis methods to develop a locally-relevant screening tool that blended Western psychiatric concepts with local idioms of distress. We conducted a validation study with a random sample of 193 pregnant women and new mothers to test the diagnostic accuracy of this scale along with the EPDS and PHQ-9.

Results: The sensitivity/specificity of the EPDS and PHQ-9 was estimated to be 0.70/0.72 and 0.70/0.73, respectively. This compared to sensitivity/specificity of 0.90/0.90 for a new 9-item locally-developed tool called the Perinatal Depression Screening (PDEPS). Across these three tools, internal consistency reliability ranged from 0.77 to 0.81 and test-retest reliability ranged from 0.57 to 0.67. he prevalence of depression ranges from 5.2% to 6.2% depending on the clinical reference standard.

Conclusion: The EPDS and PHQ-9 are valid and reliable screening tools for perinatal depression in rural Western Kenya, the PDEPS may be a more useful alternative. At less than 10%, the prevalence of depression in this region appears to be lower than other published estimates for African and other low-income countries.

1. Introduction

Depression is a leading cause of disability worldwide, yet access to timely assessment and treatment is very limited in many low-income settings, especially in rural communities. Depression affects men and women, young and old, but women who experience depression during pregnancy or in the year after childbirth are a particularly underserved population. The prevalence of perinatal depression among women living in poor countries ranges widely, possibly exceeding 30% in rural settings (Villegas et al., 2011).

Depression among pregnant women and new mothers has been linked to increased maternal morbidity and mortality (Oates, 2003; Khalifeh et al., 2016), poor infant health (Field et al., 2004; Rahman et al., 2016; Grigoriadis et al., 2013; Surkan et al., 2016; Gelaye et al., 2016), and poor early childhood outcomes—such as developmental, cognitive, and emotional delays (Beck, 1998; Junge et al., 2017; Gentile, 2017)—making it a significant public health concern. Few public health systems currently have the resources to treat perinatal depression, but recent work has shown that cognitive behavioral interventions delivered by lay health workers are efficacious (Rahman et al., 2008; Joshi et al., 2014). Before such treatments can be delivered at scale, however, it is essential to overcome many barriers, including barriers to screening for depression.

Routine screening for perinatal depression is not common in most primary health care settings. The U.S. Preventive Services Task Force only recently updated their recommendation on depression screening to

* Corresponding author.
E-mail address: eric.green@duke.edu (E.P. Green).

https://doi.org/10.1016/j.jad.2017.11.027
Received 3 February 2017; Received in revised form 25 July 2017; Accepted 7 November 2017
Available online 02 December 2017
0165-0327/

specifically recommend screening during the pre- and postpartum periods (Siu and the US Preventive Services Task Force, 2016). While practitioners in high-income countries can respond to this new recommendation by implementing one of several existing depression screening tools developed in Western contexts, such as the Edinburgh Postnatal Depression Scale (EPDS) or the Patient Health Questionnaire-9 (PHQ-9), these tools lack strong evidence of cross-cultural equivalence, validity for case finding, and precision in measuring response to treatment in developing countries (Sweetland et al., 2014; Tsai et al., 2013). Thus, there is a critical need to develop and validate new screening tools for perinatal depression that can be used by lay health workers, primary health care personnel, and patients. Our study contributes to this effort by attempting to validate the EPDS and PHQ-9 in rural Kenya, while at the same time developing and validating a new instrument that blends items from existing screening tools with local idioms of distress (Kohrt et al., 2011).

2. Methods

2.1. Setting and participants

We conducted this prospective study in Bungoma, Kenya. This rural county is situated in what used to be known as Western Province. When the 2010 Constitution of Kenya was enacted in 2013, 47 counties in a new devolved system of government replaced the existing 8 provinces. Bungoma is one of the largest counties in this new system. It is home to more than 1.6 million residents, nearly half of whom live in poverty (Wiesmann et al., 2014).

We recruited participants for two main study activities: (i) eight focus group discussions to develop a locally-anchored set of screening items and (ii) individual assessments to narrow the set of items and validate the new measure and two existing screening tools. A purposive sample of 12 women were invited to participate in the focus group discussions; women were eligible to participate if they were at least 18 years old and receiving maternity services from a particular primary health clinic (public dispensary) in Bungoma East subcounty. All 38 community health volunteers (CHVs) serving the clinic's catchment area were invited to participate in separate discussion sessions.

For the validation study, we drew a random sample of 210 pregnant women and new mothers (from a sampling frame of 276) from the 27 villages wholly or partially located within a 2-kilometer radius from the clinic. All women had to be at least 18 years of age. Pregnant women in their second or third trimesters were eligible, as were new mothers 1–6 months postpartum. Women who miscarried or experienced a stillbirth or infant death linked to their most recent pregnancy were excluded from the study. 193 women completed questionnaires and semi-structured clinical interviews.

