Research report

Postpartum depressive symptoms across time and place: Structural invariance of the Self-Reporting Questionnaire among women from the international, multi-site MAL-ED study

CrossMark

Laura L. Pendergast [a,*], Rebecca J. Scharf [b], Zeba A. Rasmussen [c], Jessica C. Seidman [c], Barbara A. Schaefer [d], Erling Svensen [e,f], Fahmida Tofail [g], Beena Koshy [h], Margaret Kosek [i], Muneera A. Rasheed [j], Reeba Roshan [h], Angelina Maphula [k], Rita Shrestha [l], Laura E. Murray-Kolb [d], The MAL-ED Network Investigators

[a] School Psychology Program, Temple University, Ritter Annex 265, 1301 Cecil B. Moore Avenue, Philadelphia, PA 19122, USA
[b] University of Virginia, USA
[c] Fogarty International Center, USA
[d] The Pennsylvania State University, USA
[e] University of Bergen, Norway
[f] Haydom Lutheran Hospital, Tanzania
[g] International Centre for Diarrhoeal Disease Research, Bangladesh
[h] Christian Medical College, Vellore, India
[i] The Johns Hopkins University Bloomberg School of Public Health, USA
[j] Aga Khan University, Pakistan
[k] University of Venda, South Africa
[l] Institute of Medicine, Tribuhvan University, Kathmandu, Nepal

ARTICLE INFO

ABSTRACT

Background: The Self-Reporting Questionnaire (SRQ) is a screening instrument that has been shown to be an effective measure of depression in postpartum women and is widely used in developing nations.
Methods: The SRQ was administered to 2028 mothers from eight nations at two time points: one and six months postpartum. All data were obtained from the Interactions of Malnutrition and Enteric Infections: Consequences for Child Health and Development (MAL-ED) study. The sample included women from MAL-ED sites in Bangladesh, Brazil, India, Nepal, Pakistan, Peru, South Africa, and Tanzania. This study examined three aspects of validity of SRQ scores including (a) structural validity, (b) cross-cultural invariance, and (c) invariance over time.
Results: A 16-item, one-factor structure with items reflecting somatic symptoms removed was deemed to be superior to the original structure in this postpartum population. Although differential item functioning (DIF) across sites was evident the one-factor model was a good fit to the data from seven sites, and the structure was invariant across the one- and six-month time points.
Limitations: Findings are based on data from self-report scales. No information about the clinical status of the participants was available.
Conclusions: Overall, findings support the validity of a modified model of the SRQ among postpartum women. Somatic symptoms (e.g., headaches, not sleeping well) may not reflect internalizing problems in a postpartum population. Implications for researchers and practitioners are discussed.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Postpartum depression is a common condition that results in considerable impairment for affected women and families (see Miller, 2002, for a review). The consequences of postpartum depression are increasingly being recognized worldwide. However, most of the research on postpartum depression has been conducted in English-speaking, Westernized nations. International research examining the correlates and effects of postpartum depression in novel settings is needed but is dependent upon the validation of measures of depressive symptomology among postpartum populations throughout the world.

The Self-Reporting Questionnaire (SRQ; Beusenberg and Orley, 1994; World Health Organization, 1994, 1980; WHO) is a self-report measure of internalizing symptoms that was developed by the World Health Organization as a tool for improving mental health screening worldwide. The SRQ has been used extensively throughout the world, and some studies have examined the validity of the scale within single nations (e.g., Ethiopia; Hanlon et al., 2008a; Stewart et al., 2009). However, to our knowledge, the structure of the SRQ has never been examined in any setting using contemporary best practice factor analytic techniques (i.e., principal axis factoring, oblique rotations for correlated factors, use of multiple criteria for determining the number of factors to retain, and the use of exploratory and confirmatory factor analyses on independent subsamples). This study sought to examine the structure and invariance of SRQ scores across eight international sites (Dhaka, Bangladesh; Fortaleza, Brazil; Vellore, India; Bhaktapur, Nepal; Loreto, Peru; Naushahro Feroze, Pakistan; Venda, South Africa; and Haydom, Tanzania) and at two time points (one and six months postpartum).

1.1. Overview of postpartum depression

1.1.1. Assessment and diagnosis

According to O'Hara and McCabe (2013), "postpartum depression is often defined as an episode of major depressive disorder (but sometimes including minor depression) that occurs in the postpartum period." For women with postpartum depression, symptoms of sadness, tearfulness, anxiety, irritability or worry often overshadow the early months of motherhood (e.g., Gjerdingen and Yawn, 2007; O'Hara and Swain, 1996).Within major diagnostic systems (e.g., *DSM-V*; American Psychiatric Association, 2013, and *ICD-10*; World Health Organization, 1990), episodes of depression that occur during pregnancy or the postpartum period are diagnosed using the same criteria as those of depressive episodes occurring at other times. If depressive symptoms begin shortly after delivery (4–6 weeks) postnatal specifiers can be used in the diagnostic label, but the same criteria are used regardless of whether or not the onset occurs in the postpartum period (see Austin, 2010, for a review).

There is disagreement in the literature regarding whether postpartum depression should be considered distinct from other forms of depression – particularly in regard to somatic symptoms (e.g., appetite changes and difficulty sleeping). Somatic symptoms commonly occur during the early months of childrearing – even among non-depressed mothers (see O'Hara and McCabe, 2013, for a review). Some researchers argue that somatic symptoms complicate the assessment of internalizing disorders in postpartum women and that somatic items should not be included on rating scales (e.g., Cox et al., 1987). However, others contend that somatic symptoms are valid indicators of depression in postpartum populations (O'Hara et al., 2014). Although there is much speculation in the literature, very little research has been conducted to evaluate whether somatic symptoms are valid reflections of postpartum depression.

1.1.2. Child outcomes

Proper assessment, identification, and treatment of postpartum depression is essential given some of the harmful outcomes associated with the condition (Marshall and Thompson, 2014; O'Hara and McCabe, 2013; Patel et al., 2002; Wisner et al., 2002). Maternal depressive symptoms are known risk factors for poor child development outcomes (e.g., Beck, 1998; Cooper and Murray, 1998). Various studies have documented impairments in children of depressed parents on an array of factors such as growth (Duarte

et al., 2012; Santos et al., 2010), mental health (Muzik and Borovska, 2010), illness (Casey et al., 2004; Turney, 2011), feeding (Casey et al., 2004; Ndokera and MacArthur, 2011; Rahman et al., 2004) and cognitive development (Azak, 2012). Because there is an emerging literature documenting links between postpartum depression and differences in child development (Conroy et al., 2012; Foss et al., 2004; Korja et al., 2008; Paulson et al., 2006, 2009; Podestá et al., 2013; Quevedo et al., 2012; Walker et al., 2007), international research in this area is timely and important. However, international research on postpartum depression is dependent upon the existence of measures of depressive symptoms that produce reliable and valid scores among postpartum women in the countries in which they are used. This study focused on the validity of the SRQ in assessing maternal depressive symptoms in eight nations around the world.

