Hammouda *et al. BMC Pediatrics* (2025) 25:309
https://doi.org/10.1186/s12887-024-05286-w

BMC Pediatrics

RESEARCH Open Access

Development and validation of an Arabic questionnaire to assess the knowledge of neonatal danger signs (AQ-KNDS) and its possible determinants among Arab population

Esraa Abdellatif Hammouda[1,3*], Suzan Abdel-Rahman[2], Marwa Abdelwahab Hassan[1,3], Marwa Abdelkarim[4], Engy Saad Abdelmoneim Elkaragy[4*], Ayed A. Shati[5] and Ramy Mohamed Ghazy[6]

Abstract

Background The neonatal period is the most dangerous time during which the newborn challenges the highest risk of death. This study aimed to develop a validated questionnaire to assess the knowledge of the population about neonatal danger signs (NDS) in Arab-speaking communities.

Methods A cross-sectional survey was conducted; participants were recruited from the population visiting El-Raml Pediatric Hospital and El-Shatby Hospital in Alexandria in 2023.The study used Item Response Theory (IRT) techniques to evaluate the validity of the Arabic questionnaire. Various IRT models (1PL, 2PL, 3PLS) were employed to improve the scale's accuracy. To assess how the item response model fits the data, we used the M2 index and other fit indices (Comparative Fit Index (CFI), Tucker–Lewis index (TLI), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC)). Additionally, item fit indices including Pearson's $\chi 2$ and Root Mean Square Error of Approximation (RMSEA) were measured to determine how well each item fits the final model. Reliability was assessed using Cronbach's alpha and test-retest reliability. Additionally, we utilized a logistic regression model to identify the predictors of knowledge of the NDS.

Results A total of 283 participants were included, their mean age was 32.3 ± 8.62 years, 96.5% were female, 94.7% were currently married, 27.6% had a university or higher education, and 83.7% were of urban residence. The final questionnaire consists of 16 items. The best model was retained where its M2 statistics were comparatively low indicating that there was no significant difference between the model and the data (M2 = 84 with 88 degrees of freedom and a $P = 0.688$) with the following fit measures RMSEA = 0.001, AIC = 2650.04, BIC = 2825.7, and TLI and CFI were 1.0. Full information factor analysis indicated that the total proportion of variance extracted by the model was

*Correspondence:
Esraa Abdellatif Hammouda
Esraa.alatif@gmail.com
Engy Saad Abdelmoneim Elkaragy
engy.saad@alexmed.edu.eg
Full list of author information is available at the end of the article

63.7%. Multiple logistic regression indicated that the explanatory variables for the level of knowledge toward danger signs were female sex (AOR = 5.54, 95% CI:1.25–31.0, P = 0.034), age (AOR = 1.04. 95%CI:1.01–1.08, P = 0.025), and working outside the medical field (AOR = 3.26, 95%CI: 1.14–9.73, P = 0.034).

Conclusions The developed questionnaire is valid and reliable in informing public health policymakers about community awareness regarding NDS and implementing interventions to improve neonatal health and reduce newborn morbidity and mortality rates.

Keywords Neonatal danger signs, Knowledge, Validation, Reliability

Introduction

The neonatal period refers to the first 28 days of life. Neonates go through a lot of physiological and environmental changes during this most vulnerable time to adjust to extra-uterine life [1, 2]. So, this period displays the most dangerous time during which the newborn challenges the highest risk of death [1, 3]. The majority of neonatal deaths occur within the first week of life due to several causes like birth asphyxia, sepsis, respiratory distress syndrome, preterm birth, low birth weight, low socioeconomic status, and cesarean section delivery associated complications. [3–6].

Neonatal health is a key indicator of the general health of countries determined by the total survival of neonates per live births. In 2021, the global average neonatal mortality rate was estimated to be 18 deaths per 1,000 live births [7]. Notably, the third goal of the sustainable development goals (SDG) to be accomplished by 2030 is to improve the health of neonates and children on a national and international level [8], aiming to reduce neonatal mortalities to at least as low as 12 per 1,000 live births [9]. The United Nations International Children's Emergency Fund (UNICEF) recommended achieving this goal through multiple policies including investigating the factors affecting knowledge regarding neonatal illness and health-seeking practices [10, 11]. Despite the global attempts to improve neonatal health, the decline in neonatal deaths between 1990 and 2021 was slower than that in infants and children aged 1–59 months (40.0% of the global under-5 mortality in 1990 while 47.0% in 2021). The World Health Organization (WHO) attributed it to the lack of postnatal care, early discharge from hospitals (before 24 h) or home delivery, late identification of neonatal illness, and delay in seeking care, in the regions with high mortality rates and limited resources [3].

Children who were born in Low and Middle-Income Countries (LMIC) were found to be ten times more likely to die in the first month of life compared to those who are born in developed countries [3, 12]. In 2019, Sub-Saharan Africa had the highest neonatal mortality rate (27 deaths per 1,000 live births) accounting for 43.0% of global neonatal deaths, followed by Central and Southern Asia (24 deaths per 1,000 live births) with 36.0% of global neonatal deaths [3]. Globally, in 2020, India had the highest neonatal mortality rate with 490,000 deaths, this figure is almost double that in Nigeria with 271,000 deaths in the same year [3]. The under five and neonatal mortalitiy rates in Egypt in 2021 were 18 and 9 per 1,000 live births respectivley which is lower than the global figures [13, 14].

It is worth noting that 75% of neonatal deaths could be avoided with simple identification and cost-effective management tools [15]. Knowledge and early detection of neonatal danger signs (NDS) are key to reducing neonatal mortality. The WHO stated that the NDS are poor feeding, convulsion, fast breathing, chest drawing, high-grade fever, hypothermia, jaundice in the 1st 24 h involving the palm and sole, unconsciousness, lethargy, and umbilical redness orpus drainage [16–19]. Former studies conducted in different countries reported inconsistent results regarding the level of mothers' knowledge and the possible related factors about NDS. The level of good knowledge among mothers ranged from 10 to 81% [10, 15, 18, 20–24]. This wide range in knowledge could be due to the lack of a validated tool and different scoring systems that assess knowledge about NDS. Therefore, this study aimed to develop a validated and reliable Arabic questionnaire to assess population knowledge regarding the NDS (AQ-KNDS).

Methodology

Study design and setting

A cross-sectional survey was used to recruit the study participants from El-Raml Pediatric Hospital and the Alexandria Main University Hospital for Children (El-Shatby) between September 1 and November 30, 2023. In Alexandria, these hospitals are the largest pediatric facilities affiliated with the Ministry of Health and Population (MoHP) and Alexandria University, respectively. They serve nearly 450,000 patients each month across three governorates in the Egyptian Western North coast (Alexandria, Behira, and Matruh) through outpatient clinics, inpatient departments, and intensive care units for children (PICU) and newborns (NICU).

Data collection tool

Participants were interviewed by trained healthcare professionals. The questionnaire consisted of two sections.

The first section included the sociodemographic data (age, sex, education level, residence, marital status, and occupation). The second section included the Arabic questionnaire to assess knowledge about the NDS (AQ-KNDS).

