

# Psychometric Classical Theory Test and Item Response Theory Validation of Patients' Knowledge, Attitudes and Practices of Uncomplicated Malaria Instrument

Nanloh Samuel Jimam<sup>1,2</sup>, Sohail Ahmad<sup>1</sup>, Nahlah Elkudssiah Ismail<sup>1\*</sup>

<sup>1</sup>Department of Clinical Pharmacy, Faculty of Pharmacy, MAHSA University, Selangor, MALAYSIA.

<sup>2</sup>Department of Clinical Pharmacy and Pharmacy Practice, Faculty of Pharmaceutical Sciences, University of Jos, Plateau State, NIGERIA.

## ABSTRACT

**Objective:** To develop and validate Patients' Knowledge, Attitudes and Practices Instrument for Uncomplicated Malaria (PKAPIUM) through Classical Test Theory (CTT) complemented by Items Response Theory (IRT).

**Methods:** A draft 31-items' scale was developed using relevant variables from literature and initially screened by six experts before it was used to collect data from 300 patients receiving treatment for uncomplicated malaria in Primary Health Care (PHC) facilities in Plateau state, Nigeria. An orchestrated classical and modern psychometric approach including CTT and IRT was then used to validate the draft instrument through IBM® Statistical Package for Social Sciences (SPSS®) version 23 and Analysis of Moment Structures (AMOS™) software version 22 and Bond and Fox software®, respectively. **Results:** The 31-items' draft scale showed good Item's Content Validity Index (I-CVI) (> 0.8) with good Universal Agreement (UA) level of Scale Content Validity Index (S-CVI/UA) (0.9 – 1) and average CVI (S-CVI/Ave) (0.98 – 1). The CTT and Rasch analyses resulted in retention of twenty one items distributed under Knowledge, Attitude and Practice (KAP) constructs, with Average Variance Extracted (AVE), square root AVE,

chi-square, Standardized Root Mean square Residual (SRMR), Root Mean Square Error Approximation (RMSEA), items' infit Mean Square (MNSQ), Infit Standardized Z-scores (infit Zstds), Point-Measure Correlation Coefficients (PTMEA Corr), Cronbach's alpha, items' and person's reliability indices within accepted limits. **Conclusion:** The new scale was considered valid and reliable for assessing patients' knowledge, attitudes and practices on uncomplicated malaria.

**Key words:** Cronbach's alpha, Factor analysis, Infit and outfit indices, Persons and items reliability, Point-measure correlation coefficients, Standardized Z-score.

## Correspondence

**Prof. Dr. Nahlah Elkudssiah Ismail**, Department of Clinical Pharmacy, Faculty of Pharmacy, MAHSA University, Selangor, MALAYSIA.

Phone: +60 162731677

Email: nahlah@mahsa.edu.my

DOI: 10.5530/jyp.2019.11.39

## INTRODUCTION

Patients' levels of adherence to treatment largely depend on many factors, including among others, their knowledge on the disease and regimen, attitude and practice during the drug therapy.<sup>1-3</sup> Studies have shown the influence of patients' knowledge of a disease on their beliefs regarding the aetiology of the disease,<sup>3,4</sup> the seriousness they attached to the ailment<sup>5,6</sup> and their perceptions on the effectiveness of the medication for such ailment.<sup>6-7</sup> Assessing these patients' characteristics during medications might help in the achievement of quality malaria treatment outcomes. Although literatures showed that some Knowledge, Attitude and Practice (KAP) studies have been conducted in Plateau state, Nigeria, the survey instruments might not be reliable and valid because they were not subjected to rigorous validation process to ascertain their reliability and fitness for the study,<sup>3,8-10</sup> hence, validation through statistical test models such as Classical Test Theory (CTT) and Items Response Theory (IRT), also known as Modern Mental Test Theory (MMTT), could be considered an initial step to obtain a test instrument that is both reliable and valid, containing appropriate items for the study.<sup>11-14</sup> While CTT approach focused more on the test measurement's reliability, the IRT which is considered as an extension to CTT, goes beyond the underlying traits which are producing the test's instrument performance.<sup>11-13</sup> It provides many statistical outputs that give more important details regarding test instruments, including items' and persons' reliability and separation and items' fitness to the individual constructs that constitute the test model.<sup>12,14</sup> This method is good because apart from measuring

the levels of consistency of the test, it also gives the respondents' levels of ability to respond to such questions and levels of difficulty of the questions/items to respondents.<sup>14,15</sup> The purpose of this study was to develop and validate a 31-items' draft Patients' Knowledge, Attitude and Practice Instrument for Uncomplicated Malaria (PKAPIUM) scale for assessing patients' KAP on uncomplicated malaria using CTT which was then complemented by IRT (Rasch Analysis Method).