1.1.3. Self-Reporting Questionnaire

The version of the Self-Reporting Questionnaire (SRQ; Beusenberg and Orley, 1994) used in this study comprises 20 items that assess psychological disturbances related to depressive symptoms occurring within the past four weeks. Questions are answered with a simple "yes" or "no" and the instrument may be self-administered or interviewer-administered.

1.1.4. Development of the SRQ

The SRQ was developed by the WHO as part of a collaborative study. In 1975, teams of psychiatrists, public health workers, and researchers from Colombia, India, Senegal, and Sudan began the study and later teams from Brazil, Egypt and the Philippines joined. The original items on the SRQ were drawn from several existing rating scales.

1.1.5. SRQ convergent and predictive validity research

The convergent and predictive validity of SRQ scores has been examined in several countries around the world. For example, studies conducted in Malawi (Akena et al., 2012), Ghana (Weobong et al., 2009), Ethiopia (Hanlon et al., 2008), and Brazil (Mari and Williams, 1986) have reported moderate to high rates of sensitivity and specificity (generally .70–.85). Additionally, the SRQ has been used successfully in many studies screening for maternal depressive symptoms (e.g., Stewart et al., 2009; Ghubash and Abou-Saleh, 1997; Nakku et al., 2007).

1.1.6. SRQ structural validity research

The structural validity of SRQ scores has been investigated in several settings. In Rwanda, Scholte et al. (2011) applied exploratory factor analysis (EFA) with principal axis factoring and varimax rotation and identified a five-factor structure with correlated factors. Multiple group confirmatory factor analysis (CFA) was used to test factorial invariance over a three-month interval, and the findings supported temporal invariance. However, in the Scholte et al. study, the EFA and CFA were conducted on the same sample which may have led to conflated results. In India, Sen (1987) examined the SRQ using principal components analysis with varimax rotation and reported a seven-factor structure ($n=202$). In Brazil, Santos et al. (2009; $n=3190$) and Iacoponi and Mari (1989; $n=1182$) separately identified a four-factor structure for the SRQ. Both Brazilian studies used principal components analysis, varimax rotation, and eigenvalue rule of one for retention. Finally, in Afghanistan, Ventevogel et al. (2007) applied principal axis factoring ($n=116$) and varimax rotation and identified a two-factor structure. To date, findings related to the

factor structure of the SRQ have varied widely across studies with solutions ranging from two to seven factors.

Unfortunately, many prior studies of the validity of SRQ scores inappropriately relied on small sample sizes (more than 5 cases per item are needed; e.g., Velicer and Fava, 1998). Moreover, all prior studies used at least one technique that has been shown to increase the risk of over-factoring (e.g., principal components analysis, applying varimax rotation with correlated factors, and the eigenvalue > 1 factor retention rule; see Henson and Roberts (2006) and Kline (2005) for reviews of current best practices in EFA and CFA, respectively). Best practices in factor analysis suggest that researchers can reduce the risk of over-factoring and increase the replicability of their findings by using principal axis factoring, oblique rotations with correlated factors, and multiple methods for determining the number of factors to retain (e.g., scree plot, parallel analysis, and minimum average partials; e.g., Fabrigar et al., 1999). The SRQ has never been examined using contemporary best practice techniques in factor analysis and, consequently, the true structure of the scale in any setting is unknown. Given the ubiquity with which the SRQ has been utilized in the empirical literature, further international research of the validity of the scale across cultures using these techniques is essential.

1.1.7. Present study

The primary objective of this study was to use best practice factor analytic techniques to examine the SRQ structure in the international MAL-ED (2009) sample, the invariance of the scale across the 8 international study sites, and the stability of the structure over time. We sought to address three research questions: (a) What is the factor structure of the SRQ in this sample?, (b) Is the factor structure of the SRQ invariant across cultural groups[1] (as determined based on study site)?, and (c) Is the factor structure of the SRQ invariant over time (one and six months postpartum)?

2. Method

2.1. Overview of MAL-ED study

The *Interactions of Malnutrition and Enteric Infections: Consequences for Child Health and Development* (MAL-ED; http://mal-ed.fnih.org/) study is a multi-disciplinary, observational, prospective, clinical/field study conducted at 8 international sites. The MAL-ED study aims to identify the periods during the first two years of life where malnutrition and specific enteric infections are associated with the greatest effect on growth and development. Factors evaluated for their effects include enteric and other infections, micronutrient levels, dietary intake, socioeconomic status, maternal depressive symptoms, and the home environment.

2.2. Participants

A total of 2028 women across the eight sites were included in this study. Demographic information is provided in Table 1. All study procedures were approved by the Institutional Review Boards at each institution affiliated with a study site and each partnering institution.

2.3. Scale translations

At six sites, harmonized scale translation procedures were used. Teams of bilingual and culturally knowledgeable researchers translated the scale. The translated versions were sent to bilingual individuals who were unaffiliated with the MAL-ED study and back-translated into English. Discrepancies between the original and back-translated versions were addressed on a case-by-case basis. Two sites (Pakistan and Tanzania) did not participate in the scale translation process and instead used versions that had been translated, back-translated, and validated in prior research (Svenson and Nordgreen, 2014).

2.4. Procedures

The SRQ was administered to each mother by a trained interviewer at one month and six months postpartum and took approximately 15 min to complete. The collection window for each assessment was ± 15 days (e.g., one month ± 15 days). Instructions were read aloud to mothers verbatim in the appropriate local language.

3. Results

3.1. Preliminary analyses

Means and standard deviations for each site and each time point are reported in Table 2. Inspection of frequencies from the Brazil site revealed an extremely high rate of zero-scores for the items. For example, 11 of the 20 items were endorsed by fewer than 10 Brazilian mothers (229 Brazilian mothers participated in this study). The pattern of responses for data from the Brazil site was clearly and dramatically different from that of the other seven sites, and it was not possible to factor analyze the data from the Brazil site because of the lack of variance. Therefore, data from the Brazil site were not included in subsequent analyses.

After exclusion of the Brazilian cases, 1799 cases remained in the sample. The sample from the one-month follow-up was randomly divided into two subsamples: one for exploratory analyses (EFA; $n = 200$) and one for confirmatory analyses (CFA, structural invariance, and longitudinal invariance; $n = 1599$). Data from the six-month time point ($n = 1713$ after deletion of 209 Brazilian cases) were used along with the CFA subsample from the one-month time point ($n = 1599$) for the temporal invariance analyses.