Validation of the questionnaire

Item generation

The authors conducted extensive literature search and revised previous studies in the specific field of interest [18, 25–29]. Then, discussions were held with individuals and experts possessing relevant knowledge of NDS to manage the questionnaire's items, wording, and overall flow. An expert panel for the questionnaire validation was selected according to their experience and comprised seven experts (one public health professional, one methodologist, and 5 pediatricians and neonatologists).

Item formatting

The authors generated the pool of items (17 danger signs) and eliminated any potential ambiguity, technical vocabulary, or bias. Authors avoided lengthy sentences and negative words. Then, we ensured the logical flow of items and assessed their face validity. Face validity was applied to guarantee that the questionnaire accurately measured what it was intended to measure. For the knowledge scores, the values assigned were as follows: 0 for "no" and "don't know," and 1 for "yes." This allowed for a scoring system that quantifies participants' knowledge based on their responses.

Content validity and expert evaluation

Content validity is defined as "the degree to which an assessment instrument is relevant to, and representative of the targeted construct it is designed to measure" [30]. Based on the item-content validity index (I-CVIs), the experts evaluated individually the items and scored them from 0 to 1. I-CVI was calculated by dividing the number of experts by giving a rating of "very relevant" for each item by the total number of experts where the item was considered relevant when scoring > 0.79, and between 0.70 and 0.79, it needed revision. If the value was below 0.70 the item was removed. The expert panel changed the wording of items, added 13 questions, and removed 2 questions and the number of items of danger signs was 28. (Supplementary file 2)

Pilot testing and cognitive interviews

The research team conducted cognitive interviews with 30 participants from both sexes representing different educational levels stratified according to their working sector into medical and non-medical field subjects, all representing the targeted population, to evaluate the language clarity, accuracy of questions, relevance, cultural compatibility, the optimal order of the questions, and participant comprehension of the questions. Based on the feedback received, minor modifications were made to the questionnaire items.

Sample size & sampling method & population of the study

Based on the sample size recommendations of having 10 participants respond to each item for validating a questionnaire (ratio 10:1) [31, 32], the minimum required sample was 280 participants. The survey was conducted in confidential areas in both hospitals by a trained research team after receiving training on how to conduct interviews, and basics of neonatal life and diseases.

Psychometric evaluation of the Arabic questionnaire

Item response theory (IRT) techniques

We used IRT models to gain insights into the validity and reliability of the items. We used IRT to create a smaller scale and develop an item tool with a high degree of validity and reliability. IRT models explain the relationship between a latent ability (denoted θ) and its observable items. IRT models focus mainly on the pattern of responses and their probabilistic terms, in contrast to classical test theory (CTT) which focuses on composite variables and regression theory. IRT accounts for 3 parameters: the item's ability to distinguish between respondents with different proficiency levels (item discrimination), the likelihood of correct response (item difficulty), and the guessing probability [14]. IRT creates powerful scales that can differentiate between items professionally. We conducted the appropriate IRT models to estimate latent variables from binary items, named 1PL, 2PL, and 3PLS models where P refers to the number of item parameters assigned for modeling the relationship between θ and item response. We conducted IRT analysis using package "mirt" [7].

IRT analysis yields factor solution (factor loading F1 and communalities h_2) like factor analysis. To assess how the item response model fits the data, we used M2 index and other fit indices (Comparative Fit Index (CFI), and Tucker–Lewis index (TLI), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC)). Lower M2, AIC, and BIC values, along with higher values of CFI and TLI, suggest a more suitable model. Item fit indices including Pearson's $\chi 2$ (S-$\chi 2$) and corresponding degrees of freedom (df), Root Mean Square Error of Approximation (RMSEA), and P-values were measured to determine how well each item fits the final model. RMSEA is a statistic used to assess the goodness of fit of a model, indicating how well the model aligns with the data. Values less than 0.05 are considered very good, while values between 0.05 and 0.08 fall within the acceptable range. If the RMSEA value exceeds 0.10, it suggests issues with the model. Respondent fit indices were also measured to assess whether the participants's response pattern was consistent with the model [8]. The Person Item Map (PIM) was used to assess the extent to which items cover the range of latent ability. PIM plots each item's location and distribution of person parameters on the same θ scale. During the IRT analysis, we tested two key assumptions: uni-dimensionality and local independence. Uni-dimensionality means that a set of items measures only one underlying trait. Local independence indicates that the responses to individual items are independent of one another when considering the latent trait. To assess uni-dimensionality, we used IRT fit indices, while we evaluated local independence by examining the residuals of item responses and using fit statistics. Assessing these assumptions is essential to ensure the validity of the IRT model results and the accuracy of the parameter estimates [33].

Reliability analysis

Cronbach's alpha was used to evaluate the internal consistency of the AQ-KNDS. Generally, Cronbach's alpha of 0.70 to 0.80 is considered acceptable, while a value exceeding 0.80 is considered well acceptable. The split-half reliability analysis determined the questionnaire's reliability by dividing responses into two sections and comparing the scores from each part. The Guttman Split-Half coefficient measured the AQ-KNDS reliability by splitting the items and comparing the resultant scores. Meanwhile, the Spearman-Brown coefficient estimates the measure's reliability when its length is altered.

To assess the consistency of our survey results over time, we utilized the test-retest reliability method. This involves giving the same test to the same respondents at two different points in time and then correlating the results. This approach helped us determine whether the instrument yields stable results under consistent conditions [34].

Initially, the survey was administered as a baseline or "pretest" assessment. At a later stage, after 15 days, a "posttest" was applied. We chose the 15-day interval because shorter intervals could lead participants to remember their previous answers, which might inflate the correlation coefficients. By using a 15-day interval, we minimize this risk while keeping the underlying characteristics stable. This duration is long enough to capture any genuine changes in what we're measuring, ensuring that our analysis reflects the instrument's true reliability rather than just temporary variations [35, 36]. We applied the test-retest on 100 participants of our sample. In terms of sample size, a sample size of 100 is often considered sufficient to achieve a reliable statistical analysis. It increases the likelihood that the results reflect the true characteristics of the larger population. It also allowed for greater diversity in responses, which enhanced the generalizability of the results. With this number of respondents, the confidence intervals for reliability estimates became narrower, providing more accurate estimates of the instrument's reliability [37–39].

As the scores from the initial survey and the follow-up were highly correlated with consistent scores and minimal error variances, the test-retest reliability was demonstrated to be reliable. The Pearson's correlation test was used to quantify this reliability. The intraclass correlation coefficient (ICC) was also utilized to assess the extent to which overall knowledge scores congruent between the two distinct time points [40]. ICC ranges from 0 to 1, where higher values explained greater reliability in repeated testing.

Statistical analysis

The quantitative variables were described using the mean and standard deviation, while the categorical variables were summarized with frequencies and percentages. The Chi-square test was employed to examine the relationships between the categorical variables and their associations with the knowledge level. Knowledge scores were categorized based on the median value. Participants with scores above the median were classified as having good knowledge, while those below the median were classified as having poor knowledge. Assumptions of normal distribution were explored with the Kolmogorov-Smirnov test and Shapiro-Wilk test. We examined the factors influencing knowledge of the NDS using a logistic regression model. The level of significance was set at $P < 0.05$. Data were collected and entered on google drive and the statistical analysis was performed using R software version R 4.3.2.