## MATERIALS AND METHODS

A draft 31-items' PKAPIUM was developed from variables that best represented patients' KAP-related to uncomplicated malaria management which were extracted from relevant literatures.<sup>3,8-10</sup> Face and content validity were checked by six experts in the field of study through the method described by Devon *et al.*<sup>15</sup>

Convenient sampling method was used to select 6 Primary Health Care (PHC) facilities in Plateau state, Nigeria, from where three hundred patients receiving treatment for uncomplicated malaria were recruited by simple random sampling to participate in the study. The patients' sample size was estimated by the Monte Carlo simulation principle.<sup>16</sup> Prior to data collection, permission was granted by the Joint Research Review and Ethics Committee, Research Management Centre (RMC), MAHSA University, Malaysia (Ref. number: RMC/EC01/2016; Dated 25/11/2016), with subsequent approval by Commissioner of Health, Plateau State Ministry of Health, Jos, Nigeria and the directors of PHCs of

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

selected local government areas (LGAs). Signed informed consent was also obtained from all individual participants included in the study.

The collated data were coded and entered into Microsoft Excel for statistical analysis using IBM Statistical Package for Social Sciences (SPSS) version 23 and Analysis of Moment Structures (AMOS) software version 22 and Bond & Fox software.

## Data analysis

The statement construction for each of the 31 items of the draft scale relating to wordings, structures, orderliness and scoring formats were subjectively checked to ensure their appropriateness based on the purpose for which the instrument was developed. The Item-Level Content Validity Index (I-CVI) was estimated from experts' rating on Likert's scale and the Scale Content Validity Index (S-CVI). Items with I-CVI  $\geq 0.79$  were accepted as valid to be used as part of the instrument,<sup>17</sup> similarly, S-CVI/Ave values of  $\geq 0.9$  indicated the relevance, clarity, simplicity and comprehensiveness of the scale.<sup>15,17</sup> An orchestrated classical and modern psychometric approach similar to the one used by Ismail *et al.*<sup>18</sup> was then used to determine the constructs validity of the draft instrument.

Preliminary screening was conducted using the patients' data for acceptability, floor, ceiling effects, relationships among the items and Kaiser-Meyer-Olkin (KMO) test for sample size adequacy.<sup>15,16</sup> This was followed by Exploratory Factor Analysis (EFA) following the 'Kaiser's eigenevalue-greater-than-one rule'<sup>15</sup> and any factor with eigenevalues above one was retained using varimax orthogonal rotation method.<sup>11,15</sup> The initial extraction yielded many factors extracts that could not be of any useful meaning after interpretation, hence, the extraction was again carried out by fixing three factors to be extracted. All items with loading strength of  $< \pm 0.3$  and communalities of  $< 0.5$  indicated poor patterned interrelationships between a factor and items.<sup>11,15</sup> The cross-loading and Fornell-Larcker criterion were used to estimate the convergent and discriminant validity of the scale,<sup>15,19</sup> which was further confirmed by goodness-of-fit statistics (Chi-square Minimum Discrepancy, Divided by its Degrees of Freedom) (CMIN/DF), the Standardized Root Mean square Residual (SRMR) and Root Mean Square Error Approximation (RMSEA)).<sup>15,19</sup> The Average Variance Extracted (AVE) for a construct greater than 0.50 indicated construct's convergent validity,<sup>15,19</sup> while the square root of the AVE of construct greater than its correlation coefficients with other constructs would indicate discriminant validity.<sup>15</sup> Similarly, chi-square values  $\leq 3$  and p-value greater than 0.05, RMSEA and SRMR  $< 0.08$  and  $\leq 0.09$  would indicate the fitness of the model.<sup>15,20</sup> The internal consistency reliability of the PKAPIUM scale was assessed using the Cronbach's alpha coefficient ( $\alpha$ ) and values of between 0.6 and 0.79 were acceptable, with values falling between 0.8 –  $< 0.9$  as more preferred and higher values of  $> 0.9$  were indication of possible items' redundancy.<sup>15,20</sup> Similarly, Rasch measurement analysis was conducted via Bond and Fox' software to complement the CTT approach.<sup>14,21,22</sup> Mean Square (MNSQ) infit and outfit values range of between 0.6 and 1.5 were recommended limits, Point-Measure Correlation Coefficients (PTMEA Corr) value range of 0.3 to 0.7 and Standardized Z-score (ZStd) values of  $\pm 2$  were also acceptable ranges for validity assessment of survey scales.<sup>14,23</sup> Items whose at least two of its fit index values (MNSQ infit and outfit, Zstds and PTMEA Corr) were above or below the estimated limits<sup>14,23</sup> and were also marked for deletion after EFA, were tagged as having fewer possibilities of contributing in the same constructs compared to the other items in the instrument,<sup>14</sup> hence, they were identified as not fitting the model and marked for removal and the Rasch analysis process was continued until there was no further improvement in fit requirements.

