TEST-RETEST RELIABILITY OF PURE TONE AUDIOMETRY FROM 250 Hz TO 16 kHz

RHYDHM GUPTA

20AUD024

A Dissertation Submitted in Part of Fulfilment for the Degree of

Master of Science

(Audiology)

University of Mysore

All India Institute of Speech and Hearing

Manasagangothri, Mysuru-570006

August 2022

CERTIFICATE

This is to certify that this dissertation entitled **"Test-Retest Reliability of Pure Tone Audiometry From 250 Hz To 16 kHz"** is a bonafide work submitted in part of fulfilment for the Degree of Master of Science in Audiology of the student with Registration Number **20AUD024**. This has been carried out under the guidance of the faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru

August, 2022

Dr. M. Pushpavathi

Director

All India Institute of Speech and Hearing,

Manasagangothri, Mysuru- 570006

CERTIFICATE

This is to certify that this dissertation entitled **"Test-Retest Reliability of Pure Tone Audiometry From 250 Hz To 16 kHz"** has been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru

August, 2022

Dr. Niraj Kumar Singh

Guide

Associate Professor in Audiology

Department of Audiology

All India Institute of Speech and Hearing,

Manasagangothri, Mysuru- 570006

DECLARATION

This is to certify that this dissertation entitled **"Test-Retest Reliability of Pure Tone Audiometry From 250 Hz To 16 kHz"** is the result of my own study under the guidance of Dr. Niraj Kumar Singh, Associate Professor, Department of Audiology, All India Institute of Speech and Hearing, Mysuru, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru

August 2022 **Registration No. 20AUD024**

Dedicated To My Parents, Grandparents, Dakshita And Niraj Sir, My Guide, For Their Constant And Everlasting Support In Pursuing My Degree

ACKNOWLEDGEMENT

First and foremost, a heartfelt bow and thanks to H.H. Shri Mataji Nirmala Devi for always keeping her grace on her child. I don't have enough words to thank you, Maa. Whatever I have achieved to date or whatever I'll succeed; it's all because of your mercy. Thank you for giving me this life and this chance to do something good in life.

I want to thank and express my heartfelt gratitude to Niraj sir, my guide, my mentor, and my guru, for guiding me, not just for my dissertation but throughout my Bachelor's and Master's Degree. It was you sir, who brought my interest in Audiology. I have always admired how you teach and the efforts and passion you put into your teaching. It is said that teachers who love teaching make children fall in love with learning. Not just academically, I have learnt so many things from you, be it hard work, punctuality, or the zeal to learn. I would try my best to imbibe these qualities in me and to follow this throughout. Thank you sir, for being so patient and kind, for not getting irritated with my stupid doubts. Thank you for everything sir. I will forever be indebted to you. I feel proud to say that I'm a student of Dr. Niraj Kumar Singh.

I would like to thank the Director of our institute, Dr. M. Pushpavathi, for providing academic support and the HOD, Dr. Pravin Kumar, for granting permission to conduct the research.

Maa-Papa, I don't know how to thank you. You are the parents all kids hope to have because you are the perfect example of noble and pure souls. "I respect dad's hard work, I remember mom's flair. I can't forget dad's sacrifice, I can't forget mom's care".

You taught me to live with dignity, kindness, honesty and love. You both are the reason behind whatever I became as a person today. Only because of your endless blessings and prayers I'm standing here today. I can't thank enough to god for giving me, such wonderful parents. Thank you my guardian angels. I love you the most and hope I can make you proud one day.

Amma, Baba, Nana, Dakshita and Manas, losing you was the most painful time in my life. But my heart knows you're always watching me from above and protecting me. I dedicate all my success to you because, to be honest, it is incomplete without you. I miss you.

Iti and Ryan, my beloved siblings, thank you for always being there and making me laugh in all serious situations.

I want to thank my school teachers and UG faculties for building the foundation of learning.

Thank you Kriti mam and Nirmala mam for all the help.

A huge thank you to my dissertation partner, Pooja, for all the support and help. I won't forget the endless talks, discussions and laughs we shared. Thank you for taking care of me when I was sick. You made me feel at home. I'll miss you.

A big thank you to my special one for always being there, motivating me, and encouraging me that I could do better. Thank you for scolding me too whenever I became lazy; otherwise, I wouldn't have been able to complete my dissertation on time.

You were my stress buster throughout this journey. Thank you for listening to my blabbering even when you didn't understand any of it, making me calm whenever I was angry and bringing a smile to my face whenever I was sad. Thank you for believing in me and giving me all the love.

If I'm forgetting anyone, please forgive me. Thank you from the core of my heart.

ABSTRACT

Pure tone audiometry, especially the extended high-frequency pure tone audiometry, is a promising and valuable tool for the early detection of ototoxicity in adults. However, extended high-frequency audiometry is not used extensively in standard clinical practices, probably because of the lack of information about its test-retest reliability. Several applications, as in cases with ototoxicity, require two or more measurements separated in time by days to weeks, hence requiring removal and placement of transducer again. However, there is no study evaluating the inter-session test-retest reliability, especially estimated across multiple sessions. Therefore, the present study aimed to evaluate the inter-session test-retest reliability of pure tone audiometric thresholds measured in the frequency range of 250 Hz-16 kHz.

Auditory thresholds in the frequency range from 250 Hz to 16 kHz were measured in one ear of 160 otologically healthy subjects in the age range of 18-35 years. There was no significant difference among the sessions, and the average measure ICC for all the frequencies were >0.9. Therefore, pure-tone audiometry, including the extended high-frequency audiometry, lends itself to applications requiring multiple measurements, such as monitoring the damage due to noise exposure and ototoxic drugs and the outcomes of surgeries. Given the present study had a big data set, the results are generalizable.

