Development and Validation of a Survey to Evaluate Mental Health Knowledge: The Case of Indonesian Pregnant Women

Yuli Kusumawati[1,2]*, Widyawati Widyawati[3], Fatwa Sari Tetra Dewi[4]

[1]Doctoral Program, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia; [2]Public Health Program, Faculty of Health, Universitas Muhammadiyah Surakarta, Central Java, Indonesia; [3]Department of Pediatric and Maternity Nursing, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia; [4]Department of Health Behavior, Environment and Social Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

Abstract

BACKGROUND: Internationally, many instruments have been designed to evaluate mental health knowledge; however, in pregnant women is very limited.

AIM: Therefore, this study aimed to develop and validate a survey to measure the mental health knowledge of pregnant women.

METHODS: In this cross-sectional study, 13 midwives attended the FGD and 10 pregnant women were invited for in-depth interviews to develop an item pool. The content validity was carried out by a panel of 6 experts. The face validity was performed with 5 pregnant women. Next, the construct validity test involved 150 pregnant women who were selected by stratified sampling from 13 public health centers in Surakarta, Indonesia. Analyses were conducted to check content validity, face validity, construct validity, internal consistency reliability, difficulty index, and exploratory factor analysis.

RESULTS: A final 20-item Mental Health Knowledge Scale (MHKS) has a content validity index of 0.97 and a correlation value per item greater than the r-table (i.e., 0.1603). In addition, the MHKS has a Kuder–Richardson 20 reliability coefficient of 0.717. Furthermore, the difficulty index ranged from 0.39 to 0.82 which was considered in the good and acceptable category. Construct validity was confirmed using exploratory factor analysis KMO = 0.713, Bartlett's test $p < 0.001$.

CONCLUSION: Based on the findings, the final version of the MHKS was considered a valid and reliable tool. The instrument can be applied to measure the understanding of pregnant women about pregnancy depression. Further studies require adjustment items to other participants regarding mental health knowledge.

Introduction

Mental health and disorders that may occur during pregnancy are not well understood by most mothers. Lack of knowledge about perinatal mental health (antenatal and 12 months postnatal) causes anxiety disorders and depression, which are often not properly treated. World data showed that among the maternal mental disorders, postpartum depression (13.0%) is higher than antenatal depression (10.0%). In developing countries, this condition is higher, reaching 15.6% for pregnancy depression and 19.8% for postpartum depression [1].

At present, most people's knowledge about maternal mental health and mental disorders in pregnancy and the postnatal period is still low. Pregnant women who can correctly identify that pregnancy anxiety/depression can impact a child's development are only around 26.6%. Research showed that the knowledge about mental health during pregnancy is lower than the postnatal knowledge [2]. Many people believe that anxiety during pregnancy is considered to be normal [3]. These beliefs will impact the absence of management and treatment and make a greater impact on child development [4].

In developing countries, especially Indonesia, the attempts to explore the knowledge of pregnant women about mental health during pregnancy and postnatal are still limited. In the previous studies, mental health knowledge was considered good if mothers were able to answer concerning the influence from the past history of anxiety and depression [5], prenatal mental health and adverse sequelae, postnatal mental health, and its adverse consequences [4], [6]. The knowledge and ability to utilize mental health information can indicate that a person has good mental health literacy. In general, mental health literacy is more focused on knowledge and beliefs about mental disorders rather than mental health [7]. People are more concerned if mental disorders have emerged. Meanwhile, behavior to improve mental health and welfare is not well understood. This is a gap that needs to be addressed.

The knowledge level of pregnant women about mental health needs to be measured objectively so that it can support efforts to provide suitable interventions. Psychometric instruments need to be developed in terms of tests used to measure the extent of pregnant women's knowledge and their understanding of health problems, especially mental health problems of pregnancy and in the postnatal period. In this study, this knowledge instrument explores the definition, the causes of mental disorders, particularly pregnancy depression, the risk factors, the symptoms or signs, and the attempts to prevent and overcome depressive disorders.

Quality research requires the use of a valid and reliable instrument to measure the research variables. According to Wei *et al*. [8], a well-validated tool not only helps to accurately evaluate the effects of mental health literacy interventions but also guides the improvement of new interventions. Validity shows the instrument's ability to measure an abstract concept. Meanwhile, the mental health knowledge of pregnant women is an abstract concept. Accordingly, it needs to be measured clearly, especially related to the risk of mental health disorders associated with pregnancy depression until the postnatal period.

At the development stage, items for the questionnaire were designed based on a situation analysis approach and the identification of information needed during pregnancy about mental health. The selection of mental health knowledge materials was based on a needs assessment conducted through Focus Group Discussions (FGDs) with pregnant women. Information needed by pregnant women included the understanding of depression, the causes and risk factors, the symptoms or signs, as well as the efforts that must be made to prevent and to overcome any interference. In addition, the selection of items for the questionnaire was also based on current literature concerning community views on perinatal mental health [2].

A measurement instrument needs to be validated because this is an important factor in implementing the instrument. The instrument validity consists of content validity determined by the value of the Content Validity Ratio (CVR) and Content Validity Index (CVI) with the Lawshe formula [9], [10], [11]. Using this instrument, pregnant women are expected to understand the mental health and mental disorders concepts during pregnancy and to be able to analyze and cope with the problems that may be the risk factors or depression triggers. All of the information once provided will increase their understanding, reduce misperceptions, and help the mothers to be able to deal with their problems independently. This knowledge instrument used closed-ended questions and was developed into a good instrument through validity, reliability, and difficulty indices. The main objective of this study was to develop and validate a questionnaire for measuring mental health knowledge in Indonesian pregnant women.

