Women and Birth 35 (2022) e118–e124

Contents lists available at ScienceDirect

Women and Birth

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

Supporting routine psychosocial assessment in the perinatal period: The concurrent and predictive validity of the Antenatal Risk Questionnaire-Revised

Nicole Reilly[a,b,c,*], Dusan Hadzi-Pavlovic[c], Deborah Loxton[d,e], Emma Black[f,g], Victoria Mule[b], Marie-Paule Austin[b,c,h]

[a] Centre for Health Service Development, Australian Health Services Research Institute, University of Wollongong, NSW, Australia
[b] Perinatal and Women's Mental Health Unit, St John of God Burwood Hospital, Australia
[c] School of Psychiatry, UNSW Medicine, NSW, Australia
[d] Research Centre for Generational Health and Ageing, University of Newcastle, NSW, Australia
[e] Australian Longitudinal Study on Women's Health, University of Newcastle, NSW, Australia
[f] Drug and Alcohol Services, South Eastern Sydney Local Health District, Surry Hills NSW and Discipline of Addiction Medicine, University of Sydney, NSW, Australia
[g] School of Public Health and Community Medicine, Faculty of Medicine, University of New South Wales, NSW, Australia
[h] Royal Hospital for Women, NSW, Australia

ARTICLE INFO

Article history:
Received 20 November 2020
Received in revised form 7 March 2021
Accepted 4 April 2021

Keywords:
Psychosocial assessment
Psychosocial risk
Pregnancy
Postnatal
Validation

ABSTRACT

Background: Australian clinical practice guidelines support comprehensive psychosocial assessment as a routine component of maternity care.
Aim: To examine the concurrent and predictive validity of the Antenatal Risk Questionnaire-Revised (ANRQ-R) when used across the perinatal period.
Methods: Women completed the ANRQ-R and a diagnostic reference standard (SAGE-SR) in the second and third trimesters and at 3-months postpartum. ANRQ-R test performance for cut-off scores at each time-point was assessed using Receiver Operator Characteristic (ROC) analysis.
Findings: Overall sample sizes were N = 1166 (second trimester), N = 957 (third trimester) and N = 796 (3-month postpartum). 6.5%, 5.6% and 6.2% of women met SAGE-SR criteria for any depressive or anxiety disorder at these time-points ('cases'), respectively. ROC analysis yielded acceptable areas under the curve (AUC) when the ANRQ-R was used to detect current (AUC = 0.789–0.798) or predict future (AUC = 0.705–0.789) depression or anxiety. Using an *example* cut-off score of 18 or more, the ANRQ-R correctly classified 72–76% of concurrent 'cases' and 'non-cases' (sensitivity = 0.70–0.74, specificity = 0.72–0.76) and correctly predicted 74–78% of postnatal 'cases' and 'non-cases' (sensitivity = 0.52–0.72, specificity = 0.75–0.79). Completion of the ANRQ-R earlier in pregnancy yielded greater positive likelihood ratios for predicting depression or anxiety at 3-months postpartum (cut-off ≥18: second trimester = 3.8; third trimester = 2.2).
Conclusion: The ANRQ-R is a structured psychosocial assessment questionnaire that can be scored to provide an overall measure of psychosocial risk. Cut-off scores need not be uniform across settings. Such decisions should be guided by factors including diagnostic prevalence rates, local needs and resource availability.
© 2021 The Authors. Published by Elsevier Ltd on behalf of Australian College of Midwives. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Statement of significance

Problem or issue

A revised version of the Antenatal Risk Questionnaire has been developed; however, its psychometric properties have not yet been established.

* Corresponding author at: Centre for Health Service Development, Australian Health Services Research Institute (AHSRI), University of Wollongong, NSW, 2022, Australia.
E-mail address: nreilly@uow.edu.au (N. Reilly).

http://dx.doi.org/10.1016/j.wombi.2021.04.003
1871-5192/© 2021 The Authors. Published by Elsevier Ltd on behalf of Australian College of Midwives. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

What is already known

Australian clinical practice guidelines support comprehensive psychosocial assessment as a routine component of maternity care. The Antenatal Risk Questionnaire-Revised (ANRQ-R) is a brief measure that assesses for a broad range of psychosocial factors known to be associated with poorer maternal mental health and parenting outcomes.

What this paper adds

The ANRQ-R has acceptable global performance when used to identify current, or predict future, depression or anxiety during pregnancy and at 3-months postpartum. It can be used with or without scoring to provide an overall measure of psychosocial risk.

1. Introduction

The potential for social adversity and untreated maternal stress, distress and mental health morbidity to impact on women and families is well established. An increasing number of reviews report an association between perinatal mental health disorders and poorer socio-emotional and developmental outcomes in neonates and children [1–4]. Recent research has also shown that maternal mental health is more strongly associated with smoking, alcohol and other substance use in pregnancy than other known risk factors, including socioeconomic status or maternal age [5].

