A STUDY ON THE ACCURACY OF PEN CLICK TEST AS A HEARING SCREENING TOOL AMONG NEWBORNS SEEN IN TWO TERTIARY GOVERNMENT HOSPITALS

GENEVIEVE A. ABUAN, MD, AILEEN GRACE T. MEMBRERE, MD,
MICHAEL M. RESURRECCION, MD

ABSTRACT

BACKGROUND: Hearing impairment has a great impact on the functional, social and emotional aspects of a child. Thus, early detection and management is crucial for optimal development of the child. The Newborn Hearing Screening Act was approved in the Philippines to "institutionalize measures for prevention and early diagnosis of congenital hearing loss among newborns". A simple, accurate and readily available hearing screening tool is necessary in less privileged communities.

OBJECTIVES: To determine the accuracy of Pen Click Test as compared to otoacoustic emission test as a hearing screening tool among newborns seen in two tertiary government hospitals. The accuracy of Pen Click Test was measured for its sensitivity, specificity, positive and negative predictive value.

METHODS: The study is an experimental design consisting of three phases: Phase I is a randomized complete block design; Phase II involves inter-rater and intra-rater variability randomized block design and Phase III is a cross sectional design. The study was done in two government tertiary hospitals. The subjects are term newborns with both ears analyzed independently from each other.

RESULTS: Phase I of the study identified Acroball retractable pen as the study stimulus based on its accessibility and its capability to produce high decibel. In phase II, all health workers produced a sound stimulus of more than 70 decibels. Majority of the health workers had no significant difference among each other which means there is minimal deviation from the mean. Phase III showed that pen click test has a high specificity of 98% and a sensitivity of 43%. Based on disease prevalence, the test showed a positive predictive value of 77% and negative predictive value of 93%. Kappa agreement showed moderate result with a Kappa coefficient of 0.54.

CONCLUSION: The study showed high specificity in identifying hearing impairment and a positive association of Pen Click test to the standard hearing audiometer. The application of this test in the community may be done as a hearing screening tool. This study provides an accessible, easily reproducible and accurate tool for hearing screening that may be applied in communities without facilities.

KEYWORDS: Newborn Hearing Screening, Hearing impairment, Accuracy

INTRODUCTION

The World Health Organization (WHO) reported increasing incidence of hearing impairment worldwide from 42 million in year 1985 to 360 million people in year 2012. Approximately 32 million (9%) of which were children with disabling hearing impairment.[1]

Two thirds of the population with hearing impairment come from Southeast Asian countries.[1] The most common causes associated with hearing impairment were congenital hearing impairment, otitis media and noise induced hearing loss. In a study by Naeem and Newton, sensor-neural hearing loss was prevalent among Asian children with rates ranging from 5.09 to 9.61per 1000. The relative risk of having hearing impairment among Asian children was two to three times greater than non-Asian children.[1] Of these children, one to two of 1000 newborns have hearing impairment that warrants observation or treatment.

In the Philippines approximately 1,443 thousand persons of the 92.1 million household

population (1.57 percent) had disability. The Philippine disability survey was categorized as follows: moving disability (39%); hearing (33%); speaking (10%); mental (10%) and seeing (8%). The survey revealed that hearing disability ranked as second highest form of disability with 1.10% prevalence rate.[2]

Hearing impairment has great impact on the functional, social and emotional aspects of a child which may include speech delay, poor academic performance and poor personal-social development. Thus, early detection and management is crucial for optimal development of the child. The National Institute of Health in 1993 and the Joint Committee on Infant Hearing in 2000 recommended that the diagnosis of hearing impairment must be done before three months of age to provide early intervention before six months of age. This age is considered as the critical period for the development of central auditory pathway.[4]

There are various audiology tests used for hearing screening that is age appropriate. The most commonly used procedure is the Otoacoustic emission test (OAE) which measures cochlear response to stimuli. It is appropriate for all ages and may be done in a few minutes. It is independent on the sleep-wake status of the patient. Automated Auditory Brainstem Response (ABR) is another audiologic tool that may be used for newborns. This test measures auditory nerve and brainstem activity. The OAE and ABR tests are considered only as a primary screening tool since they cannot assess cortical sound processing. In older children, play audiometry and conventional audiometry may be done. These tests assess auditory perception and would require participation of the child.[4]

