Development, validity and repeatability of mothers' knowledge, attitude and practice of a Universal Newborn Hearing Screening measurement tool

1. Christine Graham*, Department of ENT, Nelson Mandela School of Medicine, University of KwaZulu-Natal, 719 Umbilo Road, Berea 4001. Durban, South Africa.
2. Ayanda Gina, Department of Audiology, School of Health Sciences
3. Janet Seeley, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT. UK
4. Yougan Saman, Ear, Nose and Throat Department, University Hospitals of Leicester, Groby Road, Leicester. LE3 9QP. UK

*Corresponding author: Email: cgtinagrm333@gmail.com

ABSTRACT

Background: The assessment of the validity and reliability of measurement tools in research provides quality data. However, evidence of the validity and reliability of parental knowledge and attitude regarding childhood hearing loss and newborn hearing screening is scarce.

Objective: To design a Knowledge, Attitudes and Practice (KAP) survey tool regarding childhood hearing loss and a Universal Newborn Hearing Screening Programme of the rural Amajuba district, KwaZulu-Natal, South Africa, and then test for validity and test-retest repeatability.

Methods: Face validity was conducted with 20 mothers and a content validity index was determined by two rounds of assessments, the first by 7 experts and the second by 3 experts. The kappa statistic was used to measure the stability of the tool using data from 160 mothers where repeated measurements were applied at two-week intervals. The feasibility of developing a tool was assessed by applying the criteria of science, population and resources.

Results: The KAP tool was developed with twenty-nine items. For face validity, 97% of the participants reported that the items were clear, wording was appropriate and easy to read and the language was natural. Content validity produced excellent results with a scale and content validity index of 1. Test-retest repeatability for the KAP tool was good with a Cohen's kappa coefficient of 0.87 (95% CI: 0.87, 0.87). Individually, the knowledge scale had a kappa of 0.86 (95% CI: 0.77, 0.95); the attitude scale had a kappa of 0.87 (95% CI: 0.76, 0.99): the practice scale had a kappa of 0.86 (95% CI: 0.75, 0.97) and the awareness scale had a kappa of 0.92 (0.83, 1.00). The development of a KAP tool was shown to be feasible, given sufficient time, funds, motivation and a study population.

Conclusion: The study produced a valid and reliable tool that can be useful in generating quality evidence of a community's KAP with respect to childhood hearing loss and newborn hearing screening. Evidence gathered could be used to tailor health education and health promotion material of for a Universal New-born Hearing Screening (UNHS) programme in a culturally sensitive manner to promote service uptake.

Keywords: newborn hearing screening, reliability, validity, KAP

1 Introduction

A Universal Newborn Hearing Screening Programme (UNHSP) is a public health initiative established for the prevention of childhood hearing loss (CHL). Permanent childhood hearing impairment (PCHI) is a significant cause of disability [1]. Endorsed by the World Health Organisation (WHO) for early hearing detection, UNHSP attempts to reduce the impact on the family and the child through accessibility of services and management of the condition [2]. However, the foremost challenge in the delivery of UNHSP is the diagnostic follow-up and effective compliance with the intervention as it requires a pragmatic partnership between the health service and families [3]. It is fair to say that the success of the programme depends on the full participation of UNHSP service users at the level of screening, follow-up, diagnostic procedures and further intervention [4].

It is believed that more effective UNHSPs will result from a better understanding of the wider context of the community's knowledge and perspective about ear health. The existing literature however, refers to the maternal knowledge and attitude to UNHSPs which has been obtained from well-established programmes that are part of mainstream health care services [5–11]. In developing countries research comes predominantly from urban areas, either in immunisation clinics or community settings [12–16]. The recognition that the perspectives of families and/or communities regarding CHL and NHS is important as it can produce evidence that can improve a child's hearing health outcomes through health promotion strategies [13, 17].

Nevertheless, in poorly resourced settings, with many competing health priorities, there must be good evidence that the program can deliver good compliance before policy makers will be prepared to invest [18]. To obtain good quality evidence there is frequently a need to determine the reliability and validity of measurement tools. In a quantitative approach, meeting this requirement will demonstrate the tool's stability for reliability and its ability to measure what it is supposed to measure for validity [19, 20]. Ideally, any new or adapted measurement tool if applied to a new population needs to indicate how the reliability and validity were established [19, 21].