However, as noted by Stein and colleagues, such adverse outcomes are not inevitable [4]. Over recent decades, countries around the world have been striving to prioritise emotional health care alongside the physical heath care of pregnant women and new mothers. A number of clinical practice guidelines now recommend that depression screening be offered as a routine aspect of maternity care, with more comprehensive guidelines also supporting universal or targeted assessment of risk factors known to be associated with poorer perinatal mental health and parenting outcomes [6–10]. The importance of exploring and understanding a woman's broader psychosocial context, in order to appropriately support her and her family, is fundamental to these models of integrated psychosocial care.

Current Australian national guidelines support this more comprehensive approach and include an evidence-based recommendation that the Antenatal Risk Questionnaire (ANRQ) [11] be used for the assessment of psychosocial risk [6]. These guidelines also recommend that psychosocial assessment be undertaken in conjunction with routine depression and anxiety screening as early as practical in pregnancy and again after the birth. The ANRQ has recently been revised in response to feedback from clinicians, and in light of changes in requirements under the Australian Medicare Benefits Schedule for obstetric services [12]. Updates in the Antenatal Risk Questionnaire-Revised (ANRQ-R) include the inclusion of new items that ask about a woman's experience of recent or current domestic violence or substance misuse.

We have previously reported on the development, use and test-retest reliability of the ANRQ-R in a large sample of pregnant women [13]. The purpose of the current study was to extend our examination of the psychometric performance of the ANRQ-R by determining its concurrent and predictive validity when used across the perinatal period.

2. Methods

2.1. Participants and procedure

This prospective study was conducted as part of a larger project examining the effectiveness of two alternative models of integrated psychosocial care during pregnancy [14]. Women who completed a psychosocial assessment as part of routine antenatal care at the participating site were consecutively recruited to the current study between March 2017 and May 2019. Women were eligible to participate if they (a) agreed to be contacted by the research team, (b) could complete study measures in English, and (c) were not clients of the hospital's Indigenous midwifery service (due to differences in the models of care and service delivery).

Eligible women were contacted by the research team (phone and/or SMS) approximately 10–14 days after their appointment and provided with information about the study. Those who agreed to participate gave informed consent and were asked to complete study measures online in the second trimester, third trimester and 3-months following birth. Women were given the option to opt in or out of participation at each individual time point.

2.2. Measures

Women completed the following measures in the second trimester, third trimester and 3-months following birth.

2.2.1. Antenatal Risk Questionnaire-Revised (ANRQ-R)

The ANRQ-R asks about domains of psychosocial health known to be associated with perinatal depression and anxiety, including significant history of depression or other mental health condition, relationship with partner, social support, distress associated with recent stressful life events, worry and perfectionism, adverse childhood experiences (emotional, physical or sexual abuse), emotional support from own mother while growing up, and current or recent experience of partner violence and substance misuse. It has eleven scored items and a total score range of 5–55, with higher scores indicating higher psychosocial risk. Additional *unscored* items provide valuable clinical and contextual information relating to any previous treatment or professional support received for mental health issues, whether adverse childhood experiences included childhood sexual abuse and the nature of recent stressful life events. It can be completed using pen-paper, verbally or embedded in computer-based or existing data platforms [13]. The ANRQ-R total score has been shown to have good test-retest reliability in a community sample (intraclass correlation coefficient = 0.79) [13].

2.2.2. Series of Assessments for Guiding Evaluation-Self-Report (SAGE-SR; reference standard)

The SAGE-SR is a short, structured self-report assessment that provides both diagnostic and symptom severity information. It was developed as a self-report alternative to the SCID and NetSCID-CV for use in clinical and research settings and is compatible with the ICD-10 and the DSM-5 [15]. Only SAGE-SR depression and anxiety information was used in this study. Test-retest reliability in a nonclinical sample is good to excellent for all depression and anxiety disorder modules (intraclass correlation coefficient [ICC] or kappa coefficients ranging from 0.60 to 0.90) [16].

Data for each ANRQ-R were collected via the web-based Key Survey (for data collected to 4 February 2019) or Qualtrics platforms (for data collected from 5 February 2019). SAGE-SR data were collected using infrastructure provided by TeleSage, to whom the online SAGE-SR is copyrighted.

Additional demographic and clinical information, including maternal age, parity, country of birth, date, gestation and Edinburgh Depression Scale (EDS) score at the antenatal booking-in visit was extracted from eMaternity, the administrative database of the participating hospital. Postcode data were linked with the Socio-Economic Indexes for Areas (SEIFA) Index of Relative Disadvantage, with quintiles ranging from 1 (least advantaged) to 5 (most advantaged) [17]. This index is derived from national census data for all individuals living in a postcode and considers income, education, occupation, wealth and living conditions.

2.3. Ethics

Ethical and operational approval for this study was granted from the Local Health District's Human Research Ethics Committee and Research Governance Office (SESLHD HREC Ref: 14/117, 16/229; SSA Ref: 14/336, 16-G-247).

3. Statistical analysis

Descriptive statistics (mean, standard deviation, range, percentage) were used to examine the demographic and psychosocial profiles of participants. The characteristics of eligible women who did and did not participate in this study were compared using chi-square, independent sample t-tests or Mann–Whitney *U* test, as appropriate. Data relating to the depressive (major depressive disorder, major depressive episode, persistent depressive disorder, other specified depressive disorder) and anxiety (generalised anxiety disorder, panic disorder, social anxiety disorder, obsessive compulsive disorder, agoraphobia) disorder modules of the SAGE-SR were included in the analyses, with women who met criteria for one or more of these disorders considered a 'case' and women who did not meet criteria a 'non-case'.

