Electronic Journal of General Medicine
2023, 20(6), em534
e-ISSN: 2516-3507
https://www.ejgm.co.uk/

MODESTUM

Original Article

OPEN ACCESS

Validation of the knowledge on the Noor evidence-based medicine questionnaire for healthcare professionals: Rasch analysis

Mohd Noor Norhayati [1*], Mat Nawi Zanaridah [2], Ruhana Che Yusof [3]

[1] Department of Family Medicine, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, Kubang Kerian, Kelantan, MALAYSIA
[2] Klinik Kesihatan Putatan, Jalan Pasir Putih, Putatan, Sabah, MALAYSIA
[3] Department of Medicine, Faculty of Medicine, Universiti Malaya, Wilayah Persekutuan Kuala Lumpur, MALAYSIA
*Corresponding Author: hayatikk@usm.my

Citation: Norhayati MN, Zanaridah MN, Che Yusof R. Validation of the knowledge on the Noor evidence-based medicine questionnaire for healthcare professionals: Rasch analysis. Electron J Gen Med. 2023;20(6):em534. https://doi.org/10.29333/ejgm/13535

ARTICLE INFO

Received: 20 Dec. 2022

Accepted: 25 Jul. 2023

ABSTRACT

Aim: This study aimed to examine the construct verification of the knowledge on the Noor evidence-based medicine (EBM) questionnaire using the Rasch measurement model.

Methods: A cross-sectional study was conducted among ninety healthcare professionals working in a government hospital. The Rasch model was used to investigate the distribution of statistics, unidimensionality, polarity, misfit, and Wright map.

Results: The Rasch analysis showed that the 15 items had high reliability of items at 0.96, while reliability for persons were 0.81. Item K15 is a misfit (PtMea Corr=0.22, outfit MnSq=1.95, outfit z-std=5.00, infit MnSq=1.63, infit z-std=3.60); the item does not contribute to the construction of scale but not degrading. Three items (K1, K14, and K9) can easily be answered even by low ability respondents.

Conclusions: The Noor EBM questionnaire knowledge is robust with excellent psychometric properties that can be used for both research and clinical purposes.

Keywords: evidence-based medicine, knowledge, healthcare, Rasch model

INTRODUCTION

Evidence-based medicine (EBM) is defined as using current best evidence conscientiously, specifically, and judiciously in making decisions about the treatment of individuals [1]. The process of systematic observations, assessment and use of current research findings is often understood as the basis for clinical decisions [2]. It combines clinical experience and patient expectations with the best scientific knowledge currently available. It aims to expand the use of clinical research of high quality in clinical decision-making [3]. It provides a basis for applying the appropriate scientific findings to the condition of the patient based on the expectations of the patient using the professional judgement of the clinician to tailor the care for the patient [4].

Since the introduction of EBM, several studies have been conducted worldwide to assess the knowledge, attitude, and practice of medical staff and medical students. Based on a systematic review of barriers, facilities, knowledge, and attitude toward EBM, some healthcare providers have relatively low awareness of EBM's technical terms (44.2%). The awareness, knowledge and evidence-based performance level was less than 50.0%. Textbooks have been considered the most significant source of obtaining information for clinical practice [5]. The barriers faced for successful implementation of EBM were also studied via questionnaires. However, there is a lack of validated questionnaires assessing the applicability of EBM in general medicine with a composite outcome score among healthcare professionals. A well-developed EBM scale can provide a relevant evaluation of the knowledge of EBM among healthcare professionals.

Several questionnaires are available worldwide for assessing EBM among healthcare professionals, such as Fresno, Berlin, Baum, McColl, and EBM questionnaire [6-14]. The most frequently used questionnaire was the one developed by [15], which assessed self-reported awareness, attitudes, and barriers to EBM [13, 16]. Baum's questionnaire also measures the attitude of healthcare professionals towards and self-reported ability in EBM [12]. The Berlin questionnaire, meanwhile, tests information and skills in EBM [10]. Fresno assesses EBM skills and competencies.

Nonetheless, further research is required for its use in other settings with adaptation to the target population [6]. The assessing competency in EBM (ACE) tool assesses user performance but has been limited across different patient scenarios [7]. In addition, various EBM tests resulted in a separate evaluation of general EBM domains in a sample due to the heterogeneity of the items on EBM tests [9].

These questionnaires were used widely and were adapted in many countries. For example, the Fresno test was used among emergency medicine residents [17] and medical technology students [18] and was adapted into European Portuguese [19]. Berlin questionnaire was adapted and validated in dentistry [20] and used to assess the critical thinking skills in oncology training [21]. Meanwhile, the McColl questionnaire evaluated the perceptions and attitudes towards evidence-based urology among urologist trainees in India [22], and the ACE tool was translated into many versions, such as Persian [23], Spanish [24], and Brazilian [25].

EBM learning linked with clinical activities such as morning reports, inpatient hospital rounds, and outpatient clinics may achieve the objectives of EBM training, primarily on information acquisition or behavioural changes, such as enhancing doctors' abilities for critical analysis and making decisions [26]. In addition, the assessment of EBM was evaluated mostly by quantitative measurement demonstrated improvement of scientific knowledge among healthcare professional based on current information available as reported by previous review [27], which stated that EBM training could enhance current skills and knowledge but not much solid proof that long-term knowledge, attitudes, and clinical practice have changed.