TABLE OF CONTENTS

Contents	Page number
List of tables	i
List of Figures	ii
Chapter I Introduction	1-5
Chapter II Review of Literature	6-13
Chapter III Methods	14-20
Chapter IV Results	21-33
Chapter V Discussion	34-37
Chapter VI Summary and Conclusion	38-41
References	42-47

LIST OF TABLES

No.	Title	Page number
4.1.1	Audiometric thresholds at various measurement points	22-24
4.1.2	The outcome of the analysis of individual data for each frequency	26
4.1.3	The outcome of Friedman's test for comparison of thresholds between the measurement points	27
4.1.4	The outcome of the Wilcoxon Signed Rank test for pair-wise comparison at each measurement point	29
4.1.5	The outcome of the Intraclass Correlation coefficient test analysis	31-33

LIST OF FIGURES

No.	Title	Page number
4.1.1	The mean of thresholds obtained at various measurement points	25

CHAPTER I

INTRODUCTION

Hearing is one of the essential senses in our body because hearing makes it possible for humans to communicate, engage and participate in daily events, which helps in enhancing the quality of life. The organ responsible for hearing is the ear. The human ear is divided into the outer ear, middle ear, and inner ear. The outer ear collects the sound from the outside environment and funnels it down to the external auditory canal. This sound vibrates the tympanic membrane and gets transmitted to the middle ear, which is amplified through the three tiny ossicles. These ossicles transmit the sound to the inner ear, converting acoustic energy into electrical energy. The auditory nerve sends this electrical energy to the brain. The brain then decodes this electrical energy as sound, which we hear.

Humans have a hearing range from 20 Hz to 20 kHz (Shim et al., 2009). This range encompasses the frequencies essential for speech perception and aid in better communication. A dysfunction in any part of the ear can alter the hearing mechanism, resulting in hearing loss. The dysfunction can happen due to pathologies like otitis media, otosclerosis, ototoxicity, presbycusis, etc. (Isaacson & Vora, 2003).

There are several well-established clinical tools for assessing the functioning of the ear. These include pure tone audiometry (PTA), speech audiometry, immittance measures, otoacoustic emissions, auditory brainstem response (ABR) etc. Although all these tests can assess the integrity of the ear structures, pure tone audiometry is one test used widely for routine audiological evaluations. PTA is considered a 'gold standard' test

for the assessment of an individual's hearing abilities (Sideris and Glattke, 2006). The possible reasons for this could be (1) its high test-retest reliability and validity (Ishak et al., 2011) and (2) it gives frequency-specific information about the hearing status.

Pure tone audiometry is a procedure to obtain frequency-specific hearing thresholds using tonal stimuli (Sliwinska-Kowalska, 2015). A pure tone is presented through headphones or earphones, and the lowest level at which the stimulus can be heard 50% of the time is taken as the threshold. While the conventional pure tone audiometry involves frequencies ranging from 250 - 8000 Hz, the extended high-frequency audiometry measures the audiometric thresholds in the frequency range of 8000 – 16000 Hz (Dhrruvakumar, Shambhu, & Konadath, 2021).

Despite the wide range of frequencies that humans can hear, routine testing is usually performed from 250 Hz to 8 kHz. The possible reasons for this could be (1) most of the everyday sounds that we hear fall into this range, and (2) variability of thresholds below 8 kHz is less (Beiter and Talley, 1976). Complex physical interactions of high-frequency pure tones in the ear canal result in standing waves that increase the intra and intersubject variability of hearing thresholds for the high-frequency range (Schmuziger, Probst, & Smurzynski, 2004).

The extended high-frequency audiometry measures the audiometric thresholds in a frequency range of 8 kHz to 16 or 20 kHz. It is a suitable method for the early detection of ototoxicity in adults and children when supplemented with an objective test (Fausti, Henry, Hayden, Phillips, & Frey, 1998; Beahan, Kei, Driscoll, Charles, & Khan, 2012). However, such testing should not be used alone in isolation because a lack of change in

thresholds for the extended high-frequency range does not imply unchanged hearing thresholds in the conventional audiometric range (Campbell & Durrant, 1993). In ototoxicity, significant threshold shifts are first seen in the extended high-frequency region and then later spread to the conventional frequency region (Fausti et al., 1992, 1993). Similarly, the early effects of the occupational noise and impact noise exposure also occur in the extended high-frequency region which can be observed even before the otoacoustic emissions show a discernible impact (Mehrparvar et al., 2014). Thus, extended high-frequency audiometry makes a promising and valuable tool for detecting early hearing damage due to medication and noise trauma.

However, test-retest reliability of the thresholds in extended high-frequency audiometry remains less investigated and more uncertain compared to conventional audiometry for several reasons (Margolis et al., 1993; Hunter et al., 1996). This could be because of the lack of normative data available for the same. Another reason is the less frequent use of extended high-frequency audiometry than the conventional pure tone audiometry.

Need of the study

As mentioned above, several studies have demonstrated the high utility of extended high-frequency audiometry, especially for the early detection of cochlear damage in persons exposed to ototoxic drugs and noise (Knight, Kraemer, Winter, & Neuwelt, 2007; Mehrparvar et al., 2014). However, despite being a promising and valuable tool for detecting the early effects of high-intensity noise, ototoxic drugs and dietary problems (Downs & Northern, 1971), extended high-frequency audiometry is not

used extensively in standard clinical practices. One among the several reasons for this could be the lack of information about its test-retest reliability.