2.2. Measures

2.2.1. Screening survey (Index Tests)

We identified 17 measures commonly used to assess perinatal depression (see Table A1 in the Appendix), created a database of 365 items, assigned every screening item a short cover term (e.g., crying, unhappy, heart racing), and reviewed each cover term for exact and approximate duplicates. Out of the initial 365 screening items, we identified 171 unique cover terms and wrote an index card (with English and Kiswahili writing) for each term in preparation for our focus group discussions. Through these discussions we created a 60-item survey that included several index tests: the Edinburgh Postnatal Depression Scale, the Patient Health Questionnaire-9, items from other existing screening tools, and new items generated by the focus groups. In addition to the screening items, the survey also included demographic questions from the Phase 6 and Phase 7 Demographic and Health Household and Woman's questionnaires (DHS Program. (n.d.-a), 2016).

Edinburgh Postnatal Depression Scale. The most commonly used screening instrument for perinatal depression is the 10-item self-report Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987). The first validation study was conducted with 84 postnatal women in the United Kingdom and reported sensitivity of 86%, specificity of 78%, and a positive predictive value of 73%. A systematic review of 37 EPSD validation studies conducted between 1987 and 2008, however, revealed great heterogeneity in diagnostic sensitivity and specificity between studies for all cutoff points (Gibson et al., 2009).

Tsai et al. (2013) recently extended this evaluation of the EPDS with a new systematic review of 25 studies that screened for perinatal depression in Africa; 16 of the 25 studies included in this review used the EPDS. The authors noted that the median estimated coefficient alpha of the EPDS was 0.84, and they calculated a pooled sensitivity and specificity of 0.94 and 0.77 (cut-off ≥ 9) from 14 studies that assessed criterion validity. None of the included studies was conducted in Kenya.

Patient Health Questionnaire-9. Another brief depression screening that is often used to assess perinatal depression is the Patient Health Questionnaire-9 (PHQ-9). At least two studies have investigated Kiswahili translations of the PHQ-9 in Kenya. Omoro et al. (2006) demonstrated an association between PHQ-9 scores, TNM stage (Classification of Malignant Tumors), and scores on a cancer-specific quality of life scale. Monahan et al. (2009) found a correlation between scale scores and patient responses to the question, “In general how would you rate your overall health right now?” Neither study assessed validity by comparing results to a gold-standard, such as a clinical structured interview.

2.2.2. Criterion reference: Structured Clinical Interview for DSM-5

We used the Structured Clinical Interview for DSM-5, Research Version to diagnose cases of depression (SCID-5-RV; First et al., 2015). The SCID-5-RV is designed to be customized, and we opted to administer the non-patient overview, Module A on mood episodes with specifiers, Module BC for psychotic screening, and Module D for the differential diagnosis of mood disorders—all translated into Kiswahili prior to use. Table A2 in the Appendix details the modifications we made to each module.

The target condition was Major Depressive Episode (MDE). To meet criteria for a current MDE according to the DSM-5 (American Psychiatric Association, 2013), a woman had to experience at least 5 of 9 symptoms—including depressed mood (A1) or diminished interest or pleasure (A2)—during the same 2-week period within the past 1 month (Criterion A) and report that these symptoms caused clinically significant distress or impairment in functioning (Criterion B). Four Kenyan counselors (2 Bachelorâs-level, 2 Masterâs-level) investigated all cases in which a general medical condition, substance abuse, or medication could be the etiological factor (Criterion C). Counselors also used Module B/C to determine if psychotic symptoms were primarily accounted for by a DSM-5 Psychotic Disorder (Criterion D). Counselors did not assess Criterion E of Module D (i.e., rule out manic or hypomanic episode); therefore, we could only diagnose MDE not Major Depressive Disorder.

If a woman's symptoms suggested a depressive disorder but the woman did not meet Criterion A for MDE, the counselor assessed Criteria B-D to possibly diagnose “Other specified depressive disorder”.

2.2.3. Alternate criterion reference: local diagnosis

In addition to using the SCID-5-RV to diagnose depression as defined by the DSM-5, we also asked counselors to use their clinical judgment and asked women to self-report on their well-being.

Clinical judgment of diagnosis and functioning. Counselors responded to the following prompt to record a ‘local’ diagnosis that was not tied to the DSM criteria: “In your clinical judgment, do you think that this woman is ‘depressed’?” Counselors also rated each woman's social and occupational functioning using the SOFAS rating scale included in the SCID-RV-5. SOFAS ratings can range from 0 to 100, with 100