Two theories are generally used in the assessment of a scale, namely, classical test theory and item response theory. Classical test theory forms the measuring basis for most assessment tools [28] and uses the Likert rating scale to collect data for latent trait assessment. Respondents were asked to score their answers to the statements issued, which suggested the rank of the order of the different categories. The mark used to rank the categories does not imply the degree to which the other category is greater than others [29]. It is because the data positioned in the ranking order are not linear but are of a continuum nature and have no equivalent interval that is essential for statistical analysis [30].

However, Rasch analysis focuses on the pattern of the responses of an item that predicts the interaction between an item and a person on a mutual latent trait [31]. Rasch analysis predicts how a person of different levels of ability responds to an item with a particular level of difficulty [29]. Rasch analysis was considered a better choice because it considers the attributes of both persons and items [32]. The purpose of this study is to examine the construct verification of the knowledge on the Noor EBM questionnaire among healthcare professionals in Selangor, Malaysia, using the Rasch rating scale model.

METHODS

Population and Sample

A cross-sectional study was conducted. This study follows the STROBE guideline for reporting cross-sectional studies. Healthcare professionals working in a government hospital in Selangor, Malaysia, were recruited. House officers were excluded. The sample size was based on ½ logit at 95% confidence with best to poor targeting sample size between 64 and 144 [33], and 90 healthcare professionals were recruited for this study. Only the subscale for knowledge was analysed for the sources of misfit to the model in this study.

Convenient sampling was applied. Permission was obtained from the Medical Research Ethics Committee, Ministry of Health, to recruit participants in a hospital. The researcher explained the study and written informed consent were obtained. Participants were then given self-administered questionnaires. After completing the questionnaires, they were checked for completeness, and the participants were thanked for their cooperation.

Research Tool

A literature search was undertaken in Medline in order to ensure good content of the scale; keywords such as 'evidence-based medicine', 'evidence-based practice', 'healthcare' 'doctors', 'questionnaire' and 'scale was used. The questionnaire was developed in English, as it is used to train doctors in medical schools and is also taught as a second language in all Malaysian public schools. The questionnaires were designed using a modified Delphi technique to allow members of the research team to engage in the final round to explain the issues and present arguments explaining their views [34]. Four experts were involved in the development phase, which included a public health physician, a family medicine specialist, an expert in EBM, and a biostatistician. Each item was addressed in depth to make sure all respondents would appreciate it in the same manner. It avoided vague content, difficult or ambiguous terms, multiple thoughts, or notions in one item or a double-barreled item.

For each item, two sets of questions were prepared to ask for similar meanings but in different ways. The 30-item scale underwent cognitive debriefing with 10 respondents, including experts in the field and healthcare professionals. Each item was assessed for clarity, appropriateness, and relevance. The wording of the items was revised accordingly. For example, the initial wording for item K7 was 'randomized-control trial is more superior that prospective cohort study' was rephrased to 'case-control studies are superior to meta-analysis in evidence-based medicine'; item K11 'whenever, here is a doubt in any aspect of clinical management, evidence-based medicine can be used' was rephrased to 'evidence-based medicine can be used to answer doubts in any aspect of clinical management'; and item K12 'Improving access to summaries of evidence would be an appropriate method of encouraging evidence-based practice' was rephrased to 'improving access to summaries of evidence is appropriate to encourage evidence-based practice'. The final knowledge domain consists of 15 items on EBM. A five-point Likert scale (strongly agree=5, agree=4, neutral=3, disagree=2, and strongly disagree=1) was used. The scoring is reversed for negative-worded items.

Rasch Analyses

The model of the Rasch rating scale was used to evaluate the respondents' information on EBM using a Likert scale of five points. In response to knowledge of EBM, the Rasch rating scale model estimates the level of individual potential and performance of items. The Rasch measurement model transforms the raw data to log odd ratios (logits). The person's likely score is defined by the interaction between person ability, items difficulty and person measure location. The knowledge on EBM data was analyzed using Winsteps version 3.72.3 [35], which is a computer program for Rasch analysis.

Winsteps records global fit statistics and estimated chi-square statistics of the global log-likelihood. The level of significance is p less than or equal to 0.05. The value of the Chi-square is approximate. This is based on the latest estimates published that can differ significantly from the estimates of the 'true' maximum probability for these results. If p less than 0.001, the data reveals, as is almost always expected, a misfit for the Rasch model. Global root-mean-square residual values lower than expected indicate a better fit (or overfit) for the Rasch model [36].

Reliability values for persons and items greater than 0.8 suggest acceptable reproducibility, between 0.6 and 0.8 suggest less acceptable and less than 0.6 suggest unacceptable reproducibility [29]. Rasch goodness of fit test based on item polarity and misfit, reported in mean squared (MnSq) and z-standardized (z-std), provides an indication of how well an item fits the model. The item polarity, and point measure correlation (PtMea Corr) indicates properly functioning items [29]. PtMea Corr value less than 0.30 indicates that the items do not fulfil the criteria set. The outfit MnSq is an unweighted index that includes the differences for all items irrespective of how far away the item difficulty is from the item measures [37].