3.2. EFA

Common factor analysis (principal axis extraction (PAF) and promax rotation) was selected instead of principal components analysis because the purpose of this study was to identify the latent factor structure of the SRQ (Fabrigar et al., 1999). Due to the dichotomous nature of the data, all EFA analyses were also run using weighted least squares estimation for comparative purposes. The findings were nearly identical; thus, findings from PAF analyses are reported here. Several procedures were used to determine the number of factors to retain for rotation, including parallel analysis (Horn, 1965; Watkins, 2006), minimum average partials (MAP; Velicer, 1976), and the visual scree test (Cattell, 1966). Factor pattern coefficients ≥ .30 were considered salient. Factors with a minimum of three salient pattern coefficients and internal consistency of scores ≥ .70 were considered adequate.

The scree plot and MAP analysis both indicated that one factor should be retained, while parallel analysis suggested three factors. Therefore, both the one factor and three factors solutions were evaluated. The three factors solution was discarded because the

---

[1] In this paper, the term "cultural group" is used interchangeably with "study site." For example, women from MAL-ED study site in Haydom, Tanzania, were considered to be Tanzanian in regard to cultural group. The study sites did not comprise samples that were demographically representative of the countries in which they were located and should not be interpreted as such.

Table 1
Demographics of study sample, by site.

	Bangladesh	Brazil	India	Nepal	Peru	Pakistan	South Africa	Tanzania
N	262	229	247	238	299	274	282	259
Mean age	24.9 (5.0)	24.8(5.5)	23.9(4.1)	26.5(3.7)	24.2(6.0)	28.0 (5.9)	26.6 (7.0)	28.5 (6.7)
Mean parity	2.0 (1.1)	2.3 (1.4)	2.2 (1.4)	1.7 (.8)	2.5 (1.7)	3.6 (2.5)	2.4 (1.4)	4.3 (2.6)
Mean # of child deaths	.1 (.3)	0 (.2)	.1 (.5)	.1 (.2)	.1 (.3)	.3 (.9)	.1 (.3)	.3 (.7)
Mean years of schooling	4.5 (3.2)	9.2 (2.9)	6.9 (3.9)	8.2 (4.0)	7.7 (4.0)	3.0 (4.0)	10.2 (1.9)	5.1 (2.8)
Never married (%)	0	13	0	0	10	0	39	2

Note: *N* = unique women in study sample, included all women in BRF who completed an SRQ at either one or six months. Standard deviations are in parentheses. All figures reflect data at enrollment.

Table 2
Descriptive statistics for SRQ total scores at the one and six-month time points.

	One-month postpartum			Six-month post-partum		
	n	*M*	*SD*	*n*	*M*	*SD*
Bangladesh	262	4.24	3.38	237	5.15	3.87
India	245	3.94	3.58	236	4.53	3.67
Nepal	232	2.41	2.59	236	4.61	3.26
Peru	299	2.28	2.24	270	2.60	2.41
Pakistan	273	5.63	3.43	263	5.49	3.86
South Africa	240	2.78	2.50	234	2.61	2.60
Tanzania	248	2.63	2.83	237	2.19	2.70
Total	1799	3.43	3.19	1713	3.88	3.48

Note: Study sites did not comprise samples that were demographically representative of the countries in which they were located and should not be interpreted as such.

Table 3
Pattern coefficients from one-factor EFA solution of SRQ scores.

Items	Pattern coefficients	h^2
Trouble thinking clearly	.57	.33
Feeling nervous/tense/worried	.54	.32
Feeling worthless	.54	.38
Thinking of ending life	.48	.37
Crying more often	.48	.30
Loss of interest	.48	.33
Feeling unhappy	.47	.26
Difficulty enjoying activities	.45	.29
Tiring easily	.45	.36
Hands shake	.43	.32
Daily work suffering	.38	.30
Unable to play a useful part in life	.37	.33
Difficulty making decisions	.37	.37
Always tired	.35	.38
Easily frightened	.34	.26
Loss of appetite	.32	.24
Has headaches	.27	.16
Uncomfortable stomach feelings	.25	.21
Has digestive problems	.20	.21
Has difficulty sleeping	.20	.15
Eigenvalue	4.15	
Reliability	.77	

reliability estimates of the identified factors were unacceptably low (< .41). The one-factor solution was retained. The one-factor solution seemed to tap internalizing symptoms in general (e.g., depression and anxiety) and was named the Internalizing factor. Pattern coefficients and communalities are provided in Table 3. Of the 20 items on the SRQ, 16 had salient pattern coefficients on the Internalizing factor (α=.79). The four items that did not load on the factor reflected somatic symptoms including headaches, stomachaches, digestive difficulties, and sleep difficulties. In a sample of recently postpartum women caring for young infants, it is not surprising that somatic symptoms (e.g., not sleeping well; uncomfortable feelings in the stomach) are not reflective of internalizing symptoms. As such, four items related to somatic symptoms with loadings < .30 were excluded from subsequent analyses.

3.3. Baseline CFA

All CFAs and invariance analyses were conducted on covariance matrices of raw data using MPlus 6.12. Due to deviations from normality and the dichotomous nature of the data, robust weighted least squares estimation and the Satorra–Bentler χ^2 statistic were used. A baseline CFA with no covariates was conducted on the one-factor, 16-item model identified in EFA. Multiple criteria were used to evaluate fit (Tanaka, 1993) including comparative fit index (CFI) ≥ .90 and root mean square error of approximation (RMSEA) values ≤ .06 (Hu and Bentler, 1995; Kline, 2005). The baseline CFA model was determined to have a good fit. All fit indices are reported in Table 6. Factor loadings ranged from .42 to 1.

3.4. Cross-cultural invariance

Multiple indicators multiple causes (MIMIC; Jöreskog and Goldberger, 1975) modeling was utilized to assess differential item functioning (DIF) across seven cultural groups. MIMIC modeling is an extension of CFA that allows for the influence of multiple factors (e.g., nationality, age, etc.) to be evaluated when determining whether

latent factors (e.g., internalizing symptoms) function similarly across groups (Muthén, 1989). In this study, we used MIMIC models to determine whether cultural group influenced the extent to which participants were more or less likely to endorse particular SRQ items relative to others with similar levels of depressive symptoms overall. Data from the one-month follow-up were used for examination of cross-cultural invariance. A depiction of the MIMIC model tested in this study is provided in Fig. 1.