The universal hearing screening program was fully implemented. However, a major problem in adopting the program is the insufficient facilities of hearing screening tool such as OAE and ABR especially in developing countries. Various methods on improvised hearing screening were done among children using tuning fork and pen. These methods are easily reproducible. However, the accuracy of each test may vary as a hearing screening tool. It is important to establish a hearing screening tool that is simple, readily available, affordable and accurate. This study will answer the research question "How accurate is pen click test as a hearing screening tool among full term newborns seen in two tertiary government hospitals?

METHODOLOGY

The study is an experimental design consisting of three phases: randomized complete block design, randomized block design (inter and intra-rater variability) and cross-sectional design. The study was conducted among newborns seen in two tertiary government hospitals.

The study included 97 newborns, with a total of 194 ear subjects that underwent hearing screening in compliance with the Universal Hearing Screening Program. One hundred forty-six ear subjects were enrolled from two tertiary government hospitals in Quezon City. The sample size was computed based on the formula shown in Figure 1 with a total of 194 ear subjects.

$$n = \frac{2\, Z^2\alpha/2\, pq}{\theta^2}$$

n= desired sample size
z^2= critical value for confidence interval
α = alpha
p = specificity or sensitivity
q = 1- specificity or 1- sensitivity
θ^2 = margin of error

Thus,
For number of cases z^2=1.96; p= 0.71; q=0.21; θ^2 = 0.01
For number of controls z^2=1.96; p= 0.95; q=0.05; θ^2 = 0.02

Figure1. Formula to compute for the sample size

STUDY PROCEDURE

Phase I. Randomized Complete Block Design

Phase I of the study determined the variability of the pen click test. Three different types of retractable pens were used as the experimental factor and trials as the block. Phase 1 measured the inter-rater variability which was conducted by the principal investigator. Three types of retractable pens (Pilot Acroball, BIC and HBW) which are all easily accessible were used as the study stimulus. Three pens were tested per brand of pen included. The Pen click test was tested with a sound level meter (TES-1350A). The investigator produced the pen click stimulus thirty times at two inches distance from the sound level meter. The pen click stimulus was repeated for three trials with fifteen minutes interval. All sound measurements from

the selected pens were recorded in decibels(dB) and were written in a data form (Appendix 2). The type of ballpen with the highest mean decibel was used in Phase II and Phase III of the study.

Phase II. Randomized Block Design: Intra-rater and Inter-rater Variability

Phase II of the study determined the variability of the pen click using the type of ballpen with the highest mean decibel produced (Pilot Acroball). In Phase II, the healthcare worker was the experimental factor and trials as the block. Phase 2 measured the inter-rater and intra-rater variability among the health care workers. This phase of the study showed the reproducibility of the test among healthcare workers in the community at any given time.

A total of ten participants were chosen from the qualified healthcare workers through fishbowl sampling. Phase II was conducted in an audiology room. The health care workers clicked the pen thirty times at two inches distance from the sound level meter. Three trials were done with fifteen minutes interval per trial. All sound measurements from the selected pens were recorded in a data form. The difference in the values obtained by the health care workers were determined.

Phase III. Cross Sectional Design or Validation study

Phase III was a hospital-based trial. Inter and intra rater variability tests were conducted among health care workers selected to do the pen click test. Three videos of an actual pen click test were presented to three health care workers to determine inter rater variability and one video of an actual pen click test was shown twice to a healthcare worker for the intra-rater variability. This part of the study showed the reliability of health workers from the two government institutions in arriving with a correct assessment.

The principal investigator oriented the parents or guardians of the newborn on the importance, objectives and the benefits of the study. A written informed consent (Appendix 1) was obtained from the parents or guardians prior to the commencement of the study. Inclusion criteria were all healthy newborns with normal otoscopic examination findings. Excluded in the study were preterm newborns, sick neonates and those with ear deformities.