A number of statistics were used to evaluate the diagnostic performance of the ANRQ-R when used to (1) detect a current depressive or anxiety disorder, when completed in the second trimester, third trimester and at 3-months postpartum, and (2) predict a future depressive or anxiety disorder from second trimester to 3 months postpartum, and from third trimester to 3 months postpartum. The area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive and negative likelihood ratios, and percent correctly classified were determined for each scenario. These outcomes were investigated for a range of ANRQ-R cut-off scores. These included the lowest cut-off where sensitivity approached 80% and the highest cut-off where specificity approached 80%.

Women must have completed each ANRQ-R and SAGE-SR within 14 days of each other to be included in the concurrent analyses. Postpartum predictive analyses were restricted to women who did not meet criteria for a SAGE-SR depressive or anxiety disorder during pregnancy, in order to focus solely on those women who experienced postnatal depression or anxiety subsequent to completing an antenatal ANRQ-R. Analyses were conducted using R [18] and IBM SPSS Statistics version 24 [19].

4. Results

7183 women attended the participating site and completed a psychosocial assessment at their antenatal booking in-visit during the study period. Of these, 1565 (21.8%) women consented to participate in the current study and completed the study measures at one or more of the second trimester (N = 1166), third trimester (N = 957) or 3 months postpartum (N = 798).

Overall, there were significant differences between participants and non-participants in terms of maternal age (mean 33.5 years vs 32.2 years, $p < 0.001$, eta squared = 0.01), gestational age at booking-in visit (15.4 weeks vs 16.4 weeks, $p < 0.001$, $r = -0.08$), EDS score of 13 or more at booking-in visit (2.2% vs 3.9%, $p = 0.002$, phi = −0.036), having a partner (98.8% vs 97.3%, $p = 0.001$, phi = 0.04), being born in Australia (45.6% vs 41.6%, $p = 0.004$, phi = 0.034) and primiparity (64.2% vs 56.3%, $p < 0.001$, phi = −0.066). However, the magnitude of these differences were small or not clinically significant.

Overall, the proportions of women meeting SAGE-SR criteria for any depressive or anxiety disorder in the second trimester, third trimester and three months postpartum were 6.5%, 5.6% and 6.2%, respectively. Women who were a SAGE-SR 'case' had higher mean

Table 1
Demographic and clinical characteristics of study participants at each data collection point.

	Second trimester (N = 1166)	Third trimester (N = 957)	3 months postpartum (N = 796)
	M (SD, range)	**M (SD, range)**	**M (SD, range)**
Gestation/infant age (weeks)	17.4 (2.1, 13–26)	33.5 (1.9, 27–40)	12.6 (1.4, 7–20)
ANRQ-R total score	13.7 (7.0, 5–39)	14.2 (7.3, 5–42)	14.2 (7.2, 5–37)
EDS total score	4.6 (4.0, 0–26)	5.0 (4.0, 0–25)	4.9 (3.9, 0–21)
	N (%)	**N (%)**	**N (%)**
SAGE-SR depression or anxiety	76 (6.5)	54 (5.6)	49 (6.2)
Australian born	528 (45.3)	466 (48.7)	373 (46.9)
First child	780 (66.9)	606 (63.3)	477 (59.9)
Partnered	1145 (98.2)	943 (98.5)	784 (98.5)
Missing	7 (0.6)	5 (0.5)	3 (0.4)
SEIFA quintile			
1 (greatest disadvantage)	4 (0.3)	2 (0.2)	2 (0.3)
2	121 (10.4)	101 (10.6)	79 (9.9)
3	165 (14.2)	118 (12.3)	88 (11.1)
4	284 (24.4)	219 (22.9)	179 (22.5)
5 (least disadvantage)	592 (50.8)	517 (54.0)	398 (50.0)
Missing	–	–	50 (6.3)
Highest level of education			
High School	63 (5.4)	40 (4.2)	31 (3.9)
TAFE/Certificate	269 (23.1)	214 (22.4)	161 (20.2)
University Degree or higher	815 (69.9)	658 (68.8)	548 (68.8)
Missing	19 (1.6)	45 (4.7)	56 (7.0)

ANRQ-R: Antenatal Risk Questionnaire-Revised; EDS: Edinburgh Depression Scale.
SAGE-SR: Series of Assessments to Guide Evaluation - Self Report.

ANRQ-R scores than non-cases at each time point: second trimester: M = 21.5 (SD = 8.0) vs 13.1 (SD = 6.6); third trimester: M = 22.7 (SD = 8.7) vs 13.7 (SD = 6.9); 3 months postpartum: M = 22.0 (SD = 7.5) vs 13.7 (SD = 6.8). Additional sociodemographic and clinical characteristics of participating women are summarised in Table 1.

4.1. Concurrent validity of the ANRQ-R

The majority of women included in the analysis of concurrent validity completed all measures on the same day (second trimester: same day = 91.7%, 1–5 days = 6.6%, 6–12 days = 1.7%; third trimester: same day = 90.8%, 1–5 days = 7.8%, 6–12 days = 1.4%; 3-months postpartum: same day = 87.9%, 1–5 days = 9.8%, 6–12 days = 2.3%).