Although MnSq outfit values are often used as an indicator of item misfit in the Rasch analysis [36], both MnSq infit and outfit together with z-std were considered as a valid tool for public use in this study to offer a deeper understanding of the Noor EBM questionnaire of knowledge. The MnSq values between 0.5 and 1.5 logits indicate the measurement scale is productive, less than 0.5 indicates less productive but not degrading, between 1.5 and 2.0 indicates unproductive but not degrading and more than 2.0 distorts or degrades the scale [38]. The z-std indicates the standardized sum of all differences between observed and expected values summed over all persons [37]. The z-std values are expected to be zero and any values less than -2, and more than two is considered a misfit [38].

The spread for the persons and items in Rasch are measured by the separation index. Separation index values can range from zero to infinity, and the higher the value indicates, the better the measurement and the power of a set of items to discriminate between individuals. Separation index value more than one is considered useful, and more than two as good for the scale [29, 39, 40].

Unidimensionality is a crucial element in determining the presence of secondary latent variables. A construct requires at least five items to weight a factor before the construct is viewed as another dimension [40]. The percentage of unexplained variance by the 1st contrast value of less than 3.0% indicates excellent, 3.0-5.0% indicates very good, 5.0-10.0% indicates good, 10.0-15.0% indicates fair , and more than 15.0% indicates poor [41]. Unidimensionality threshold minimum of 40.0% variance explained by measures [41] indicates a strong measurement dimension but is it better to exceed 60.0% [35]. Multicollinearity or local item dependency identified from the largest standardized residual correlation of more than 0.7 is considered redundant.

Wright's map represents locations for item measurement and the distribution of person measurements along the logit scale. The respondents at the top of the map had the highest scores, while the items at the top were the most difficult. The respondents at the bottom of the map had the lowest scores, while the items at the bottom were the easiest [29].

The DIF size measures the size of the item DIF compared to the overall 'baseline' item difficulty for the person-classification. The acceptable values range from -0.5 to 0.5. The DIF t-value provides a simple estimated t-test of the DIF item against the overall item difficulty. The acceptable values are between -2 and 2.0 [36].

Table 1. Characteristics of respondents (n=90)

Variable	n (%)
Age (year)	
25-30	40 (44.4)
31-35	35 (38.9)
36-40	12 (13.3)
41-45	3 (3.3)
Sex	
Male	32 (35.6)
Female	58 (64.4)
Race	
Malay	65 (72.2)
Others	30 (27.8)
Marital status	
Single	56 (62.2)
Married	34 (37.8)
Education	
First degree	67 (74.4)
Second degree	23 (25.6)

Note. Working experience's (year) median (IQR)=2.9 (2.03)

Table 2. Person & item summary statistics: Initial analysis & after removal of misfit responses

	Initial analysis (n=90)		After identifying misfit respondents (n=81)	
	Person (n=90)	Item (n=15)	Person (n=81)	Item (n=15)
Cronbach's alpha (α)	0.81		0.82	
Reliability index (μ)	0.81	0.96	0.84	0.96
Separation index	2.03	5.00	2.29	5.00
Mean	1.25	0.00	1.45	0.00
Max measure	4.36	1.55	5.08	1.74
Min measure	-0.59	-1.64	-0.17	-1.94
Spread	4.95	3.19	5.25	3.68
Standard deviation	1.02	0.84	1.17	0.97
Outfit				
Mean square	1.00	1.00	0.98	0.98
z-standard	-0.20	0.00	-0.10	-0.20
Infit				
Mean square	1.05	0.97	1.01	0.96
z-standard	-0.10	-0.20	0.00	-0.20

RESULTS

The study involved 90 respondents with a response rate of 100%. **Table 1** shows the characteristics of respondents. The mean (standard deviation) age was 31.8 (4.0) years old. Most of the respondents had a medical bachelor or equivalent as their highest degree of education qualification.

The summary statistics of 90 persons and 15 items measured regarding knowledge on EBM questionnaire was shown in **Table 2**.

There were 1,350 data points with a log-likelihood Chi-square (degree of freedom) of 2,643.72 (1,243) (p<0.001). The mean square root residual was 0.68. The Cronbach's alpha was 0.81. The item reliability (standard error, SE) was 0.96 (0.22), and the person reliability (SE) was 0.81 (0.11).

The item separation index was excellent at 5.00 (**Table 2**). However, the items are capable of separating the persons into three strata, indicated by the person separation index of 2.03. **Table 3** shows the assessment of unidimensionality using the principal component analysis.

Table 3. Standardized residual variance using principal component analysis

Standardized residual variance (Eigenvalue units)	Empirical(%)
Total raw variance in observations	100
Raw variance explained by measures	42.4
Raw variance explained by persons	15.8
Raw variance explained by items	26.6
Raw unexplained variance (total)	57.6
Unexplained variance in 1st contrast	10.9

Local dependence tests for the largest standardized residual correlation yield a very good outcome, where none of the items breaches the 0.70 limits indicating item independence in the instrument. A further investigation into the item misfit statistics was conducted.