Results

Table 1 shows the demographic characteristics of the respondents. The average age of the respondents was 32.3 ± 8.6 years. Most respondents were females (96.5%), currently married (94.7%), urban residents (83.7%), over a third (37.1%) had a secondary education, and 27.6% had university or higher education. More than half (55.8%) of the respondents believed their income was sufficient, 12.0% of respondents worked in the medical field, 16.3% worked outside the medical sector, and the majority 71.7% were not working. About 49.0% of respondents who were not working in the medical field had previously been hospitalized with one of the family's children. Supplementary Table 1 showed that lethargy or convulsion were the most frequently reported NDS for 97.9% of the respondents, whereas low body temperature was the least frequently reported sign (61.1%).

Table 1 Characteristics of the study participants ($N=283$)

Characteristics	Labels	Frequency	Percentage (%)
Sex	Male	10	3.5
	Female	273	96.5
Age	Mean ± SD	32.3 ± 8.6	
Marital status	Currently married	268	94.7
	Currently not married	15	5.3
Education level	Illiterate or primary education	54	19.1
	Preparoty education	46	16.3
	Secondary education	105	37.1
	University degree or postgraduate	78	27.6
Place of residence	Urban	237	83.7
	Rural	46	16.3
Income	Enough	158	55.8
	Not enough	125	44.2
Occupation sector	Medical field (physicians, pharmacists, dentists, and nurses)	34	12.0
	Out of the medical field	46	16.3
	Not working or housewife	203	71.7
Hospitalized with one of your family's children?	Yes	138	48.8
	No	111	39.2
	Not applicable	34	12.0
Knowledge assessment	Good	168	59.4
	Poor	115	40.6

Note. Knowledge score was categorized according to the median value

Table 2 Fit indices for comparing different IRT models

Models	M2 (df)	*P*-value	RMSEA	TLI	CFI	AIC	BIC	proportion of variance
Model 1	482.1(322)	< 0.001	0.042	0.892	0.908	5022.5	5328.7	0.51
Model 2	339.4(228)	< 0.001	0.041	0.916	0.931	3905.35	4167.8	0.59
Model 3	246.2(187)	0.002	0.031	0.082	0.960	3605.07	3845.6	0.63
Model 4	84.0(88)	0.688	0.001	1	1	2650.04	2825.7	0.68

Item response theory (IRT) results

We conducted IRT analysis based on different-item models until we reached the best one based on the M2 statistics and other fit indices. Model 4 was the best model where M2 statistics were comparatively low and insignificant ($p > 0.05$), indicating no significant difference between the model and the data. Lower RMSEA, AIC, and BIC values and higher values of TLI and CFI supported the same conclusion. Moreover, the full information factor analysis indicated that the total proportion of variance extracted by one factor is 63.7% (Table 2). Therefore, we retained 16 items and eliminated 12 items based on the model 4 results.

Factor solution: most items had higher loading values and had a substantive relationship with the latent component, confirming uni-dimensionality assumption). Loading values ranged between 0.54 and 0.98. **IRT Parameters**: Most items also had steeper slopes (larger values of a) and differentiate respondents more precisely, showing a strong relationship with the factor. The Slope parameter (a) ranged from 1.33 to 67.76. The location parameter (b) indicated that items covered a wide range of latent abilities. **Item fit indices**: all included items had a good fit as P-values corresponding to S- X^2 were greater than 0.05. Moreover, infit and outfit statistics for most items were between 0.5 and 1.5 showing that items had a better fit, (Fig. 1) (Table 3).

Infit and outfit statistics indicated that a low proportion of respondents (0.07%) had infit and outfit values outside the range of ± 1.96, indicating that very few respondents had misfitting responses (Fig. 2). As the proportion of non-fitting respondents is less than 5%, we can conclude that respondents' responses align with the fit model, supporting the validity of the model.

We compared the final 3PL model with other restricted versions (2PL and 1PL) using the same items. The 3PL model that accounts for guessing parameters didn't improve the model fit comparing 1PL and 2P. Fit indices and $\chi 2$ test indicated that 1PL and 2PL fit the data were slightly better but with non-significant differences between them.

Predictors of knowledge towards NDS

The mean total knowledge score was 14.25 ± 2.20. The histogram showed that the knowledge score wasn't