As shown in Table 2, we found acceptable areas under the curve for the ANRQ-R when completed proximal to the SAGE-SR in the second trimester (AUC 0.789, 95%CI: 0.734−0.844), third trimester (AUC 0.797, 95%CI: 0.743−0.851) and 3-months postpartum (AUC 0.798, 95%CI: 0.739−0.857). Using an *example* cut-off score of 18 or more, the ANRQ-R correctly classified between 75.5% of concurrent 'cases' and 'non-cases' during the second trimester (sensitivity = 0.697, specificity = 0.759, positive likelihood ratio = 2.89, negative likelihood ratio = 0.40) and 72.2% of concurrent 'cases' and 'non-cases' in the third trimester (sensitivity = 0.704, specificity = 0.723, positive likelihood ratio = 2.54, negative likelihood ratio = 0.41). Applying this cut-off at 3-months postpartum yielded similar results: 73.4% correct classification of cases and non-cases, with sensitivity and specificity values of 0.735 and 0.734, respectively (positive likelihood ratio = 2.76, negative likelihood ratio = 0.36).

4.2. Predictive validity of the ANRQ-R

The results of the predictive analyses are presented in Table 3. ROC curves showed that the overall capacity of the ANRQ-R to predict a SAGE-SR depressive or anxiety disorder at three-months postpartum was greater when completed in the second trimester (AUC = 0.789, 95%CI: 0.681, 0.896) than in the third trimester (AUC = 0.705, 95%CI: 0.600, 0.810). Table 3 also shows that when our *example* cut-off ANRQ-R score of 18 or more is applied, a greater balance between sensitivity and specificity is achieved in the second trimester than in the third trimester (sensitivity = 0.772, specificity = 0.786 vs. sensitivity = 0.542, specificity = 0.749, respectively). Using this cut-off of 18 or more across pregnancy, the ANRQ-R had a similar capacity to predict non-cases (NPV = 97.2–98.5%) although was marginally better at predicting cases when used *earlier* in pregnancy (second trimester: PPV = 13.0%, positive likelihood ratio = 3.4) vs third trimester: PPV = 9.3%, positive likelihood ratio = 2.2). At this cut-off, the sensitivity and specificity of the ANRQ-R were largely comparable when used in the second trimester to either identify current, or predict future, SAGE-SR depressive or anxiety disorders (Fig. 1).

5. Discussion

Perinatal psychosocial risk measures such as the ANRQ-R are used by clinicians to gain an insight into a woman's psychosocial circumstances and to help guide decisions about appropriate care options. Although a number of barriers to implementing routine mental health screening and psychosocial assessment have been reported [20], studies have also shown these issues can be overcome when there is a commitment to adequate planning and systems in place for staff training and clinical supervision, follow-up, management and referral of women and families in need, collaborative and inter-disciplinary partnerships and appropriate resourcing [21,22].

The ANRQ-R was designed to support primary health care providers to explore the key psychosocial circumstances of women under their care and can be used as a structured questionnaire as part of a broader clinical interview, as well as being scored to provide an overall measure of psychosocial risk. Although the ANRQ-R was not specifically intended to be used to identify current, or predict future, perinatal depression or anxiety, we recognise that there are circumstances, including during the development of clinical practice guidelines for example, when the psychometric performance of psychosocial risk measures are evaluated in a similar way to symptom-based tools such as the EDS [23,24]. We have responded to this with respect to the ANRQ-R, using data collected from a large community-based sample of pregnant and postpartum women.

Table 2
Concurrent validity of the Antenatal Risk Questionnaire-Revised (ANRQ-R) against SAGE-SR depressive and anxiety disorders, when completed in the second trimester, third trimester and 3-months postpartum.

Timing of ANRQ-R & SAGE-SR administration	Total N	Cases	AUC (95%CI)	Cut-point	Se	Sp	LR+	LR-	PPV	NPV	Correctly classified[a] (%)
Second trimester	1166	6.52%	0.789 (0.734, 0.844)	≥14	0.803	0.605	2.03	0.33	0.124	0.978	61.7
				≥15	0.776	0.649	2.21	0.34	0.133	0.977	65.7
				≥16	0.737	0.689	2.37	0.38	0.142	0.974	69.2
				≥17	0.724	0.728	2.66	0.38	0.156	0.974	72.7
				≥18	0.697	0.759	2.89	0.40	0.168	0.973	75.5
				≥19	0.645	0.794	3.12	0.45	0.168	0.973	78.4
				≥20	0.632	0.820	3.51	0.45	0.197	0.970	80.8
Third trimester	957	5.64%	0.797 (0.743, 0.851)	≥14	0.796	0.580	1.90	0.35	0.102	0.979	59.2
				≥15	0.741	0.622	1.96	0.42	0.105	0.976	62.9
				≥16	0.722	0.656	2.10	0.42	0.111	0.975	65.9
				≥17	0.704	0.695	2.31	0.43	0.121	0.975	69.6
				≥18	0.704	0.723	2.54	0.41	0.132	0.976	72.2
				≥19	0.648	0.756	2.66	0.47	0.137	0.973	75.0
				≥20	0.593	0.786	2.77	0.52	0.142	0.970	77.5
3-months postpartum	796	6.16%	0.798 (0.739, 0.857)	≥15	0.796	0.640	2.21	0.32	0.127	0.980	64.9
				≥16	0.755	0.671	2.29	0.37	0.131	0.977	67.6
				≥17	0.735	0.707	2.51	0.38	0.141	0.976	70.9
				≥18	0.735	0.734	2.76	0.36	0.153	0.977	73.4
				≥19	0.714	0.772	3.14	0.37	0.171	0.976	76.9
				≥20	0.673	0.794	3.27	0.41	0.176	0.974	78.6