3.5. Overview of DIF analyses

Preliminary analyses were conducted to identify a subset of DIF-free items to define the factor in subsequent analyses using the method illustrated by Woods et al. (2009). Subsequently, DIF analyses occurred in a stepwise fashion as described by Jones (2006) involving estimation of multiple interim MIMIC models. First, a baseline model was examined containing only the identified anchor items and each cultural variable was entered as a covariate. Then, the remaining (non-anchor) items were added one at a time and tested for DIF. The modification indices were inspected after each analysis, and if it indicated that freeing a path between an item and a cultural variable would improve model fit (i.e., significantly increase the χ^2 value) then the aforementioned path was freed. If the path between the cultural variable and the item was statistically significant, then the item was considered to have DIF. Subsequently, a final model was tested wherein DIF was allowed when deemed appropriate based on the analyses described above. A listing of items and cultural groups for which DIF was allowed can be found in Table 4. The final model, which

allowed for DIF, demonstrated good fit to the data (e.g., CFI=.96; RMSEA=.04; findings are reported in Table 5).

To allow for estimation of effect sizes associated with DIF, the final MIMIC model was re-estimated using maximum likelihood estimator with robust standard errors (MLR). Using MLR estimation allowed us to calculate proportional odds ratios (OR) estimates. The OR estimates were interpreted according to the guidelines suggested by Cole et al. (2000) whereby OR values > 2.0 or < .5 were considered indicative of meaningful DIF. For example an OR estimate of 2.0 would indicate that a given cultural group is twice as likely to endorse a particular item, after controlling for overall internalizing symptoms, and would be considered indicative of meaningful DIF. OR estimates are reported in Table 4.

3.5.1. Relationships between cultural group and internalizing symptoms total scores. The final model contained statistically significant paths between the internalizing symptoms factor and four of the seven cultural groups: the Bangladesh site (λ=.51; SE=.06); the India site (λ=.18; SE=.07); the Pakistan site (λ=.45; SE=.06); and the Peru site (λ= −.22; SE=.06). In the Tanzania, South Africa,

Fig. 1. MIMIC model examining one-factor SRQ structure with MAL-ED study site as a covariate.

and Nepal sites, there were no significant site differences in SRQ total scores.

3.5.2. Item-level DIF findings. Every cultural group had some items with meaningful DIF ranging from one item (the Tanzania site) to six items (the Bangladesh, Pakistan, and South Africa sites). Detailed information regarding item-level DIF findings for all sites are reported in Table 4, and an overview is provided here. In the Bangladeshi subsample, significant and meaningful DIF was identified on six items, and the total SRQ score was significantly higher than was found for the other groups ($p < .001$). After controlling for overall levels of internalizing symptoms, Bangladeshi participants were less likely to endorse five items (e.g., *thinking of ending life*) and more likely to endorse one item (*unable to play a useful part in life*).

In the Indian subsample the total SRQ score was higher than the mean of the other groups ($p < .001$), and three items had significant and meaningful DIF. After controlling for overall levels of internalizing symptoms, Indian respondents were more likely to endorse three items (*feeling worthless*, *thinking of ending life*, and *hands shaking*).

Among Nepali participants, the total SRQ score did not significantly differ from the overall mean, but three items did exhibit significant and meaningful DIF. After controlling for the total internalizing symptoms score, Nepali participants were more likely to endorse three items (e.g., *feeling unhappy*).

With regard to the Peruvian subsample the total SRQ score was significantly lower than the overall mean ($p < .005$), and five items had significant, meaningful DIF. Peruvian respondents were more
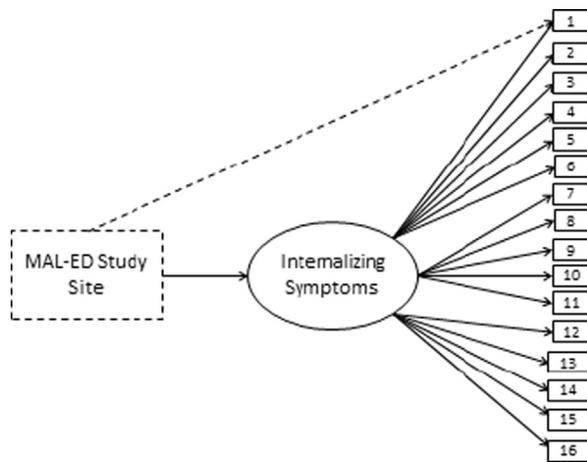
Table 5
Fit statistics from single-site CFAs.

Model	χ^2	df	p	CFI	RMSEA
Bangladesh	168.78	102	.000	.93	.07
India	245.59	102	.000	.97	.07
Nepal	156.63	102	.001	.93	.05
Peru	235.44	102	.000	.91	.07
Pakistan	180.15	102	.000	.96	.05
South Africa	199.98	102	.000	.92	.07
Tanzania	125.84	102	.000	.98	.03
Overall SRQ 1 month	564.51	173	.000	.96	.04

Table 4
Parameter estimates and differential item functioning of SRQ items across seven sites.

Items	Bangladesh PE	SE	India PE	SE	Nepal PE	SE	Peru PE	SE	Pakistan PE	SE	South Africa PE	SE	Tanzania PE	SE
Trouble thinking clearly	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Feeling nervous/tense/worried	–	–	–	–	–	–	**.46(3.23)**	.08	–	–	–	–	–	–
Feeling worthless	–	–	**.46 (6.83)**	.08	–	–	–	–	–	–	–	–	–	–
Thinking of ending life	−.39 (.26)	.12	**.60 (6.10)**	.09	–	–	–	–	–	–	–	–	–	–
Crying more often	−.26 (.57)	.09	–	–	–	–	–	–	**.34(2.89)**	.09	**.47(3.34)**	.10	–	–
Loss of interest	–	–	–	–	–	–	**−.54(.22)**	.12	**.55 (3.43)**	.08	–	–	–	–
Feeling unhappy	**−.42 (.40)**	.08	–	–	**.49(4.15)**	.09	**−.67(.16)**	.12	–	–	–	–	–	–
Difficulty enjoying activities	–	–	–	–	–	–	–	–	–	–	**−.66 (.39)**	–	**.18(.40)**	.08
Tiring easily	–	–	–	–	**.44 (2.76)**	.07	**−.48(.23)**	.12	.27 (1.90)	.07	–	–	–	–
Hands shake	–	–	**.35(2.30)**	.09	–	–	–	–	**.44(2.92)**	.09	**−.75(.28)**	.15	–	–
Daily work suffering	**−.54(.22)**	.10	.25(1.77)	.08	–	–	–	–	–	–	–	–	–	–
Unable to play a useful part in life?	**.38(2.60)**	.08	–	–	–	–	–	–	–	–	**.58(5.75)**	.09	–	–
Difficulty making decisions	**−.66(.30)**	.09	–	–	–	–	–	–	**−.55(.35)**	.09	**.36(3.57)**	.09	**.30(2.54)**	.08
Always tired	–	–	–	–	–	–	**.50(2.68)**	.09	–	–	**−.31(.47)**	.06	–	–
Easily frightened?	**−.56(.41)**	.10	.27(1.88)	.09	–	–	.31(1.90)	.09	–	–	–	–	–	–
Loss of appetite	.34(1.58)	.09	–	–	**.39(2.06)**	.10	–	–	**−.51(.40)**	.11	–	–	.43(1.94)	.09
Total number of meaningful DIF items	6		3		3		5		6		6		1	

Note: – indicates that there was no DIF and the parameters were not estimated. PE=standardized parameter estimate; SE=standard error; odds ratios are reported in parentheses, and items deemed to have meaningful DIF (OR > 2.0 or < .50) are in bold. Study sites did not comprise samples that were demographically representative of the countries in which they were located and should not be interpreted as such.