Items' and persons' separation and reliability parameters were evaluated by interpretation of the Rasch model analysis outputs based on the recommended limits of persons and items separation index values of  $\geq 1$  as

acceptable and  $\geq 2$  as good; and reliability values of 0.8 as good,  $> 0.6$  and  $< 0.8$  as fair and acceptable and values  $< 0.6$  to be rejected.<sup>14,23</sup>

## RESULTS

Based on the observations made by the panel of experts, compound and complex statements were reframed into simple formats capturing basic ideas to ensure easy reading, understanding and comprehension by the prospective respondents. All the 31-items had I-CVI values greater than 0.8 for relevance, clarity, simplicity and comprehensiveness. The Universal Agreement (UA) level (S-CVI/UA) were 0.97 (Relevance), 1 (Clarity), 0.9 (Simplicity) and 0.9 (Comprehensiveness) and these implied 97% of relevance, 100% clarity, 90% simplicity and comprehensiveness levels of percentage agreement among the panel of experts on the acceptability of the content of the draft PKAPIUM scale. Similarly, the results of the average CVI (S-CVI/Ave) approaches were 0.99 (Relevance), 1 (Clarity), 0.98 (Simplicity) and 0.98 (Comprehensiveness), respectively.

The 300 patients attempted all the items (100%) in the instrument, indicating the acceptability of the draft scale and none of the items had more than 9% of the patients selecting the highest or lowest possible options which showed absence of floor and ceiling effects. Furthermore, the percentage of missing data were less than 10%, hence, all the 31 items were retained for further analysis.

There were weak inter-item correlations between some of the items with others, with respective relationships were statistically significant ( $r \geq 0.3$ ) between an item with at least two others in the scale and the existence of these relationships were further confirmed by the Bartlett's test of sphericity (Chi-square = 1968.82, df = 210,  $p = 0.000$  for PKAPIUM scale).<sup>15,24</sup> The absence of multicollinearity was seen, as none of the items' correlation strength was  $\geq 0.8$  and this was confirmed by the high determinant scores of the R-matrix of 0.001, which was above 0.00001 according to the rule of thumb.<sup>15,24</sup> Similarly, the overall Cronbach's alpha coefficient for the draft scale was 0.59, with corrected item-total correlations of between 0.01 – 0.5.

The KMO value for the sample population of 300 patients was 0.76. Ten items with low loading strengths ( $< 0.3$ ) and communalities of  $< 0.5$  were excluded in the initial stage of the CTT approach (Table 1). The second EFA resulted in 21 items arranged under their respective constructs (knowledge = 7, attitudes = 10 and practice = 4 items) (Table 1).

Table 2 indicated the validity of the new scale which the Average Variance Extracted (AVE) for all the constructs were  $> 0.5$  and the square root of the AVE of each construct greater than its correlation coefficients with other constructs, indicating convergent and discriminant validity. This was confirmed by the goodness-of-fits results which showed the chi-square value for the 3-factors' scale as 3.1, SRMR of  $\leq 0.09$  and RMSEA = 0.09 (Table 2). Similarly, the Cronbach's alpha reliability coefficients of the scale were 0.59 before and 0.74 after deletion.

The outcome of Rasch analysis of the draft scale showed 9 misfitting items (Items 1, 9, 10, 12, 23, 24, 25, 26 and 31) and were deleted reducing the scale to 22 items (Table 3). The subsequent analysis resulted in total 21 items with majority of the fit index values within the suggested limits described by Bond and Fox and Linacre.<sup>14,23</sup> Few items with at most one misfit index value outside the suggested ranges were also retained based on the minimum standard set for items to be deleted in the present study (Table 3).