SAGE-SR: Series of Assessments to Guide Evaluation – Self Report (reference standard); AUC: area under the receiver operating characteristic (ROC) curve; Se: sensitivity; Sp: specificity; LR+: positive likelihood ratio; LR-: negative likelihood ratio; PPV: positive predictive value; NPV: negative predictive value.
[a] Percent correctly classified = true positives + true negatives/total N.

Table 3
Predictive validity of the Antenatal Risk Questionnaire-Revised (ANRQ-R) completed in the second and third trimesters, against SAGE-SR depressive and anxiety disorders completed at 3-months postpartum.

Timing of ANRQ-R administration	Timing of SAGE-SR administration	Total N	Cases	AUC (95%CI)	Cut-point	Se	Sp	LR+	LR-	PPV	NPV	Correctly classified[a] (%)
Second trimester	3-months postpartum	425	4.24%	0.789 (0.681, 0.896)	≥13	0.833	0.604	2.11	0.28	0.085	0.988	61.4
					≥14	0.722	0.641	2.01	0.43	0.082	0.981	64.5
					≥15	0.722	0.686	2.30	0.41	0.092	0.982	68.7
					≥16	0.722	0.715	2.53	0.39	0.101	0.983	71.5
					≥17	0.722	0.759	3.00	0.37	0.117	0.984	75.8
					≥18	0.722	0.786	3.38	0.35	0.130	0.985	78.4
					≥19	0.667	0.813	3.57	0.41	0.136	0.982	80.7
					≥20	0.611	0.838	3.77	0.46	0.143	0.980	82.8
Third trimester	3-months postpartum	530	4.53%	0.705 (0.600, 0.810)	≥11	0.792	0.468	1.49	0.44	0.066	0.979	48.3
					≥12	0.750	0.530	1.59	0.47	0.070	0.978	54.0
					≥13	0.750	0.581	1.79	0.43	0.078	0.980	58.9
					≥14	0.625	0.619	1.64	0.61	0.072	0.972	61.9
					≥15	0.625	0.662	1.85	0.57	0.081	0.974	66.0
					≥16	0.542	0.684	1.71	0.67	0.075	0.969	67.7
					≥17	0.542	0.725	1.97	0.63	0.086	0.971	71.7
					≥18	0.542	0.749	2.16	0.61	0.093	0.972	74.0
					≥19	0.542	0.779	2.45	0.59	0.104	0.973	76.8
					≥20	0.542	0.802	2.74	0.57	0.115	0.974	79.1

SAGE-SR: Series of Assessments to Guide Evaluation – Self Report (reference standard); AUC: area under the receiver operating characteristic (ROC) curve; Se: sensitivity; Sp: specificity; LR+: positive likelihood ratio; LR−: negative likelihood ratio; PPV: positive predictive value; NPV: negative predictive value.
[a] Percent correctly classified = true positives + true negatives/total.

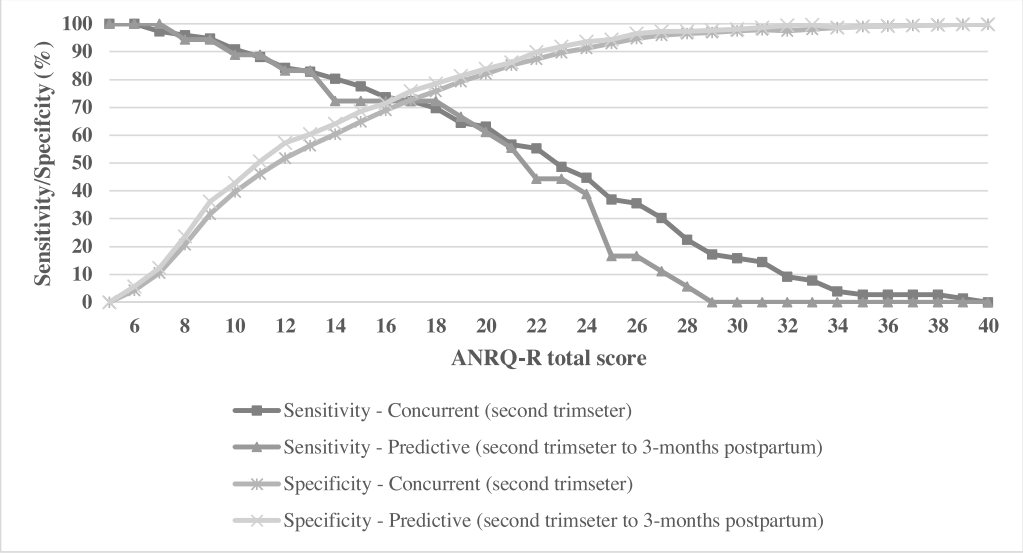

Fig. 1. Concurrent and predictive validity of the ANRQ-R for SAGE-SR depressive or anxiety disorders, when used in the second trimester.