It was therefore necessary to develop a tool that will be acceptable to a rural community. As a component of Amajuba, KwaZulu-Natal, South Africa, UNHSP research programme, the main goal of the current study was to assess whether the newly developed knowledge, attitude and practice (KAP) survey tool, regarding childhood hearing loss and UNHSP, is valid and reliable. The first objective of the current study was to design an

appropriate KAP tool, followed by the second objective of validating the content as reviewed by a panel of experts from several disciplines and face validity as evaluated by participants from the community of the study. The third objective measured the test-retest repeatability designed to assess the reliability of the KAP tool. We then demonstrated the achieved validity and repeatability of the KAP measurement tool.

2 Methods

2.1 *KAP survey tool construct development*

The questionnaire was developed in English using significant findings from our previous qualitative study [22] and also incorporated key theoretical aspects from the KAP literature [23]. The questions were designed to identify the key concepts with respect to CHL and UNHSP, as commonly shared by the community, to deepen our understanding of the issues [24, 25]. It was thus constructed according to the broader concepts of KAP which are based on the premise that we can measure the existing knowledge, perspectives and actions taken with respect to CHL and UNHS. This can then offer space to increase the provision of information that can change their current KAP and ultimately, their behaviour [26]. Hence, the questionnaire also included the concept of behaviour.

The initial draft tool was developed with twenty-five items. However, after content validation it was revised to twenty-nine items including two contingency items (filter questions), which reduced the numbered items to twenty-seven. These were then divided into four scale constructs and one demographic section as described below:

2.1.1 Demographic: 6 items

2.1.2 Knowledge scale: 6 items in total; 3 items have three response options (Yes/No/I don't know) and the other 3 have multiple response options.

2.1.3 Attitudes and Behaviour scale: 6 items in total: 1 item with yes/no response; 3 items have multiple responses and 1 has a rating scale (very seriously to not seriously) and 1 has one choice response from different statements.

2.1.4 Health care seeking (Practice) and Behaviour scale: 6 items in total: 3 items have multiple responses; 2 items have one choice from several statements and 1 item has a dichotomous response option.

2.1.5 Awareness scale: 5 items in total; 3 items have multiple response options; 1 item has a choice from several statements and 1 item has a dichotomous option.

2.2 *Participants and procedures*

In assessing validity and reliability the sampling was approached differently. The data was collected between November 2016 and March 2017.

2.2.1 *Content validity*

For content validity, the tool was evaluated by a team of seven experts from the disciplines of audiology, otorhinolaryngology and public health. The experts were first asked if they would like to participate in the study. After acceptance, a formal letter of invitation with the evaluation form was sent to the expert. These experts reviewed the questionnaire for comprehensiveness as well as relevancy of the scale's content and content domain [21, 27]. They came back with comments in relation to the wording and added two questions. However, after the content validity index (CVI) analysis was done it was found that the practice scale construct did not achieve the required CVI. Therefore, the questions were revised, with the input from the experts. Then a second team of three experts was invited to evaluate the relevancy of the questionnaire with regard to the scale's content and content domain.

2.2.2 *Translation of the tool*

Thereafter, we engaged a professional from the linguistic department at University of Kwa-Zulu Natal to translate the questionnaire into the Zulu language as this was the medium of communication used by most participants. The translated questionnaire was then taken back to the community where another expert translated it back into English. The principal investigator, working with two research assistants (recruited nurses for the larger study who are Zulu speakers from the same community), then reviewed each item for the appropriateness of spoken language in everyday settings as well as the structure of questionnaire.

2.2.3 *Face validity*

A face validity exercise was conducted by recruiting twenty participants from three ante-natal clinics (Madadeni 1, Stafford, Osizweni 3) at the study site. Convenience sampling was applied to identify seven, seven and six participants from each clinic respectively, chosen from those waiting for consultation. We asked the pregnant women selected if they would

like to review the questionnaire and participate in the study. Those who accepted were first given a consent form, then a review sheet and a questionnaire.