Table 6
Fit statistics for invariance of SRQ scores across time points.

Model		χ^2	df	p	$\Delta\chi^{2*}$	Δ df	p	CFI	Δ CFI	RMSEA
1	SRQ 1 month	564.51	173	.00	–	–	–	.961	–	.038
2	SRQ 6 month	711.17	168	.00	–	–	–	.933	–	.043
3	Configural invariance	1575.89	346	.00	–	–	–	.932	–	.046
4	Metric invariance	1623.77	362	.00	47.88	16	.397	.930	.002	.046
5	Scalar invariance	1636.45	363	.00	60.56	17	.285	.930	.002	.046
6	Residual invariance	1710.89	370	.00	136.00	24	.019	.926	.006	.047

likely to endorse two items (e.g., *feeling nervous/tense/worried*) and less likely to endorse three items (e.g., *loss of interest*) after controlling for overall internalizing symptoms.

Overall, the Pakistani subsample had the highest mean level of internalizing symptoms ($p < .001$). Controlling for overall internalizing symptoms, Pakistani respondents were more likely to endorse four items (e.g, *crying more often*) and less likely to endorse three items (e.g., *difficulty enjoying activities*).

With regard to the South African sample, the SRQ total scores were comparable to the overall mean, but six items exhibited significant and meaningful DIF. After accounting for the total internalizing symptoms score, South African participants were more likely to endorse three items (e.g., *crying more often*) and less likely to endorse three items (e.g., *difficulty enjoying activities*).

In the Tanzanian subsample the SRQ total scores did not significantly differ from the overall, multi-group mean, but one item had significant and meaningful DIF. Specifically, participants from Tanzania were more likely to indicate that they had *difficulty making decisions* even after controlling for internalizing symptoms overall.

3.5.3. Concluding analyses. In summary, DIF was identified for some items within each site. Therefore, CFAs were conducted separately for each site. The final, 16-item, one-factor model was found to be a good fit at each site. Site-specific fit statistics are reported in Table 5. The final model was accepted and used for temporal invariance testing, allowing for DIF wherever indicated.

3.6. Temporal invariance

Temporal invariance analyses were conducted to determine whether changes in the data are true or the result of a change in the construct over time (Chan, 1998). Prior to beginning analyses examining temporal invariance, the one-factor model was tested separately with data from each time point: one month postpartum and six months postpartum. Subsequently, multi-group CFAs were used to evaluate structural invariance across time points. Invariance of the one-factor structure was assessed by applying increasingly restrictive constraints across time points to examine (a) configural invariance (all parameters were free to vary across groups), (b) metric invariance (factor loadings constrained to be equal), (c) scalar invariance (intercepts of item parcels also constrained to be equal), and (d) residual invariance (residuals also constrained to be equal; e.g., Dimitrov, 2010; Meredith, 1993). Change in Satorra–Bentler chi-square ($\Delta\chi^2$) and change in CFI (Δ CFI) were used to compare nested models.

Fit statistics are provided in Table 6. Findings indicated that the one-factor SRQ model was a good fit to the data at both the one- and six-month time points. Moreover, configural, metric, scalar, and residual invariance across time points was supported (e.g., Δ CFI $\leq$ .02).

4. Discussion

4.1. Overview of findings

In this study, a one-factor model that tapped an overall internalizing symptoms factor was identified and showed a good fit to the data. Four of the original 20 items were excluded due to low factor loadings. The four excluded items may have had low factor loadings because they reflect somatic symptoms and are not suitable for use with a postpartum population. In other words, symptoms such as "not sleeping well" and having "uncomfortable feelings" in the stomach may be typical for women caring for newborns and recovering from childbirth and may not reflect internalizing problems.

This was the first study to identify a one-factor SRQ solution. However, the findings from prior studies have been highly inconsistent in regard to the number of factors identified – even among sites within the same geographic region. All prior studies used one or more now obsolete factor analytic technique that is known to lead to over-factoring (e.g., using eigenvalue rule of one for factor retention). This was the first study to examine the SRQ using contemporary, best practice factor analytic techniques along with multi-site analyses. Therefore, the one-factor structure identified in this study may be more generalizable than those found in prior research.

Findings from MIMIC analyses indicated that the 16-item, one-factor structure was largely invariant across sites. Each site had at least one item that was non-invariant, but the majority of the items were DIF-free at each site. Finally, findings from analyses examining temporal invariance suggest that the factor structure is invariant across the one- and six-month time points.

4.2. Implications

4.2.1. Measurement implications

These findings suggest that a 16-item, one-factor model of the SRQ may be appropriate for use with postpartum women in many countries. Research on other aspects of validity (e.g., predictive validity) will be needed to buttress these findings. Although the model fits the data relatively well at all seven sites, because significant and meaningful DIF was identified, direct comparisons across sites without accounting for DIF would be inappropriate.

4.2.2. SRQ in Brazil

Because of limited variability among participants from the Brazilian subsample in this study, these findings raise concerns about the validity of SRQ scores among postpartum women in the MAL-ED Brazilian site and perhaps for postpartum women in the Northern region of Brazil in general. The Brazilian MAL-ED site is located at the Clinical Research Unit at the Federal University of Ceara in Fortaleza. Fortaleza is the capital city of the state of Ceará

in the Northeastern region of Brazil, has a population of approximately 2.1 million, and is one of the poorest regions of Brazil.

Postpartum depression prevalence rates in Brazilian studies have varied widely (e.g., 1–7%; Chavez, 2012; Cantilino et al., 2010) by region and might possibly be associated with environment (e.g., climate, demographics), measurement, and cultural issues. Although other researchers have examined the validity of the SRQ among postpartum women in Brazil, most prior studies were based on samples from the Southern region of Brazil or hospital-based samples. It is possible that sampling differences or differences associated with the different geographic regions may have contributed to the discrepant findings, but more research is needed to better understand these findings.

4.2.3. Theoretical and practical implications

Overall, symptoms of postpartum depression appear to be similar across these seven diverse sites. Difficulty thinking clearly, feeling nervous and tense, and feeling worthless appear to be the most prominent features of depression in this postpartum sample as evidenced by the relatively high factor loadings. Conversely, somatic symptoms that are often considered to be indicative of depression and anxiety in non-postpartum populations (i.e., difficulty sleeping, uncomfortable feelings in the stomach, digestive difficulties, and headaches) were not reflective of internalizing symptoms in this population. As such, it may be advisable for researchers and practitioners assessing internalizing disorders in postpartum populations to interpret somatic symptoms with extreme caution – particularly if our findings are replicated in future research.