Methods

The cross-sectional study was conducted from August to December 2019. In this study, the Mental Health Knowledge Scale (MHKS) was developed using Boateng *et al*.'s model; item development, scale development, and scale evaluation [12]. All stages were described in detail as follows (Figure 1).
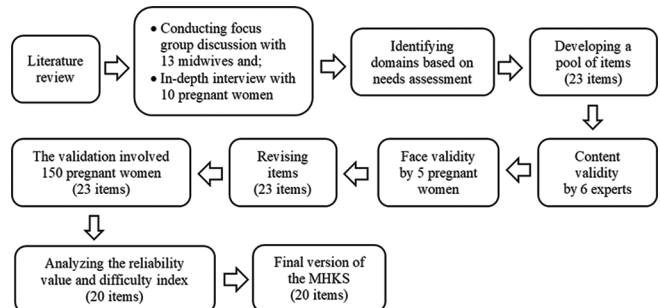

Figure 1: Diagram of the current study procedures

Item development

Step I: Identification of domain and item generation

This phase consisted of two activities, namely domain identification and item development (constructing a statement form). The domain identification step was done based on a thorough literature review about mental health literacy [2], [3]. To produce an initial scale in the current study, we conducted FGDs with 13 midwives. They were 25–54 years old with 1–34 years of work experience. Besides, we performed in-depth interviews with 10 pregnant women (14–39 weeks gestation). In this study, FGDs were conducted with midwives to explore information that pregnant women should know about changes in mental conditions that may occur during pregnancy and efforts to adapt to them. Furthermore, in-depth interviews were conducted with pregnant women with the aim of obtaining the personal information needed about mental health and disorders, as well as personal experiences that they had never experienced before. The first focus group involved 8 midwives to obtain information about mental health and the second focus group recruited 5 midwives to generate insights on mental health management. The in-depth interviews aimed to explore mental health and perinatal disorders among pregnant women. The FGDs were conducted in the community health center's meeting room and were guided by researchers. Examples of questions included: *"What are the symptoms of pregnancy depression?"* and *"What is the impact of depression on the fetus?"* The results of the FGDs were triangulated with responses from the mothers' husbands to confirm the suitability of the information needed. This step involved reviewing the items of each domain and item generation.

To ensure that items are consistent with a systematic understanding of mental health and maternal depression, we analyzed the literature about people's views of postpartum depression [2]. Based on this analysis, 23 items were constructed (Table 1). All items developed from the results of previous studies on barriers to prenatal mental health screening [13] and the development of an operational definition of mental health literacy [14]. The statements represented the domains of the definition, the causes, the signs and symptoms, the risk factors, time of depression, the impact, and the ways to prevent or treat the depression during pregnancy and in the postpartum period [3], [15], [16], [17]. The instrument used closed questions with a "true" or "false" answer. Participants answered by giving a checkmark in each statement. This scale was created to produce two subscale scores which included the following items and domains: Definition (2 items), causes of depression (2 items), signs and symptoms of depression (3 items), risk factors (5 items), time of depression (3 items), the impact of pregnancy depression (3 items), and how to treat and prevent (5 items). The minimum and maximum scores were 0 and 20, respectively. It should be noted that we used the true or false question (TFQ) since it was able to measure the respondent's understanding of the information obtained in a short time. Thus, TFQ is a good instrument for measuring facts, particularly those related to memory [18], [19]. All items in this questionnaire were written in the Indonesian language (as the mother tongue).

Step 2: Content validity test

The content validity of the items was determined by selected experts (n = 6) in the domain's content areas. They were recruited by purposive sampling. The inclusion criteria for participants were educational experience (psychometric), a mental health researcher, obstetrician, clinical psychologist, and midwife. The panelists had at least 5 years of professional experience. The main purpose of developing this instrument was explained previously to the experts, and then the expert team was asked to review the accuracy of each item with the instrument objectives. Specifically, the experts were asked to review each item based on the following criteria: (a) Question relevance in each item; (b) material representation in each item, and (c) language clarity in each item. The experts reviewed each item using a four-point ordinal scale, namely: 1 for not appropriate, 2=item needs major revision (some appropriate), 3=appropriate, but needs minor revision, and 4 for totally appropriate [20]. The expert panel gave feedback on the questionnaire and also provided explanations for their responses on each item. Suggestions and advice from the experts were used to improve the instrument.

All experts evaluated and gave suggestions for improvement in each item. Then, the research team calculated the Content Validity Ratio (CVR) with the Lawshe formula [9], [21]: $CVR = \dfrac{n_e - (\frac{N}{2})}{N/2}$. Content Validity Index (CVI) can be calculated individually for each item-level (I-CVI) [22], [23], which was considered as the total number of experts' who gave a score of 3 or 4 for a relevant item and 2 or 1 for an irrelevant item. Furthermore, the total number of expert judgments was divided by the total number of experts. I-CVI expresses the proportion of expert approval on relevant items, which has a value between zero and one [24]. Furthermore, Lynn recommends that an I-CVI value is >0.70 to be considered valid [23].

Scale development

Step 3: Face validity test

This step involved testing the questionnaire on a small scope to assess the readability of questionnaire. The number of participants to measure face validity is not set in advance but is decided as needed until the questionnaire is considered clearly legible and easily

Table 1: Scores of CVR and CVI in the MHKS

Items (n=20)	N of agreement		I-CVI	Interpretation	CVR		
	Relevant (Rating 3 or 4)	Not relevant (Rating 1 or 2)			Relevancy	Clarity	Simplicity
Q1	5	1	0.83	Appropriate	1.00	0.60	0.60
Q2	6	0	1.00	Appropriate	1.00	0.60	1.00
Q3	6	0	1.00	Appropriate	1.00	1.00	0.60
Q4	6	0	1.00	Appropriate	1.00	1.00	1.00
Q5	6	0	1.00	Appropriate	1.00	1.00	1.00
Q6	6	0	1.00	Appropriate	1.00	1.00	1.00
Q7	6	0	1.00	Appropriate	1.00	1.00	1.00
Q8	6	0	1.00	Appropriate	1.00	1.00	1.00
Q9	6	0	1.00	Appropriate	1.00	1.00	1.00
Q10	6	0	1.00	Appropriate	1.00	1.00	1.00
Q11	6	0	1.00	Appropriate	1.00	1.00	1.00
Q12	6	0	1.00	Appropriate	0.48	1.00	1.00
Q13	6	0	1.00	Appropriate	1.00	1.00	0.60
Q14	4	2	0.67	Need revision	1.00	1.00	1.00
Q15	5	1	0.83	Appropriate	0.60	1.00	1.00
Q16	6	0	1.00	Appropriate	1.00	1.00	1.00
Q17	6	0	1.00	Appropriate	1.00	1.00	1.00
Q18	6	0	1.00	Appropriate	1.00	1.00	1.00
Q21	6	0	1.00	Appropriate	1.00	1.00	1.00
Q22	6	0	1.00	Appropriate	1.00	1.00	1.00
	I-CVI		0.96		0.96	0.97	0.96