The ANRQ-R was found to have acceptable global performance when used to identify *current* depression or anxiety across the perinatal period, as demonstrated by AUC values greater than 0.7. When applying an example cut-off of 18 or more, the sensitivity of the ANRQ-R ranged from 70% to 74%, specificity from 72% to 76%, and it correctly classified between 72% and 76% of current cases and non-cases. Studies that have examined the clinical utility of other psychosocial risk measures to identify current (rather than predicting future) perinatal mental health morbidity have reported comparable test performance. For example, van Heyningen and colleagues examined the test characteristics of a risk screening tool to detect major depressive and anxiety disorders in a sample of pregnant women from a socio-economically disadvantaged community in South Africa [25]. This tool combined the five best-performing items from a battery of measures (perceived lack of partner support, current domestic violence, history of abuse, self-reported history of mental health issues and experience of major, adverse life events in the past year), and had an AUC of 0.73, sensitivity of 77% and specificity of 64% when a cut-point of two or more risk factors was used. In the Netherlands, short, intermediate and extended triage tools developed to screen for a broad spectrum of DSM-IV axis I and axis II disorders diagnosed during pregnancy using the Structured Clinical Interview (SCID) yielded AUC values ranging from 0.63 to 0.82 [26]. These tools included questions relating to mental health history, previous traumatic experiences, substance use, social support, financial stress and relationship problems. After identifying optimal cut-offs, sensitivity and specificity of the tools (administered without a depression screening measure) was shown to vary from 63%–83% and 77%–86%, respectively. The PPVs for these triage tools were much higher than those identified for the ANRQ-R, although the prevalence of the broader set of conditions identified by the authors was also much greater than the overall prevalence of depression or anxiety among women in our study (40% versus 5.6–6.5%). This is a key

difference to note, as it is well established that predictive values of a test are highly dependent on the population prevalence of the condition in question, with higher prevalence corresponding with higher PPV [27]. Increased recognition of this has resulted in other test characteristics, including positive and negative likelihood ratios, being given enhanced priority in validation studies and evidence reviews [23].

The overall *predictive* performance of the ANRQ-R (AUC 0.71–0.79) was found to be reasonably similar to the original ANRQ (0.69) [11] and its longer precursor, the Pregnancy Risk Questionnaire (0.79) [28], although there was greater variation between the measures in other reported test characteristics. When using a cut-off of 18 or more, the ANRQ-R performed marginally better in predicting postpartum depression or anxiety when used in the second trimester (positive likelihood ratio = 3.4) than in the third trimester (positive likelihood ratio 2.2). Applying this cut-off during the second trimester accurately predicted 78% of postpartum cases and non-cases. Similarly, the antenatal version of the Postpartum Depression Predictors Inventory-Revised (PDPI-R) has recently been shown to correctly classify up to 86% of Portuguese women, using a second trimester cut-off score of 4.5 out of a possible 32 for the prediction of major or minor depression at 3–4 months postpartum [29]. While considerable heterogeneity across these studies, in terms of population groups, risk domains assessed, outcomes of interest and study design, makes more direct comparisons difficult, taken together these results provide support for clinical practice frameworks that recommend an initial routine psychosocial assessment be offered earlier, rather than later, in pregnancy [6]. The predictive validity of other perinatal psychosocial risk measures, including Antenatal Psychosocial Health Assessment Tool [30,31], and the Antenatal Routine Psychosocial Assessment [32], remains unknown [23].

Several limitations of this study need to be acknowledged. We recruited from a single site, and the majority of women lived in socioeconomically advantaged areas. While the clinical and sociodemographic characteristics of participants was similar to the local maternity population, our sample differs from the general population of women giving birth in Australia in a number of ways, including an overrepresentation of women born outside of Australia and underrepresentation of younger women [33]. The ANRQ-R covers a comprehensive range of psychosocial domains, but does not include specific questions relating to all known risk factors, such as food insecurity [34] or family history of severe mental health issues [8]. Our sample also had a low overall prevalence of depressive and anxiety disorder, which may have introduced specificity bias [35]. Study strengths include its large community based sample and its inclusion of both pregnant and postpartum women. We were able to report on the psychometric features of the ANRQ-R as a function of perinatal status (antenatal versus postnatal), intended use (identification of risk for current or future depressive or anxiety disorder) and at various cut-points. This information will help researchers and health care professionals make decisions about which cut-off score is most suitable for their particular context. Future research should explore how the ANRQ-R performs when used in different clinical settings and with different subgroups of the perinatal population. The contribution of additional postnatal-specific items to the ANRQ-R, and the psychometric 'value-add' of using a combination of symptom-based and risk-based measures as recommended in Australian clinical practice guidelines, also requires more comprehensive evaluation.