Phase III of the study involved the pen click as the sound stimulus. The Acroball pen had the highest mean decibel and was used as the sound stimulus based on the results generated in phase I. Acroball pen is readily available, accessible and produced the least variable level of sound intensity. A trained health care worker facilitated the pen click test. The study was done at the audiology room of two tertiary government hospitals. The audiology room was quiet with a comfortable environment.

The newborns included in the study were on their 2nd day to 30th day of life. The newborns were in a comfortable condition prior to the study. A comfortable condition was described as asleep, in a fed state or awake but has no abrupt movements and is not crying. The healthcare worker stood on the side of the patient's ear to be examined. An ear plug was applied on the contralateral ear prior to the test. The same procedure was done to the other ear. The health care worker positioned the pen two inches away from the test ear. Pen click test was administered to the newborns for three trials with fifteen minutes interval per trial. The results of the test were recorded as "pass" if blinking, startling or crying was observed from the newborn. The test was labelled "fail" in the absence of the blinking, startling and crying in two out of three trials. The pen click test was recorded with a video camera. An assistant investigator determined or observed the reaction of the newborns. An assistant investigator unknowledgeable of the study objective performed an otoacoustic emission test at the hearing unit of the two tertiary government hospitals. The pens were replaced every 10 patients. All pens used in this study were calibrated every two days.

Descriptive statistics were used to summarize the clinical characteristics of the patients. Frequency and proportion were used for nominal variables and range for ordinal variables. Coefficient of variation was used to determine inter-rater agreement. Intra-class coefficient was used to determine reliability of intra-rater agreement. Sensitivity, specificity, Negative Predictive Value, Positive Predictive

Value, and likelihood ratios were used to determine the diagnostic quality and accuracy of the Pen Click Test compared to otoacoustic emission test as a hearing screening tool among newborns in two tertiary government hospitals. Kappa agreement was also used to test inter-rater reliability.

All valid data was included in the analysis. Missing variables were neither replaced nor estimated. Null hypothesis was rejected at 0.05 α-level of significance. STATA 15.0 was used for data analysis.

The study protocol was submitted and approved by the Institutional Review Board – Ethics Committee of two tertiary government hospitals. A written informed consent was obtained from the parents prior to inclusion to the study. The primary investigator explained in Filipino the nature, objectives and possible risks of the study to the parents or guardians of the newborns in Filipino language. The parents or guardians were informed that a video of the pen click test was obtained during the procedure. Personal information of the participants was kept confidential. The results of the tests were disclosed to the parents or guardians. The babies with their mother were requested to stay for forty-five more minutes from the usual procedure of hearing screening. The research assured that no harm was inflicted to the participants. Throughout the procedure, the investigator avoided unnecessary noise, pain or stimuli that can disturbed the patient. Normothermia was always maintained .

RESULTS

Phase I. Randomized Complete Block Design

Phase I of the study determined the variability of the pen click test using Pilot Acroball, BIC and HBW. Acroball pen showed the highest decibel obtained from TES sound level meter with the mean of 79.1 decibels. In comparison, HBW ball pen showed the lowest decibel with a mean of 77.4.

We sought to compare the intra-rater agreement using three pen types. In all pens, the intra-rater coefficient of variations was all below 5%, which is deemed acceptable for repeatability testing. However, the 95% Confidence Interval for Coefficient of Variation for Acroball 2, Acroball 3, BIC 1 included values slightly higher than 5%. (Table 1)

	%CV (95% CI)	ICC (95% CI)	P-Value
Pilot Acroball			
Acroball 1 (n=3)	1.76% (0.35% - 3.16%)	0.167 (0.07 – 0.89)	0.390
Acroball 2 (n=3)	3.4% (0.68% - 6.12%)		
Acroball 3 (n=3)	3.0% (0.6% - 5.42%)		
HBW			
HBW 1 (n=3)	1.94% (0.39% - 3.49%)	0.432 (0.05 – 0.78)	0.283
HBW 2 (n=3)	2.27% (0.45% - 4.09%)		
HBW 3 (n=3)	1.0% (0.2% - 1.85%)		
BIC			
BIC 1 (n=3)	3.97% (0.79% - 7.14%)	0.443 (0.23 – 0.69)	0.278
BIC 2 (n=3)	2.13% (0.43% - 3.83%)		
BIC 3 (n=3)	1.0% (0.2% - 1.8%)		