2.2.4 *Repeatability*

Repeatability is a test-retest reliability exercise which demonstrates the consistency of the measurement tool that has been administered at two or more points with short intervals between tests [23]. The questionnaire was thus used to conduct a test-retest repeatability study with a sample of 160 participants, recruited randomly from the Newcastle hospital ante-natal clinic. We first established the total number of pregnant mothers attending the clinic that day, enumerated them separately on pieces of paper before shuffling them in a box. We then picked twenty numbers at random and correlated these numbers with the names on the registered list of the day. Participants were first briefed about the purpose of the study and were then notified that they would be required to repeat the same exercise after two weeks. Thereafter, they were asked for written consent and to self-administer the questionnaire independently without discussing with anybody. However, some participants were accompanied by their mother, sister etc. and we felt it acceptable to allow them to complete the questionnaire together. Normally, repeatability requires participants to repeat the same exercise at a later date. Data collection took 6 weeks, with the first test data collected over a two week period in early-mid February 2017. Refreshments and snacks were given to the participants as a token of appreciation. Thereafter, we had an interval of two weeks before the re-test data was collected in early-mid March 2017. During the second phase of data collection we devised a mechanism to encourage participants to come to the hospital for the study arranging four time slots every day, over a two week period. Since we had participants' contact details, we called each participant and asked them to choose a day and a time slot when they would be available for the study. This was an arduous task that required persistent phone calls and follow-ups for those who did not turn up on the first call. We used transport subsidies, refreshments and snacks to encourage them to come.

2.3 ***Data analysis***

2.3.1 *Content validity and face validity*

Content Validity addresses issues of the comprehensiveness and representativeness of the content domain. Experts rated the items as follows: 1- not relevant; 2 – somewhat relevant; 3 – relevant and 4 – very relevant. To analyse the data, we used the CVI measurement. CVI

refers to the extent to which an instrument covers the content it is supposed to measure [28]. This measurement provides two results: Item Content Validity Index (I-CVI) which measures the efficacy of the item and Scale Content Validity Index average (S-CVI/Ave) which measures the efficacy of the scale. The criteria used for analysis is that of I-CVI of .78 and S-CVI/Ave of .90 or higher for 6 to 10 experts and I-CVI of 1 for 3 to 5 experts [21]. By using Excel, the I-CVI was calculated as the number of experts who rated relevant or very relevant divided by the total number of experts. Whereas, the S-CVI was calculated by averaging the proportion of items 3 and 4 amongst experts.

For face validity, we used descriptive analysis, where participants evaluated (1-strongly disagree; 2 – disagree; 3 – agree and 4 – disagree) the tool with respect to clarity, wording, readability, layout and language [29].

2.3.2 *Repeatability*

To analyse the data for test-retest repeatability, SPSS version 24 was used. Repeatability is a measure of reliability and since the measurement scales were nominal, Cohen's Kappa was considered as an appropriate statistic [30, 31]. Kappa is a measure which determines the amount of agreement between measurements that is greater than the amount expected by chance alone. Kappa allowed us to calculate observed agreement between the two measurements and adjust for agreement expected by chance then normalise the values to create a coefficient from -1 to 1. The negative value demonstrates that the observed agreement is less than that expected by chance and when the value is 0 the observed agreement can be justified by chance and when it is 1 there is a perfect agreement. As suggested by Landis and Koch, Cohen's Kappa strength of agreement will be interpreted as follows ≤ 0 as poor; 0.00 – 020 as slight; 0.21 – 0.40 - as fair; 0.41- 0.60 – as moderate; 0.61 – 0.80 as substantial and 0.81 – 1.00 as almost perfect agreement [32]. Any kappa that is below 0.60 implies insufficient agreement, suggesting slight confidence in the study results. For all items that had multiple responses we used dichotomous options by scoring 'yes' for one and 'no' for zero and calculated the Cohen's kappa. We analysed each variable separately and then a pooled kappa for the item. For the rest of the items we calculated kappa to the item directly. A pooled kappa was also used for knowledge, attitude and practice as they are individual constructs and were later used for the full instrument. A pooled kappa is the averaging of all observed agreements and of all the expected agreements which were then set into the kappa formula [33]. A standard error was also computed for each item, scale

construct and the total instrument which allowed us to understand the degree of uncertainty in the kappa estimate results. This gave meaning to the kappa by providing 95% confidence intervals.

2.3.3 *The feasibility of developing a KAP survey tool*

To assess the feasibility of developing the KAP survey tool we followed the guidelines of the 'research study feasibility tool' which focusses on three assessment criteria, science, population and resources [34]. We then selected components that were applicable to the assessment of the tool, as follows:

Science: Whether 1) the tool will make a contribution to the existing body of knowledge, 2) the research team was motivated during the process of developing the tool and 3) the procedures of developing the tool were realistic

Population: Whether 4) it was easy to access the study population during the process and 5) the incentives for participants were sufficient

Resources: Whether 6) time was sufficient for the whole process of developing the tool – designing, data collection, capturing and analysing and 7) funds were sufficient and did not delay the study.