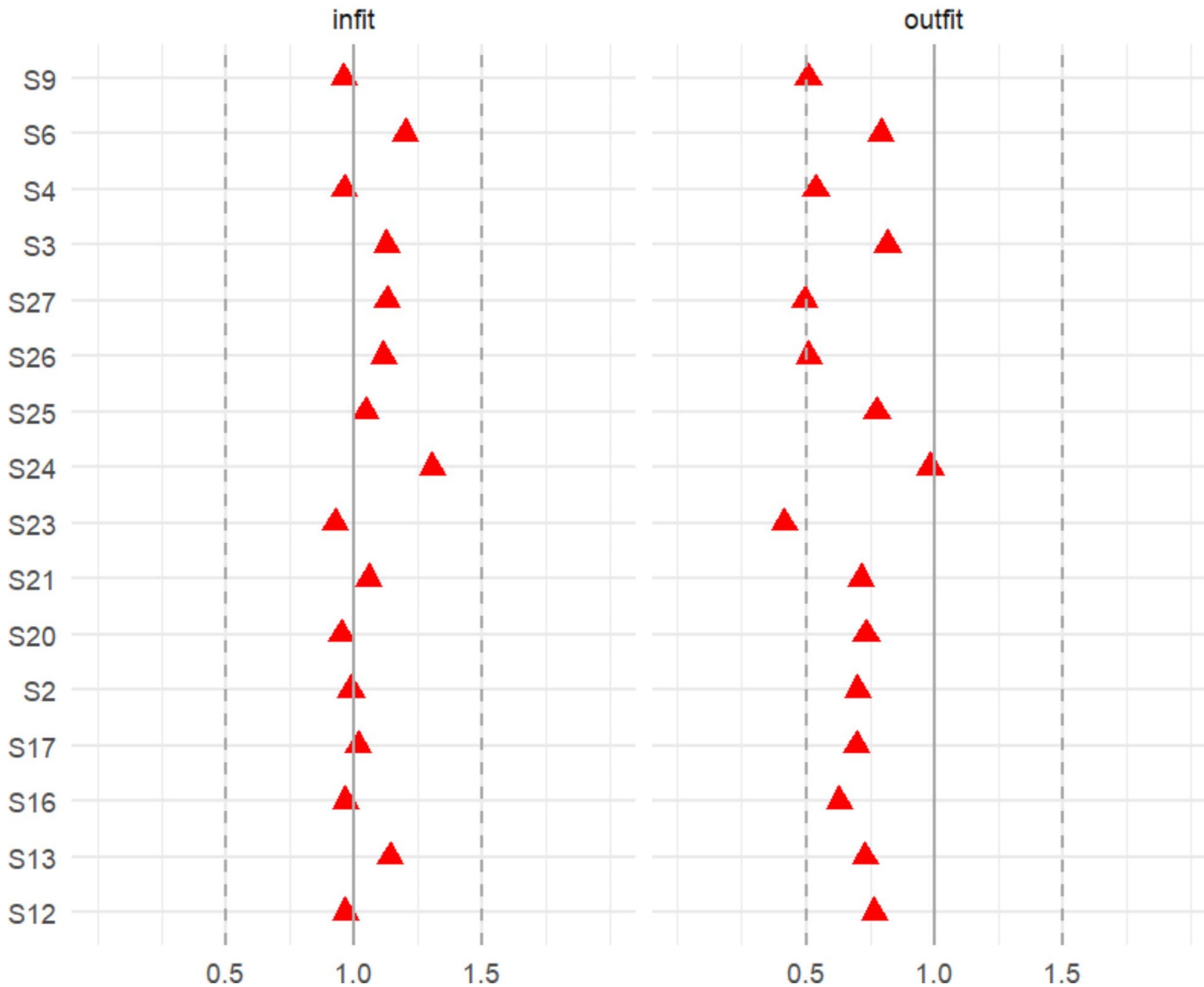

Fig. 1 Item infit and outfit statistics

Table 3 Results of IRT analysis

	Factor solution		**IRT Parameters**				**Item fit indices**			
Items	F1	h_2	a	b	c	u	S_ X^2	df	RMSEA	*P*-value
S2	0.67	0.45	1.53	-1.95	0.00	1	10.76	5	0.06	0.06
S3	0.78	0.62	2.14	-0.62	0.62	1	2.06	5	0.00	0.84
S4	0.79	0.63	2.22	-1.69	0.00	1	9.34	5	0.06	0.09
S6	085	0.72	2.74	-0.77	0.71	1	1.43	5	0.00	0.92
S9	0.82	0.67	2.45	-1.75	0.00	1	4.15	4	0.01	0.39
S12	0.62	0.38	1.33	-1.02	0.00	1	4.56	4	0.02	0.34
S13	0.69	0.48	1.62	-2.41	0.20	1	2.06	3	0.00	0.56
S16	0.67	0.45	1.54	-1.71	0.00	1	3.97	5	0.00	0.55
S17	0.97	0.94	6.92	-0.17	0.63	1	6.01	5	0.03	0.31
S20	0.61	0.38	1.32	-1.31	0.00	1	1.50	4	0.00	0.83
S21	0.63	0.9	1.37	-3.01	0.00	1	1.90	2	0.00	0.39
S23	0.91	0.83	3.74	-1.56	0.00	1	8.51	2	0.11	0.09
S24	0.97	0.95	7.16	0.04	0.87	1	2.95	5	0.00	0.71
S25	0.99	0.98	12.35	0.07	0.57	1	7.26	5	0.04	0.21
S26	0.99	0.99	49.03	-0.52	0.89	1	5.25	2	0.07	0.07
S27	1	1	67.76	-0.49	0.85	1	4.35	4	0.02	0.36

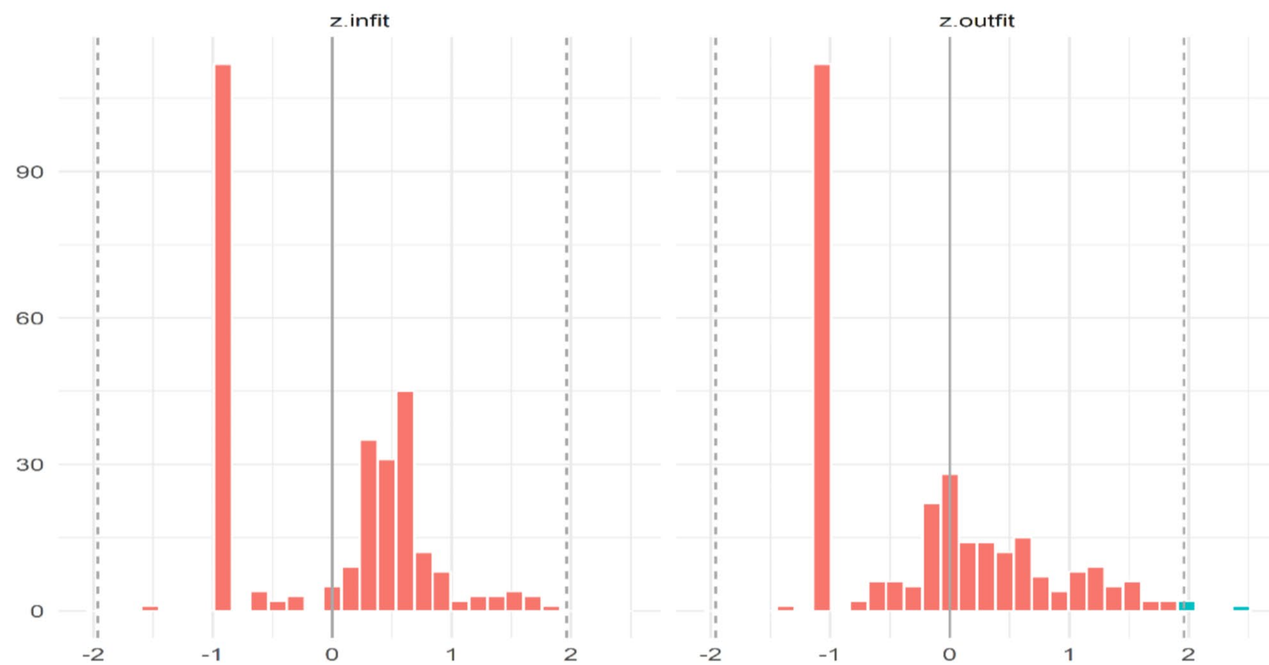

Fig. 2 Person item infit and outfit statistics

perfectly normally distributed, (Fig. 3**).** Shapiro-Wilk test and Kolmogorov-Smirnov test also demonstrated that the score wasn't normally distributed ($P < 0.001$). Knowledge score was categorized according to the median value where participants having a score above 15 were considered to have good knowledge and those below the median value were considered to have poor knowledge.

Table 4 shows the distribution of respondents according to their level of knowledge of danger signs. The majority of female, married, and urban residents had good knowledge of the danger signs. Most respondents with secondary education and outside the medical field were also aware of danger signs. Based on the chi-square values (χ^2), there was only a significant association between educational level and knowledge level.

Table 5 provides crude odds ratios (COR) and adjusted odds ratios (AOR) of knowledge level toward NDS. Respondents with university degree or post-graduate education had COR of 2.37, indicating that they are 137% more likely to have greater knowledge of danger signs than respondents with less than university education. After adjusting for other factors, the multiple logistic regression showed that female sex (AOR = 5.54, 95%CI:1.25-31.0, $P = 0.032$), age (AOR = 1.04. 95%CI:1.01–1.08, $P = 0.025$), and working outside the medical field (AOR = 3.26, 95%CI: 1.14–9.73, $P = 0.034$) were significant predictors for knowledge. However, education level, place of residence, and income didn't have a significant effect on the knowledge level.

Reliability analysis The item-to-mean score correlations were all positive and significant at the 0.001 level, showing a strong relationship between each item's score and the total knowledge score, Table 6. With Cronbach's alpha of 0.77, the knowledge score showed acceptable internal consistency. The Spearman-Brown coefficient standed at 0.864, indicating that the measure's reliability would remain constant even if the length of the questionnaire were changed. Additionally, the Guttman Split-Half coefficient of 0.812 reflected a strong internal consistency between the two halves of the questionnaire Supplementary Table 2. Regarding test-retest reliability, the number of respondents with agreed answers (pretest/posttest) for each item is provided in Supplementary Table 3. The results showed evidence of test-retest reliability where Pearson's r coefficient is significantly above 0.7 at a 0.05 significant level. ICC between scores equals 0.74 showing better test-retest reliability Supplementary file 3.