%CV – Coefficient of Variation percentage; ICC – Intraclass correlation coefficient

Table 1. Intra-rater and inter-rater agreement using three pen types in Phase 1 of the study

Phase II. Inter-rater and Intra-rater Randomized Block Design

Phase II of the study measured the inter-rater and intra-rater variability among health workers regardless of pen type. This phase of the study showed the reproducibility of the test among health care workers in the community. Table 3 outlines the 95% confidence interval of decibels produced by each health worker. The mean decibel of each health worker was within the 95% Confidence Interval. In most of the health workers, the intra-rater coefficient of variations were all below 5%, which is deemed acceptable for repeatability testing. However, the 95% Confidence Interval for Coefficient of Variation for Health Worker 6 and Health Worker 9 included values slightly higher than 5%.

In the inter-rater agreement, the Intraclass Coefficient (ICC) point estimate was at 0.543 ($p = 0.031$), halfway between 0 to 1, which indicated that neither intra-rater nor inter-rater variation dominates. Result showed a wide 95% CI ranges for ICC (ranging from 0.04 to 0.87), However, the relatively low coefficients of variation indicated a good intra-rater agreement (Table 2).

	%CV (95% CI)	ICC (95% CI)	P-Value
Health worker 1	1.87% (0.37% - 3.37%)		
Health worker 2	0.64% (0.13% - 1.15%)		
Health worker 3	2.14% (0.43% - 3.86%)		
Health worker 4	1.84% (0.37% - 3.3%)		
Health worker 5	1.81% (0.36% - 3.25%)	0.543 (0.04 – 0.87)	**0.031**
Health worker 6	3.03% (0.61% - 5.46%)		
Health worker 7	1.89% (0.38% - 3.39%)		
Health worker 8	2.28% (0.46% - 4.11%)		
Health worker 9	3.11% (0.62% - 5.59%)		
Health worker 10	1.95% (0.39% - 3.51%)		

%CV – Coefficient of Variation percentage; ICC – Intraclass correlation coefficient

Table 2. Intra-rater and inter-rater agreement among ten health workers in Phase II of the study

Phase III. Cross Sectional Design or Validation study

A total of 97 newborns, 2 to 30 days of age with a mean age of 6 days was enrolled in this study. The male to female ratio was 1:1.3. A total of 194 ear subjects were tested with pen click test and OAE.

The accuracy of pen click test as compared to OAE was determined through computation of the sensitivity and specificity of the test. Results showed that 43.48% of patients with hearing impairment have a positive pen click test. In comparison, 98.25% of patients without hearing impairment have a negative pen click test. (see Table 3)

The estimation of the probability of the presence or absence of the disease was determined through the computation of the positive predictive value and negative predictive value. When pen click test shows a positive result, we have 76.92% probability that a positive result will be obtained in OAE. In comparison, we have 92.82% probability that the patient will have a negative OAE result when pen click test is negative. (Table 3)

A useful measure in the interpretation of diagnostic test is the likelihood ratio. Patients who are OAE positive are 24.78 times more likely to yield a positive pen click test compared to patients who are OAE negative (LR+) and are 58% less likely to yield a negative pen click test result (LR-) (Table 3). Overall, the accuracy of the pen click test, when compared against OAE in screening for hearing loss, is 91.75% (Table 3)

Pen click test	Otoacoustic Emission test		Total
	Positive/Refer	Negative/Pass	
Positive/Refer	10 (5.15)	3 (1.55)	13 (6.7)
Negative/Pass	13 (6.7)	168 (86.6)	181 (93.3)
Total	23 (11.86)	171 (88.14)	194 (100)