CVR = (Ne-N/2)/(N/2) with a panel of experts (n = 6).

understood by a lay respondent. The questionnaire was piloted on a small sample (n = 5) of pregnant women. They were asked to provide additional feedback on the wording and clarity of the instrument. For the record, these respondents were not included in the main study. Furthermore, their feedback for instrument improvement was discussed by the researchers.

In measuring the face validity, the researchers also asked the first pregnant woman to read and correct each item in the initial 23 item scale from the results of content validity by experts. Then, the researchers revised the sentence items according to the sample suggestions of the first pregnant woman. Then, the researchers revised the questionnaire items according to the suggestion of the first pregnant woman. In the same way as the first respondent, the second to fifth respondents commented sequentially. This aimed to minimize writing errors and ambiguity of each item [25].

Step 4: Construct validity test

A total of 13 outpatient public health centers (PHC) were used to conduct the construct validity test. There were 165 pregnant women who were eligible and approached to participate in the study, and 150 respondents actually participated. All participants included the consideration that the minimum sample size is five samples for each item [23], [24].

The inclusion criteria of the participant were: Women with singleton pregnancies, without complications of pregnancy, and can read the instrument questionnaire in the Indonesian language. Furthermore, the exclusion criteria were: Pregnant women who are in the process of giving birth and a pregnant woman who was waiting for the referral process of childbirth.

Pregnant women who were identified by the PHC were listed based on their health records. Pregnant women who did not have health records were identified through the ANC service schedule card. The researchers explained to the participants about this research and provided a data confidentiality guarantee with an informed consent form, and also asked for their commitment to participate until the end of the study. To mitigate the conflicts of interest, it should be noted that researchers are not employees or clinicians of the health service centers where participants were recruited.

The final step of instrument validity was measuring the construct validity, which was carried out on 150 pregnant women. Participants who were willing to participate in the main study were asked to answer all of the questions by giving a checkmark. The respondent's answer was coded 1 for the correct answer and 0 for the incorrect answer. The maximum and minimum scores for the entire questionnaire were 23 and 0, respectively. The data were then transferred to SPSS, analyzed, and interpreted to determine the level of pregnant women's knowledge.

Scale evaluation

Step 5: Item reduction and extraction of factors

Construct validity was analyzed using Exploratory Factor Analysis (EFA), which began with an analysis of the adequacy of the sample with the Kaiser-Meyer-Olkin Index (KMO) > 0.6 in the mediocre category [26], then followed by a test of the feasibility data with MSA coefficient mean higher than 0.5. Furthermore, researchers apply factor analysis to see the dominant factors and use the Varimax Rotation Method (with Kaiser Normalization) to see the component rotation. The percentage of variation explained and the Bartlett test was used to assess the adequacy of the model. The communalities value indicates the amount of variance explained in each item. If the value of each item's communalities was < 0.3, it indicates that the item does not fit with the model [27]. Consequently, the item is eliminated. The Kaiser Eigenvalue (Eigenvalue ≥ 1) criteria and scree plot were used to determine the amount of extracted factors [28].

Construct validity was determined by looking at the loading factor (Table 2). If the loading factor was greater than 0.30, it means that the construct validity is fulfilled (item instrument is valid). To provide evidence that supports the construct validity of this maternal mental health knowledge questionnaire, the number of corrected items should be > 0.30, the average variance extracted (AVE) for each subscale should be at least 0.50, and composite reliability (CR) for each subscale should be higher than 0.70. Convergent and discriminant validity are also calculated to support the construct validity evidence of this new instrument. Construct validity describes a strong correlation among items that measure the same construct, whereas discriminatory validity describes a weak correlation among items that do not support being related (measuring the different constructs). CR is used to measure convergent validity; however, AVE is used to measure the discriminant validity of the instrument.

Sample used in the validation

Respondents consisted of 150 pregnant women (aged 17–42 years; mean age: 28, 21) randomly drawn from five districts in Surakarta, Indonesia. They lived in urban (85.0%) and rural (15.0%) areas. Their educational levels were varied, namely elementary school (5.3%), secondary school (18.0%), high school (60.7%), and university (16.0%) They come from low (45.0%), middle (48.0%), and high (7.0%) economic status. Additional details of respondent characteristics are presented in Table 3. Each respondent completed MHKS in about 10 to 15 min.

Table 2: Factor loadings based on a principal component analysis extraction with Varimax rotation