2.4 Ethics

Ethical approval was obtained from the Biomedical Research Ethical Committee (BREC) University of KwaZulu-Natal - No. BFC261/16 (sub-study of BFC421/15). Voluntary informed written consent was obtained for participation which included maintaining confidentiality and anonymity within possible bounds.

3 Results

3.1 *Content validity*

The results of the second stage of rating all items were rated very relevant by the experts with a total agreement on all 23 items, resulting in an I-CVI of 1. All scale constructs had a S-CVI/Ave of 1, demonstrating that the measurement tool was valid in content.

3.2 *Face Validity*

The rating for the assessment of face validity was established at one to four and all participants rated the questionnaire three or four. Ninety percent indicated that the instructions were clear and understandable. Ninety five percent indicated that the wording

was appropriate and that the readability was suitable. All of them indicated that the questions were easy to answer, the language natural and the layout was good.

3.3 *Test-retest Repeatability*

The repeatability study consisted of 160 participants, all of whom were expectant mothers from the ante-natal clinic. The demographic characteristics of the participants are shown in Table 1.

Table 1: Characteristics of participants and descriptive analysis (N=160)

Characteristics		**Frequency**	**Percentage**
Age	18 – 20	27	17
	21 – 30	80	50
	31 – 40	44	27
	Over 40 years	9	6
Marital Status	Married	22	14
	Single	135	84
	Living with a partner	3	2
Religious Belief	Muslim	1	1
	Christian	103	64
	Hindu	2	1
	African Ancestral	54	34
Level of Completed education	No school	4	2
	Primary	31	19
	High school	81	51
	College	36	23
	Higher Education (University)	8	5
Current employment status	Employed	27	17
	Unemployed	108	67
	Student	25	16

The reliability of the questionnaire was determined by test-retest repeatability. Item-specific results are shown in each scale construct, with item numbering according to the numbering in the questionnaire as follows:

3.3.1 *Knowledge*

The kappa values for 5 items show almost perfect agreement, which indicates the clear structure of the items. One item, however, indicated only a substantial level of agreement, demonstrating an inconsistency by the participants in the two point assessment when compared to other items of knowledge.

Table 2: Cohen's Kappa Coefficient of the Knowledge scale construct (N=160)

		Test-Retest Repeatability

Item	Assessment Criterion	Level of agreement	Cohen's Kappa Coefficient	Standard Error - Kappa	95% Confidence Interval (Lower limit – Upper Limit)
7	Baby born with hearing loss	Almost perfect	0.86	0.035	(0.79 – 0.93)
8	Causes of hearing loss	Almost perfect	0.89*	0.045	(0.80 – 0.98)
9	Detection in a newborn	Substantial	0.78	0.044	(0.69 – 0.86)
10	Develop HL after passing test	Almost perfect	0.89	0.031	(0.83 – 0.95)
11	Identifying a child with HL	Almost perfect	0.81*	0.050	(0.71 – 0.91)
12	Treatment for a child with HL	Almost perfect	0.86*	0.056	(0.75 – 0.97)

*pooled kappa

3.3.2 *Attitude and Behaviour*

The results of the kappa in the attitude scale reveal an almost perfect agreement. However, the CI width (margin of error – 0.36) of item 13a is so wide that it shows a large disagreement at the lower limit, even though the overall level of agreement shows reliability.

Table 3: Cohen's Kappa Coefficient of the Attitude scale construct (N=160)

Item	Assessment Criterion	Test-Retest Repeatability			
		Level of agreement	Cohen's Kappa Coefficient	Standard Error - Kappa	95% Confidence Interval (Lower limit – Upper Limit)
13	Screening acceptance	Almost perfect	1.00	0.00	(1.00 – 1.00)
13a	Reasons for not accepting screening	Almost perfect	0.83*	0.186	(0.47 – 1.00)
14	Reaction if baby found with HL	Almost perfect	0.83*	0.066	(0.70 – 0.96)
15	Gravity of HL impact on family and community	Almost perfect	0.97	0.018	(0.93 – 1.00)
16	Descriptions of the impact of HL	Almost perfect	0.84*	0.058	(0.73 – 0.96)
17	Community attitudes to deaf people	Almost perfect	0.98	0.013	(0.96 – 1.00)

*pooled kappa

3.3.3 *Practice (Health care seeking and behaviour)*

The level of agreement in the items of the practice scale construct varied from moderate to substantial as shown in Table 4 below. Please note item 20 where the level of agreement was moderate with the lowest kappa indicating that about half of the participant's responses disagreed with respect to the acceptance of further examination.