The test-retest reliability is an essential parameter for the clinical utility of any test, especially if the test has to be used repeatedly for monitoring the hearing status. The use of extended high-frequency audiometry for hearing status monitoring represents a similar scenario where the post-treatment or post-noise exposure thresholds are compared with the pre-treatment or pre-exposure baseline. In such a case, it is crucial to know the extent of variation in the thresholds, even without an intervention, to judge whether or not the changes being observed are due to the intervention (use of medication or exposure to noise). Therefore, knowing the test-retest reliability of the extended high-frequency audiometry is pertinent. However, there is a paucity of studies on the test-retest reliability of the thresholds obtained in the extended high-frequency region of the human hearing. A study reported high variability of the extended high-frequency audiometry, especially at 14 and 16 kHz (Schmuziger, Probst, &Smurzynski, 2004). They reported intrasession threshold changes exceeding 10 dB HL at 16 kHz in 6% of the otologically healthy subjects. However, this is the only study to have reported the test-retest reliability of the extended high-frequency audiometric measures using adult subjects. There is a need for more studies to ensure better generalization of the findings.

Further, Schmuziger et al. (2004) evaluated only the intrasession test-retest reliability. However, as in cases with ototoxicity, several applications require two or more measurements separated in time by days to weeks, hence requiring removal and placement of transducer again. However, no study evaluated the intersession test-retest reliability, especially estimated across multiple sessions.

Therefore, due to the absence of literature reporting the test-retest reliability of extended high-frequency audiometry along with conventional audiometry, there is a strong need to take a fresh look at the test-retest reliability of pure tone audiometry from 250 Hz to 16 kHz.

Aim of the study

This study aims to evaluate the inter-session test-retest reliability of pure tone audiometric thresholds measured in the frequency range of 250 Hz -16 kHz for a group of otologically healthy subjects.

Objectives of the study

To fulfil the above-stated aim, the following objectives were taken up:

1. To compare pure-tone thresholds obtained at different time intervals.
2. To find the variations in pure tone thresholds among the measurement points.

CHAPTER II

REVIEW OF LITERATURE

Pure tone audiometry (PTA) is a behavioural measure for assessing an individual's hearing abilities (Kutz, Mullin, & Campbell, 2018). PTA assesses the integrity of the auditory system through two pathways: air conduction and bone conduction. It obtains frequency-specific hearing thresholds using tonal stimuli (Sliwinska-Kowalska, 2015). A pure tone of a specific frequency is routed through the appropriate transducers, and the minimum level in decibels Hearing Level (dB HL) at which the stimulus can be heard 50% of the time is considered the threshold of the particular frequency. In Air Conduction testing, pure tone stimulus is delivered through the headphones or earphones, and it reaches the cochlea by entering through the external auditory canal and passing through the tympanic membrane and ossicular chain. In contrast, in bone conduction testing, stimuli bypass the outer and middle ear and reach the cochlea through vibrations of the bone vibrator placed on the mastoid process or sometimes on the forehead (Davies, 2016).

Routine testing involves measuring thresholds at octaves and mid-octaves from 250 Hz to 8000 Hz and 250 Hz to 4000 Hz for air and bone conduction testing, respectively (ASHA, 2005; BSA, 2018). Testing the threshold at 125 Hz is recommended only in case of a low-frequency hearing loss (ASHA, 2005). The rationale for the use of the frequency range mentioned above in audiometric testing could be (1) this range encompasses the frequencies essential for speech perception, and (2) variability of thresholds below 8000 Hz is less (Beiter& Talley, 1976).

Thresholds obtained in pure tone audiometry provide the necessary diagnostic information about the degree, type and configuration of the hearing loss, which in turn assist healthcare professionals in determining the possible aetiology, the line of treatment and prognosis (Musiek, Shinn, Chermak, & Bamiou, 2017). In addition, pure tone thresholds are useful (1) as a baseline measure for hearing conservation programmes, (2) for monitoring the effects of treatment or exposure to noise, (3) for determining candidacy for a cochlear implant or hearing aid, (4) for selecting the frequency gain characteristics of a hearing aid, and (5) for providing a reference level for presentation of speech testing, and acoustic reflex testing (Schlauch & Nelson, 2015).

The extended high-frequency audiometry (EHFA) measures the audiometric thresholds in a frequency range of 8 kHz to 16 or 20 kHz. It might be a suitable and more sensitive method for detecting early signs of hearing impairment than conventional audiometry (Somma et al., 2008). Although no data is currently available for humans, it has been found in animal experiments that damage caused by broad-spectrum noise is first seen in the outer hair cells of the basal region of the cochlea (Hamernik et al., 1989) and that even the slightest damage to these hair cells can cause hearing loss (Prosen et al., 1990). When supplemented with an objective test, EHFA has been found useful for the early detection of ototoxicity in adults and children (Fausti et al., 1998; Beahan et al., 2012). In ototoxicity, significant threshold shifts are first seen in the extended high-frequency region, which later spread to the conventional frequency region (Fausti et al., 1992, 1993).

Furthermore, in patients with tinnitus and normal pure tone audiometry results, EHFA provides valuable additional information. It might be of therapeutic value in

counselling patients with tinnitus who have normal conventional audiometric thresholds but impaired high-frequency thresholds (Vielsmeier et al., 2015). In addition, EHFA is used to differentiate between noise-induced hearing loss and other high-frequency sensorineural hearing loss such as presbycusis (Laukli & Mair, 1985). Despite all these advantages, extended high-frequency audiometry is not used extensively in standard clinical practices due to a lack of information about its test-retest reliability, lack of international standards, difficulty in calibration and large intersubject variability (Borchgrevink et al., 1996).

