Development and psychometric evaluation of triage questionnaire (QTrix): Exploratory factor analysis and item response theory analysis

Nurul Fateha Abd Rahman¹(✉), Tuan Hairulnizam Tuan Kamauzaman¹, Wan Nor Arifin² and Abu Yazid Md Noh¹

Abstract
Introduction: Triage errors can occur in all emergency departments, regardless of the type of triage system being used. One way to minimize triaging errors is by enhancing the triage officers’ knowledge and attitude on triage. The assessment of knowledge and attitude can be carried out by questionnaire assessment. This study aims to perform content, face, and construct validation on a newly developed triage questionnaire, QTrix, which is designed for healthcare personnel in a tertiary teaching hospital in Kelantan, Malaysia, that uses the three-tier Malaysian Triage Category system.

Methods: This study consisted of two phases: the first phase was the questionnaire development phase, which included the content validity with the expert panel and the face validity using 30 respondents; and the second phase was the psychometric assessment phase, which included the item response theory and the exploratory factor analysis using 139 respondents.

Results: The knowledge section with 12 remaining items was considered unidimensional by item response theory after removing items with extreme difficulty coefficients (outside the range of −3 to +3) and items with very low discrimination values (<0.35). After exploratory factor analysis, two items in the attitude section were removed due to low factor loadings (<0.3) and high item complexity. The reliability of the remaining 13 items in the attitude section was very good as shown by Cronbach’s alpha values of more than 0.8.

Conclusion: The QTrix questionnaire is a well-validated and reliable tool to assess the knowledge and the attitude on triage. Its use among healthcare personnel can help minimize triaging errors in emergency departments that utilize the Malaysian Triage Category system.

Keywords
Triage, emergency departments, knowledge, attitude, questionnaire

Introduction
Triage is a French word, meaning to sort out. In the past, this term was being used for agricultural products.¹ In medicine, to triage means to prioritize patients according to their severity.² Triaging is necessary in a situation where there are many patients to be treated at the same time, for example, in the Emergency Department (ED). In 2013, there were nearly 10,000,000 outpatient visits in ED in Malaysian hospitals and about three-quarters of these visits were in government hospitals.³
A triage officer in ED is expected to perform every ED triage accurately despite having very limited time and information. This is not easy even with experience, and a wrong triage can put the patient at risk. A wrong triage can be in the form of either under-triaging a patient or over-triaging a patient. Under-triaging leads to assigning a less urgent code to a patient that can cause the patient to deteriorate while waiting for treatment, while over-triaging leads to assigning a more urgent code to a patient that can lead to unnecessary wasting of resources and man power. The most important factors to improve one’s triage are still education, experience, and empathy.

The Malaysian Triage Category (MTC) is a three-tier triaging system, in which the patients can be triaged into three categories: red, yellow, and green. Upon primary triage, all unstable patients who need urgent care will be triaged to red to be managed immediately, while other stable patients will undergo secondary triage. After a quick evaluation in secondary triage, the triage officer will further decide to triage the patient to yellow or green, according to the respective clinical urgency. In all hospitals in Malaysia, the assistant medical officers (AMOs) and staff nurses (SNs) are the main triage officers, and in selected hospitals such as the tertiary hospitals and teaching hospitals, doctors are also posted as triage officers.

In the literature, knowledge and attitude have been assessed using questionnaires, and the results are reliable and reproducible. Exploratory factor analysis (EFA) is a form of factor analysis for polytomous data that explores the presence of factors in relation to the number of items being analyzed and has been widely in psychology and in health education. Item response theory (IRT) is another tool that has been used for test validations by assessing the difficulty and the discriminative level of a test.

So far, there has been no single questionnaire to assess the knowledge and the attitude among triage officers in Malaysia. The aims of this study are to perform content, face, and construct validations on a recently developed questionnaire, the QTrix questionnaire, to assess the knowledge and the attitude on ED triage among healthcare personnel in the ED of Hospital Universiti Sains Malaysia (EDHUSM), where the three-tier MTC system is used. Psychometric evaluation of the questionnaire is carried out using EFA and IRT to assess its validity and reliability.

Methods

Study design

This study consisted of two phases: (1) the development phase of the QTrix questionnaire, including content and face validation; and (2) the psychometric evaluation phase, which focused on internal structure validation using IRT and EFA. The internal reliability of this questionnaire is further evaluated using Cronbach’s alpha coefficient and the corrected item-total correlation (CITC). The study was conducted between 16 February 2017 and 15 February 2018. Written informed consent was not necessary because no patient data have been included in this article.