Discussion

Neonatal mortality is still a major health challenge that hinders improvement in healthcare systems. That is attributed to many factors; including the presence of a wide range of awareness levels regarding NDS globally [10, 18, 21, 25]. Poor knowledge is associated with delays in healthcare-seeking behaviors and higher rates of morbidity and mortality in such vulnerable groups. Additionally, caregivers may prefer traditional medicine to seek medical advice in conservative communities due to cultural norms. This may lead to further delays in applying the required medical intervention. Therefore,

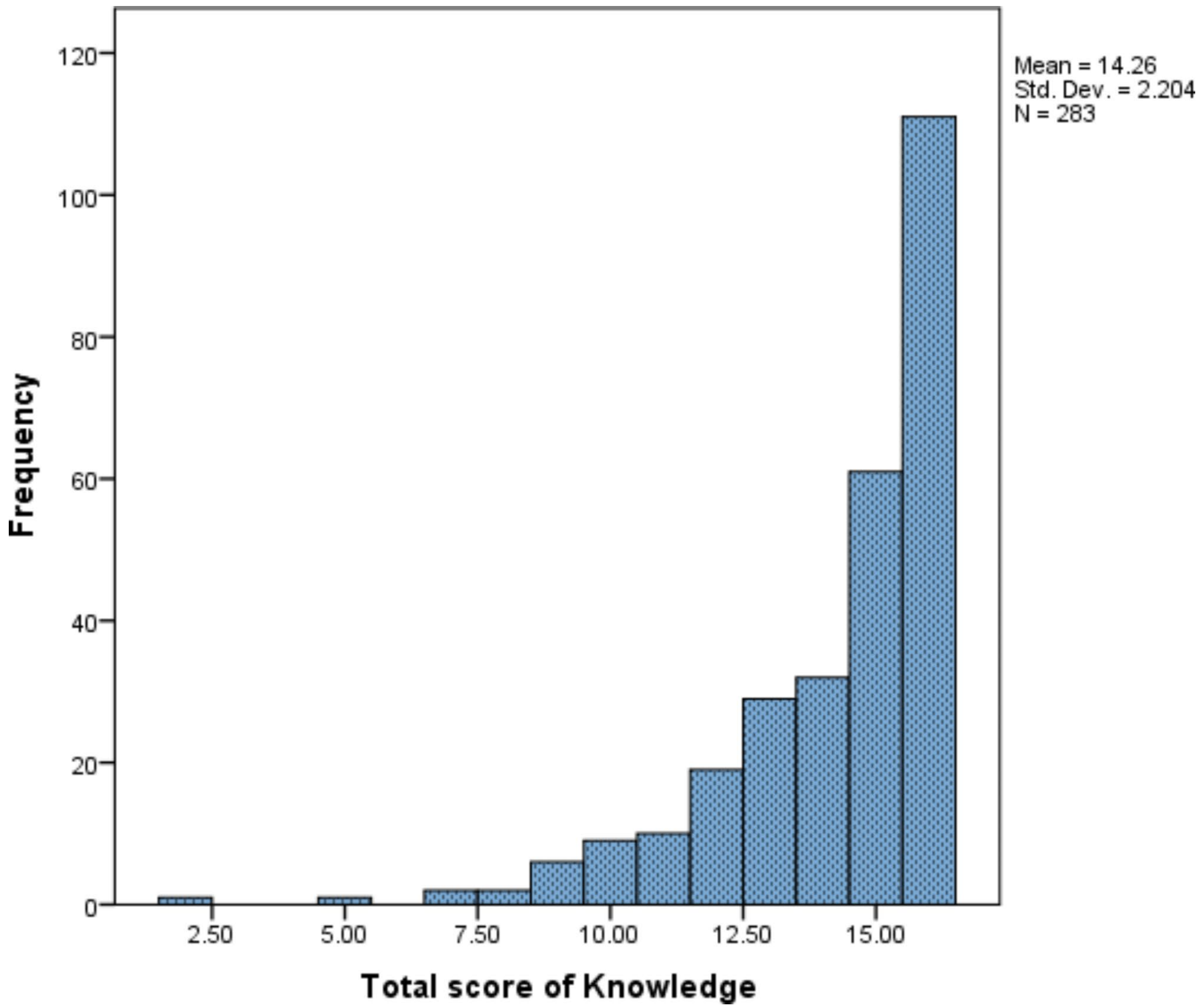

Fig. 3 Histogram of the total knowledge score towards the NDS

implementing preventional and interventional strategies to increase awareness of NDS has become important to reduce its impact on neonatal health. This necessitates the adoption of new approaches like one health to improve child health and development [41]. Despite former trials in the literature [10, 18, 21, 22, 24, 42, 43] that revealed poor knowledge, there is no validated questionnaire to measure the NDS, and most studies have relied on self-reported experiences and knowledge of mothers regarding these signs. This urged the need to develop and validate a questionnaire that covers the most significant danger symptoms using a consistent scoring system. So, this study aimed to develop and validate a reliable questionnaire in the Arabic language for the assessment of the Arabic-speaking population's knowledge regarding the NDS.

Reliability analysis

Reliability refers to the degree of consistency with which the same results are obtained when the same indicators or measurement tools are used to measure the same parameter repeatedly [44]. In this study, we considered several methods to assess the reliability of the questionnaire: Internal consistency was assessed using Cronbach's Alpha. The internal consistency has a high coefficient (0.766). Cronbach's Alpha value above 0.7 is considered acceptable, with higher values indicating stronger internal consistency. In this case, the value of 0.766 supports the correlation of the questionnaire's items. Additionally, test-retest reliability was assessed by re-administering the questionnaire to 100 participants after an interval of 15 days. The correlation between the respondents' initial responses and their responses after 15 days was statistically significant at the 0.05 level, proving the tool's stability over time and reducing the risk of systematic variation

Table 4 Distribution of respondents according to demographic characteristics and knowledge level

Variables	Levels	Knowledge level		χ^2 *P*-value
		Poor knowledge	Good knowledge	
Sex	Male	7 (70.0)	3 (30.0)	3.7
	Female	108 (39.6)	165 (60.4)	0.054
Age	Mean ± SD	31.2 ± 8.2	33.0 ± 8.8	t = 1.7 0.821
Marital status	Currently not married	9 (60.0)	6 (40.0)	6.6
	Currently married	106 (39.6)	162 (60.4)	0.087
Place of residence	Urban	92 (38.8)	145 (61.2)	2.0
	Rural	23 (50.0)	23 (50.0)	0.158
Education level	Illiterate or primary education	22 (40.7)	32 (59.3)	13.3
	preparatory education	25 (54.3)	21 (45.7)	0.021*
	Secondary education	30 (28.6)	75 (71.4)	
	University degree or postgraduate	38 (48.7)	40 (51.3)	
Income status	Not enough	54 (43.2)	71 (56.8)	0.6
	Enough	61 (38.6)	97 (61.4)	0.435
Occupation	Medical field	19 (55.9)	15 (44.1)	5.3
	Out of the medical field	14 (30.4)	32 (69.6)	0.072
	Not working or housewife	82(40.4)	121 (59.6)	
Hospitalized with one of your family's children?	Yes	56 (40.6)	82 (59.4)	4.3
	No	40 (36.0)	71 (64.0)	0.112
	Not applicable	19 (55.9)	15(44.1)	