4.3. Strengths, limitations, and future directions

This study has several notable strengths, including the use of a large, diverse, international sample representing women from eight international sites. To our knowledge, this is the first study to examine the structural validity of the SRQ using contemporary factor analytic criteria and to explicitly test structural invariance across international sites. Finally, the longitudinal nature of this study allowed for examination of the validity of SRQ scores at two time points: one and six months postpartum.

One limitation of this study is that the SRQ was the only measure of internalizing symptoms that was administered. Future research that incorporates a criterion measure would be very valuable. Additionally, future research utilizing an IRT approach to evaluate both uniform and non-uniform DIF would be interesting and informative.

The findings from this study raise several important questions for future research. Further study examining the reasons why DIF emerged on some items would be a fruitful avenue for scholarly inquiry. For example, in this study, women in the Indian subsample were more than 6 times more likely than their international counterparts to respond affirmatively when asked "Has the thought of ending your life been on your mind?" Future research examining DIF on this item in particular will be crucial to inform interventions and to better understand the distress these mothers appear to be experiencing. Finally, the findings from this study suggest that somatic symptoms may not be reflective of internalizing problems among postpartum women. Presently, somatic symptoms are included in the diagnostic criteria for depression in postpartum populations in *DSM-V.* It will be crucial for future investigators to conduct research to better understand the relationship between somatic symptoms and depression in postpartum women (or lack thereof) and to evaluate the appropriateness of current diagnostic criteria.

5. Conclusions

In this study, a one-factor, 16-item model of SRQ scores was identified. Overall, this study underscores the validity of SRQ scores among recently postpartum women across geographically and culturally diverse settings. Although studies comparing SRQ scores across sites may need to account for DIF, these findings provide support for the international use of the alternative one-factor model (with somatic symptoms omitted) among postpartum women.

Anita K.M. Zaidi (Aga Khan University, Naushahro Feroze, Pakistan), Sajid Soofi (Aga Khan University, Naushahro Feroze, Pakistan), Ali Turab (Aga Khan University, Naushahro Feroze, Pakistan), Didar Alam (Aga Khan University, Naushahro Feroze, Pakistan), Shahida Qureshi (Aga Khan University, Naushahro Feroze, Pakistan), Aisha K. Yousafzai (Aga Khan University, Naushahro Feroze, Pakistan), Asad Ali (Aga Khan University, Naushahro Feroze, Pakistan), Imran Ahmed (Aga Khan University, Naushahro Feroze, Pakistan), Sajad Memon (Aga Khan University, Naushahro Feroze, Pakistan), Muneera Rasheed (Aga Khan University, Naushahro Feroze, Pakistan). North America: in the United States: Michael Gottlieb (Foundation for the NIH, Bethesda, MD, USA), Mark Miller (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Karen H. Tountas (Foundation for the NIH, Bethesda, MD, USA), Rebecca Blank (Foundation for the NIH, Bethesda, MD, USA), Dennis Lang (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Stacey Knobler (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Monica McGrath (Fogarty International Center/ National Institutes of Health, Bethesda, MD, USA), Stephanie Richard (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Jessica Seidman (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Zeba Rasmussen (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Ramya Ambikapathi (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Benjamin McCormick (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Stephanie Psaki (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Vivek Charu (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Jhanelle Graham (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Gaurvika Nayyar (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Viyada Doan (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Leyfou Dabo (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Danny Carreon (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Archana Mohale (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Christel Host (Fogarty International Center/National Institutes of Health, Bethesda, MD, USA), Dick Guerrant (University of Virginia, Charlottesville, VA, USA), Bill Petri (University of Virginia, Charlottesville, VA, USA), Eric Houpt (University of Virginia, Charlottesville, VA, USA), Jean Gratz (University of Virginia, Charlottesville, VA, USA), Leah Barrett (University of Virginia, Charlottesville, VA, USA), Rebecca Scharf (University of Virginia, Charlottesville, VA, USA), Laura Caulfield (Johns Hopkins University, Baltimore, MD, USA), William Checkley (Johns Hopkins University, Baltimore, MD, USA), Margaret Kosek (Johns Hopkins University, Baltimore, MD, USA), Pablo Penataro Yori (Johns Hopkins University, Baltimore, MD, USA), Gwenyth Lee (Johns Hopkins University, Baltimore, MD, USA), Ping Chen (Johns Hopkins University, Baltimore, MD, USA), Robert Black (Johns Hopkins University, Baltimore, MD, USA), Laura Murray-Kolb (Pennsylvania State University, University Park, PA, USA), Barbara Schaefer (Pennsylvania State University, University Park, PA, USA), William Pan (Duke University, Durham, NC, USA).

References

Akena, D., Joska, J., Obuku, E.A., Amos, T., Musisi, S., Stein, D.J., 2012. Comparing the accuracy of brief versus long depression screening instruments which have been validated in low and middle income countries: a systematic review. BMC Psychiatry 12, 187, http://dx.doi.org/10.1186/1471-244X-12-187.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed.. American Journal of Psychiatry, retrieved from ⟨http://ajp.psychiatryonline.org/article.aspx?articleID=158714⟩.

Austin, M.P., 2010. Classification of mental health disorders in the perinatal period: future directions for DSM-V and ICD-11. Arch. Women's Ment. Health 13, 41–44, http://dx.doi.org/10.1007/s00737-009-0110-5.

Azak, S., 2012. Maternal depression and sex differences shape the infants' trajectories of cognitive development. Infant Behav. Dev. 35, 803–814, http://dx.doi.org/10.1016/j.infbeh.2012.07.017.

Beck, C.T., 1998. The effects of postpartum depression on child development: a meta-analysis. Arch. Psychiatr. Nurs. 12 (1), 12–20.

Beusenberg, M., Orley, J., 1994. A User's Guide to the Self Reporting Questionnaire (SRQ). World Health Organization, Geneva, Switzerland.

Casey, P., Goolsby, S., Berkowitz, C., Frank, D., Cook, J., Cutts, D., Meyers, A., 2004. Maternal depression, changing public assistance, food security, and child health status. Pediatrics 113, 298–304, http://dx.doi.org/10.1542/peds.113.2.298.

Cattell, R.B., 1966. The scree test for the number of factors. Multivar. Behav. Res. 1, 245–276, http://dx.doi.org/10.1207/s15327906mbr0102_10.

Chan, D., 1998. The conceptualization and analysis of change over time: an integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). Organ. Res. Methods 1 (4), 421–483.