A further investigation was conducted into the item misfit statistics. The statistical parameters for the 15 items recorded measurements between 1.55 and -1.64 *logit.* The outfit MnSq was 1.95 to 0.60 *logit,* the outfit z-std was 5.0 to -2.70 *logit,* the infit MnSq was 1.63 to 0.58 *logit,* the infit z-std was 3.6 to -2.70 *logit,* and the PtMea Corr was 0.66 to 0.22 *logit.*

Based on item misfit order, PtMea Corr had 14 items between 0.44 and 0.66, which indicated that the items measured the corresponding constructs. Only item K15 (*'application of evidence-based practice is cost-effective to healthcare system')* is misfitting (PtMea Corr= .22, outfit MnSq=1.95, outfit z-std=5.0, infit MnSq=1.63, infit z-std=3.6).

A further investigation into the statistics on the person misfit was conducted to ensure that the 90 people were in fit conditions. The parameters for person fit reported measures between 4.36 and -0.59 *logit.* The outfit MnSq was 3.83 to 0.18 *logit,* the outfit z-std was 4.30 to -3.30 *logit,* infit MnSq was 4.43 to 0.19 *logit,* the infit z-std was 5.0 to -3.40 *logit,* and the PtMea Corr was 0.91 to -0.09 *logit.*

Nine individuals were found to show most unfit response strings based on item K15 (*application of evidence-based practice is cost-effective to healthcare system*) following further investigation. By the pattern of the responses, the possible causes might be guessing or carelessness. The persons were temporarily deleted because they added noise to the measurement process [39]. The knowledge on EBM questionnaire was fitted twice to assess the impact of misbehaved persons; however, extremely trivial variations in the overall parameter estimates were identified (**Table 2**). Therefore, the persons were retained in the model.

Figure 1 displays on the logit scale the number of endorsements issued by the respondent and the item measures. The person mean (1.25 *logit*) is around the same position for both measures, suggesting that the items for this sample are well-focused. With the 3.19 *logit* spread, the Item difficulty measures from +1.55 to -1.64 *logits*. Meanwhile, with the spread of 4.95 *logit* the ability of the person estimates from +4.36 to -0.59 *logit*. The distribution of items has a much lower spread compared to persons.

Respondents with better knowledge were at the top of the map, while those with poorer knowledge were at the bottom. The most complicated item is upper scale K3 (1.55 *logit*). Though on the lower scale are item K1 (-1.64 *logit*) (*'difficulty in understanding statistical terms is the major setback in applying evidence-based medicine'*), K14 (-0.79 *logit*) *('evidence-based medicine improves clinicians' understanding on research methodology'*) and K9 (-0.79 *logit*) *('evidence-based medicine improves clinicians' understanding on research methodology'*), which can be easily answered by respondents. Two items, i.e., K10 (0.42 *logit*) (*'clinicians who practice evidence-based*