Note. *Physicians: 4, Pharmacists: 19, Nurses: 5, Technicians: 1, Administration: 2, Coworkers: 3

Table 5 Multivariate logistic regression of knowledge level toward neonatal danger signs

Independent variables	Labels	COR	AOR
Sex	Male	-	-
	Female	3.35 (0.90-14.08, *P* = 0.07)	5.54 (1.25–13.10, *P* = 0.034)
Age	Mean (SD)	1.01 (0.99–1.05, *P* = 0.84)	1.04 (1.01–1.08, *P* = 0.025)
Place of residence	Urban	-	-
	Rural	0.63 (0.34–1.19, *P* = 0.16)	0.70 (0.35–1.41, *P* = 0.312)
Education level	Illiterate or primary education	-	-
	Preparatory education	1.38 (0.68–2.78, *P* = 0.37)	0.55 (0.24–1.25, *P* = 0.151)
	Secondary education	0.79 (0.38–1.65, *P* = 0.54)	1.56 (0.75–3.25, *P* = 0.233)
	University degree or postgraduate	2.37 (1.28–4.38, *P* = 0.006)	0.67 (0.29–1.52, *P* = 0.342)
Income status	Not enough	-	-
	Enough	1.21 (0.75–1.94, *P* = 0.43)	1.70 (0.99–2.95, *P* = 0.052)
Occupation	Medical field	-	-
	Out of the medical field	0.53 (0.25–1.11, *P* = 0.53)	3.26 (1.14–9.73, *P* = 0.034)
	Not working	1.55 (0.48–3.01, *P* = 0.21)	2.17 (0.86–5.59, *P* = 0.102)

Table 6 Item-total correlation of the NDS items with total knowledge score

Items	Item- total correlation (ITC)
S2	0.518**($P < 0.001$)
S3	0.568**($P < 0.001$)
S4	0.546**($P < 0.001$)
S6	0.381**($P < 0.001$)
S9	0.539**($P < 0.001$)
S12	0.521**($P < 0.001$)
S13	0.326**($P < 0.001$)
S16	0.585**($P < 0.001$)
S17	0.496**($P < 0.001$)
S20	0.535**($P < 0.001$)
S21	0.442**($P < 0.001$)
S23	0.592**($P < 0.001$)
S24	0.351**($P < 0.001$)
S25	0.588**($P < 0.001$)
S26	0.354**($P < 0.001$)
S27	0.358**($P < 0.001$)

between the initial and repeated administrations, further indicating the tool's reliability. This is vital for ensuring that the tool is not prone to random fluctuations or measurement errors when used repeatedly with the same respondents [43, 45].

Validity of the questionnaire

An IRT analysis was employed to identify the most suitable model for the final form. This process involved evaluating multiple indicators, including RMSEA, AIC, BIC, CFI, and TLI. IRT was preferred over factor analysis to validate the questionnaire since it offers more reliable and valid measures based on several pathways; (1) IRT models the relationship between latent traits and item responses, allowing for an understanding of how various factors influence responses across different populations. (2) IRT provides item-specific parameters that reflect item difficulty and discrimination, while factor analysis (FA) assumes equal contribution of items to the underlying construct. This capability lowers bias by facilitating the detection of differential item functioning (DIF) thereby enhancing the validity of the instrument. (3) IRT enables more fit score equating, allowing for comparisons across different tests or scales by placing scores on a common metric, which is not achievable through traditional FA methods [46]. Overall, IRT analyzed the questionnaire through different item models based on M2 statistics and retained 16 out of 28 items that showed the highest discrimination power and lowest value of RMSEA, AIC, and BIC. Using these items will effectively distinguish between individuals according to their knowledge level [47, 48].

Model fit indices are crucial in IRT as they assess how well the selected IRT model aligns with the observed data. Moreover, the proper fit model increases the generalizability of the questionnaire and guarantees application to different populations. The current study used factor solution and item fit indices to assess the quality and validity of individual responses. The high loading values from the factor solution (0.54 to 0.98) suggest that most items have a relationship with the latent component, demonstrating the robustness of the underlying factor structure.

The steep slopes of the IRT parameters (ranging from 1.33 to 67.76) demonstrate that the items are highly discriminative, meaning they effectively differentiate between individuals with different ability levels. The non-significant *P*-values for the S-X^2 statistic ($P > 0.05$) show the least discrepancy between the expected and observed response patterns, supporting well-fitting.

Factors associated with knowledge

Almost two-thirds (59.4%) of our sample demonstrated good knowledge surpassing the results reported in Saudi Arabia [20] and Sudan [15] (37.0% and 7.0% respectively). This difference could be attributed to the variation in data collection tools as the mentioned studies considered good knowledge to answer only three danger signs out of the nine signs mentioned by the WHO. The political conflict in Sudan may influence health programs and neonatal care [49, 50]. Furthermore, family planning initiatives in Egypt play a crutial role in emphasizing neonatal health [51–53]. The most known signs as a danger signs in our study were "convulsion and lethargy (97%) while Abu-Shaheen et al., [20] reported that the most known sign was "yellow soles" (48.2%). This underscores the need to reassess knowledge regarding NDS using a validated tool and apply interventional programs for mothers and caregivers.

We found that the proportion of good knowledge in females was double that in males (60.4% vs. 30%). This finding was in line with previous research [54, 55]. Roney et al., [55] compared the knowledge between mothers and their partners and reported that women identified several danger signs while the majority of men identified one sign only. This finding could be due to the social norms that put the responsibility of neonatal care on mothers and limit the father's role to providing appropriate financial care, and physical and emotional support for mothers [56, 57]. Another determinant for men's lower involvement could be their lower education level and consequently lower awareness regarding neonatal care [58]. The analysis of the participants' knowledge levels revealed a disparity between those employed within the medical field and those employed in other fields and not working (44.5%, 59.6%, and 69.6% respectively). This observation broadly supports the former findings [59, 60]. The small sample size of healthcare providers in the current study, predominantly including non-physicians, may contribute to this finding. The extant literature suggests knowledge deficiencies among healthcare professionals in areas beyond their specific expertise. Hasnain et al., [61] reported that only 16.4% of Pakistani healthcare providers had good knowledge of basic life support.

Strengths and limitations

To the best of our knowledge, this study represents the first attempt to develop a validated tool to evaluate the knowledge of community members regarding NDS in the Arabic language which is the third most spoken language worldwide. We used several statistical methods to analyze the validity and reliability of the questionnaire. The findings of this study serve as a crucial foundation for future research and can inform public health policymakers on how to improve the community knowledge related to NDS. However, the study has limitations as the data was collected only from Egypt and the majority of respondants were females. That's attributed to the nature of the Egyptian society which mothers are considered the main caregivers for their children.