Various methods are used to measure pure tone thresholds. The procedure used for obtaining audiometric thresholds is based on the psychophysical Method of Limits (Fechner,1860). In the method of limits, the tester has control over the stimulus. Both ascending and descending trials are employed, and intensity increments are fixed and equal. A comprehensive description of psychophysical measures in audiometry and the effects these procedures have upon audiometric thresholds was provided by Hirsh (1952). He suggested that testing should start with a 1000 Hz tone, and the frequencies higher or lower than 1000 Hz should be tested after the threshold for 1000 Hz is obtained. Compared to the method of limits, the Hughson-Westlake (1944) method always employs ascending run for stimulus presentation. Carhart and Jerger (1959) revised the Hughson-Westlake method and used 5 dB increments and 10-15 dB decrements for ascending and descending runs, respectively, because it yielded consistent measures of the unadapted level of hearing acuity (Carhart & Jerger, 1959).

In contrast to the methods used for manual pure tone audiometric evaluation, Bekesy (1947) developed an audiometer that enabled the subjects to track their own

thresholds. Bekesy audiometry is based on the classical psychophysical Method of Limits and Method of Adjustments and the adaptive procedure. However, several factors can affect the results of pure tone audiometry.

2.1 Factors affecting Pure Tone Audiometry

Although the pure tone thresholds provide valuable information, the audiometric results may be affected by several variables. These variables include the acoustic, patient, environmental, instrumentation and clinician variables.

2.1.1 Patient variables

In some patients, particularly the elderly population, when earphones are placed on the ears, due to pressure from the headphones and decreased skin elasticity of the external auditory canal, it causes a collapsed ear canal resulting in a high-frequency conductive loss and lack of test-retest reliability (Ventry et al., 1961). Another factor which affects the measurement of thresholds is the standing waves. It results from the complex interaction of high-frequency pure tones in the ear canal that increases the intra and intersubject variability of hearing thresholds for the high-frequency range (Schmuziger et al., 2004).

False responses (false positive or negative) can also affect the threshold measurement. False-positive responses are hard to distinguish from the actual response when they occur in the contiguity of the stimulus and may jeopardize the determination of valid thresholds (Dancer et al., 1976). The persistence effect (errors of anticipation and habituation) can also serve as a source of variation in the threshold. The subject may

change the response before it is applicable (errors of anticipation) or continue to respond even when the response is no longer applicable (errors of habituation).

Some patients with retrocochlear pathology with marked adaptation can display large intrasubject variability in thresholds. Carhart and Jerger (1959) recommended using brief tonal durations to minimize the adaptation, maximize the on-effect phenomenon and enhance the threshold reliability.

2.1.2 Instrumentation and Environmental Variables

Inadequate calibration is one of the most common variables affecting the accuracy of the pure tone thresholds. According to the Professional Services Board of the American Speech-Langauge-Hearing Association (ASHA, 1978b), electroacoustic calibration of tones, masking noise, force levels for headphones and bone vibrators should be done quarterly. Furthermore, ambient noise is another variable affecting audiometric measurement accuracy. Ambient noise above the maximum permissible noise levels can interfere with the threshold measurement.

Inaccurate thresholds can be obtained if the earphones are not plugged or incompletely plugged into the jacks or if there is an excessive winding of the earphone wires. In addition, high temperatures exceeding 85ºF and lack of ventilation inside the booth can affect the accuracy of the pure tone thresholds, especially in children (Wilber, 1979).

2.1.3 Clinician variables

Inappropriate placement of headphones, earphones or bone vibrators can lead to a discrepancy in the threshold measurement. Furthermore, providing the subject with visual cues every time the sound is presented can affect the accuracy of thresholds. In addition, instructions that are not language-appropriate for the subject may result in a discrepancy in the audiometric measurements.

2.2 Test-retest reliability of Pure Tone Audiometry

As discussed above, pure tone thresholds are highly affected by extraneous variables; hence the test-retest reliability of pure tone audiometry becomes a critical affair to be discussed. Therefore there followed a period of reliability and validity studies of pure tone audiometry. Various studies have measured the test-retest reliability of pure tone audiometry results using different transducers. Few studies have suggested using insert earphones to overcome the limitations of headphones (König, 2009; Barbara et al., 2009; Stuart et al., 1991). However, few studies have shown a high test-retest consistency for both transducers and that ER-3A insert earphones produced similar results as TDH-50 headphones (Larson et al., 1998).

Another similar study by Schmuziger et al. (2004) reported an excellent intrasession test-retest repeatability from 0.5 to 12.5 kHz at each frequency for both transducers. Ishak et al. (2011) also reported excellent inter-session test-retest variability of hearing thresholds across different individual frequencies from 0.25 to 8.0 kHz in normally hearing people. When a typical sound-treated room and a natural room for pure tone audiometric measurements were used, it showed average threshold differences within specific test-retest reliability limits (Maclennan-Smith et al., 2013).

The above-stated literature shows that pure tone audiometry has high test-retest reliability. Hence it is considered a 'gold standard' test for assessing auditory abilities (Sideris and Glattke, 2006). However, this statement holds good when the measurement is done till 8 kHz, as thresholds below 8 kHz show less variability (Beiter and Talley, 1976).

However, the test-retest reliability of extended high-frequency audiometry remains less investigated and more uncertain than conventional audiometry for several reasons (Margolis et al., 1993; Hunter et al., 1996). This could be because of the lack of normative data available for the same. Another reason might be the less frequent use of the EHFA than conventional audiometry.