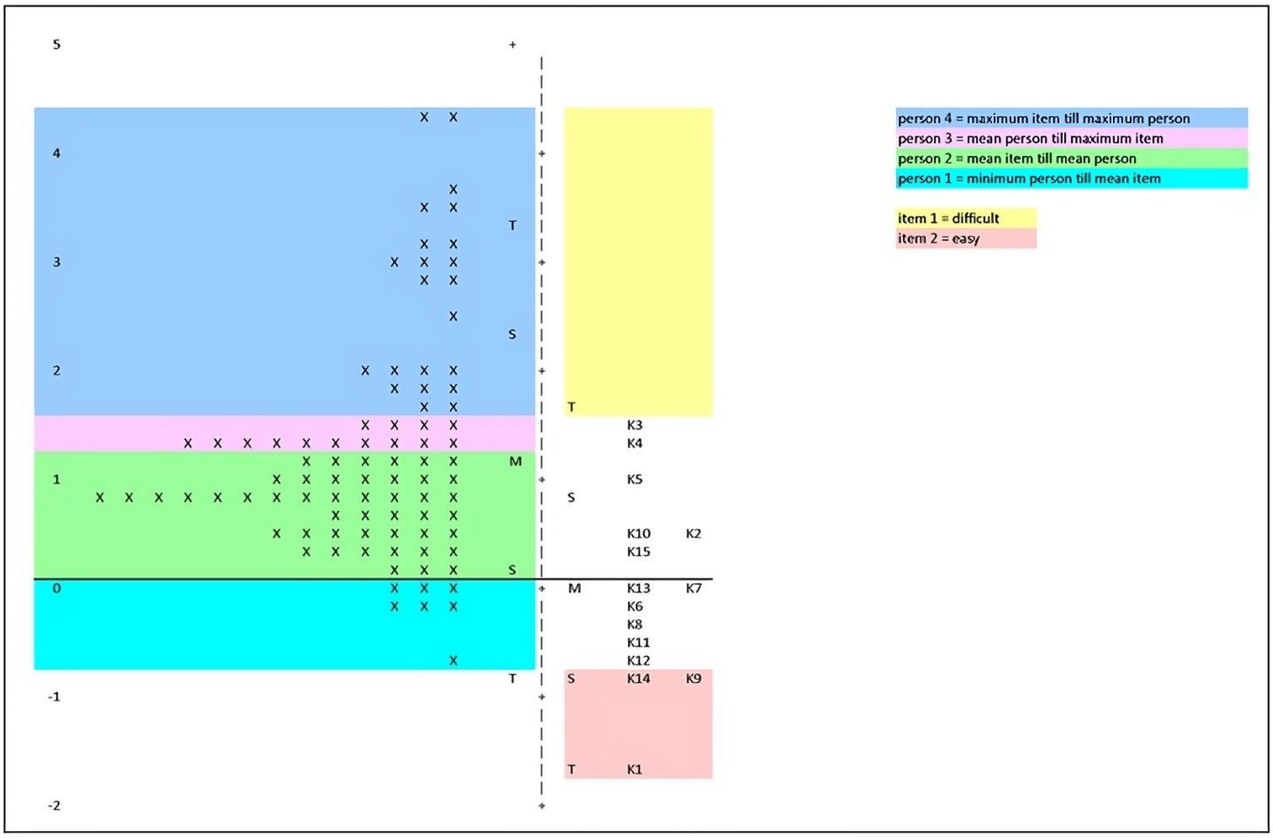

Figure 1. Wright map of knowledge on Noor EBM questionnaire (Source: Authors' own elaboration)

medicine become less critical in using data in systemic reviews') and K2 (0.44 *logit*) (*'evidence-based medicine focuses on the best current available research without considering clinical experience'*) were closely located.

Other two items, i.e., K13 (-0.08 *logit*) (*'Increasing number of systematic reviews that are applicable to general practice can be found in the Cochrane Library'*) and K7 (0.00 *logit*) (*'meta-analysis is superior to case-control studies in evidence-based medicine'*) were closely located in the Wright map. There are 22 respondents who are highly knowledgeable (above 1.55 *logit*).

The DIF indicates the potential measurement bias of an item. The DIF was analyzed for age and sex. The age was divided into four groups (25-30, 31-35, 36-40, and 41-45 years old). The DIF was found for item K8 *('four essential components structured in the PICO format will make a good clinical question'*) for age and item K7 (*'meta-analysis is superior to case-control studies in evidence-based medicine'*) for sex. For item K8, there is age response bias in the age group 31-35 years old (DIF size=-0.76, t=-2.19) and 41-45 years old (DIF size=-0.29, t=-0.21). There is male response bias for item K7 (DIF size=-0.79, t=-2.27).

The rating scale diagnostics suggest collapsing category 1 (observed count=15, 1.0%; outfit MnSq=1.91) and category 5 (observed count=221, 16.0%; outfit MnSq=1.01) with sensible adjacent category categories. A three-point Likert scale (correct=3, not sure=2, and wrong=1) with reverse scoring for negative-worded items were suggested.

DISCUSSION

We have developed the knowledge of the Noor EBM questionnaire using a literature search followed by testing for its content and face validity. This study reports the first construct verification testing of knowledge items regarding EBM among healthcare professionals using the Rasch measurement. We evaluated the application of this scale and provided suggestions for future refining of the scale in assessing the knowledge of EBM in healthcare professionals. Most of the 15 items conformed well to the Rasch model. The person reliability on the scale in measuring a single latent trait or construct was appropriate. The high reliability of the items also suggests that if this instrument is offered to another sample of respondents, there is a 96.0% probability that the items will be replicable [29]. The Item fit statistics indicated the instrument's goodness of fit and that it measured what was to be measured. The person fit statistics revealed that the Rasch model fits the overall person data.

The knowledge on EBM questionnaire showed high item reliability at 0.96 and good person reliability at 0.81. The reliability of items depends primarily on variance in the item difficulty and the sample size [36]. A wide range of item difficulties and a large sample size would produce high reliability. It is independent of test length and generally uninfluenced by the model fit. In comparison, reliability depends primarily on variation in sample performance, test length, and sample-item targeting [36]. These suggest that the number of samples in this study is adequate to locate the items accurately and to differentiate the sample into high and low performers.

Among the questionnaires that relate to EBM, only the Berlin questionnaire [10], Fresno test [6], and ACE tool [7] assess knowledge and skills on EBM. Whereas questionnaires such as McColl [15] and Baum [12] assess awareness, attitudes, competencies, and barriers to EBM.

The Berlin questionnaire tests applied knowledge through its multiple-choice structure, with a selective assessment of particular skills, such as constructing a clinical question or a search strategy [42]. The Berlin questionnaire consists of 15 items with appropriate internal coefficients of reliability between 0.75 and 0.82 [10]. No further psychometric properties for the Berlin questionnaire were assessed. The items also cited the setting of hypothetical studies (Bolivian, Argentinean, and Chilean), which in this context were deemed non-essential.