Table 4: Cohen's Kappa Coefficient of the Practice scale construct (N=160)

		Test-Retest Repeatability

Item	Assessment Criterion	Level of agreement	Cohen's Kappa Coefficient	Standard Error - Kappa	95% Confidence Interval (Lower limit – Upper Limit)
18	Habitual health seeking behaviour	Substantial	0.79*	0.079	(0.64 – 0.95)
19	Action taken if child identified with HL	Almost perfect	0.87*	0.075	(0.72 – 1.00)
20	Acceptance of further examination	Moderate	0.50	0.277	(-0.04 – 1.00)
20a	Willingness to take a child for further examination	Almost perfect	0.97	0.019	(0.93 – 1.00)
21	Challenges that may hinder frequent visits to a health facility	Almost perfect	0.95*	0.033	(0.88 – 1.00)
22	Usual support when women take a child to the health facility	Substantial	0.66	0.051	(0.56 – 0.76)

*pooled kappa

3.3.4 *Awareness of childhood hearing loss and newborn hearing screening*

In terms of the awareness scale construct, five items were assessed by Cohen's kappa (Table 5).

Table 5: Cohen's Kappa Coefficient of the Awareness scale construct (N=160)

Item	Assessment Criterion	Test-Retest Repeatability			
		Level of agreement	Cohen's Kappa Coefficient	Standard Error - Kappa	95% Confidence Interval (Lower limit – Upper Limit)
23	First heard about newborn hearing screening	Substantial	0.73	0.057	(0.62 – 0.85)
24	Whether well informed about NHS programme	Substantial	0.79	0.061	(0.67 – 0.91)
25	Current places to get health information	Almost perfect	0.91*	0.061	(0.79 – 1.00)
26	Information they would like to get if a child is at risk of HL	Almost perfect	0.96*	0.026	(0.91 – 1.00)
27	Effective sources of information that can reach the community regarding NHS programme	Almost perfect	0.96*	0.029	(0.91 – 1.00)

*pooled kappa

3.3.5 *Knowledge, Attitude, Practice and Awareness*

A pooled Cohen's kappa was assessed for the four scale constructs. The pooled kappa result for knowledge was 0.86 (95% CI: 077, 0.95); for attitude it was 0.87 (95% CI: 0.76, 0.99); for practice it was 0.86 (95% CI: 0.75, 0.97) and for awareness it was 0.92 (95% CI: 0.83,

1.00) all indicating an almost perfect agreement.

The measured, pooled Cohen's kappa for all 23 items of the KAP survey tool was 0.87 (95% CI: (0.87, 0.87) indicating an almost perfect agreement. Hence, the test-retest repeatability demonstrates a reliable KAP survey tool.

3.4 *The feasibility of developing a KAP survey tool*

It was important to develop this tool as it would encourage the community to share their perspectives regarding CHL and UNHSP. The development of the tool followed a rigorous scientific approach which consisted of designing, refining, validating and assessing the reliability of the tool [23]. Since the study was part of the Amajuba UNHS research programme, the recruitment of participants at each stage of the development of the tool was possible, supported by subsidised incentives. This process also demanded a great deal of time, from the initial design stage to the validation of the tool and although tedious, most of the procedures were realistic and achievable. Another important factor was funding, which provided a budget to cover accommodation, transport and research team expenses.

4 Discussion

The development of the KAP survey tool was feasible given sufficient time, funds, motivation and a study population as demonstrated in this study. This study presents the stages of the development and validation of the KAP tool regarding childhood hearing loss (CHL) and UNHS. The question was whether the proposed KAP measures were measuring what they were supposed to measure, in terms of accuracy or stability [19, 23]. The results suggest that the developed tooth is both theoretically sound and a valid measure of KAP regarding childhood hearing loss and UNHS.

The results of the validity assessments in the questionnaire indicated that it is an applicable measure for the phenomena of NHS and CHL, as it went through appropriate validation processes. Although face validity is understood as the weakest approach to validity due to its subjective nature [35], it is pragmatic in the context of acceptability [36]. It has provided significant information that allowed the tool to be more understandable to participants in this study. Content validity results, on the other hand, demonstrated the KAP tool scale relevancy to the phenomena of NHS and CHL [21].