The items reliability and separation value for the three constructs of the 31-items' scale were all  $> 0.8$ , while those for persons' indicators were on the lower side ( $< 0.8$ ) (Table 4). The re-run analysis on the reduced scale resulted in slight decreased in the items reliability and separation values for attitude and practice constructs compared to values in the 31-items'

**Table 1: Factor Structures and Loading of 31-Items in the Draft PKAPIUM Scale before Deletion and 21-Items after Deletion of Misfitting Items (n=300).**

Item No	31-items			21-items		
	Factors			Factors		
	1	2	3	1	2	3
1						
2		0.7			0.76	
3		0.77			0.81	
4		0.7			0.77	
5		0.63			0.65	
6		0.49			0.39	
7		0.41			0.3	
8		0.3			0.5	
9						
10						
11	0.47			0.45		
12						
13	0.55			0.54		
14	0.6			0.6		
15	0.56			0.58		
16	0.6			0.62		
17	0.56			0.6		
18	0.67			0.69		
19	0.69			0.72		
20	0.74			0.78		
21	0.7			0.73		
22			0.4			0.32
23						
24						
25						
26						
27			0.76			0.84
28			0.77			0.84
29	- 0.32		0.53			0.32
30						
31						

**Table 2: Confirmatory Factor Analysis (CFA) and Reliability of PKAPIUM by Classical Test Theory (CTT) (n=300).**

Factors	Convergent and discriminant validity				Cronbach's alpha (α) for 31-items (before deletion)		Cronbach's alpha (α) for 21-items (after deletion)		
	*AVE	1	2	3	PKAPIUM	Number of items	A	Number of items	α
Attitudes	0.58	<b>0.76</b>				13	0.77	10	0.83
Practices	0.61	0.52	<b>0.78</b>			10	0.43	4	0.73
Knowledge	0.66	0.49	- 0.25	<b>0.81</b>		8	0.62	7	0.61
PKAPIUM scale	0.61				<b>0.78</b>	31	0.59	21	0.74

\*AVE = Average variance extracted, Bold diagonal = The square root of AVE, Off-diagonal = Correlations between constructs.

Chi-square (CMIN/DF) = 3.10 at p = 0.000, RMSEA = 0.09, SRMR = 0.08.