While the Fresno test consists of 18 open-ended items regarding the four steps of evidence-based practice [42]. It is built around a clinical scenario, with questions requiring a demonstration of core EBM skills, including the development of evidence-based clinical questions, identifying an appropriate search design to answer a research question, knowledge of literature search, and issues regarding validity and relevance [6]. It does not reflect the common clinical conditions in Malaysia.

The ACE tool is 15 dichotomous-response items that test EBM knowledge in four domains: question formulation, literature search, evidence appraisal , and evidence application [7]. It has a lower potential to distinguish between the levels of EBM knowledge compared to the Berlin questionnaire and the Fresno Test. With the binary answering format, an individual who does not know the answer still has a 50.0% chance to guess the correct answer on average. While all three questionnaires are standardized measures of basic EBM knowledge, the outcomes in the same population are not similar [9]. It means a standardized EBM questionnaire should be available that is ideally suited for assessing EBM.

The item separation index indicates the ability of the person to discriminate between the 15 items in seven strata or concordance levels. Overall, this shows a good spread of the 15 items in measuring the level of knowledge on EBM. In addition, it indicates the instrument's goodness of fitness, measuring what is meant to be measured. The higher the value of the items separation index, the better the instrument of measurement since the items are separated by varying difficulty levels. The separation index will improve if the reliability of items is improved, and misfit individuals are identified and excluded from the study or added more people. These indices show that the separation between persons can measure the ability of persons [29, 43].

The Rasch principal component analysis is low but has reached the minimum unidimensionality threshold of 40.0%, suggesting a good measuring dimension [44]. It is because of the noise in the item, which is 10.9%. The unexplained variance between 10.0-15.0% in the 1st contrast is a fair indicator of unidimensionality [41]. Therefore, one can conclude that the current 15 items can be treated as one-dimensional when measuring the level of knowledge on EBM. None of the items demonstrated local item dependency. The items are unidimensional with good internal consistency.

The goodness of fit test is a fundamental step to measure the quality of the items. The high MnSq for item K15 (*application of evidence-based practice is cost-effective to healthcare system*) indicates that the item does not contribute to the construction of scale but does not degrading. The PtMea Corr is positive, which suggests that the item is poorly correlated with the construct. The item is maintained, but it will probably require further refining.

Item K10 and K2 as well as items K13 and K7, seem to be overlapping in the Wright map. However, they were located at different logit measures and were of different contexts; therefore, they were retained in EBM knowledge scale. Three items, items K1, K14, and K9, can easily be answered even by low ability respondents. These items should be removed or replaced to improve the scale in the future. Item K15 is suggested for refining, which does not contribute to the construction of a scale but does not degrade. New more difficult items (above 1.55 logit) should be added. Identifying DIF is a critical step for improving measurement. However, in this study, the violation of the measurement invariance is limited to a single item for a particular sociodemographic character studied, and the underlying mechanisms of DIF in these person groups are unclear.

As highlights of the discussion, this study developed a new EBM questionnaire based on a literature search with high psychometric properties to be applied in a similar population. This questionnaire contained 15 items similar to The Berlin questionnaire and ACE tool, but fewer items if compared with the Fresno test. This questionnaire assessed knowledge of EBM compared to the Berlin, Fresno test, and ACE tool, which assessed the knowledge and skills. Meanwhile, McColl and Baum assessed the awareness, attitudes, competencies, and barriers to EBM.

There were some limitations to this study. First, the items were tested only among healthcare professionals in one hospital in this study. Therefore, the fitness of the Rasch model in another setting is not known. Second, these items might function differently if translated or used in other populations.

CONCLUSIONS

The knowledge of the Noor EBM questionnaire tested among healthcare professionals has shown to have a good person and a high item reliability index. It is a robust scale with excellent psychometric properties that can measure knowledge of EBM for both research and clinical purposes.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: Authors stated that the study was approved by the Human Research Ethics Committee, Universiti Sains Malaysia (USM/JEPeM/18040195) and Medical Registration and Ethics Committee, Ministry of Health (NMRR-18-349-40727).