Item	Component								Communalities
	1	2	3	4	5	6	7	8	
1	−0.015	0.097	0.087	−0.049	0.055	**0.849**	−0.052	0.053	0.749
2	**0.484**	0.123	0.038	0.093	0.291	0.433	−0.080	0.209	0.582
3	0.030	0.053	0.184	−0.033	**0.821**	0.003	−0.075	−0.056	0.722
4	0.292	0.237	0.135	−0.305	**0.463**	−0.060	0.338	0.014	0.585
5	0.042	0.214	0.091	**0.583**	0.323	−0.038	0.193	0.112	0.551
6	0.155	−0.005	−0.069	0.141	−0.097	0.466	**0.554**	−0.387	0.732
7	0.193	0.101	0.111	**0.727**	−0.103	−0.035	−0.059	−0.003	0.604
8	0.048	**0.663**	0.152	0.122	−0.042	0.328	0.107	−0.043	0.603
9	0.286	**0.612**	0.141	0.059	−0.036	−0.065	−0.011	0.200	0.526
10	−0.039	−0.003	0.053	0.014	0.022	−0.106	**0.829**	0.242	0.762
11	**0.547**	0.146	0.472	0.111	0.095	−0.083	0.164	−0.274	0.674
12	0.203	0.041	**0.521**	0.438	−0.013	0.146	−0.009	0.221	0.577
13	0.018	−0.032	−0.138	0.421	**0.574**	0.195	0.044	0.001	0.567
15	−0.137	0.218	**0.531**	0.238	0.097	0.090	0.010	−0.286	0.504
16	−0.031	−0.023	**0.785**	−0.049	0.082	0.018	0.018	0.147	0.649
17	0.065	**0.731**	−0.093	0.081	0.181	0.005	−0.050	−0.134	0.607
18	−0.042	−0.007	0.057	0.115	−0.032	0.070	0.160	**0.795**	0.682
21	**0.755**	0.189	−0.113	0.143	0.042	0.033	0.053	−0.079	0.650
22	**0.801**	0.035	0.014	0.053	−0.006	0.015	−0.031	0.030	0.649
Eigenvalue after rotation	**3.403**	1.566	1.320	1.251	1.226	1.133	1.037	1.029	
% explained Variance	**18.068**	8.242	6.948	6.582	6.452	5.964	5.457	5.415	
% Cumulative	**18.086**	26.226	33.161	39.746	46.201	52.159	57.606	63.012	
Cronbach alpha									
CR	**0.89**								
AVE	**0.52**								

*Bold-faced type indicates the highest loading factor.

Data analysis

Respondents' answers on the final 20-item MHK Scale were scored as correct (1) or incorrect (0), with a higher score indicating a higher level of mental health knowledge. Descriptive statistics were used to obtain mean scores, standard deviations, and frequencies and to assess respondents' sociodemographic data. Reliability and validity tests were carried out according to the final scale. The validity of each item was performed using point-biserial correlation and exploratory factor analysis (EFA). To check the internal consistency reliability of the items, Kuder–Richardson formula 20 (KR-20) was employed. In this study, the KR-20 was appropriately utilized for a questionnaire with dichotomous choices (i.e., true or false) [29]. To interpret the findings, a reliability coefficient greater than 0.70 was considered satisfactory [30]. The difficulty index

ranged from 0 (no respondent answered correctly) to 1 (all respondents answered correctly). In this study, the item difficulty index was classified based on the criteria of Kheyami *et al.*: Difficult (<0.30), acceptable (0.30–0.70), and easy (>0.70) [31]. Then, the preliminary results of Indonesian pregnant women's knowledge about mental health and antenatal depression were classified into three categories; low (0–7 pts), medium (8–15 pts), and high (16–23 pts).

Ethical approval

This study was conducted after the researchers obtained ethical approval from the Ethics Commission of the Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada Yogyakarta [Ethical approval code: KE/FK/1208/EC/2018]. As an ethical procedure, we informed all respondents of the study purpose, and the confidentiality of their data was assured. Then, they signed a consent form before participating.

Results

Table 4 shows that 23 items have been constructed and 3 of them were dropped based on the validity results of the KR-20 test. Based on the point-biserial correlation coefficient, of the initial 23 items, there were 20 valid items (r-value > r-table; [r-table = 0.1603]). A total of 3 items were declared invalid (Items 19, 20, and 23) and eliminated as a result of application (e.g., antenatal depression does not affect milk production [the correct answer is "false"]) (Table 4). Furthermore, 20 valid items were tested for reliability using the KR-20 formula because the items used a true-false dichotomy answer. As a result, the reliability coefficient of KR-20

Table 3: Participant characteristics

Variable	n	%	Variable	n	%
Mother's age (in years)			Husband's education		
<20	6	4.0	Elementary school	8	5.4
20-35>35	129	86.0	Secondary school	26	17.4
	15	10.0	High school	96	68.4
			University	19	12.8
Gravida			Husband's occupation		
Primigravida	55	36.7	Laborer	24	16.1
Multigravida	95	63.3	Farmer/seller	12	8.1
			Private employee	94	63.1
			Government employee	2	1.1
			Other	17	11.4
Parity			Resident		
0	55	36.7	Own house	67	44.7
1	52	34.7	Parent's house	48	32.0
2>2	29	19.3	In-law house	35	23.3
	14	9.4			
Planned pregnancy			Health support		
Yes	123	82.0	Yes	142	94.7
No	27	18.0	No	8	8.3
Mother's education			The use of contraception before pregnant	39	26.0
Elementary school	8	5.3	Yes	111	74.0
Secondary school	27	18.0	No		
High school	91	60.7			
University	24	16.0			
Mother's occupation			Bad pregnancy history		
Unemployed/housewife	101	67.3	Yes	20	13.3
Work from home	12	8.0	No	130	86.7
Work outside	37	24.7			