Cole, S.R., Kawachi, I., Maller, S.J., Berkman, L.F., 2000. Test of item-response bias in the CES-D scale: experience from the New Haven EPESE Study. J. Clin. Epidemiol. 53, 285–291, http://dx.doi.org/10.1016/S0895-4356(99)00151-1.

Conroy, S., Pariante, C.M., Marks, M.N., Davies, H.A., Farrelly, S., Schacht, R., Moran, P., 2012. Maternal psychopathology and infant development at 18 months: the

impact of maternal personality disorder and depression. J. Am. Acad. Child Adolesc. Psychiatry 51, 51–61, http://dx.doi.org/10.1016/j.jaac.2011.10.007.

Cooper, P.J., Murray, L., 1998. Postnatal depression. Br. Med. J. (Clin. Res. Ed.) 316 (7148), 1884–1886.

Cox, J.L., Holden, J.M., Sagovsky, R., 1987. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br. J. Psychiatry: J. Ment. Sci. 150, 782–786, http://dx.doi.org/10.1192/bjp.150.6.782.

Dimitrov, D.M., 2010. Testing for factorial invariance in the context of construct validation. Meas. Eval. Couns. Dev. 43, 121–149, http://dx.doi.org/10.1177/0748175610373459.

Duarte, C.S., Shen, S., Wu, P., Must, A., 2012. Maternal depression and child BMI: longitudinal findings from a US sample. Pediatric Obes. 7, 124–133, http://dx.doi.org/10.1111/j.2047-6310.2011.00012.x.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J., 1999. Evaluating the use of exploratory factor analysis in psychological research. Psychol. Methods 4, 272–299, http://dx.doi.org/10.1037/1082-989X.4.3.272.

Foss, G.F., Chantal, A.W., Hendrickson, S., 2004. Maternal depression and anxiety and infant development: a comparison of foreign-born and native-born mothers. Pub. Health Nurs. 21 (3), 237–246, http://dx.doi.org/10.1111/j.0737-1209.2004.21306.x.

Ghubash, R., Abou-Saleh, M.T., 1997. Postpartum psychiatric illness in Arab culture: prevalence and psychosocial correlates. Br. J. Psychiatry: J. Ment. Sci. 171, 65–68, http://dx.doi.org/10.1192/bjp.171.1.65.

Gjerdingen, D.K., Yawn, B.P., 2007. Postpartum depression screening: importance, methods, barriers and recommendations for practice. J. Am. Board Fam. Med., 280–288 (Retrieved from)⟨http://www.medscape.com/viewarticle/558440⟩.

Hanlon, C., Medhin, G., Alem, A., Araya, M., Abdulahi, A., Hughes, M., Prince, M., 2008. Detecting perinatal common mental disorders in Ethiopia: validation of the self-reporting questionnaire and Edinburgh Postnatal Depression Scale. J. Affective Disord. 108, 251–262, http://dx.doi.org/10.1016/j.jad.2007.10.023.

Henson, R.K., Roberts, J.K., 2006. Use of exploratory factor analysis in published research: common errors and some comments on improved practice. Educ. Psychol Meas. 66, 393–416, http://dx.doi.org/10.1177/0013164405282485.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30, 179–185, http://dx.doi.org/10.1207/s15327906mbr0801_8.

Hu, L.T., Bentler, P.M., 1995. Evaluating model fit. In: Hoyle, R.H. (Ed.), Structural Equation Modeling: Concepts, Issues, and Applications. Sage Publications, Thousand Oaks, CA, pp. 76–99.

Iacoponi, E., Mari, J.J., 1989. Reliability and factor structure of the Portuguese version of Self-Reporting Questionnaire. Int. J. Soc. Psychiatry 35, 213–222, http://dx.doi.org/10.1177/002076408903500301.

Jones, R.N., 2006. Identification of measurement differences between English and Spanish language versions of the Mini-Mental State Examination. Detecting differential item functioning using MIMIC modeling. Med. Care 44, S124–S133, http://dx.doi.org/10.1097/01.mlr.0000245250.50114.0f.

Jöreskog, K.G., Goldberger, A.S., 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. J. Am. Stat. Assoc. 70 (351a), 631–639.

Kline, R.B., 2005. Principles and practice of structural equation modeling. Methodol. Soc. Sci. 2, 366, http://dx.doi.org/10.1038/156278a0.

Korja, R., Savonlahti, E., Ahlqvist-Björkroth, S., Stolt, S., Haataja, L., Lapinleimu, H., Lehtonen, L., 2008. Maternal depression is associated with mother–infant interaction in preterm infants. Acta Paediatr. 97, 724–730, http://dx.doi.org/10.1111/j.1651-2227.2008.00733.x.

MAL-ED, 2009. Interactions of Malnutrition and Enteric Infections: Consequences for Child Health and Development. Retrieved from ⟨http://mal-ed.fnih.org/⟩.

Mari, J.J., Williams, P., 1986. A validity study of a psychiatric screening questionnaire (SRQ-20) in primary care in the city of Sao Paulo. Br. J. Psychiatry: J. Ment. Sci. 148, 23–26, http://dx.doi.org/10.1192/bjp.148.1.23.

Marshall, E.J., Thompson, A.P., 2014. Shedding light on the difficulties and challenges experienced by mothers of infants. Aust. Psychol. 49 (1), 44–53, http://dx.doi.org/10.1111/j.1742-9544.2012.00093.x.

Meredith, W., 1993. Measurement invariance, factor analysis, and factorial invariance. Psychometrika 58, 523–543, http://dx.doi.org/10.1007/BF02294825.

Miller, L.J., 2002. Postpartum depression. J. Am. Med. Assoc. 287, 762–765.

Muzik, M., Borovska, S., 2010. Perinatal depression: implications for child mental health. Ment. Health Fam. Med. 7 (4), 239–247.

Muthén, B.O., 1989. Latent variable modeling in heterogeneous populations. Psychometrika 54, 557–585, http://dx.doi.org/10.1007/BF02296397.

Nakku, J.N., Nakasi, G., Mirembe, F., 2007. Postpartum major depression at six weeks in primary health care: prevalence and associated factors. Afr. Health Sci. 6 (4), 207–214.

Ndokera, R., MacArthur, C., 2011. The relationship between maternal depression and adverse infant health outcomes in Zambia: a cross-sectional feasibility study. Child: Care Health Dev. 37, 74–81, http://dx.doi.org/10.1111/j.1365-2214.2010.01129.x.

O'Hara, M.W., McCabe, J.E., 2013. Postpartum depression: current status and future directions. Annu. Rev. Clin. Psychol. 9, 379–407, http://dx.doi.org/10.1146/annurev-clinpsy-050212-185612.