It could be argued that some of the generic concepts and measures of the KAP tool overlap with the previous KAP tool and these can be compared to yield additional evidence of the validity of the tool. For example, these could include measures that assessed

knowledge about (1) a baby born with HL, (2) risk factors (e.g. noise, ear-discharge, medication, hereditary, traditional medicine), (3) hearing loss identified at birth, (4) treatment of CHL and (5) cultural beliefs (ancestral sins, bewitchment) as would measures that assessed attitude towards screening and whether parents would like more information. Clearly, we could have compared these measures at face value but we did not as context is important and varies between communities. The meaning of concepts can be unclear if they are interpreted within specific socio-cultural contexts and language differences as these factors can influence the outcome [37, 38]. Our tool differed from the previous tool as the wording of the questions and the scale constructs captured the specific context and the concepts that were defined and which could readily be understood by the community of study. Further, this study also revealed that the interpretation of concepts in any content domain can be ambiguous [39], as our own experience, through the repetitive process of content validity varied amongst the experts.

For various reasons, these results were not comparable to previous studies that measured similar variables. Some studies did not report validity [9, 14, 40]. Other studies adapted previous tools [6, 11, 13], while some studies modified these tools and conducted a pilot study but provided no evidence of validity [15, 41]. Nevertheless, it is hoped that the tool developed in this study will help initiate a new line of research which integrates and validates community perspectives of KAP with regard to CHL and NHS.

On the other hand, the test-retest repeatability exercise was undertaken to investigate whether or not the developed KAP survey tool of NHS and CHL was consistent and stable enough to be of value and to quantify its agreement and repeatability. The repeatability assessment of a measurement tool requires that it is undertaken at two points in time [19]. In research practice, the degree of agreement between the two assessments is an indication of the quality of a single measurement, suggesting test-retest reliability for consistency and stability across time [36, 42]. The results of test-retest repeatability showed a Cohen's Kappa coefficient of 0.87 with almost perfect agreement indicating the consistency and stability of the tool and its constructs. The majority of items (22 of 23, 96%), with kappa values greater than 0.61, suggested a substantial to almost perfect agreement. However, there were two items that were incongruent with other items in their respective scale construct. In the knowledge scale construct, the level of agreement of one item was lower than the other five items. This could be interpreted to mean that items which assessed the general knowledge of childhood hearing loss such as causes, treatment etc., had clearly achieved better agreement than the early detection items. With the practice scale construct, the item "acceptance of

further examination when offered to the child" achieved the lowest value of reliability leading to a negative value in confidence intervals. The kappa estimate claimed a moderate level of agreement with a 95% confidence interval that the true estimate was between -0.04 – 1. We can conclude from the negative CI in lower limits, based on the 95% confidence interval, that there is a disagreement with regard to the likely acceptance of further examination of a child if offered. This evidently demonstrates the limitation of kappa as the estimates of CI includes negative values of poor agreement to almost perfect agreement. In this context, statistical significance signifies nothing when so much error exists in the results [31].

These results are not in line with previous studies as the mode of analysis is different. As previously stated, the current study used Cohen's Kappa test-retest repeatability to account for chance agreement in order to achieve reliability of the KAP survey tool [30]. The assumptions of Cohen's kappa coefficient is that the nominal scales with an agreement are independent, mutually exclusive and exhaustive, showing stability at those two points in time. Previous studies that reported reliability used internal consistency with the Cronbach alpha coefficient, which reflects the coherence of the components of the scale of the measurement tool [6, 12, 14, 15]. Although, the procedures undertaken to obtain reliability were not elaborated in these studies, the alpha coefficient is one way of assessing the internal consistency of a measuring scale [23]. This usually refers to the degree of homogeneity or the inter-relatedness of a set of items within a scale.

Overall, it can be argued that the validated KAP survey tool will be resourceful and versatile in addressing the needs of this community and other communities with similar characteristics.

4.1 Limitations and recommendations

There is a need for further validation of this tool using predictive validity to examine subsequent performance with regard to knowledge and attitude after UNHS programme implementation and health education.

To demonstrate further stability of the scale constructs, we recommend a cross-validation of the questionnaire across independent samples. This will strengthen the rigor of the questionnaire and broaden the generalisability.