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. BMJ. 1996;312(7023):71-2. https://doi.org/10.1136/bmj.312.7023.71 PMid:8555924 PMCid:PMC2349778
2. Rosenberg W, Donald A. Evidence based medicine: An approach to clinical problem-solving. BMJ. 1995;310(6987):1122-6. https://doi.org/10.1136/bmj.310.6987.1122 PMid:7742682 PMCid:PMC2549505
3. Masic I, Miokovic M, Muhamedagic B. Evidence based medicine–New approaches and challenges. Acta Inform Med. 2008;16(4):219-25. https://doi.org/10.5455/aim.2008.16.219-225 PMid:24109156 PMCid:PMC3789163
4. Tenny S, Gossman WG. Evidence based medicine (EBM). StatPearls. Treasure Island, FL: StatPearls Publishing; 2017.
5. Ghojazadeh M, Azami-Aghdash S, Pournaghi Azar F, Fardid M, Mohseni M. A systematic review on barriers, facilities, knowledge and attitude toward evidence-based medicine in Iran. J Anal Res Clin Med. 2015;3(1):1-11. https://doi.org/10.15171/jarcm.2015.001
6. Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. British Med J. 2003;326(7384):319-21. https://doi.org/10.1136/bmj.326.7384.319 PMid:12574047 PMCid:PMC143529
7. Ilic D, Nordin RB, Glasziou P, Tilson JK, Villanueva E. Development and validation of the ACE tool: Assessing medical trainees' competency in evidence based medicine. BMC Med Educ. 2014;14:114. https://doi.org/10.1186/1472-6920-14-114 PMid:24909434 PMCid:PMC4062508
8. Smith AB, Semler L, Rehman EA, et al. A cross-sectional study of medical student knowledge of evidence-based medicine as measured by the Fresno test of evidence-based medicine. J Emerg Med. 2016;50(5):759-64. https://doi.org/10.1016/j.jemermed.2016.02.006 PMid:26961178
9. Buljan I, Jerončić A, Malički M, Marušić M, Marušić A. How to choose an evidence-based medicine knowledge test for medical students? Comparison of three knowledge measures. BMC Med Educ. 2018;18(1):290. https://doi.org/10.1186/s12909-018-1391-z PMid:30514288 PMCid:PMC6278026
10. Fritsche L, Greenhalgh T, Falck-Ytter Y, Neumayer HH, Kunz R. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. British Med J. 2002;325(7376):1338-41. https://doi.org/10.1136/bmj.325.7376.1338 PMid:12468485 PMCid:PMC137813
11. Zwolsman SE, Wieringa-de Waard M, Hooft L, van Dijk N. Measuring evidence-based medicine knowledge and skills. The Dutch Berlin questionnaire: Translation and validation. J Clin Epidemiol. 2011;64(8):928-30. https://doi.org/10.1016/j.jclinepi.2011.02.005 PMid:21640555
12. Lavizeh M, Kouhpayeh Zadeh J. Validity and reliability of the attitudes towards and self-reported ability in evidence-based medicine questionnaire. Int J Med Educ. 2015;14(11):998-1006.
13. Zwolsman SE, van Dijk N, te Pas E, Wieringa-de Waard M. Barriers to the use of evidence-based medicine: Knowledge and skills, attitude, and external factors. Perspect Med Educ. 2013;2(1):4-13. https://doi.org/10.1007/S40037-013-0039-2 PMid:23670651 PMCid:PMC3576485
14. Hisham R, Ng CJ, Liew SM, et al. Development and validation of the evidence based medicine questionnaire (EBMQ) to assess doctors' knowledge, practice and barriers regarding the implementation of evidence-based medicine in primary care. BMC Fam Pract. 2018;19(1):98. https://doi.org/10.1186/s12875-018-0779-5 PMid:29935527 PMCid:PMC6015654