Phase 1: questionnaire development

The QTrix questionnaire was written in English to minimize the misinterpretation of medical terms. It has three sections: the sociodemographic, knowledge, and attitude sections. Items in the knowledge section were derived from the book titled Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care. Content validation was conducted by discussing the relevance of each item in each section of the questionnaire with the expert panels. The expert panels were selected based on their expertise in ED triage and in questionnaire validation. Five consultants from the Emergency Medicine Department and one lecturer from the Unit of Biostatistics and Research Methodology were recruited during this phase.

The content-validated questionnaire then underwent pre-testing among 30 subjects to verify the applicability of each item and to evaluate respondents’ understanding of the questionnaire. This process is also known as the face validation. The participants were recruited via convenience sampling among medical officers, assistant medical officers, and staff nurses from EDHUSM who had more than 5 years of working experience in ED. The variability of their responses and their understanding of the questions, readability, and presence of any ambiguity were recorded and evaluated. The results were used to produce the preliminary version of the questionnaire, which was subjected to the remainder of the evaluation in the study.

Phase 2: psychometric evaluation

The internal structure validation of this questionnaire was performed among the healthcare personnel in EDHUSM via construct validation or psychometric evaluation. The sample population for this phase included house officers, assistant medical officers, staff nurses, and health assistants. Those who were already involved during pre-testing of the questionnaire were excluded. The preliminary version of the QTrix questionnaire was administered via an online form, and the results were recorded into the Statistical Package for the Social Science (SPSS) and analyzed using the R studio.

Psychometric evaluation of this questionnaire involved the use of IRT for dichotomous data as in the knowledge section and EFA for polytomous data as in the attitude section. Most recommendations in the literature agreed to a sample size of at least 100 for both EFA and IRT, and others advised respondents-to-item ratio that ranged from 3:1 to 20:1. The flowchart for the whole process of IRT and EFA is summarized in Figure 1.
Phase 2a: psychometric evaluation using IRT

IRT had two postulations in each analysis: that all items in a test were unidimensional; and that the chance of one respondent having one answer correct was independent of the chance that the same trait of respondent had the next answer correct.  

Using two-parameter logistic IRT (2-PL IRT), each item from the knowledge section was given its difficulty and discrimination parameters. Items with extreme difficulty coefficients—either too positive (>3) or too negative (<−3) and very low discriminative coefficients (<0.35)—were removed to ensure unidimensionality. 

A test was deemed unidimensional once items with extreme difficulty coefficients and non-discriminative items were removed. Subsequently, the performance of a respondent in each test item can be predicted by the respondent’s ability, and the relationship between the respondent’s ability and the item’s discrimination and difficulty can be presented in item characteristic curves (ICC) and item information curves (IIC). 

Phase 2b: psychometric evaluation using EFA

Analysis of the attitude section using EFA required assessment of the normality of the data to determine the mode of analysis: maximum likelihood (ML) if the data were normally distributed or principal axis factoring (PAF) if the data were not normally distributed. In this section, the rotation method is carried out using oblimin, considering that the factors are correlated with each other. The data were assessed for its suitability for analysis using the Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy (MSA) of at least 0.7 and using Bartlett’s test of sphericity, where the p value needs to be significant. 

The number of extracted factors for the attitude section was determined using Kaiser’s eigenvalue of more-than-one rule, where factors with eigenvalue more than 1 are kept. The number of extracted factors was confirmed with Cattell’s scree test, parallel analysis, very simple structure, and Velicer’s minimum average partial (MAP). A good factor must have at least three items, and a good item must have the factor loading of more than 0.3 with no cross loadings, the item communality of more than 0.4, and the item complexity as close to 1. Cronbach’s alpha coefficient and CITC. A satisfactory Cronbach’s alpha coefficient ranges from 0.7 to not more than 0.9, while an ideal CITC is defined as more than 0.5. Cronbach’s alpha coefficient, if an item is deleted, is only considered significant if there is a marked improvement.

Results

Phase 1: questionnaire development

The QTrix questionnaire has three sections: the sociodemographic, knowledge, and attitude sections, which are summarized in Table 1. In the content validation of the knowledge section, knowledge question 10 (K10) in Table 2 was identified as unfit for its intended red domain; therefore, the item was modified by altering the patient’s heart rate to “190 beats/min” so that the number of items remained equal for each domain to maintain equal coverage of items for each domain. In the attitude section, which is presented in Table 3, the expert panel unanimously agreed that all items were consistent with the intended domains in terms of relevance, coverage, and representativeness.