5 Conclusion

Although the development of the tool was laborious it proved to be feasible and may offer valuable information for future interventions around childhood hearing loss and early detection. The KAP scale constructs showed a good validity with high I-CVI and S-CVI. The reliability of the KAP survey tool was good as the three constructs achieved an almost perfect agreement between the participants' two point results, after taking chance agreement into account. However, estimates of kappa can be ambiguous in certain contexts when the confidence intervals comprise the whole scale of kappa interpretation. Overall, the developed KAP survey tool may be useful in understanding rural communities that are similar to the community of study.

Acknowledgements

The authors wish to thank the women who participated in this study. Our gratitude to the experts, for their time and contributions to the development of the tool. We also thank the Department of Health and the Newcastle hospital staff for theirs support during the study.

Funding sources

This research did not receive any specific funds from funding agencies in the public, commercial, or not-for-profit organization.

References

[1] WHO. Childhood Hearing Strategies for prevention and care. *WHO Libr Cat Dat*, apps.who.int/iris/bitstream/10665/204632/1/9789241510325_eng.pdf (2016).

[2] WHO. Prevention of deafness and hearing loss. *WHO Libr Cat Data*, http://apps.who.int/gb/ebwha/pdf_files/WHA70/A70_R13-en.pdf?ua=1 (2017, accessed 2 April 2018).

[3] Shulman S, Besculides M, Saltzman A, et al. Evaluation of the universal newborn hearing screening and intervention program. *Pediatrics* 2010; 126 Suppl: S19-27.

[4] Engle PL, Black MM, Behrman JR, et al. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *Lancet* 2007; 369: 229–242.

[5] Young A, Tattersall H. Universal newborn hearing screening and early identification of deafness: Parents' responses to knowing early and their expectations of child communication development. *J Deaf Stud Deaf Educ* 2007; 12: 209–220.

[6] Crockett R, Wright AJ, Uus K, et al. Maternal anxiety following newborn hearing screening: the moderating role of knowledge. *J Med Screen* 2006; 13: 20–5.

[7] Park AH, Warner J, Sturgill N, et al. A survey of parental views regarding their child's hearing loss: A pilot study. *Otolaryngol - Head Neck Surg*; 134. Epub ahead of print 2006. DOI: 10.1016/j.otohns.2006.01.013.

[8] Fox R, Minchom S. Parental experiences of the newborn hearing screening programme in Wales: A postal questionnaire survey. *Health Expectations* 2008; 11: 376–383.

[9] MacNeil JR, Stone S. Evaluating Families' Satisfaction With Early Hearing Detection and Intervention Services in Massachusetts. *Am J Audiol* •; 16.

[10] Fitzpatrick E, Graham ID, Durieux-Smith A, et al. Parents' perspectives on the impact of the early diagnosis of childhood hearing loss. *Int J Audiol* 2007; 46: 97–106.

[11] Suppiej A, Cainelli E, De Benedittis M, et al. Failure of hearing screening in high-risk neonates does not increase parental anxiety. *J Matern neonatal Med* 2013; 26: 932–5.

[12] Olusanya BO, Luxon LM, Wirz SL. Maternal views on infant hearing loss in a developing country. *Int J Pediatr Otorhinolaryngol* 2006; 70: 619–623.

[13] Swanepoel D, Almec N. Maternal views on infant hearing loss and early intervention in a South African community. *Int J Audiol* 2008; 47 Suppl 1: S44-8.

[14] Mohd Khairi MD, Noor Rafidah K, Affizal A, et al. Anxiety of the mothers with referred baby during Universal Newborn Hearing Screening. *Int J Pediatr Otorhinolaryngol* 2011; 75: 513–517.

[15] Ravi R, Yerraguntla K, Gunjawate DR, et al. Knowledge and attitude (KA) survey regarding infant hearing loss in Karnataka , India. *Int J Pediatr Otorhinolaryngol* 2016; 85: 1–4.

[16] Mazlan R, Lee T, Sarah SZ, et al. A questionnaire-based study on parental satisfaction with a universal newborn hearing screening program in Kuala Lumpur , Malaysia. *Int J Pediatr Otorhinolaryngol* 2014; 78: 348–353.

[17] Ravi R, Gunjawate DR, Yerraguntla K, et al. Knowledge and attitude of parents/caregivers towards hearing loss and screening in newborns – a systematic review. *Int J Audiol* 2016; 2027: 1–8.