A study by Schmuziger (2004) reported significantly increased intrasession variability as frequency increased especially the frequency range of 14 to 16 kHz showing large variability. They reported intrasession threshold changes exceeding 10 dB HL at 16 kHz in 6% of the otologically healthy subjects. They also concluded that measurements could be done in the extended high-frequency range of 8 to 14 kHz but not up to 16 kHz. However, another study by Beiter and Talley (1976) evaluated the test-retest reliability of EHFA from 8 to 20 kHz and reported no significant difference between the mean threshold values for the right and left ear at all the individual frequencies tested. They also reported that the average rate of change of threshold change is greater from 14 to 20 kHz compared to 8 to 14 kHz. They concluded that test-retest reliability measures indicated accurate repeatability of thresholds even after an extended period of time. However, the test-retest reliability measures were only done for 8 out of 41 subjects due to the unavailability of subjects.

Another study by Beahan et al. (2012) investigated the test-retest reliability of EHFA in children. It revealed good test-retest reliability of HFPTA with no significant difference in mean HFPTA thresholds across test and retest conditions for all age groups. However, an age effect was seen in the test-retest reliability; the 4 to 6 years group showed a significantly lower percentage of normal variability at 14 kHz than 7 to 9 years and 10 to 13 years group.

All the above-mentioned studies have evaluated only the intrasession test-retest reliability, and the study which tested intersession test-retest reliability used a small sample size (n=8). In cases with ototoxicity, several applications require two or more measurements separated in time by days to weeks, hence requiring removal and placement of transducer again. However, no study evaluated the inter-session test-retest reliability, especially estimated across multiple sessions. Therefore, the test-retest reliability of the extended high-frequency audiometry remains relatively unknown, especially over several sessions.

CHAPTER III

METHODS

To accomplish the objectives of this study, the data was taken from the AIISH Research Fund(ARF) project titled "Impact of Acoustic Stimuli Used for Various Measures of VEMP on the Auditory System" with the sanction number SH/CDN/ARF-61/2016-17 dated 14.09.2016 and Dr. Niraj Kumar Singh as the principal investigator. The project's primary goal was to check the effects of stimuli used for evoking vestibular-evoked myogenic potentials (VEMP) on hearing function. The participants had undergone a unilateral VEMP test. However, audiometric measurements were done in both ears at several time points before and after VEMP. Hence, the data of the non-VEMP ears were analyzed to report the test-retest reliability of pure tone thresholds in the present study.

Participants

The secondary data of 160 ears (160 subjects) in the age range of 18 to 35 years were taken in the study. Each participant in the study was explained about the experiment and the option of dropping out of the study at any point in time if they wished to. Before recruiting them to the study, informed written consent was obtained, and they were not paid for participation.

Participant selection criteria

A detailed case history was taken to rule out the history or complaint of ear discharge, ear pain, itching sensation, tinnitus, vertigo, migraine, headache or any other medical or surgical history related to the audio-vestibular disorders. All study

participants had a normal audio-vestibular system which was ascertained through an audio-vestibular test battery. An otoscopic examination was done to check for visible abnormalities of the outer ear and tympanic membrane. The tests used to evaluate the auditory system included pure tone audiometry, immittance evaluation and transient evoked otoacoustic emissions. Their hearing thresholds were measured from 250 Hz to 8000 Hz and were within 15 dB HL at each octave and mid-octave frequencies. The speech recognition threshold (SRT) was within ±12dB HL of the four-frequency pure tone average threshold, and the speech identification scores (SIS) were at least 95% for all participants. Further, all participants underwent immittance evaluation and had an 'A' or 'As' type tympanogram with both ipsilateral and contralateral acoustic reflexes present at 100 dB HL at 1000Hz. All participants had the presence of otoacoustic emissions, indicating normal outer hair cell functioning and also normal middle ear functioning.

Test environment

All the tests were carried out in well-illuminated, air-conditioned sound-treated rooms with the ambient noise levels within the acceptable limits of the specifications given by the American National Standard Institute (ANSI S3.1, 1999, R2013). Among the tests mentioned above, pure tone audiometry was carried out in a double room set-up, whereas the remaining tests were performed in a single room set-up.

Instrumentation

A calibrated Madsen Astera clinical audiometer with impedance matched Telephonic TDH-50 supra-aural headphones and a set of calibrated Sennheiser

HDA 200 circum-aural headphones were used for conventional pure tone audiometry and extended high-frequency audiometry, respectively. Radioear B-71 bone vibrator was used to obtain bone conduction thresholds. A calibrated GSI-tympstar Pro clinical immittance meter was used for tympanometry and reflexometry. A calibrated Mimosa Acoustics otoacoustic emission system was used for transient evoked otoacoustic emissions.

Procedure

Initially, a detailed case history in the form of an interview was taken from all the participants before the commencement of the audiological evaluation. During this, the participants were asked about the history of auditory problems such as otitis externa, occlusion due to earwax, and otitis media. The structural abnormalities, such as the presence of stenosis, atresia etc., were ruled out through visual inspection. The participants were also asked if they had ear pain, itching sensation, and the presence of tinnitus or any ear-related surgeries. The participants were also enquired about a blocked nose andfullness in the ear due to a cold present during testing. If so, the participants were either excluded from the study or asked to report back once the cold had resolved. An otoscopic examination was done to check for visible abnormalities of the outer ear and tympanic membrane.