**Table 3: Items Fit and Misfit Indices for Draft PKAPIUM Scale (n=300).**

Item No	31-items' draft PKAPIUM scale					22-items' PKAPIUM scale					21-items' PKAPIUM scale (practice construct containing 4 items)				
	Infit MNSQ	Outfit MNSQ	Infit Zstd	Outfit Zstd	PTMEA	Infit MNSQ	Outfit MNSQ	Infit Zstd	Outfit Zstd	PTMEA	Infit MNSQ	Outfit MNSQ	Infit Zstd	Outfit std	PTMEA
Knowledge															
1	1.05	<b>1.55</b>	0.7	<b>2.6</b>	0.63										
2	0.98	0.79	0	- 0.2	0.46	1.05	1.2	0.3	0.5	0.42	1.05	1.20	0.3	0.5	0.42
3	1.12	0.89	0.5	0	0.47	1.11	0.65	0.5	- 1.2	0.45	1.11	0.65	0.5	- 1.2	0.45
4	1.2	0.91	0.8	0	0.43	1.20	0.74	0.8	- 0.3	0.42	1.2	0.74	0.8	- 0.3	0.42
5	1.48	0.78	1.8	- 0.3	0.43	1.47	1.47	1.9	0.2	0.41	1.47	1.47	1.9	0.2	0.41
6	0.92	0.81	- 1	- 1.7	<b>0.73</b>	1.01	<b>2.8</b>	0.2	1.9	0.7	1.01	<b>2.8</b>	0.2	1.9	0.7
7	0.78	0.88	- 2	- 0.9	0.64	0.66	<b>2.9</b>	- 2	1.9	0.69	0.66	<b>2.9</b>	- 2	1.9	0.69
8	1.06	1.47	0.8	<b>2.1</b>	0.47	1.09	<b>3.9</b>	1.2	2	0.55	1.09	<b>3.9</b>	1.2	2	0.55
Attitudes															
9	0.56	0.68	- <b>4.9</b>	- 3.3	0.21										
10	<b>1.62</b>	<b>1.74</b>	<b>4.9</b>	<b>5.7</b>	- <b>0.17</b>										
11	0.77	0.93	- <b>2.1</b>	- 0.9	0.51	0.9	1.41	- 1.3	<b>3.8</b>	0.5	0.9	1.41	- 1.3	<b>3.8</b>	0.5
12	1.14	<b>1.64</b>	<b>2.5</b>	0.3	0.46										
13	1.27	1.27	2	2.1	0.58	1.38	1.4	1.9	1.7	0.56	1.38	1.4	1.9	1.7	0.56
14	0.88	0.77	- 1.3	- 2	0.62	0.97	0.85	- 0.3	- 1.3	0.56	0.97	0.85	- 0.3	- 1.3	0.56
15	1.14	1	1.4	0.1	0.57	1.19	0.92	1.7	- 0.6	0.57	1.19	0.92	1.7	- 0.6	0.57
16	0.93	0.69	- 0.6	- 1.8	0.61	0.99	0.65	0	- <b>2.9</b>	0.57	0.99	0.65	0	- <b>2.9</b>	0.57
17	1.34	1.03	<b>2.1</b>	0.3	0.6	1.31	0.93	1.7	- 0.5	0.6	1.31	0.93	1.7	- 0.5	0.6
18	1.12	1.04	1.2	0.3	0.67	1.15	1.02	1.3	0.2	0.63	1.15	1.02	1.3	0.2	0.63
19	0.84	0.62	- 1.6	- 1.5	0.64	0.85	0.58	- 1.4	- <b>2.2</b>	0.6	0.85	0.58	- 1.4	- <b>2.2</b>	0.6
20	0.78	0.6	- <b>2.1</b>	- 0.7	0.67	0.77	0.64	- 2	- 2	0.63	0.77	0.64	- 2	- 2	0.63
21	0.82	0.6	- 1.8	- 1.7	0.65	0.86	0.59	- 1.2	- <b>2.5</b>	0.6	0.86	0.59	- 1.2	- <b>2.5</b>	0.6
Practices															
22	1.28	1.27	2	2.7	0.51	1.26	1.39	1.8	<b>3.8</b>	0.61	1.38	1.41	1.9	<b>2.7</b>	0.66
23	1.45	1.54	4.6	5.2	0.36										
24	1.11	1.72	1.3	2.5	0.26										
25	<b>1.84</b>	0.87	- <b>2.6</b>	- 1.7	<b>0.22</b>										
26	0.84	0.8	- <b>2.8</b>	- <b>3.2</b>	0.62										
27	0.67	0.85	- <b>2.2</b>	- 1.5	0.61	0.55	0.61	- 2	- <b>3</b>	0.7	0.85	0.9	- 1.7	- <b>2.8</b>	0.7
28	0.66	0.56	- 1.6	- 0.5	0.54	0.65	0.61	- 1.7	- <b>3.9</b>	0.69	0.63	0.65	- 1.9	- <b>2.6</b>	0.67
29	1.1	1.04	1.1	0.4	0.63	1.5	1.48	1.6	<b>4</b>	0.63	1.36	1.1	1.1	<b>3.8</b>	0.68
30	1.08	1.07	0.9	0.9	0.3	1.06	<b>1.55</b>	1.1	1.8	<b>0.29</b>					
31	1.21	1.21	<b>2.5</b>	<b>2.5</b>	0.39										

draft scale, with no change in the values for knowledge construct (Table 4).

## DISCUSSION

The purpose of this study was to develop and validate a scale for assessing patients' knowledge, attitudes and practices on uncomplicated malaria management using classical test and modern test theory. Although the conventional method of analysis has been in used for a long time for instrument validation, the model has some limitations including having only the test instrument as its basis and the outcomes of the analysis varies with variation in the sample population due to differences in re-

spondents' levels of understanding. In addition, IRT gives more details on both the items difficulties and respondents' abilities<sup>17</sup> and this form the basis of complementing the validation with Rasch analysis which was in agreement with reports of similar previous studies.<sup>24,25</sup>

Face and content validity had been considered as the first approach in instrument validation, because it revealed the appropriateness of the draft scale by detection and exclusion of items with the potentials of decreasing the quality of the instruments.<sup>17</sup> The outcome of the experts' viewpoints in the present study was an indication that all the retained items for further validation were necessary and relevant for assessing patients' KAP on uncomplicated malaria and this was consistent with result of