[18] Olusanya BO, Newton VE. Global burden of childhood hearing impairment and disease control priorities for developing countries. *Lancet* 2007; 369: 1314–1317.

[19] Coughlan M, Cronin P, Ryan F. Step ' by-step guide to critiquing research . Part 1 : quantitative research. *Br J Nurs* 2007; 16: 658–664.

[20] Jack L, Hayes S, Jeanfreau S, et al. Appraising Quantitative Research in Health Education: Guidelines for Public Health Educators. *Health Promot Pract* 2010; 11: 161–165.

[21] Polit DF, Beck CT. *Essentials of Nursing Research: Appraising Evidence for Nursing*

Practice 7th Ed. Seventh Ed. Lippincott Williams & Wilkins, https://web2.aabu.edu.jo/tool/course_file/lec_notes/1001391_Essentials of Nursing Research Appraising Evidence for Nursing Practice Essentials of Nursing Research Polit.pdf (2010).

[22] Graham C, Seeley J, Gina A, et al. Mapping the content of mothers ' knowledge , attitude and practice towards universal newborn hearing screening for development of a KAP survey tool. *PLoS One* 2019; 14: 1–20.

[23] Kumar A. Review of the Steps for Development of Quantitative Research Tools. *Adv Pract Nurs* 2015; 1: 1–4.

[24] WHO. A Guide To Developing Knowledge , Attitude and Practice Surveys. *WHO Libr Cat Data*, http://whqlibdoc.who.int/publications/2008/9789241596176_eng.pdf (2008).

[25] Launiala A. How much can a KAP survey tell us about people ' s knowledge , attitudes and practices ? Some observations from medical anthropology on malaria in pregnancy in Malawi Background : KAP surveys. *Anthropol Matters J* 2009; 11: 1–13.

[26] WHO. *Health education: theoretical concepts, effective strategies and core competencies*. Cairo: Regional Office for the Eastern Mediterranean, https://apps.who.int/iris/handle/10665/119953?show=full (2012).

[27] Waltz CF, Strickland O, Lenz ER. *Measurement in nursing and health research*. Fourth ed. Springer Pulbishing Company LLC, 2010.

[28] Polit DF, Beck CT. The Content Validity Index : Are You Sure You Know What ' s Being Reported ? Critique and Recommendations. *Res Nurs Health* 2006; 29: 489–497.

[29] DeVon H a., Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability [Electronic Version]. *J Nurs Scholarsh* 2007; 39: 155–164.

[30] Watson PF, Petrie A. Method agreement analysis : A review of correct methodology. *Theriogenology* 2010; 73: 1167–1179.

[31] McHugh ML. Interrater reliability : the kappa statistic. *Biochem Medica* 2012; 22: 276–282.

[32] Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977; 33: 159.

[33] DeVries H, Europe R, Elliott MN, et al. Using Pooled Kappa to Summarize Interrater Agreement across Many Items. *Field methods* 2008; 20: 272–282.

[34] Institute of Translation Health Sciences. Research Study Feasibility Tool. *Institute of Translational Health Sciences*, https://www.iths.org/wp-content/uploads/ResearchStudyFeasibilityTool_V12017Feb6.pdf (2017, accessed 1 August 2020).

[35] Drost EA. Validity and Reliability in Social Science Research. *Educ Res Perspect*

2011; 38: 105–123.

[36] Bannigan K, Watson R. Reliability and validity in a nutshell. *J Clin Nurs* 2009; 18: 3237–3243.

[37] Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Heal VoI* 2005; 27: 281–291.

[38] Boateng GO, Neilands TB, Frongillo EA, et al. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Heal* 2018; 6: 149.

[39] Bollen KA. *Structural equations with latent variables*. Epub ahead of print 1989. DOI: 10.1002/9781118619179.

[40] Scheepers LJ, Swanepoel DW, Roux T le. Why parents refuse newborn hearing screening and default on follow-up rescreening-A South African perspective. *Int J Pediatr Otorhinolaryngol* 2014; 78: 652–658.

[41] Rajagopalan R, Selvarajan HG, Rajendran A, et al. Grandmothers ' perspective on hearing loss in children and newborn hearing screening. *Indian J Otol* 2014; 20: 20–24.

[42] Bartlett JW, Frost C. Reliability, repeatability and reproducibility: Analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol* 2008; 31: 466–475.