Immittance evaluation (tympanometry & reflexometry) was done to rule out middle ear pathology. Tympanometry was carried out using a probe tone frequency of 226 Hz. For this, the pressure in the ear canal was varied from -400 daPa to +200 daPa at 50 daPa/s. Using the same probe-tone frequency, both ipsilateral and

contralateral acoustic reflexes were obtained for a stimulus frequency of 1000Hzpresented at 100 dB HL.

For recording TEOAE, non-linear clicks were delivered at 80 dB peSPL through the probe assembly placed in the ear canal. The parameter noted was the global signal-to-noise ratio. The global signal-to noise ratio of >6dB was considered for the presence of TEOAEs.

Measuring pure tone thresholds

For pure tone thresholds measurement, the participant was seated on a comfortable chair in an upright position. The participant was then instructed to sit quietly without speaking and were told that headphones would be positioned over their ears. Further, the participant was informed to respond with their finger raised every time a tone was heard, regardless of how faint it was, and to forestall responding whilst the tone goes off (ASHA, 2005). Instructions were given in a language appropriate for the participant.

For conventional pure tone audiometry (250 Hz-8000 Hz), pure tones were presented using a calibrated Madsen Astera clinical audiometer with impedance-matched Telephonic TDH-50 supra-aural headphones. Before starting with the actual threshold measurement procedure, the familiarization task was done using a 1000 Hz pure tone continuously on but attenuated by gradually increasing the pure tone level until it was heard. Then, a 1000 Hz pure tone at 30 dB HL was presented, and if a clear response was elicited, threshold measurement was begun (ASHA, 2005).

A modified Hughson-Westlake procedure was used for threshold measurement (Carhart & Jerger, 1959). A pure tone of 1000 Hz was presented at a level well above the threshold at 40 dB HL for a duration of 1-2 seconds. If a response was obtained, the intensity of pure tone was decreased by 10 dB steps. When inaudibility was reached, the intensity was increased in 5 dB steps until a response was obtained. When a response was obtained, the intensity was again decreased in 10 dB steps, and another ascending trial (5 dB increments) began. The criterion for threshold was the lowest intensity at which three responses were obtained on ascending runs. The stimuli were separated by toneless intervals. Only the right ear was tested for all the participants. The sequence of frequencies tested was 1000 Hz, followed by 1500, 2000, 3000, 4000, 6000, and 8000 Hz, followed by a retest at 1000 Hz before testing 750, 500 and 250 Hz (ASHA, 2005).

After obtaining the thresholds for conventional pure tone audiometry, TDH-50 supra-aural headphones were removed, and Sennheiser HDA 200 circum-aural headphones were placed to obtain thresholds for extended high-frequency audiometry (8-16 kHz). Again a modified Hughson-Westlake procedure was used for threshold measurement. The frequencies tested were 8000, 9000, 10000, 11000, 12500, 14000 and 16000 Hz. The rest of the procedure remained the same as used for conventional pure tone audiometry.

Bone conduction threshold measurement was also done using a modified Hughson-Westlake procedure for the octave frequencies from 250 to 4000 Hz. SRT was obtained using a standard spondee word list, and SIS was obtained using a standard phonetically balanced word list in the participants' native language.

Hearing thresholds were obtained five times for each participant with both TDH-50 supra-aural headphones and Sennheiser HDA 200 circum-aural headphones. The first three measurements were done with 30 minutes intervals between the first two and 90 minutes between the first and third. The fourth measurement was done on day 2 (24 hours after the first one), and the fifth measurement after seven days from the first measurement. After each measurement, earphones were removed and placed again in order to mimic such a scenario likely to happen in clinical settings.

Statistical analyses

To check for the required sample size, G*Power software version 3.1.9.4 was used. Effect size (0.37), α error probability (0.005) and power (0.95) were used, which calculated a sample size of 156 subjects to be used. For the statistical analyses, Statistical Package for Social Sciences (SPSS)softwareversion20 was used. The Shapiro-Wilk's normality test was used to decide about the use of parametric or non-parametric procedures. It was found that the data distribution was not normal ($p > 0.05$) for several frequencies across the measurement points; hence, non-parametric tests were carried out. Friedman's test was administered for the within-group comparison of thresholds between the measurement points. If data showed significant differences among the measurement points, a Wilcoxon signed-rank test was administered for pair-wise comparison between the measurement points. As multiple comparisons were made, α correction factor was used to adjust the individual p-value and reduce the chances of committing type 1 error. It was calculated by dividing the original α level (0.05) by the number of comparisons

made (10) which gave a value of p=0.005. Additionally, the intra-class correlation coefficient was obtained to categorize the test-retest reliability of thresholds at various frequencies.

CHAPTER IV

RESULTS

The present study aimed to investigate the test-retest reliability of pure tone audiometry from 250 Hz to 16 kHz in otologically healthy individuals across multiple measurement points. In order to achieve the above-stated aim, each of the 160 participants underwent pure tone audiometric measurements from 250 Hz to 16 kHz for the right ear. Hearing thresholds were obtained at five-time points for the frequency range mentioned above. The first three measurements were done with 30 minutes intervals between the first two and 90 minutes between the first and third. The fourth measurement was done on day 2 (24 hours after the first one), and the fifth measurement after seven days (1 week) from the first measurement.

4.1 The within-group comparison of Thresholds between the measurement points

The thresholds obtained at all five measurement points were subjected to descriptive statistics in order to obtain the mean, median, standard deviation and interquartile range. The outcome of the descriptive statistics is shown in Table 4.1.1. The mean of thresholds at all five measurement points for all the frequencies is shown in Figure 4.1.1.