6. Conclusion

This study reports on the concurrent and predictive validity of the ANRQ-R and presents test characteristics for a range of cut-points and scenarios. At its simplest, the tool can be used – with or without scoring – as an aid to streamline exploration of a woman's psychosocial circumstances so that risk factors can be identified early and progression to poorer emotional health slowed or prevented. The option to score the ANRQ-R might be considered in particular circumstances, such as when a specific level of risk (or combination of risk factors) is required before a referral is initiated or in settings where automatic calculation of scores can be used to eliminate manual scoring errors, for example. In this paper, we have used an *example* cut-off score of 18 or more to demonstrate the value of the ANRQ-R when used to identify both concurrent and predictive 'caseness', however use of a single cut-off need not be uniform across settings. As noted previously [36], such decisions should be made within the clinical context in which such measures are used, including local and national guideline recommendations, and should be guided by factors including estimated or known prevalence rates, local needs and resource availability, and the societal and economic costs of false positive or false negative results.

Role of the funding source

Infrastructure funding for this project was provided by St John of God Health Care. NR thanks Australian Rotary Health and the University of Newcastle (2018–2020), and the University of Wollongong (2020–2023), for postdoctoral fellowship funding support. The funding bodies had no involvement in the design of this study, analysis or interpretation of data, writing of the manuscript or decision to submit the article for publication.

Ethical statement

Ethical and operational approval for this study has been granted from the South Eastern Sydney Local Health District Human Research Ethics Committee and Research Governance Office (SESLHD HREC Ref: 14/117, 16/229; SSA Ref: 14/336, 16-G-247).

Conflict of interest

None declared.

Author contributions

NR and MPA conceptualised the study, NR and DHP conducted the data analysis, DL assisted with manuscript and analysis planning, and EB and VM were responsible for data collection and management. NR prepared the manuscript and all authors critically reviewed and edited its content. All authors approved the final version submitted for publication.

Acknowledgements

We sincerely thank all the women who participated in the study. We thank Jo Farrell, Kristina Talcevska and Tegan Murray for their significant contributions to data collection and acknowledge the co-investigators on the larger PIPA project (Professor Virginia Schmied, Professor Georgina Chambers, A/Professor Dawn Kingston, A/Professor Stephen Matthey, Dr Andrew Bisits). We are grateful to the RHW management and staff, in particular Helen Jarman, Vanessa Medunic, Alyce Finch, Virginia Spear, Joanne Arkwright, Fiona Donovan, Michele Pezzuti and Rebecca Moore, and to clinical, nurse educator and administrative staff, who have generously contributed their time and expertise to the overarching PIPA Project.

References

[1] L.M. Howard, H. Khalifeh, Perinatal mental health: a review of progress and challenges, World Psychiatry 19 (3) (2020) 313–327.