**Table 4: Reliability and Separation Index Values Before and After Deletion of Misfit Items (n=300).**

	31 items before deletion					22 and 21 items after deletion				
		Item measure		Person measure			Item measure		Person measure	
Construct	n	Rel	Sep	Rel	Sep	n	Rel	Sep	Rel	Sep
Knowledge	8	0.99	12.1	0.7	1.52	7	0.99	12.13	0.60	1.22
Attitudes	13	0.99	12.59	0.74	1.69	10	0.95	4.19	0.79	1.97
Practices	10	0.99	13.31	0.54	1.08	5	0.95	4.44	0.60	1.23
Practices						4	0.98	6.52	0.66	1.39

Note: n = Number of items; Rel = Reliability; Sep = Separation.

similar study conducted by Rodrigues *et al.*<sup>26</sup> The observed acceptability of the draft scale by patients during the preliminary stage of exploratory factor analysis might imply that the questions were neither too easy nor too hard for them and this might be good signals of the draft instruments' quality.<sup>15,17</sup> The observed weak correlation coefficients between some of the items might be due to possibility of the software treating all the 31 items as single dimension instruments rather than considering them as multi-dimensional during the computation,<sup>15</sup> therefore, to avoid elimination of potentially important items, all the 31 items were retained for EFA. Regarding sample size, the KMO value of 0.76 indicated the adequacy of the 300-sample size of patients for factor analysis on the 31 items' -draft scale base on minimum acceptable factor loading strength of 0.3 and was also considered adequate for Rasch analysis having in mind both the advantages and disadvantages involved.<sup>15</sup> Due to variations in acceptable sample size adequacy for conducting both CTT and Rasch analysis, various absolute and simulated sample size values have been reported to be adequate, including values as low as 30 to 50 when the factor loading is  $> 0.8$ .<sup>18</sup> Some studies have reported that the use of such low sample size, especially for Rasch measurement analysis might yield good outputs relating to person's reliability and separation index values because there might be fewer lapses made by the respondents compared to larger populations.<sup>14</sup> On the contrary, it has been documented that higher sample size might result in improved items' reliability when assessed through both factor and Rasch analyses.<sup>14,27</sup> The overall outcome of this dual validation approach resulted in a valid and reliable 21-items' PKAPIUM scale for assessing patients' KAP on uncomplicated malaria.<sup>14,28</sup> The CTT results indicated validity of the instrument based on the convergent and discriminant validity and goodness-of-fit test results,<sup>19,20,29</sup> in addition, the Cronbach's alpha reliability of 0.74 (Table 2) to be within accepted limits.<sup>15</sup> Similar observations were made with the outcome of Rasch analysis, with less data departure from the Rasch measurement model because majority of the items' fit index values were either within the suggested range limits or only one of the parameters been outside the expected limits (Table 3), which did not qualify most of the items for deletion. The observed variations in the MNSQ outfit indices could be a proof for the unweighted nature of the outfit index parameter and its high sensitivity to extreme responses such as guessing in responding to items considered to be tough to respondents or in some cases, their under narration of simple items thereby giving them less attention.<sup>14</sup>

Also, the observed decreased in items reliabilities after deletion of misfit items might not be connected with sample size since there were no changes in the number of study populations. It might be possible that such decrease might be due to the deletion of some of the items that would have possibly increased the scales' precision and also decreasing its standard errors of measurement.<sup>14,26</sup> On the contrary, the deletion of misfit items resulted in an increase in the person reliability and separation values for attitude and practice and a decrease in knowledge con-

struct values (Table 4), although, they were still within the acceptable limits.<sup>14,26</sup> This implied that such reduction resulted in better levels of consistency in the behaviors of the respondents when responding to the items, thereby improving the instruments' capacity to better distinguish them into low or high-performance' abilities in answering the attitude and practice-related questions.<sup>14</sup> While the decrease in person reliability and separation index for knowledge construct could be possible that some of the retained items were too difficult or not very clear for the respondents to appropriately respond to them, which might reduce the scales' capacity to better distinguish the population into low or high-performance' abilities base on their response patterns to the knowledge-related items.<sup>14</sup> Apart from gaining deeper understanding on psychometric evaluation methods, the use of this orchestrated validation approach (CTT and IRT) have revealed additional information on the instrument which were not available through CTT, such as persons' abilities and levels of difficulties of the items which might be useful for future improvement in the quality of the instrument and also for research purpose.