Sensitivity (Sn)	43.48% (23.2% – 65.5%)	Positive LR	24.78 (73.36 – 83.5)
Specificity (Sp)	98.25% (95% – 99.6%)	Negative LR	0.58 (0.4 – 0.82)
PPV	76.92% (49.7% – 91.8%)	Accuracy	91.75% (87% – 95.2%)
NPV	92.82% (90% – 94.9%)		

PPV, positive predictive value; NPV, negative predicted value; LR, likelihood ratio

Table 3. Diagnostic accuracy of Pen Click Test as compared to Otoacoustic emission test as a hearing screening tool among newborns in two tertiary government hospitals (n = 194)

The kappa values were interpreted according to the guidelines adapted from Landis and Koch. As shown in Table 4, Pen Click Test and OAE has moderate agreement based on Kappa Coefficient.

	Kappa Coefficient			
	Agreement	(95% CI)	Interpretation	P-Value
Pen click test	91.75	0.514 (0.31 – 0.72)	Moderate	**<0.001**

Table 4. Kappa agreement (n = 194)

DISCUSSION

Newborn Hearing Screening Act or Republic Act 9709 was approved in the Philippines with the purpose to "institutionalize measures for prevention and early diagnosis of congenital hearing loss among newborns". [9] However, the availability of an accessible and cost-effective tool is a major problem among distant communities.

Hearing impairment refers to hearing loss greater than 30 decibels among children. Hearing loss may be described as mild, moderate or severe depending on how a person can hear the loudness or intensity of a sound stimulus. Determining the proper stimulus is very important for a hearing screening tool as being presented in the Phase I of the study. Criteria for the hearing screening stimulus in this study

includes the highest decibel produced and its availability in the community. Acroball recorded the highest decibel with a mean of 79 (ranges from 77.6 to 83.9 decibels). A normal ear is able to hear at a range of 25 to 75 decibels. With the use of Acroball pen as the study stimulus, we are able to identify ear subjects that may have hearing impairment. Hearing impairment is present if the ear subject is unable to identify the sound stimulus greater or equal to 77 decibels.

The reproducibility of the study is also essential in finding a diagnostic test such as a hearing screening tool. In Phase II of the study, health workers were randomly selected and were asked to produce the sound stimulus using the Acroball pen obtained from Phase I. The mean decibels ranging from 73 to 77.9 were computed from the three trials done by each health worker. The results showed that all health workers were able to produce a sound stimulus of more than 70 decibels which is a requirement in the study. Results in the study showed most of the health workers had no significant difference among each other which means that there is minimal deviation from the mean decibel. The test can be reproduced by most of the health care workers and may be applied in the community.

Validation study allows evaluation of an index test in comparison to the reference test in determining subjects with the target disease. Test validation involves the basic measure in quantifying a diagnostic accuracy of a test such as the sensitivity, specificity, positive predictive value, negative predictive value and likelihood ratios. [10]

The pen click test has a high specificity of 98. 25%, which would indicate those ears that have no hearing impairment is correctly identified by the index test. This means that 98.25% of ears tested without hearing impairment will have a negative result in pen click test. The sensitivity of the test measures the capability in identifying subjects with the disease.[9] The pen click test has a sensitivity of 43.48% which means that we would expect 43.48% of ears tested with hearing impairment to have a positive pen click test result. Pen click test is a highly specific test which can be used as a screening tool to rule in patients with hearing impairment.

The positive and negative predictive values were also computed. The positive predictive value is the probability that the disease is truly present given a positive result is obtained. [9] The test showed a positive predictive value of 76.92% and negative predictive value of 98.25%. Based on prevalence of hearing impairment, we would expect 76.92% of ears with positive results to have hearing impairment while 98.25% of ears with negative results to have no hearing impairment. Using the 2x2 table, the accuracy of the study was computed to be at 91.75% which means that the results obtained from pen click test is close to the results from the standard hearing screening test (OAE).

The likelihood ratio of positive results tells us how well the test performs in the study population. [16] The likelihood ratio for a positive test in the present study is 24.78. This means that a child with hearing impairment is 24.78 times more likely to have a positive result than a child without hearing impairment.