Table 4: Items and summary statistics for the initial 23-item MHK scale

No	Items (n = 23)	r-value	p	Interpretation
1.	Antenatal depression is sadness for more than 2 weeks in pregnancy	0.27	0.001	Valid
2.	People who have experienced insecurity are at risk of mental health disorders	0.51	0.000	Valid
3.	Hormone balance is a cause of antenatal depression	0.39	0.000	Valid
4.	The genetics factor can be reinforcing to mental disorders	0.42	0.000	Valid
5.	Difficulty concentrating is a sign that pregnant women are depressed	0.50	0.000	Valid
6.	Diet changes are not a symptom of antenatal depression	0.32	0.000	Valid
7.	Depression appears in the loss of desire to perform normal activities	0.41	0.000	Valid
8.	Pregnant women who are too young, are at risk for antenatal depression	0.47	0.000	Valid
9.	A history of past parenting can be a trigger for depression during pregnancy	0.43	0.000	Valid
10.	Poor communication with husband is not a cause of antenatal depression	0.27	0.044	Valid
11.	Unplanned and unwanted pregnancy triggers antenatal depression	0.52	0.000	Valid
12.	Early detection of depressive symptoms can be felt by pregnant women from a faster heartbeat.	0.48	0.000	Valid
13.	Depression in pregnant women can be caused by environments exposed to cigarette smoke	0.38	0.000	Valid
14.	Culture prohibit eat certain foods to pregnant women is not a factor for depression	0.27	0.000	Valid
15.	Antenatal depression is common in primigravida	0.38	0.000	Valid
16.	Antenatal depression is often at 7-9 months of gestation	0.30	0.000	Valid
17.	Depression in early pregnancy is usually due to miscarriage	0.39	0.000	Valid
18.	Depression does not affect low birth weight	0.22	0.004	Valid
19.	Fetal growth retardation not associated with antenatal depression (D)	0.05	0.227	Invalid
20.	Antenatal depression does not affect breast milk production (D)	−0.03	0.601	Invalid
21.	Breathing exercises can help prevent depression	0.48	0.000	Valid
22.	Husband's assistance in doing homework can reduce the risk of antenatal depression	0.35	0.033	Valid
23.	Depression can be prevented by exercise (D)	0.12	0.071	Invalid

D = Item was dropped; _r_-table with _df_ (148) is 0.1603.

is 0.717 (greater than the acceptance limit of 0.70). It indicates that the internal consistency value was good and the instrument is reliable. Thus, we claimed that a final 20-item scale is considered a valid and reliable instrument.

Content validity

Based on the results of the literature search, FGD, and in-depth interviews, a total of 23 items in the initial knowledge questionnaire were developed. The results of the interviews were transcribed verbatim and analyzed using thematic analysis. The researchers then received written feedback from experts and professionals, who had assessed the instrument qualitatively about the relevancy of the content, clarity, and simplicity of each item statement related to maternal depression. Some items were revised and improved by applying the expert panel's opinions quantitatively. The value of Average Congruency Percentage (ACP) = 0.97 showed that the overall coefficients of each item were greater than the valid criteria (≥0.70); therefore, each item of knowledge statement was considered very appropriate or relevant (30). In brief, all panelists rated the questionnaire with a validity index above 0.70,

indicating that the questions were relevant, clear, and unambiguous for respondents. The scores of relevancy, clarity, simplicity, CVR, and I-CVI are shown in Table 1.

Face validity

Of the five pregnant women, two persons are the first pregnancy in the second trimester. Three pregnant women are the second pregnancy in the third trimester. The results of the readability test involved revising of term _hormone_ with the more common term _fluid_ that controls bodily functions and also revising some sentences which were too long. After the instrument was corrected, it was shown back to the respondent one by one for the next readability tests. All face validity respondents were not involved in the construct validity test.

Construct validity

Participants included 150 pregnant women. The mean of the pregnant women's age was 28.21 (±5.25). More than one-third (36.7%) of mothers were primigravidae. Their gestational age was between 4 and 38 weeks. Table 3 shows the characteristics of the participants.

The first test findings on 150 participants of pregnant women have obtained the feasibility data test which the first analysis coefficient, the Kaiser-Meyer-Olkin (KMO) index > 0.7 is 0.716. It means that it is in the sufficient category but can be continued to conduct Exploratory Factor Analysis (EFA). The result of the Bartlett test of sphericity (p = 0.000 < 0.05) was in line with KMO. The correlation in each item of MSA > 0.5 with an average of MSA was 0.632, so the two factors analysis requirements were accomplished. One variable (P14) with an MSA value < 0.5 was eliminated and re-analyzed. In the last stage analysis, it is obtained by the Kaiser-Meyer-Olkin index (KMO) > 0.6, which is 0.713. It indicates that the level of variation and the sample are sufficient to conduct a factor analysis (miserable). The result of the Bartlett test of sphericity (p = 0.000 < 0.05) was still in line and showed a significant correlation among items. The correlation of each item of MSA > 0.5 with an average of MSA is 0.54 thus the requirements of both analysis factors were accomplished.

Based on the analysis of communalities, the eight factors formed showed that empirically the instruments of knowledge with 20 items were valid to measure the level of knowledge with six developed indicators. Based on the results of eigenvalue > 1, there were eight factors that exist in component numbers 1, 2, 3, 4, 5, 6, 7, and 8. The factors formed show that there are only eight factors that are best for summarizing the whole items of mental health knowledge of pregnant women. About 63.021% of the total variance of all items could be explained by 8 factors, and the communalities

items value is about 0.504 to 0.762. There were seven items that had the lowest communalities value. Based on the eight component factors, it can be interpreted into indicators of knowledge assessment of pregnant women, namely (1) the definition of depression, (2) the causes, (3) risk factors, (4) symptoms or signs, (5) the impact of depression, (6) attempts to prevent (7) efforts to overcome, and (8) family support.

From the data score of valid items (20 items), it was obtained the mean±(SD) of the pregnant women's knowledge which was 12.04 ± (3.79) (range 2–19). The results obtained from the calculation of the percentage of pregnant women who received a minimum score of 2 was 1.3%, while those who received a maximum score of 19 were 2 people (1.3%). The analysis result got the AVE value of 0.52 (> 0.5). This shows a strong correlation among items that measure the same construct so that the instrument has good discriminant validity. The CR value was 0.89 (> 0.7), this indicates that all subscales have good convergent validity. Items that have no correlation are related.

The findings of this study indicated that there were 20 items stated valid and reliable. Based on the difficulty index, each item was in the range between 0.39 and 0.82. Items were in the category of easy (15%), moderate (70%), and difficult (15%). This knowledge instrument was a good category and acceptable items. There were three questions that have a difficulty index of 0.82. That value was equal to the difficulty index which corresponds to the easy category.

Difficulty index

Based on the difficulty index, each item was in the range between 0.39 and 0.82. All items were in two main categories: easy (2 items; 10.0%) with difficulty index more than 0.80 and acceptable (18 items; 90.0%) with difficulty index 0.39–0.65. Accordingly, this knowledge instrument was considered in the good category. Based on their difficulty indexes, all final items were appropriate.