However, the study had some limitations, which included the fact that the PKAPIUM scale was developed for assessing patients' KAP on uncomplicated malaria disease and its management in PHC facilities of Plateau state, therefore implied that it cannot be used in assessing patients' treatment practices for severe malaria and other disease conditions. In addition, its applicability in secondary and tertiary health facilities would be feasible with further revalidation using data from patients in such health facilities. This study was carried out in Plateau state, therefore for the scale to be generalized, the recent study proposed to carry out the same survey across the PHC facilities in the country and beyond.

## CONCLUSION

Outcome of this study have shown the importance of using CTT and IRT in reducing misfit items to arrive at a reliable and valid PKAPIUM scale that is useful for assessing patients' KAP on uncomplicated malaria disease and its management in PHC settings. In addition, the Rasch model revealed the detailed characteristics of both the items and the respondents through index values that indicated the levels of difficulties of the items and the respondents' abilities to respond to the items, as compared to the CTT studies which only concentrated on the test items.

## ACKNOWLEDGEMENT

The authors are grateful to all the staffs of PHC facilities for their cooperation, Dr Maxwell LP Dapar, Pharmacist Joel Bulus Kaben and Micah Yunana Jingina who assisted during the field work.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ABBREVIATIONS

**AMOS™**: Analysis of moment structures; **AVE**: Average variance extracted; **CFA**: Confirmatory factor analysis; **CMIN/DF**: Chi-square minimum discrepancy divided by its degrees of freedom; **CTT**: Classical test theory; **df**: Degree of freedom; **EFA**: Exploratory factor analysis; **I-CVI**: Item's content validity index; **Infit Zstds**: Infit standardized Z-scores; **IRT**: Items response theory; **KAP**: Knowledge, attitudes and practices; **KMO**: Kaiser-meyer-olkin; **LGAs**: Local government areas; **MMTT**: Modern mental test theory; **MNSQ**: Mean square; **PHC**: Primary health care; **PKAPIUM**: Patients' knowledge, attitudes and practices instrument for uncomplicated malaria; **PTMEA Corr**: Point-measure correlation coefficients; **RMSEA**: Root mean square error approximation; **RMC/EC**: Research Management Centre/ Ethics Committee; **S-CVI**: Scale content validity index; **SPSS®**: Statistical Package for Social Sciences; **SRMR**: Standardized root mean square residual; **UA**: Universal agreement; **r**: Pearson correlation coefficients; **α**: Cronbach's alpha coefficient.