Kappa agreement showed moderate result with a Kappa coefficient of 0.541. This measures interrater variability or the consistency among individuals who underwent data collection.[14]

Based on the results of validity and association test, the Pen click test can be an accurate hearing screening tool in identifying ear subjects of newborns with high suspicion of hearing impairment. The test is readily available for communities without facilities for OAE or ABR as hearing screening tool. The procedure done in the pen click test is easily reproducible and health workers may be trained to standardize the sound stimulus. This screening may encourage more health workers to be involved in community work should this test be considered in the community. As a primary health care worker, early recognition and timely intervention if necessary is of outmost important.

CONCLUSION AND RECOMMENDATION

The study showed high specificity of 98.25% which signifies that Pen Click test is a good screening tool in ruling in patient with hearing impairment. Pen click test has a sensitivity of 43.48%. There is a positive

association of Pen Click test to the standard hearing audiometer with a positive predictive value of 76.92% and negative predictive value of 98.25% as evident in the study. Thus, a child with hearing impairment is 24.78 times more likely to have a positive result than a child without hearing impairment. Kappa agreement which measures the inter-rater variability showed moderate result with a Kappa coefficient of 0.541. Pen click test has a high accuracy rate of 91.75%. Therefore, application of this test in the community may be done as a hearing screening tool in comparison with the standard hearing screening tool.

This study provides an accessible and easily reproducible tool for hearing screening that may be applied in communities without facilities. This hearing screening tool will help in early recognition and intervention to prevent profound hearing impairment.

This study recommends involvement of a larger sample size and a longer study period. More health workers should be involved in phase II in assessing the reproducibility of the hearing screening tool. Hearing screening test should be implemented in all communities.

REFERENCES

1. World Health Organization. Newborn and infant hearing screening current issues and guiding principles for action.2010;7-23

2. Better Hearing Philippines, Prevalence of Hearing Impairment in the Philippines, Retrieved from http://bhphil.org/downloads/prevalence_of_ear.pdf

3. Ptok,M. Early detection of hearing impairment in newborns and infants. Dutsch Arztebl International.2011;108(25):426-31

4. Naeem Z and Newton V. Prevalence of sensorineural hearing loss in Asian children. 1996; 30(5):332-339

5. Gloria-Cruz T, Abes G, Abes F. The voice test for newborn hearing screening. Acta Medica Philippina.2012; 46(3): 46-51

6. Pirozzo S , Papinczak, T, Glasziou P. Whispered voice test for screening for hearing impairment in adults and children: systematic review. BMJ.2003.327:967s

7. Garcia M, Chiong C, Abes G, Carrillo R. Accuracy of reflexive behavioral (“Baah”) test in the screening for hearing impairment in infants six months old and below. Philippine Journal of Otolaryngology-Head and Neck Surgery.2012; 27(1): 6-11

8. Subroto O.et al, A simple ballpen click as a tool for screening of hearing impairment in Filipino school children in Metro Manila”.The Philippine Journal of Pediatrics. 1995; 44(4): 249-252.

9. Patrick M. Bossuyt, Johannes B. Retisma, David E. Bruns. The STARD Statement for Reporting Studies of Diagnostic Accuracy: Explanation and Elaboration. 2003; 49:7-18.

10. Hwee Bee Wong and Gek Hsiang Lim. Measures of Diagnostic Accuracy: Sensitivity, Specificity, PPV and NPV. 2011;20 (4): 316-318.

11. 11.Wrightson,Stevens. Universal Newborn Hearing Screening.American Family Physician.2007; 75(9); 1349-1352

12. Republic Act 9709.Republic of the Philippines, Fourtheen Congress, Second Regular Session, Metro Manila. Senate and House of Representatives of the Philippines.

13. Sedighi,Iran.Interprretation Diagnostic Tessts: Likelihood ratio vs Predictive value. 2013. Kowsar Medical Institute; 23(6): 717

14. McHugh, Mary. Interrater reliability: The Kappa Statistic. 2012. Biochem Med; 22(3): 276-282