Conclusions

This study successfully developed a validated and reliable tool to assess the knowledge of Arab populations towards the NDS. This questionnaire could assess the knowledge gap and empower mothers and caregivers to take an active role in their children's health. Consequently, it could help policymakers in assessing the knowledge level prioritize areas that need interventional education campaigns and track the improvement in awareness of NDS over time. This could help in the allocation of sources actively and lower the morbidity and mortality rates in such precious age group. Furthermore, the tool provides a foundation for future research, both in assessing knowledge gaps in Arab-speaking countries and through translation and validation for use in other languages.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12887-024-05286-w.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through a Large Research Project under grant number RGP2/440/45. We acknowledge all participants of the study and Dr Marina Labane for her contribution in data collection.

Author contributions

EAH. conceptualization of the research, methodology, Writing, reviewing, and editing the manuscript, SA. Formal analysis, Writing, reviewing, and editing the manuscript, MAH, Data collection, Writing the manuscript, MA & EE. Validation, Deta Curation, AS. Writing the Manuscript, RG. methodology, Writing, editing and reviewing the manuscript.

Funding

Not applicable.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the ethical committee of the Central Directorate for Research and Health Development REC (Com. No/Dec. No: 25-2023/12). After informed consent was obtained from all the study participants, the study followed the International Ethical Guidelines for Epidemiological Studies [62]. Clinical Trial Number: Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Clinical Reasearch Department, El-Raml Pediatric Hospital, Ministry of Health and Population, Alexandria, Egypt
[2]Biostatistics and Demography Department, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt
[3]Student in Biomedical Informatics and Medical Statistics Department, Medical Research Institute, Alexandria University, Alexandria, Egypt
[4]Pediatric Department, Faculty of Medicine, Alexandria University, Alexandria, Egypt
[5]Department of Child Health, College of Medicine, King Khalid University, P.O. Box 641, Abha, Saudi Arabia
[6]Family and Community Medicine Department, College of Medicine, King Khalid University, Abha, Saudi Arabia

Received: 15 August 2024 / Accepted: 26 November 2024
Published online: 21 April 2025

References

1. Rahim F, Jan A, Mohummad J, Iqbal H. Pattern and outcome of admissions to neonatal unit of Khyber Teaching Hospital, Peshawar. Pakistan J Med Sci. 2007;23(2):249.
2. Hug L, Alexander M, You D, Alkema L. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. Lancet Global Health. 2019;7(6):e710–20.
3. World Health Organization. Newborn Mortality 2021 [https://www.who.int/news-room/fact-sheets/detail/levels-and-trends-in-child-mortality-report-2021
4. Andegiorgish AK, Andemariam M, Temesghen S, Ogbai L, Ogbe Z, Zeng L. Neonatal mortality and associated factors in the specialized neonatal care unit Asmara, Eritrea. BMC Public Health. 2020;20:1–9.
5. Mahtab S, Madhi SA, Baillie VL, Els T, Thwala BN, Onyango D, et al. Causes of death identified in neonates enrolled through Child Health and Mortality Prevention Surveillance (CHAMPS), December 2016–December 2021. PLOS Global Public Health. 2023;3(3):e0001612.
6. Abd ElHafeez S, Hassan MA, Hammouda EA, Omran A, Esmail OF, Mahboob AS, et al. Geospatial distribution of under-five mortality in Alexandria, Egypt: a cross-sectional survey. Egypt Pediatr Association Gaz. 2023;71(1):75.
7. Tolossa T, Wakuma B, Mengist B, Fetensa G, Mulisa D, Ayala D, et al. Survival status and predictors of neonatal mortality among neonates admitted to