For face validation, the content-validated QTrix questionnaire was pre-tested among senior healthcare personnel and medical officers. From their responses, a few changes were made in terms of wording, terminologies, and layout to ensure that the questions were clear and easy to understand. The final number of items in the final draft of the questionnaire at this stage remained the same.
Table 1. Summary of the QTrix questionnaire.

| Sections         | Number of items | Concepts measured                                                                 | Response options                                      |
|------------------|-----------------|-----------------------------------------------------------------------------------|-------------------------------------------------------|
| Sociodemographic | 8               | Sociodemographic data, working duration in ED, any triage training in the past, and profession. | Closed-ended, multiple-choice answers.                |
| Knowledge        | 24              | Red triage code, yellow triage code, and green triage code.a                      | Red/yellow/green/I don’t know; 1 = correct answer and 0 = incorrect or no answer. |
| Attitude         | 15 (four reversed scored items) | Affective, behavioral, and cognitive component of attitude.b | 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree. |

ED: Emergency Department.

aItems were adapted from the book Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care and divided into three domains according to the Malaysian Triage Categories.

bThree components of attitude according to Essentials of Social Psychology.

Table 2. IRT parameter estimates for items in the knowledge section.

| Items | Difficulty coefficient (SE) | Discrimination coefficient (SE) |
|-------|-------------------------------|---------------------------------|
| K1    | 1.29 (0.46)                   | 0.79 (0.29)                     |
| K2    | −0.36 (0.14)                  | 2.26 (0.67)                     |
| K5    | −3.70 (1.69)                  | 1.01 (0.58)                     |
| K6    | −0.59 (0.37)                  | 0.61 (0.24)                     |
| K7    | −1.00 (0.18)                  | 2.70 (0.92)                     |
| K9    | 0.25 (0.16)                   | 1.53 (0.42)                     |
| K10   | 0.44 (0.46)                   | 0.44 (0.23)                     |
| K12   | −0.35 (0.49)                  | 0.39 (0.22)                     |
| K17   | −1.34 (0.33)                  | 1.23 (0.37)                     |
| K18   | −3.38 (1.37)                  | 1.06 (0.55)                     |
| K20   | −1.56 (0.23)                  | 3.86 (2.04)                     |
| K22   | −4.66 (2.95)                  | 0.76 (0.55)                     |

SE: standard error; HR: heart rate; BP: blood pressure; RR: respiration rate; SpO2: peripheral oxygen saturation; DM: diabetes mellitus; ED: Emergency Department.
Phase 2: psychometric evaluation

There were 139 respondents for the construct validation. The summary of the respondent’s sociodemographic data is shown in Figure 2. The sample size for this phase fulfilled the minimum requirement of 100 respondents.\(^{11-14}\) The responses in the knowledge section were coded as: 1 = correct responses and 0 = incorrect or missing responses. The responses in the attitude section were coded as: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. Any missing data were handled using the MICE function in the R studio.

Phase 2a: psychometric evaluation using IRT

Items in the knowledge section of this questionnaire were given its difficulty and discrimination parameters using 2-PL IRT. A total of 12 items were identified to have extreme difficulty values and very low discrimination values, which were K3, K4, K8, K11, K13, K14, K15, K16, K19, K21, K23, and K24. These items were removed, and the remaining 12 items were unidimensional. The result was presented in Table 2.

The result from IRT is presented in Figure 3(a) to (d): the ICC, the IIC, the test information function, and the test response function, respectively. The knowledge section of this questionnaire has a good discriminative peak with a value of more than 1.7, which is shown by the test information function curve in Figure 3(c). This section also has a good smoothly rising slope signifying an adequately discriminative test, as portrayed by the test response function in Figure 3(d).

Phase 2b: psychometric evaluation using EFA

Assessment of normality in the attitude section revealed that the data were not normally distributed. The data were deemed fit for analysis because the MSA was 0.82 which indicated good sampling adequacy. Bartlett’s test of sphericity was significant with \(p < 0.05\), indicating that there were correlations between the items in this section.

Kaiser’s eigenvalue calculation of more-than-one rule suggested three-factor solution, while Cattell’s scree test and Velicer’s MAP suggested only two factors. Analysis with two factors gave the most acceptable result, and attitude question items—three (A3a) and five (A5a)—were removed due to their low factor loadings and high item complexity. There was no factor overlap as indicated by factor correlations of less than 0.85.