The preliminary results of mental health knowledge

Of all the participants, the mean score of correct answers for the whole questionnaire was 12.03 out of 20 points (60.15%) and the standard deviation was 3.76. The highest score obtained by respondents was 19 and the lowest score was 3. According to Table 5, the item most answered correctly by the participants was item 21 (n = 123, 82.0%, i.e., breathing exercises can help prevent depression). Meanwhile, the item most answered incorrectly is Item 18 (n = 58, 39.0%, i.e., depression does not affect low birth weight). In general, the results indicate that the knowledge of Indonesian pregnant women about mental health and antenatal depression is classified as moderate (*M* = 12.03; *SD* = 3.783).

Table 5: Mental health knowledge among Indonesian pregnant women

Items (n=20)	Mean	SD	Frequency (n = 150)	
			Yes (%)	No (%)
Q1	0.593	0.493	89 (59)	61 (41)
Q2	0.660	0.475	99 (66)	51 (34)
Q3	0.467	0.501	70 (47)	80 (53)
Q4	0.553	0.499	83 (55)	67 (45)
Q5	0.600	0.492	90 (60)	60 (40)
Q6	0.600	0.492	90 (60)	60 (40)
Q7	0.640	0.482	96 (64)	54 (36)
Q8	0.660	0.475	99 (66)	51 (34)
Q9	0.473	0.501	71 (47)	79 (53)
Q10	0.453	0.499	69 (45)	82 (55)
Q11	0.660	0.475	99 (66)	41 (34)
Q12	0.700	0.460	105 (70)	45 (30)
Q13	0.647	0.480	97 (65)	53 (35)
Q14	0.613	0.489	92 (61)	58 (39)
Q15	0.507	0.502	76 (51)	74 (49)
Q16	0.553	0.499	83 (55)	67 (45)
Q17	0.647	0.480	97 (65)	43 (35)
Q18	0.387	0.489	58 (39)	92 (51)
Q21	0.820	0.385	123 (82)	27 (18)
Q22	0.800	0.401	120 (80)	30 (20)
Total	12.03	3.783		

Discussion

This research was done to develop and validate a survey to evaluate the mental health knowledge (MHK) for Indonesian pregnant women. The final version of the MHK scale consisted of 20 items related to mental health in terms of pregnancy depression disorders from definition and impacts to prevention efforts and family support. The findings of this study indicated that this questionnaire is a reliable and valid tool for measuring maternal understanding about depression during pregnancy and in the postnatal period and also about its prevention attempts. According to literature [9], [10], [11], content validity is an important step in instrument validation. This study included content validation by accepting feedback given by experts in psychology and obstetrics. Content validity aims to determine that the content in each question is sufficient to evaluate the objectives of the study [32]. The researchers asked the panelists to provide their professional opinions in terms of relevance, simplicity, and clarity in each item.

At first, 23 knowledge questions were reduced to 20 after 3 questions were omitted as a result of the application. Items 19 and 20 included in the depressive impact domain have been dropped. These two questions were not revised because they were represented by item 18 on the impact of depression on low birth weight. Question 23 included in the domain of prevention efforts. However, it has been dropped because it is represented in point 21 with breathing exercises (Table 4). Item 21 had the highest score. Most of the participants understand that exercise can make the mind refreshed, so they think exercise can prevent depression [33], [34]. Item 12 also had the highest score because the participants thought, if he experiences a stressor and anxiety or mood disorder, a faster heartbeat is a sign that is easily felt physically [35]. Conversely, item 18 had the lowest score. Pregnant women considered that low birth weight is caused by lack

of nutrition and does not impact depressive disorders. This opinion is not in line with or lacks implementation of pregnant women. The reason is, pregnant women who experience depression tend to lead unhealthy lifestyles, such as paying less attention to nutrition which results in low birth weight [36]. In general, we reported that the mental health knowledge of pregnant women in Indonesia tends to be unsatisfactory. Our findings are in line with the study of Kingston *et al*. They found that knowledge of prenatal mental health and its effects on child development should be enhanced to a satisfactory level [2]. Therefore, their understanding needs to improve by providing proper education so that pregnant women can be stressed and at risk of depression.

The development of this instrument is intended to measure the knowledge level of pregnant women about mental health and mental disorders during pregnancy and puerperium. In line with the aims of this study, the previous literature [31], [37] found that mental health literacy includes not only knowledge but also attitudes, stigma, positive mental health, and beliefs to seek help related to mental problems and mental disorders. Instruments developed based on mental health literacy theory could help practitioners, program developers, and decision-makers to make decisions in improving the mental health literacy of pregnant women at the individual and community level by measuring the knowledge accurately [38].

A valid instrument can measure concepts correctly, for example, knowledge about symptoms of pregnancy depression. According to the previous studies in South Africa, most of the pregnant women were unable to identify the signs and symptoms described in the maternal mental disorders outline [33]. The researchers believed that MHKS can accurately measure the knowledge level of pregnant women about maternal depression. These items cover the information needs which were expected from FGDs with pregnant women. This result is in line with studies in Australia that found mental health literacy about postpartum depression needs an understanding of attempts to prevent it, such as through family and group support, relaxation/time alone, and regular sleep [34], [35], [39]. Information about the causes and risk factors is also needed, including hormonal changes, younger age, lack of parenthood, obstetric factors, single-parent status, and lack of sleep [3]. Other factors that contribute to maternal depression are marital problems, less self-confidence, financial problems, and traumatic incidents [36]. This mental health knowledge instrument still prioritizes the understanding of cognitive material about mental health disorders. The researchers argue that with a good understanding of the possibility of mental health disorders during pregnancy, pregnant women can be better prepared and make efforts to overcome and prevent it. Questions on help-seeking behavior when experiencing some interference were not included in this knowledge instrument even though it belongs in efforts to increase mental health literacy [36], [40].