The analysis of individual data showed that a greater than 5 dBHL (=/>10 dBHL) was observed in ≤11 subjects at any given frequency. The outcome of the analysis of individual data is shown in Table 4.1.2.

Table 4.1.1 *Audiometric thresholds at various measurement points*

	Baseline			After 30 minutes			After 1 hour 30 minutes			After 24 hours			After 1 week		
	Mean (SD.)	**Median**	**IQR**	**Mean (SD.)**	**Median**	**IQR**	**Mean (SD.)**	**Median**	**IQR**	**Mean (SD.)**	**Median**	**IQR**	**Mean (SD.)**	**Median**	**IQR**
250 Hz	3.77 (5.14)	5.00	5.00	3.33 (4.66)	5.00	5.00	3.36 (4.65)	5.00	5.00	3.11 (4.62)	5.00	5.00	2.98 (4.17)	5.00	5.00
500 Hz	4.77 (4.58)	5.00	10.00	4.40 (4.26)	5.00	5.00	4.43 (4.35)	5.00	5.00	4.11 (4.22)	5.00	5.00	4.18 (4.24)	5.00	5.00
750 Hz	4.18 (4.90)	5.00	5.00	3.68 (4.87)	5.00	5.00	4.02 (4.38)	5.00	5.00	4.02 (4.62)	5.00	5.00	3.99 (4.56)	5.00	5.00
1 kHz	3.86 (5.12)	5.00	5.00	2.86 (5.29)	5.00	5.00	3.23 (4.90)	5.00	5.00	2.79 (5.01)	0.00	5.00	3.30 (4.59)	5.00	5.00
1.5 kHz	3.05 (5.15)	5.00	5.00	2.95 (5.13)	5.00	5.00	2.79 (4.82)	5.00	5.00	3.20 (9.09)	5.00	5.00	2.45 (4.73)	0.00	5.00
2 kHz	2.04	0.00	10.00	1.28	0.00	10.00	1.88	0.00	5.00	1.76	0.00	5.00	1.50	0.00	5.00

Table continued to next page

Table continued from previous page

	(5.98)			(5.31)			(5.55)			(5.37)			(5.35)		
3 kHz	1.79	0.00	10.00	1.32	0.00	10.00	1.19	0.00	10.00	0.75	0.00	10.00	1.25	0.00	5.00
	(5.96)			(6.25)			(5.46)			(5.28)			(5.56)		
4 kHz	1.63	0.00	10.00	1.41	0.00	10.00	.94	0.00	10.00	1.00	0.00	10.00	0.88	0.00	10.00
	(6.12)			(5.82)			(5.57)			(5.59)			(5.66)		
6 kHz	1.72	5.00	10.00	1.44	0.00	10.00	1.29	0.00	10.00	1.12	0.00	10.00	0.81	0.00	10.00
	(6.82)			(6.54)			(6.31)			(6.11)			(5.97)		
8 kHz	-3.55	-5.00	10.00	-3.20	-5.00	10.00	-2.76	-5.00	5.00	-3.33	-5.00	5.00	-2.98	-5.00	5.00
	(7.29)			(7.09)			(6.57)			(5.97)			(6.21)		
9 kHz	-5.15	-5.00	10.00	-4.96	-5.00	10.00	-4.37	-5.00	10.00	-4.49	-5.00	10.00	-4.71	-5.00	10.00
	(8.10)			(7.98)			(7.74)			(7.20)			(6.73)		
10 kHz	0.94	0.00	10.00	-0.22	0.00	10.00	-0.53	0.00	10.00	-0.40	0.00	10.00	-0.47	0.00	10.00
	(7.36)			(7.25)			(7.21)			(7.09)			(6.67)		
11 kHz	-2.20	-5.00	15.00	-2.54	-5.00	10.00	-2.38	-5.00	15.00	-2.26	-5.00	5.00	-2.64	-5.00	10.00
	(8.24)			(7.71)			(7.39)			(7.26)			(6.69)		

Table continued to next page

Table continued from previous page

12.5 kHz	-1.94 (8.16)	-5.00	15.00	-2.32 (7.87)	-5.00	10.00	-2.23 (7.58)	-5.00	15.00	-2.38 (7.41)	-5.00	5.00	-2.54 (7.54)	-5.00	5.00
14 kHz	-5.00 (11.70)	-10.00	20.00	-4.84 (11.42)	-10.00	20.00	-4.90 (10.45)	-5.00	20.00	-4.52 (10.75)	-10.00	20.00	-4.40 (10.16)	-5.00	20.00
16 kHz	-.59 (15.65)	-5.00	30.00	-1.25 (15.31)	-5.00	25.00	-1.19 (14.65)	-5.00	25.00	-1.06 (14.15)	-5.00	25.00	-1.10 (13.91)	-5.00	25.00