The two factors identified were labeled as Attitude Factor 1 (PA1) and Attitude Factor 2 (PA2). PA1 had 10 items that originated from all three domains of the attitude section, while PA2 only had three items that originated from the behavior domain. Cronbach’s alpha coefficient was very good for all items, ranging from more than 0.7 to not more than 0.9. There is no significant improvement in...
Figure 2. The sociodemographic data of respondents from the construct validation process.

Cronbach’s alpha if any item in this questionnaire was deleted. Each item’s CITC of more than 0.45 was accepted and agreed upon among the expert panel. The result from the EFA was summarized in Table 3.

**Discussion**

This study aimed to develop and validate the recently developed triage questionnaire, QTrix. This questionnaire was designed in English language, using the Malaysian three-tier triaging system to assess the knowledge and the attitude on ED triage among healthcare personnel from EDHUSM—a tertiary teaching hospital in Kelantan, Malaysia. Each item in each domain was discussed among the expert panels regarding its relevance and suitability as part of the questionnaire’s content validity. An identified limitation to this phase was the lack of psychologist input, considering that this questionnaire assessed not only the respondent’s knowledge but also the attitude on ED triage.

The expert panels and the intended group of respondents may have different understandings of the items in this questionnaire; therefore, a pre-testing was required following the content validation. Pre-testing, or also called face validity, was conducted to ensure that each item is understood the way the researcher intended. Pre-testing can enhance responses by improving respondent’s motivation and cooperation, decreasing any form of respondent’s disapproval, inviting more respondents, allowing other readers to be convinced by the result, and boosting public acceptance.10

The internal structure validation, or the construct validity, of this QTrix questionnaire involved performing psychometric evaluation of the data using IRT and EFA. IRT has been widely used since the 1970s in tests like the Scholastic Aptitude Tests (SAT) in the United States and is suitable for assessing dichotomous data.20 EFA is a form of factor analysis that is suitable for polytomous data, which was mainly used in psychology and education in the past, and now slowly gaining its glamor in the health sector.14 Both IRT and EFA can be carried out using the R studio, which is an open free program with an extensive user-friendly network called the Comprehensive R Archive Network (CRAN).21

Using IRT, each item was analyzed for its difficulty and discrimination values. A difficult question has a positive
Figure 3. A more discriminative item was represented by (a) a steeper slope curve in ICC and (b) a higher peak curve in IIC. In both ICC (a) and IIC (b), a more difficult item has its curve shifted to the right, while an easier item has its curve shifted to the left. (c) This test has a good discriminative peak as represented by the test information function. (d) This test has a good smoothly rising slope as portrayed in the test response function.
difficult value and vice versa, and a question with a high discrimination value can distinguish a brilliant respondent from a less brilliant respondent. There were three items with difficulty coefficient of less than −3, namely, K5, K18, and K22. These items were retained because of their good discriminative coefficients (>0.7), and this action did not alter the unidimensionality of the test.

Difficulty values were translated into ability in the ICC, IIC, test information function, and test response function as depicted in Figure 3. In IIC, an item that peaks at the negative ability value can discriminate respondents with lower ability, while an item that peaks at the positive ability value can discriminate respondents with higher ability. Figure 3(b) displayed the IIC of this QTrix questionnaire, where the highest peak information was represented by knowledge question 20 (K20), at the ability value of −1.6.

The test information function curve in Figure 3(c) displayed that the knowledge section of this QTrix questionnaire could best discriminate a respondent with the ability value within the range of −4 and +2. This means that the QTrix questionnaire can best discriminate respondents with a slightly lower ability and therefore is suitable for the assessment of knowledge and attitude among healthcare personnel involved in ED triage, who are mostly AMOs and SNs.

EFA, on the contrary, finds out the correlation between each item in the attitude section of this questionnaire: whether the items are (1) strongly related to each other, (2) not related to each other, or (3) can be grouped together with other items alike. The items in the attitude section of this QTrix questionnaire were initially drafted with three domains: the affective, behavioral, and cognitive components of attitude. After analysis, there were only two identified factors: Attitude Factor 1 and Attitude Factor 2. Attitude Factor 1 had 10 items that originated from all three domains of the attitude section, while Attitude Factor 2 contained three items that originated from the behavior domain only. Attitude Factor 1 is the explicit attitude, and Attitude Factor 2 is the implicit attitude.