Strengths and weaknesses

Overall, the MHKS indicates reasonable validity and high internal consistency. This study's strengths included a high response from participants. Besides, the questionnaire items which were true and false questions are considered suitable to help pregnant women answering easily based on the low and middle education level. The advantages of true and false questions are: The item can cover a large range of material needed and is simple and easy to design. Moreover, it is easy to administer and score. In addition, the instrument is better if measuring the facts and learning outcomes directly or information needed in a short time [15]. In some cases, the young adults or older subjects can more easily answer scaled items with two options, especially when the information is not well-known [41]. True and false responses may be best suited for assessing knowledge, but it assumes that the participants are considered to lack knowledge when responding [42]. A limitation of testing this instrument is the difficulty of random sampling, the possibility that sampling does not represent the distribution of all pregnant women at different gestational ages.

Conclusion and Suggestion

The instrument of mental health knowledge of pregnant women with 20 items in the structure of six factors concerning maternal mental disorders related to depression and how to overcome them was successfully designed and developed as a valid and reliable psychometric measurement. The instrument can be applied to measure the knowledge level of pregnant women regarding maternal mental health. Appropriate knowledge can be identified through the participant's correct answers for each question in demonstrating an understanding of the material through analyzing concepts, depression causes, signs and symptoms, risk factors, problems analyzing, and attempts to overcome and prevent them.

In the future, researchers who want to evaluate mental health education are expected to re-evaluate this psychology knowledge instrument. This instrument does not include knowledge about seeking first aid and professional assistance if pregnant women have mental health problems. For the health care providers for mothers and children, as well as other researchers, this instrument can be used to better know the level of understanding of maternal mental health and depression disorders more appropriately. Thus, it can be used to determine the correct attempts to improve that knowledge, for example, psychological education

activities to reduce the risks and the impacts of pregnancy depression and postpartum disorders during the puerperium period. Since this study concentrates on the level of Mental Health Knowledge (MHK) among Indonesian pregnant women, this instrument can be utilized to explore the MHK levels of pregnant women in different countries, both developed and developing countries. Furthermore, further studies need to be carried out with a large sample size considering that the current study only involves a small sample size from Surakarta, Indonesia. More importantly, further studies should be carried out to discuss the views of Indonesian pregnant women on MHK in a comprehensive manner because our study has only concentrated on developing a valid and reliable MHK scale.

References

1. World Health Organization. Maternal Health. Mental Health and Substance Use. Geneva: World Health Organization; 2018. Available from: https://www.who.int/teams/mental-health-and-substance-use/maternal-mental-health. https://doi.org/10.1037/e400972004-001. [Last accessed on 2019 Nov 01].

2. Kingston DE, Mcdonald S, Austin MP, Hegadoren K, Lasiuk G, Tough S. The Public's views of mental health in pregnant and postpartum women: A population-based study. BMC Pregnancy Childbirth. 2014;14(1):84. https://doi.org/10.1186/1471-2393-14-84
PMid:24564783

3. Highet NJ, Gemmill AW, Milgrom J. Depression in the perinatal period: Awareness, attitudes and knowledge in the Australian population. Aust N Z J Psychiatry. 2011;45(3):223-31. https://doi.org/10.3109/00048674.2010.547842
PMid:21438748

4. Ibanez G, Bernard JY, Rondet C, Peyre H, Forhan A, Kaminski M, et al. Effects of antenatal maternal depression and anxiety on children's early cognitive development: A prospective Cohort study. PLoS One. 2015;10(8):e0135849. https://doi.org/10.1371/journal.pone.0135849
PMid:26317609

5. Fadzil A, Balakrishnan K, Razali R, Sidi H, Malapan T, Japaraj RP, et al. Risk factors for depression and anxiety among pregnant women in Hospital Tuanku Bainun, Ipoh, Malaysia. Asia Pac Psychiatry. 2013;5(Suppl 1):7-13. https://doi.org/10.1111/appy.12036
PMid:23857831

6. Saeed A, Raana T, Saeed AM, Humayun A. Effect of antenatal depression on maternal dietary intake and neonatal outcome: A prospective cohort. Nutr J. 2016;15(1):64. https://doi.org/10.1186/s12937-016-0184-7
PMid:27401187

7. Bjørnsen HN, Eilertsen ME, Ringdal R, Espnes GA, Moksnes UK. Positive mental health literacy: Development and validation of a measure among Norwegian adolescents. BMC Public Health. 2017;17(1):717. https://doi.org/10.1186/s12889-017-4733-6

8. Wei Y, McGrath PJ, Hayden J, Kutcher S. Measurement properties of tools measuring mental health knowledge: A systematic review. BMC Psychiatry. 2016;16(1):297. https://doi.org/10.1186/s12888-016-1012-5
PMid:27553955

9. Baghestani AR, Ahmadi F, Tanha A, Meshkat M. Bayesian critical values for lawshe's content validity ratio. Meas Eval Couns Dev. 2019;52(1):69-73. https://doi.org/10.1080/07481756.2017.1308227

10. Hendryadi H. Validitas Isi Tahap awal pengembangan kuesioner. J Riset Manajemen Bis Fakult Ekon. 2012;2(2):169-78. https://doi.org/10.36226/jrmb.v2i2.47.