[2] D. Kingston, S. Tough, H. Whitfield, Prenatal and postpartum maternal psychological distress and infant development: a systematic review, Child Psychiatry Hum. Dev. 43 (5) (2012) 683–714.
[3] A. Smith, J. Twynstra, J.A. Seabrook, Antenatal depression and offspring health outcomes, Obstet. Med. 13 (2) (2020) 55–61.
[4] A. Stein, R.M. Pearson, S.H. Goodman, et al., Effects of perinatal mental disorders on the fetus and child, Lancet 384 (9956) (2014) 1800–1819.
[5] R.A. Brown, H. Dakkak, J. Gilliland, J.A. Seabrook, Predictors of drug use during pregnancy: the relative effects of socioeconomic, demographic, and mental health risk factors, J. Neonatal. Med. 12 (2) (2019) 179–187.
[6] M.-P. Austin, N. Highet, the Expert Working Group, Mental Health Care in the Perinatal Period: Australian Clinical Practice Guideline, COPE: Centre of Perinatal Excellence, Melbourne, 2017.
[7] DoH, Clinical Practice Guidelines: Pregnancy Care, Department of Health, Australian Government, Canberra, 2018.
[8] NICE, Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance. NICE Clinical Guideline 192, National Collaborating Centre for Mental Health, 2014.
[9] SIGN, Management of Perinatal Mood Disorders (SIGN Publication no. 127), Scottish Intercollegiate Guidelines Network, Edinburgh, 2012.
[10] PMHP, Service Development Guidelines for Facilities Providing Maternal Mental Health Care. Cape Town: Perinatal Mental Health Project, Alan J. Fisher Centre for Public Health, University of Cape Town, 2010.
[11] M.P. Austin, J. Colton, S. Priest, N. Reilly, D. Hadzi-Pavlovic, The Antenatal Risk Questionnaire (ANRQ): acceptability and use for psychosocial risk assessment in the maternity setting, Women Birth 26 (2013) 17–25.
[12] DoH, Changes to MBS Items for Obstetric Services. Effective from 1 November 2017, Department of Health, Commonwealth of Australia, Canberra, 2017.
[13] N. Reilly, D. Loxton, E. Black, M.-P. Austin. The Antenatal Risk Questionnaire-Revised: development, use and test-retest reliability in a community sample of pregnant women in Australia. J. Affec. Dis., Under review.
[14] N. Reilly, E. Black, G.M. Chambers, V. Schmied, S. Matthey, J. Farrell, D. Kingston, A. Bisits, M.-P. Austin, Study protocol for a comparative effectiveness trial of two models of perinatal integrated psychosocial assessment: the PIPA project, BMC Pregnancy Childbirth 17 (1) (2017) 236.
[15] TeleSage, SAGE-SR: Self-report Diagnostic Screen, (2017) . https://telesage.com/sage/.
[16] B. Brodey, S.E. Purcell, K. Rhea, et al., Rapid and accurate behavioral health diagnostic screening: initial validation study of a web-based, self-report tool (the SAGE-SR), J. Med. Internet Res. 20 (3) (2018) e108.
[17] ABS, Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australian Bureau of Statistics, Canberra, 2011.
[18] R Core Team, R: a Language and Environment for Statistical Computing, R Core Team, Vienna, Austria, 2020 R Foundation for Statistical Computing.
[19] IBM Corp, IBM SPSS Statistics for Windows, Version 24.0, IBM Corp, Armonk, NY, 2016.
[20] T. Connell, B. Barnett, D. Waters, Barriers to antenatal psychosocial assessment and depression screening in private hospital settings, Women Birth 31 (4) (2018) 292–298.
[21] N. Reilly, E. Brake, H. Kalra, M.-P. Austin, Insights into implementation of routine depression screening and psychosocial assessment in a private hospital setting: a qualitative study, Aust. N. Z. J. Obstet. Gynaecol. 60 (3) (2020) 419–424.
[22] N. Reilly, D. Kingston, D. Loxton, K. Talcevska, M.-P. Austin, A narrative review of studies addressing the clinical effectiveness of perinatal depression screening programs, Women Birth 33 (1) (2020) 51–59.
[23] COPE, Australian Perinatal Mental Health Guideline. Evidence Review Technical Report, Prepared by Hereco for COPE: Centre of Perinatal Excellence, Melbourne, 2017.
[24] M. Johnson, V. Schmeid, S.J. Lupton, et al., Measuring perinatal mental health risk, Arch. Womens Ment. Health 15 (5) (2012) 375–386.
[25] T. van Heyningen, L. Myer, M. Tomlinson, S. Field, S. Honikman, The development of an ultra-short, maternal mental health screening tool in South Africa, Glob. Ment. Health 6 (2019) e24.
[26] C. Quispel, T.A.J. Schneider, W.J.G. Hoogendijk, G.J. Bonsel, M.P. Lambregtse-van den Berg, Successful five-item triage for the broad spectrum of mental disorders in pregnancy — a validation study, BMC Pregnancy Childbirth 15 (1) (2015) 51.
[27] K. Hajian-Tilaki, The choice of methods in determining the optimal cut-off value for quantitative diagnostic test evaluation, Stat. Methods Med. Res. 27 (8) (2018) 2374–2383.
[28] M.-P. Austin, D. Hadzi-Pavlovic, L. Leader, K. Saint, G. Parker, Antenatal screening for the prediction of postnatal depression: validation of a psychosocial Pregnancy Risk Questionnaire, Acta Psychiatr. Scand. 112 (4) (2005) 310–317.
[29] S. Alves, A. Fonseca, M.C. Canavarro, M. Pereira, Predictive validity of the Postpartum Depression Predictors Inventory-Revised (PDPI-R): a longitudinal study with Portuguese women, Midwifery 69 (2019) 113–120.
[30] J.C. Carroll, A.J. Reid, A. Biringer, et al., Effectiveness of the Antenatal Psychosocial Health Assessment (ALPHA) form in detecting psychosocial concerns: a randomized controlled trial, CMAJ 173 (3) (2005) 253–259.
[31] E.R. Blackmore, J.C. Carroll, A.J. Reid, et al., The use of the antenatal psychosocial health assessment (ALPHA) tool in the detection of psychosocial risk factors for postpartum depression: a randomized controlled trial, J. Obstet. Gynaecol. Canada 28 (2006).
[32] S. Matthey, J. Phillips, T. White, et al., Routine psychosocial assessment of women in the antenatal period: frequency of risk factors and implications for clinical services, Arch. Womens Ment. Health 7 (4) (2004) 223–229.
[33] AIHW, Australia's Mothers and Babies 2018: in Brief. Perinatal Statistics Series No. 36. Cat. No. PER 108, Australian Institute of Health and Welfare, Canberra, 2020.
[34] H. Preis, D.J. Garry, K. Herrera, D.J. Garretto, M. Lobel, Improving assessment, treatment, and understanding of pregnant women with opioid use disorder: the importance of life context, Womens Reprod. Health 7 (3) (2020) 153–163.
[35] M.M.G. Leeflang, A.W.S. Rutjes, J.B. Reitsma, L. Hooft, P.M.M. Bossuyt, Variation of a test's sensitivity and specificity with disease prevalence, CMAJ 185 (11) (2013) e537–e544.
[36] F. Habibzadeh, P. Habibzadeh, M. Yadollahie, On determining the most appropriate test cut-off value: the case of tests with continuous results, Biochem. Med. (Zagreb) 26 (3) (2016) 297–307.