neonatal intensive care unit (NICU) of Wollega University referral hospital (WURH) and Nekemte Specialized hospital, Western Ethiopia: a prospective cohort study. PLoS ONE. 2022;17(7):e0268744.
8. Kruk ME, Pate M, Mullan Z. Introducing the Lancet Global Health Commission on high-quality health systems in the SDG era. Lancet Global Health. 2017;5(5):e480–1.
9. Affairs DoEaS. Sustainable Development Geneva: United Nations. 2024 [https://sdgs.un.org/2030agenda
10. Abdulrida HN, Hassan RJ, Sabri MM. Knowledge and health-seeking practices of mothers attending primary health-care centers in Baghdad Al-Karkh sector about danger signs in newborns. Mustansiriya Med J. 2018;17(1):29.
11. United Nations International Children's Emergency Fund. GOAL 3: GOOD HEALTH AND WELL-BEING, Geneva2024 [https://data.unicef.org/sdgs/goal-3-good-health-wellbeing/#cme_mrm0
12. United Nations Children's Fund. Neonatal mortality Geneva2023 [https://data.unicef.org/topic/child-survival/neonatal-mortality/#
13. Statistics CAoPMa. Egyptian Household Health Survey. Cairo2022.
14. Ministry of Planning and Economic Development. Egypt's 2021 voluntary National Review. Cairo; 2021.
15. Ahmed NSOM, Al-Gamar EA. Mother's knowledge and attitude regarding recognition of neonatal signs of danger. Sudan J Paediatrics. 2022;22(1):27.
16. World Health Organization. WHO recommendations on maternal and newborn care for a positive post-natal experience. Geneva; 2022.
17. Degefa N, Diriba K, Girma T, Kebede A, Senbeto A, Eshetu E, et al. Knowledge about neonatal Danger signs and Associated Factors among mothers attending immunization clinic at Arba Minch General Hospital, Southern Ethiopia: a cross-sectional study. Biomed Res Int. 2019;2019:9180314.
18. Ekwochi U, Ndu IK, Osuorah CD, Amadi OF, Okeke IB, Obuoha E, et al. Knowledge of danger signs in newborns and health seeking practices of mothers and care givers in Enugu state, South-East Nigeria. Ital J Pediatr. 2015;41:18.
19. Jemberia MM, Berhe ET, Mirkena HB, Gishen DM, Tegegne AE, Reta MA. Low level of knowledge about neonatal danger signs and its associated factors among postnatal mothers attending at Woldia general hospital. Ethiopia Maternal Health Neonatology Perinatol. 2018;4:1–8.
20. Abu-Shaheen A, AlFayyad I, Riaz M, Nofal A, AlMatary A, Khan A, et al. Mothers' and caregivers' knowledge and experience of neonatal Danger signs: a cross-sectional survey in Saudi Arabia. Biomed Res Int. 2019;2019:1750240.
21. Awasthi S, Verma T, Agarwal M. Danger signs of neonatal illnesses: perceptions of caregivers and health workers in northern India. Bull World Health Organ. 2006;84(10):819–26.
22. Dongre AR, Deshmukh PR, Garg BS. Awareness and health care seeking for newborn danger signs among mothers in peri-urban Wardha. Indian J Pediatr. 2009;76:691–3.
23. Hill Z, Kendall C, Arthur P, Kirkwood B, Adjei E. Recognizing childhood illnesses and their traditional explanations: exploring options for care-seeking interventions in the context of the IMCI strategy in rural Ghana. Tropical Med Int Health. 2003;8(7):668–76.
24. Sandberg J, Odberg Pettersson K, Asp G, Kabakyenga J, Agardh A. Inadequate knowledge of neonatal danger signs among recently delivered women in southwestern rural Uganda: a community survey. PLoS ONE. 2014;9(5):e97253.
25. Kebede AA, Cherkos EA, Taye EB. Mother's knowledge of neonatal danger signs and health-seeking practices and associated factors in Debretabor, Northwest Ethiopia: a community-based cross-sectional study. Res Rep Neonatology. 2020:47–58.
26. Kibaru EG, Otara AM. Knowledge of neonatal danger signs among mothers attending well baby clinic in Nakuru Central District, Kenya: cross sectional descriptive study. BMC Res Notes. 2016;9:1–8.
27. Mose A, Abebe H, Shitu S, Shimels H. Determinants of maternal knowledge of neonatal danger signs among postpartum mothers in Southern Ethiopia: institutional-based cross-sectional study. BMJ open. 2021;11(9):e048888.
28. Al-Gamar NSOMA. Mother's knowledge and attitude regarding recognition of neonatal signs of danger. Sudan J Paediatrics. 2022;22(1):27–35.
29. Zhou J, Hua W, Zheng Q, Cai Q, Zhang X, Jiang L. Knowledge about neonatal danger signs and associated factors among mothers of children aged 0–12 months in a rural county, Southwest of China: a cross-sectional study. BMC Pregnancy Childbirth. 2022;22(1):346.
30. Rusticus S. Content validity. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research. Dordrecht: Springer Netherlands; 2014. pp. 1261–2.
31. White M. Sample size in quantitative instrument validation studies: a systematic review of articles published in Scopus, 2021. Heliyon. 2022;8(12):e12223.
32. Gunawan J, Marzilli C, Aungsuroch Y. Establishing appropriate sample size for developing and validating a questionnaire in nursing research. Belitung Nurs J. 2021;7(5):356–60.
33. Nguyen TH, Han H-R, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. Patient-Patient-Centered Outcomes Res. 2014;7:23–35.
34. Vilagut G. Test-retest reliability. Encyclopedia of quality of life and well-being research. 2014:6622-5.
35. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol. 2003;56(8):730–5.
36. Park MS, Kang KJ, Jang SJ, Lee JY, Chang SJ. Evaluating test-retest reliability in patient-reported outcome measures for older people: a systematic review. Int J Nurs Stud. 2018;79:58–69.
37. Andrade C. Sample size and its importance in research. Indian J Psychol Med. 2020;42(1):102–3.
38. Brysbaert M. How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. J Cognition. 2019;2(1).
39. Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochemia Med. 2021;31(1):27–53.
40. Moran MC, Huo MH, Garvin KL, Pellicci PM, Salvati EA. Total hip arthroplasty in Sickle Cell Hemoglobinopathy. Clin Orthop Relat Research®. 1993;294.
41. Ghazy RM, Saidouni A, Taha SHN. Promoting child health through a comprehensive one health perspective: a narrative review. Egypt Pediatr Association Gaz. 2024;72(1):3.
42. Callaghan-Koru JA, Seifu A, Tholandi M, de Graft-Johnson J, Daniel E, Rawlins B, et al. Newborn care practices at home and in health facilities in 4 regions of Ethiopia. BMC Pediatr. 2013;13(1):1–11.
43. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53.
44. Mo Z, Li X, Zhai Y, Men Y, Tang Y, Qiao J, et al. Reliability and validity of a questionnaire measuring knowledge, attitude and practice regarding oil, salt and sugar among canteen staff. Sci Rep. 2023;13(1):20442.
45. Anvari A, Halpern EF, Samir AE. Essentials of statistical methods for assessing reliability and agreement in quantitative imaging. Acad Radiol. 2018;25(3):391–6.
46. Bortolotti SLV, Tezza R, de Andrade DF, Bornia AC, de Sousa Júnior AF. Relevance and advantages of using the item response theory. Qual Quant. 2013;47(4):2341–60.
47. Columbia University Mailman School of Public Health. Item Response Theory 2024 [https://www.publichealth.columbia.edu/research/population-health-methods/item-response-theory
48. Mars C. Item response theory: understanding statistics Measurement. New York: Oxford University Press, Inc; 2010.
49. Abdelmola A. Antenatal Care Services in Sudan before and during the 2023 war: a review article. Cureus. 2023;15(12):e51005.
50. Wharton G, Ali OE, Khalil S, Yagoub H, Mossialos E. Rebuilding Sudan's health system: opportunities and challenges. Lancet. 2020;395(10219):171–3.
51. Abd Allah RAE, Mohamed HA-E, Kamel HH. Impact of Health Educational Program for pregnant women on their selected Postpartum and Newborn Care. Minia Sci Nurs J. 2023;13(1):18–26.
52. Abd Elmoniem Syan S, Llaguno BB, Sayed Abdalla S. Effect of Instructional guidelines on Mothers' knowledge and practice regarding their neonatal care. Egypt J Health Care. 2021;12(3):615–30.
53. ElGindy HA, Mohamed MF, Kotb MA, Ezz ElDin Z, Hamdy HS, Hussein MF, et al. Neonatal auditory screening is a necessity in the neonatal intensive care unit: single Center Study. Pediatr Sci J. 2022;2(2):120–6.
54. Nasreen HE, Leppard M, Al Mamun M, Billah M, Mistry SK, Rahman M, et al. Men's knowledge and awareness of maternal, neonatal and child health care in rural Bangladesh: a comparative cross sectional study. Reproductive Health. 2012;9:1–9.
55. Roney E, Morgan C, Gatungu D, Mwaura P, Mwambeo H, Natecho A, et al. Men's and women's knowledge of danger signs relevant to postnatal and neonatal care-seeking: a cross sectional study from Bungoma County, Kenya. PLoS ONE. 2021;16(5):e0251543.
56. Muloongo H, Sitali D, Zulu JM, Hazemba AN, Mweemba O. Men's perspectives on male participation in antenatal care with their pregnant wives: a case of a military hospital in Lusaka, Zambia. BMC Health Serv Res. 2019;19(1):463.
57. Shibeshi K, Lemu Y, Gebretsadik L, Gebretsadik A, Morankar S. Gender-based roles, psychosocial variation, and power relations during delivery

and postnatal care: a qualitative case study in rural Ethiopia. Front Global Women's Health. 2023;4.

58. Abie A, Belay G, Asmare E, Ayalew N, Feyisa W, Kassie N. Male partner involvement in postnatal care service utilization and associated factors: a community-based cross-sectional study in Motta district, Northwest Ethiopia. PLoS ONE. 2023;18(1):e0276477.
59. Abrha MW, Asresu TT, Araya AA, Weldearegay HG. Healthcare professionals' knowledge of neonatal resuscitation in Ethiopia: analysis from 2016 National Emergency Obstetric and Newborn Care Survey. Int J Pediatr. 2019;2019:8571351.
60. Yosef T, Getachew D, Weldekidan F. Health professionals' knowledge and practice of essential newborn care at public health facilities in Bench-Sheko Zone, Southwest Ethiopia. Heliyon. 2021;7(11):e08369.
61. Hasnain S, Hussan J, Khan L, Muhammad S, Kamal K, Hayat U, et al. Factors affecting knowledge and attitude of healthcare workers towards basic life support in Khyber Teaching Hospital, Peshawar, Pakistan: a cross-sectional analysis. BMJ open. 2023;13(9):e073369.
62. Rose S. International ethical guidelines for epidemiological studies: by the Council for International Organizations of Medical Sciences (CIOMS). Oxford University Press; 2009.

Publisher's note