11. Lawshe CH. A quantitative approach to content validity. Pers Psychol. 1975;28(4):563-75.

12. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Front Public Heal. 2018;6:149. https://doi.org/10.3389/fpubh.2018.00149
PMid:29942800

13. Kingston D, Austin MP, Heaman M, McDonald S, Lasiuk G, Sword W, et al. Barriers and facilitators of mental health screening in pregnancy. J Affect Disord. 2015;186:350-7. https://doi.org/10.1016/j.jad.2015.06.029
PMid:26281038

14. O'Connor M, Casey L. The mental health literacy scale (MHLS): A new scale-based measure of mental health literacy. Psychiatry Res. 2015;229(1-2):511-6. https://doi.org/10.1016/j.psychres.2015.05.064
PMid:26228163

15. Austin MP, Middleton P, Reilly NM, Highet NJ. Detection and management of mood disorders in the maternity setting: The Australian clinical practice guidelines. Women Birth. 2013;26:2-9. https://doi.org/10.1016/j.wombi.2011.12.001
PMid:22217978

16. Spedding MF, Stein DJ, Naledi T, Sorsdahl K. Pregnant women's mental health literacy and perceptions of perinatal mental disorders in the Western Cape, South Africa. Ment Health Prev. 2018;11:16-23. https://doi.org/10.1016/j.mhp.2018.05.002

17. Madlala S, Kassier S. Antenatal and postpartum depression: Effects on infant and young child health and feeding practices. S Afr J Clin Nutr. 2018;31(1):1-7. https://doi.org/10.1080/16070658.2017.1333753

18. Schuwirth LW, Van Der Vleuten CP. Different written assessment methods: What can be said about their strengths and weakness? Med Educ. 2004;38(9):974-9. https://doi.org/10.1111/j.1365-2929.2004.01916.x
PMid:15327679

19. Widoyoko E. Teknik Penyusunan Instrumen Penelitian. Yogyakarta: Pustaka Pelajar; 2012.

20. Davis LL. Instrument review: Getting the most from a panel of experts. Appl Nurs Res. 1992;5(4):194-7.

21. Azwar S. Reliabilitas dan validitas. 4th ed. Yogyakarta: Pustaka Pelajar; 2018. p. 181.

22. Tsai AC, Scott JA, Hung KJ, Zhu JQ, Matthews LT, Psaros C, et al. Reliability and validity of instruments for assessing perinatal depression in African settings: Systematic review and meta-analysis. PLoS One. 2013;8(12):e82521. https://doi.org/10.1371/journal.pone.0082521
PMid:24340036

23. Lynn MR. Determination and quantificatiom of conten validity. J Exp Psychol Gen. 1986;35(6):382-6.
PMid:3640358

24. Sartori R. Face validity in personality tests: Psychometric instruments and projective techniques in comparison. Qual Quant. 2010;44(4):749-59. https://doi.org/10.1007/s11135-009-9224-0

25. Comrey A, Lee H. A First Course in Factor Analysis. 2nd ed. Hillside, NJ: Lawrence Erlbaum; 1992.

26. Kaiser HF. An index of factorial simplicity. Psychometrika. 1974;39(1):31-6.

27. Costello AB, Osborne JW. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the most from your Analysis. Vol. 10. Massachusetts: Practical Assessment, Research, and Evaluation; 2005.

28. Cohen L, Manion L, Moriison K. Research methods in education. In: Professional Development in Education. 6th ed., Vol. 38. London, United Kingdom: Routledge; 2012. p. 507-9.

29. Kuder G, Richardson M. The theory of the estimation of test reliability. Psychometrika. 1937;2(3):151-60.

30. Aday L, Cornelius L. Designing and Conducting Health Surveys: A Comprehensive Guide. 3rd ed. San Francisco: Jossey-Bass; 2006.

31. Kheyami D, Jaradat A, Al-Shibani T, Ali FA. Item analysis of multiple choice questions at the department of paediatrics, Arabian gulf university, Manama, Bahrain. Sultan Qaboos Univ Med J. 2018;18(1):e68-74. https://doi.org/10.18295/squmj.2018.18.01.011
    PMid:29666684

32. Waltz C, Strickland O, Lenz E. Measurement Nursing and Health Research. 4th ed., Vol. 53. New York: Springer Publising Company; 2013. p. 1689-99.

33. Recto P, Champion JD. Assessment of mental health literacy among perinatal hispanic adolescents. Issues Ment Health Nurs. 2017;38(12):1030-8. https://doi.org/10.1080/01612840.2017.1349224
    PMid:28766973

34. Spiker DA, Hammer JH. Mental health literacy as theory: Current challenges and future directions. J Ment Health. 2019;28(3):238-42. https://doi.org/10.1080/09638237.2018.1437613
    PMid:29436864

35. Daley AJ, Jolly K, Sharp DJ, Turner KM, Blamey RV, Coleman S, *et al*. The effectiveness of exercise as a treatment for postnatal depression: Study protocol. BMC Pregnancy Childbirth. 2012;12(1):45. https://doi.org/10.1186/1471-2393-12-45
    PMid:22682671

36. Daley AJ, Foster L, Long G, Palmer C, Robinson O, Walmsley H, *et al*. The effectiveness of exercise for the prevention and treatment of antenatal depression: Systematic review with meta-analysis. BJOG. 2015;122(1)57-62. https://doi.org/10.1111/1471-0528.12909
    PMid:24935560

37. de Souza AC, Alexandre NM, de Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. Epidemiol Serv Saude. 2017;26(3):649-59.
    PMid:28977189

38. Wei Y, McGrath PJ, Hayden J, Kutcher S. Mental health literacy measures evaluating knowledge, attitudes and help-seeking: A scoping review. BMC Psychiatry. 2015;15(1):291. https://doi.org/10.1186/s12888-015-0681-9
    PMid:26576680

39. Thorsteinsson EB, Loi NM, Moulynox AL. Mental health literacy of depression and postnatal depression: A community sample. Open J Depress. 2014;3(3):101-11. https://doi.org/10.4236/ojd.2014.33014

40. Jorm AF, Kelly CM. Improving the public's understanding and response to mental disorders. Aust Psychol. 2007;42(2):81-9.

41. Guy S, Sterling BS, Walker LO, Harrison TC. Mental health literacy and postpartum depression: A qualitative description of views of lower income women. Arch Psychiatr Nurs. 2014;28(4):256-62. https://doi.org/10.1016/j.apnu.2014.04.001

42. Evans-Lacko S, Little K, Meltzer H, Rose D, Rhydderch D, Henderson C, *et al*. Development and psychometric properties of the mental health knowledge schedule. Can J Psychiatry. 2010;55(7):440-8. https://doi.org/10.1177/070674371005500707
    PMid:20704771