Attitude has been widely studied way back in World War II by Stouffer et al., which concluded that a person’s attitude does not necessarily predict that person’s behavior and this was supported by LaPiere’s observational study. This buttressed the finding of two factors in the attitude section: the explicit and implicit attitudes, or in simpler words, conscious and subconscious attitudes. Explicit attitude requires more thinking due to its conscious nature as compared to implicit attitude, which occurs subconsciously. While explicit attitude is easily measurable by written tests, implicit attitude is measured by how fast the test is carried out. The items in Attitude Factor 1 require the respondents to think more than does the items in Attitude Factor 2; hence, it is reasonable to rename these two factors as explicit attitude and implicit attitude, respectively.

The internal reliability of this questionnaire in the attitude section was achieved by its good Cronbach’s alpha coefficient of each factor, ranging from 0.7 to 0.9. Although there were a few items not meeting a good value of CITC, these items were kept because of their good factor loadings and Cronbach’s alpha coefficient values. The removal of these items would not significantly improve the overall value of Cronbach’s alpha in the attitude section.

The QTrix questionnaire has good psychometric properties, and is a valid and reliable tool to evaluate knowledge and attitude on triage among healthcare personnel in ED that practices three-tier triaging systems like the MTC. It is thus not suitable for hospitals outside Malaysia not using the same three-tier system. Other limitations of this study were its relatively small sample size. Improvements that can be done to improve the validity and the reliability of this questionnaire are to run the study with a bigger sample size and additional analysis using confirmatory factor analysis (CFA) as part of the psychometric evaluation of this questionnaire.

Acknowledgements

First and foremost, the authors wish to express their sincere gratitude to their supervisors and specialists from Emergency Medicine Department, Universiti Sains Malaysia. Their continuous support and supervision had greatly contributed to the completion of this research. The authors also express their gratitude to all healthcare personnel in Emergency Department, Hospital Universiti Sains Malaysia, for their kind involvement in this study. A special tribute to Nur Zafirah and Wan Nur Hazrati, who have been a tremendous help in finishing this study. A very special appreciation to Aisyah Abd Jalil and her husband Christopher Nuttall for their final touch in this dissertation. The authors also express their sincere love and appreciation to their friends and family for their continuous support and infinite prayers that have kept them going.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Availability of data and materials

The data and materials are available in soft copy attached as SPSS data named as follows: 1. SOCIODEMOGRAPHICdata.sav. 2.KNOWLEDGE.sav. 3a. ATTITUDE.sav. 3b. ATTITUDEwith IMPUTEDdata.sav.

Informed consent

All the participants who involved in this study had their informed consent taken by the authors.
Ethical approval
Ethical approval was obtained from the Human Research Ethics Committee (HREC), Universiti Sains Malaysia (USM) with JEPEM Code: USM/JEPEM/16090308.

Human rights
There is no violation of human rights during the completion of this study.

ORCID iD
Nurul Fateha Abd Rahman https://orcid.org/0000-0001-9620-8815

References
1. Curtis K and Ramsden C. Emergency and trauma care for nurses and paramedics. Sydney, NSW, Australia: Elsevier, 2015.
2. Goransson KE, Ehrenberg A and Ehnfors M. Triage in emergency departments: national survey. J Clin Nurs 2005; 14: 1067–1074.
3. Sivasampu S, Foo C, Aimi N, et al. (eds). National healthcare establishment and workforce statistics (hospital) 2012–2013. The Ministry of Health Malaysia: The National Healthcare Statistics Initiative (NHSI), 2015.
4. DiDomenico PB, Pietzsch JB and Pate-Cornell ME. Bayesian assessment of overtriage and undertriage at a level I trauma centre. Philos Trans A Math Phys Eng Sci 2008; 366: 2265–2277.
5. McNair RS. It takes more than string to fly a kite: 5-level acuity scales are effective, but education, clinical expertise, and compassion are still essential. J Emerg Nurs 2005; 31: 600–603.
6. Abd Wahab M, Aik Howe T, Mohd Noor AT, et al. Emergency medicine and trauma services policy. 1st ed. Putrajaya, Malaysia: The Ministry of Health Malaysia, 2012.
7. Vatnøy TK, Fossum M, Smith N, et al. Triage assessment of registered nurses in the emergency department. Int Emerg Nurs 2012; 21: 89–96.
8. Rahmati H, Azmoon M, Kalantari Meibodi M, et al. Effects of triage education on knowledge, practice and qualitative index of emergency room staff: a quasi-interventional study. Bull Emerg Trauma 2013; 1(2): 69–75.
9. Gilboy N, Tanabe P, Travers D, et al. Emergency severity index (ESI): a triage tool for emergency department care.