15. McColl A, Smith H, White P, Field J. General practitioners; perceptions of the route to evidence based medicine: A questionnaire survey. BMJ. 1998;316(7128):361. https://doi.org/10.1136/bmj.316.7128.361 PMid:9487174 PMCid:PMC2665572
16. Barzkar F, Baradaran H, Koohpayehzadeh J. Knowledge, attitude, and practice of evidence-based medicine among physicians: A systematic review. J Evid Based Med. 2018;11(4):246-51. https://doi.org/10.1111/jebm.12325 PMid:30430759
17. Katsilometes J, Galuska M, Kraus CK, et al. Multisite assessment of emergency medicine resident knowledge of evidence-based medicine as measured by the Fresno test of evidence-based medicine. J Osteopath Med. 2022; 122(10):509-15. https://doi.org/10.1515/jom-2022-0027 PMid:35704661
18. Huang H-L, Chou C-P, Leu S, You H-L, Tiao M-M, Chen C-H. Effects of a quasi-experimental study of using flipped classroom approach to teach evidence-based medicine to medical technology students. BMC Med Educ. 2020;20(1):31. https://doi.org/10.1186/s12909-020-1946-7 PMid:32005227 PMCid:PMC6995213
19. Cardoso D, Couto F, Cardoso AF, et al. Fresno test to measure evidence-based practice knowledge and skills for Portuguese undergraduate nursing students: A translation and adaptation study. Nurse Educ Today. 2021;97:104671. https://doi.org/10.1016/j.nedt.2020.104671 PMid: 33264737
20. Imorde L, Möltner A, Runschke M, Weberschock T, Rüttermann S, Gerhardt-Szép S. Adaptation and validation of the Berlin questionnaire of competence in evidence-based dentistry for dental students: A pilot study. BMC Med Educ. 2020;20(1):136. https://doi.org/10.1186/s12909-020-02053-0 PMid:32366287 PMCid:PMC7197120
21. Vulpe H, Vanderpuyne V, Yarney J, et al. Design and implementation of a distant-learning clinical research mentorship program: The accra-Toronto collaboration. JCO Glob Oncol. 2020;6:919-28. https://doi.org/10.1200/JGO.19.00240 PMid:32603189 PMCid:PMC7328115
22. Mukherjee P, Chaudhary K, Yadav BK, Kumar S, Kekre N, Devasia A. Attitude and perceived barriers towards the practice of evidence-based urology amongst urological trainees in India. Indian J Urol. 2019;35(2):116-9.
23. Habibi MA, Amini M, Ostovarfar M, Ostovarfar J, Moosavi M, Keshavarzi MH. Reliability and validity of the Persian version of the ACE tool: Assessing medical trainees' competency in evidence-based medicine. BMC Med Educ. 2022;22(1):468. https://doi.org/10.1186/s12909-022-03540-2 PMid:35710410 PMCid:PMC9204903
24. Sgarbossa N, Marino S, Aletta L, Vázquez Peña FR, Franco JVA. Validity and reliability of the Spanish version of the ACE tool for assessing competencies in evidence-based medicine in medical students. BMJ Evid Based Med. 2023;28(2):89-94. https://doi.org/10.1136/bmjebm-2021-111888 PMid:36150894
25. da Silva Maia FGSS, de Souza AKC, de Simas BCC, Lopes IS, Bezerra MPRD, Diniz RVZ. Brazilian version of the ACE (assessing competencies in evidence-based medicine) tool: A validation study. Rev Bras Educ Med. 2022;46(03):e93. https://doi.org/10.1590/1981-5271v46.3-20220081.ing
26. Simons M, Rapport F, Zurynski Y, Stoodley M, Cullis J, Davidson AS. Links between evidence-based medicine and shared decision-making in courses for doctors in training: A scoping review. BMJ Open. 2022;12(4):e057335. https://doi.org/10.1136/bmjopen-2021-057335 PMid: 35470193 PMCid:PMC9039384
27. Simons MR, Zurynski Y, Cullis J, Morgan MK, Davidson AS. Does evidence-based medicine training improve doctors' knowledge, practice and patient outcomes? A systematic review of the evidence. Med Teach. 2019;41(5):532-8. https://doi.org/10.1080/0142159X.2018.1503646 PMid: 30328793
28. DeVellis RF. Classical test theory. Med Care. 2006;44(11 Suppl 3):S50-9. https://doi.org/10.1097/01.mlr.0000245426.10853.30 PMid:17060836
29. Bond TG, Fox CM. Applying the Rasch model: Fundamental measurement in the human sciences. Psychology Press; 2013.
30. Jamieson S. Likert scales: How to (ab)use them. Med Educ. 2004;38(12):1217-8. https://doi.org/10.1111/j.1365-2929.2004.02012.x PMid:15566531
31. Ishar NIM, Roslin RM. Customer voice retaliation (CVR) construct verification: A Rasch analysis approach. Procedia Econ Financ. 2016;37:214-20. https://doi.org/10.1016/S2212-5671(16)30116-2
32. Müller S, Kohlmann T, Wilke T. Validation of the adherence barriers questionnaire–An instrument for identifying potential risk factors associated with medication-related non-adherence. BMC Health Serv Res. 2015;15:153. https://doi.org/10.1186/s12913-015-0809-0 PMid:25884193 PMCid:PMC4422301
33. Linacre JM. Sample size and item calibration stability. Rasch Meas Trans. 1994;7:328.
34. Eubank BH, Mohtadi NG, Lafave MR, et al. Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. BMC Med Res Methodol. 2016;16:56. https://doi.org/10.1186/s12874-016-0165-8 PMid:27206853 PMCid:PMC4875724
35. Linacre J. A user's guide to Winsteps Rasch model computer programs: Program manual 3.92. 0. Chicago, IL: Mesa-Press; 2016.
36. Linacre JM. Winsteps® Rasch measurement computer program user's guide. Beaverton, OR: Winsteps.com; 2020.
37. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 2007;57(8):1358-62. https://doi.org/10.1002/art.23108 PMid:18050173
38. Streiner DL, Norman GR, Cairney J. Health measurement scales: A practical guide to their development and use. Oxford: Oxford University Press; 2015. https://doi.org/10.1093/med/9780199685219.001.0001
39. Fox CM, Jones JA. Uses of Rasch modeling in counseling psychology research. J Couns Psychol. 1998;45(1):30. https://doi.org/10.1037/0022-0167.45.1.30
40. Linacre JM. Test validity, and Rasch measurement: Construct, content, etc. Rasch.org; 2004. Avaişable at: https://www.rasch.org/rmt/rmt181h.htm (Accessed: 19 December 2022).
41. Fisher Jr WP. Rating scale instrument quality criteria. Rasch Meas Trans. 2007;21:1095.

42. West CP, Jaeger TM, McDonald FS. Extended evaluation of a longitudinal medical school evidence-based medicine curriculum. J gen Intern Med. 2011;26(6):611-5. https://doi.org/10.1007/s11606-011-1642-8 PMid:21286836 PMCid:PMC3101983
43. Wright BD, Masters GN. Rating scale analysis: MESA press; 1982.
44. Conrad KM, Conrad KJ, Passetti LL, Funk RR, Dennis ML. Validation of the full and short-form self-help involvement scale against the Rasch measurement model. Eval rep. 2015;39(4):395-427. https://doi.org/10.1177/0193841X15599645 PMid:26275980 PMCid:PMC4591028