O'Hara, M.W., Swain, A.M., 1996. Rates and risk of postpartum depression—a meta-analysis. Int. Rev. Psychiatry 8, 37–54, http://dx.doi.org/10.3109/09540269609037816.

O'Hara, M.W., Williamson, J.A., Watson, D., 2014. The structure of postpartum depression, Submitted for publication.

Patel, V., Rodrigues, M., DeSouza, N., 2002. Gender, poverty, and postnatal depression: a study of mothers in Goa, India. Am. J. Psychiatry 159, 43–47, http://dx.doi.org/10.1176/appi.ajp.159.1.43.

Paulson, J.F., Dauber, S., Leiferman, J.A., 2006. Individual and combined effects of postpartum depression in mothers and fathers on parenting behavior. Pediatrics 118, 659–668, http://dx.doi.org/10.1542/peds.2005-2948.

Paulson, J.F., Keefe, H.A., Leiferman, J.A., 2009. Early parental depression and child language development. J. Child Psychol. Psychiatry Allied Discip. 50, 254–262, http://dx.doi.org/10.1111/j.1469-7610.2008.01973.x.

Podestá, L., Alarcón, L., Muñoz, A.M., Legüe, S., C, M., Bustos, L., Barría, P.M., 2013. Psychomotor development in offspring of mothers with post partum depression. Rev. Méd. Chile 141 (4), 464–470, http://dx.doi.org/10.4067/S0034-98872013000400007.

Quevedo, L.A., Silva, R.A., Godoy, R., Jansen, K., Matos, M.B., Tavares Pinheiro, K.A., Pinheiro, R.T., 2012. The impact of maternal post-partum depression on the language development of children at 12 months. Child: Care Health Dev. 38 (3), 420–424, http://dx.doi.org/10.1111/j.1365-2214.2011.01251.x.

Rahman, A., Iqbal, Z., Bunn, J., Lovel, H., Harrington, R., 2004. Impact of maternal depression on infant nutritional status and illness: a cohort study. Arch. Gen. Psychiatry 61, 946–952, http://dx.doi.org/10.1001/archpsyc.61.9.946.

Santos, I.S., Matijasevich, A., Domingues, M.R., Barros, A.J., Barros, F.C., 2010. Long-lasting maternal depression and child growth at 4 years of age: a cohort study. J. Pediatrics 157, 401–406, http://dx.doi.org/10.1016/j.jpeds.2010.03.008 (10.1016/j.jpeds.2010.03.008).

Santos, K.O., Araujo, T.M., Oliveira, N.F., 2009. Factor structure and internal consistency of the Self-Reporting Questionnaire (SRQ-20) in an urban population. Cad. Saude Publ. 25, 214–222 (Retrieved from⟨http://www.scielo.br/scielo.php?script=sci_nlinks&ref=000147&pid=S1516-44462010000010000500010&lng=en⟩.

Scholte, W.F., Verduin, F., van Lammeren, A., Rutayisire, T., Kamperman, A.M., 2011. Psychometric properties and longitudinal validation of the self-reporting questionnaire (SRQ-20) in a Rwandan community setting: a validation study. BMC Med. Res. Methodol. 11 (1), 116, http://dx.doi.org/10.1186/1471-2288-11-116.

Sen, B., 1987. An analysis of the nature of depressive phenomena in primary health care utilising multivariate statistical techniques. Acta Psychiatr. Scand. 76, 28–32.

Stewart, R.C., Kauye, F., Umar, E., Vokhiwa, M., Bunn, J., Fitzgerald, M., Creed, F., 2009. Validation of a Chichewa version of the Self-Reporting Questionnaire (SRQ) as a brief screening measure for maternal depressive disorder in Malawi, Africa. J. Affective Disord. 112, 126–134, http://dx.doi.org/10.1016/j.jad.2008.04.001.

Svensen, E., Nordgreen, T., 2014. Detecting common mental disorder among patients in rural Tanzania using SRQ-20, HSCL-20, and locally developed items. (in preparation).

Tanaka, J.S., 1993. Multifaceted conceptions of fit in structural equation models. In: Bollen, K.A., Long, J.S. (Eds.), Testing Structural Equation Models. Sage, Newbury Park, CA, pp. 10–39.

Turney, K., 2011. Maternal depression and childhood health inequalities. J. Health Soc. Behav. 52, 314–332, http://dx.doi.org/10.1177/0022146511408096.

Velicer, W.F., 1976. Determining the number of components from the matrix of partial correlations. Psychometrika 41, 321–327, http://dx.doi.org/10.1007/BF02293557.

Velicer, W.F., Fava, J.L., 1998. Affects of variable and subject sampling on factor pattern recovery. Psychol. Methods 3, 231–251, http://dx.doi.org/10.1037/1082-989X.3.2.231.

Ventevogel, P., De Vries, G., Scholte, W.F., Shinwari, N.R., Faiz, H., Nassery, R., Olff, M., 2007. Properties of the Hopkins Symptom Checklist-25 (HSCL-25) and the Self-Reporting Questionnaire (SRQ-20) as screening instruments used in primary care in Afghanistan. Soc. Psychiatry Psychiatr. Epidemiol. 42, 328–335, http://dx.doi.org/10.1007/s00127-007-0161-8.

Walker, S.P., Wachs, T.D., Meeks Gardner, J., Lozoff, B., Wasserman, G.A., Pollitt, E., Carter, J.A., 2007. Child development: risk factors for adverse outcomes in developing countries. Lancet 369, 145–157, http://dx.doi.org/10.1016/S0140-6736(07)60076-2.

Watkins, M.W., 2006. Determining parallel analysis criteria. J. Mod. Appl. Stat. Methods 5, 344–346 (Retrieved from⟨http://www.public.asu.edu/~mwwatkin/Papers/ParallelAnalysis(2007).pdf⟩.

Weobong, B., Akpalu, B., Doku, V., Owusu-Agyei, S., Hurt, L., Kirkwood, B., Prince, M., 2009. The comparative validity of screening scales for postnatal common mental disorder in Kintampo, Ghana. J. Affective Disord. 113, 109–117, http://dx.doi.org/10.1016/j.jad.2008.05.009.

Wisner, K.L., Parry, B.L., Piontek, C.M., 2002. Postpartum depression. N. Engl. J. Med. 347, 194–199, http://dx.doi.org/10.1056/NEJMcp011542.

Woods, C.M., Oltmanns, T.F., Turkheimer, E., 2009. Illustration of MIMIC-model DIF testing with the schedule for nonadaptive and adaptive personality. J. Psychopathol. Behav. Assess. 31, 320–330, http://dx.doi.org/10.1007/s10862-008-9118-9.

World Health Organization. (1980). Major depressive disorder. In: International Statistical Classification of Diseases and Related Health Problems, 10th ed. Retrieved from ⟨http://apps.who.int/classifications/icd10/browse/2010/en#/F32⟩.