## REFERENCES

- Ahmad A, Sann LM, Rahman HA. Factors associated with knowledge, attitude and practice related to hepatitis B and C among international students of Universiti Putra Malaysia. *BMC Public Health*. 2016;16(1):611. doi:10.1186/s12889-016-3188-5.
- Eke RA, Diwe KC, Chineke HN, Duru CB, Uwakwe KA. Evaluation of the practice of self-medication among undergraduates of Imo State University (IMSU) Owerri, South-East Nigeria. *Orient J Med*. 2014;26(3-4):79-83.
- Liu J, Isiguzo C, Sieverding M. Differences in malaria care seeking and dispensing outcomes for adults and children attending drug vendors in Nasarawa, Nigeria. *Trop Med Int Health*. 2015;20(8):1081-92. http://doi.org/10.1111/tmi.12520.
- Mwenesi HA. Sociocultural and behavioural issues in the treatment and prevention of malaria. Working group on malaria, Geneva. 2003 [Cited 2017 May 23]. Available from: [http://www.tropika.net/review/030324-Malaria\\_4/article.pdf](http://www.tropika.net/review/030324-Malaria_4/article.pdf).
- Gillam S. Ethnicity and the use of health services. *Postgrad Med J*. 1990;66(782):989-93.
- Okafor EE, Amzat J. Problems of malaria menace and behavioural intervention for its management in sub-Saharan Africa. *J Human Eco*. 2007;21(2):155-62.
- Yadav SP, Yadav S, Kuma P, Yadav S. Knowledge, treatment-seeking behaviour and socio-economic impact of malaria in the desert of Rajasthan, India. *South Afr J Epidemiol Infect*. 2013;28(1):41-7.
- Adetola OT, Aishat LL, Olusola O. Perception and treatment practices of malaria among tertiary institution students in Oyo and Osun States, Nigeria. *J Nat Sc Res*. 2014;4(5):30-3.
- Edet-Utan O, Ojediran T, Usman SO, Akitayo-Usman NO, Fadero T, Oluberu OA, et al. Knowledge, perception and practice of malaria management among non-medical students of higher institutions in Osun State Nigeria. *Am J Biotech Med Res*. 2016;1(1):5-9. doi: 10.5455/ajbmr.20151226010533.
- Jimam NS, David S, Musa N, Kadir GA. Assessment of the knowledge and patterns of malaria management among the residents of Jos metropolies. *W J Pharm Pharm Sci*. 2015;4(6):1686-98.
- Gie YA, Pearce SA. Beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutor Quant Methods Psycho*. 2013;9(2):79-94.
- Thompson NA. Ability estimation with item response theory (Assessment systems corporation). 2009. Available from: [www.assess.com](http://www.assess.com).
- Wirth RJ, Edwards MC. Item factor analysis: current approaches and future directions. *Psychol Methods*. 2007;12(1):58-79.
- Bond TG, Fox CM. Applying the Rasch model: Fundamental measurement in the human sciences, 3<sup>rd</sup> ed. New York, Lawrence Erlbaum Associates, Inc. 2015.
- DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007;39(2):155-64.
- Hogarty KY, Hines CV, Kromrey JD, Ferron JM, Mumford KR. The quality of factor solutions in exploratory factor analysis: the influence of sample size, communality and over determination. *Edu Psychol Meas*. 2005;65(2):202. https://doi.org/10.1177/0013164404267287.
- Hamad EO, Savundranayagam MY, Holmes JD, Kinsella EA, Johnson AM. Toward a mixed-methods research approach to content analysis in the digital age: The combined content-analysis model and its applications to health care twitter feeds. *J Med Internet Res*. 2016;18(3):e60. http://doi.org/10.2196/jmir.5391.
- Ismail R, Nor MZM, Ismail NH, Tamil AM. A tale of two constructs validation analysis: Rasch model and exploratory factor analysis approach for three-factor eating questionnaire (TFEQ-R21) among Malaysian male workers. *Med J Malaysia*. 2015;70(3):169-76.
- Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res*. 1981;18(1):39-50.
- Kline RB. Principles and practice of structural equation modeling. 4<sup>th</sup> ed. New York: Guilford Publications. 2015.
- Cappelleri JC, Lundy JJ, Hays RD. Overview of classical test theory and item response theory for quantitative assessment of items in developing patient-reported outcome measures. *Clin Ther*. 2014;36(5):648-62. doi:10.1016/j.clinthera.2014.04.006.
- Linacre JM. Winsteps Rasch measurement computer program User's guide, Beaverton. 2012. Available from: <http://www.winsteps.com>.
- Tabachnick BG, Fidell LS. Using multivariate statistics (5<sup>th</sup> ed.). Boston, MA: Allyn and Bacon. 2007.
- Janssen G, Meier V, Trace J. Classical test theory and item response theory: Two understandings of one high-stakes performance exam. *Colombian Applied Linguist J*. 2014;16(2):167-84.
- Rodrigues IB, Adachi JD, Beattie KA, MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. *BMC Musculoskeletal Disorders*. 2017;18(1):540. doi: 10.1186/s12891-017-1914-1915.
- Kjellström S, Golinob H, Hamer R, Rossum EJV, Almers E. Psychometric properties of the Epistemological Development in Teaching Learning Questionnaire (EDTLQ): An inventory to measure higher order epistemological development. *Frontline Learn Res*. 2016;4(5):1-33. doi: <http://dx.doi.org/10.14786/flr.v4i5.239>.
- Mohamad MM, Sulaiman NL, Sern LC, Saleh KM. Measuring the validity and reliability of research instruments. *Procedia-Soc Behav Sci*. 2015;204:164-71. http://doi.org/10.1016/j.sbspro.2015.08.129.
- Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci*. 2015;43(1):115-35.

**Article History:** Submission Date : 15-10-2018; Revised Date : 03-01-2019; Acceptance Date : 14-02-2019.

**Cite this article:** Jimam NS, Ahmad S, Ismail NE. Psychometric Classical Theory Test and Item Response Theory Validation of Patients' Knowledge, Attitudes and Practices of Uncomplicated Malaria Instrument. *J Young Pharm*. 2019;11(2):186-91.