Note: 'SD'- standard deviation, 'IQR'- interquartile range

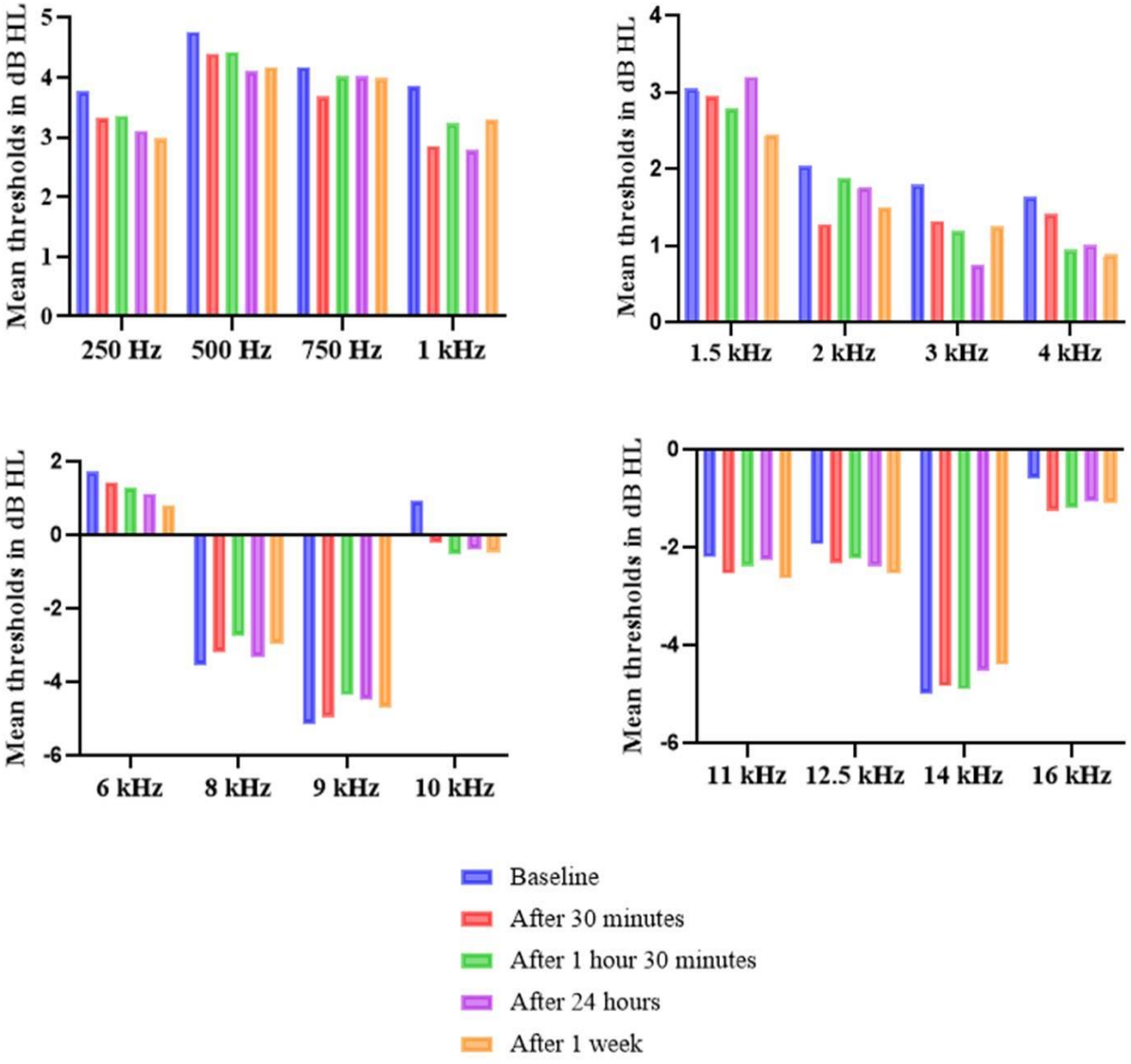

Figure 4.1.1 The mean of thresholds obtained at various measurement points

Table 4.1.2 *The outcome of analysis of individual data for each frequency*

Measurement points	**250 Hz**	**500 Hz**	**750 Hz**	**1 kHz**	**1.5 kHz**	**2 kHz**	**3 kHz**	**4 kHz**	**6 kHz**	**8 kHz**	**9 kHz**	**10 kHz**	**11 kHz**	**12.5 kHz**	**14 kHz**	**16 kHz**
Baseline-30 minutes	3	1	5	5	5	7	8	8	11	6	5	6	9	7	3	9
Baseline-1 hour 30 min	3	4	0	4	1	3	1	8	9	2	4	7	5	3	7	3
Baseline-24 hours	1	4	2	2	2	4	5	1	2	5	0	3	4	5	5	5
Baseline- 1 week	2	3	0	0	1	2	0	3	4	3	6	5	4	0	8	3
30 minutes-1 hour 30 minutes	1	2	1	2	3	1	2	1	2	1	2	5	1	3	3	2
30 minutes- 24 hours	2	1	1	1	0	1	2	4	2	1	3	2	0	1	5	2
30 minutes- 1 week	3	0	5	3	2	4	1	3	0	2	4	3	1	8	0	3
1 hour 30 minutes-24 hours	3	0	2	1	0	2	0	2	1	1	1	1	0	2	1	1
1 hour 30 minutes-1 week	1	1	1	5	1	3	2	1	1	1	3	2	1	2	1	1
1 day- 1 week	1	1	1	2	0	0	0	0	0	1	1	1	2	2	1	0

Further, the within-group comparison of thresholds between the measurement points was made. This comparison was made between the measurement points using Friedman's test. Table 4.1.3 shows the outcome of Friedman's test.

Table 4.1.3 *The outcome of Friedman's test for comparison of thresholds between the measurement points*

	χ^2 **value**	**p-value**
250 Hz	8.39	0.07
500 Hz	7.58	0.10
750 Hz	4.96	0.29
1 kHz	20.24	<0.001
1.5 kHz	9.76	0.04
2 kHz	9.08	0.05
3 kHz	13.30	0.01
4 kHz	9.82	0.04
6 kHz	13.71	0.00
8 kHz	8.18	0.08
9 kHz	11.45	0.02
10 kHz	3.93	0.41
11 kHz	1.88	0.75
12.5 kHz	3.88	0.42
14 kHz	4.35	0.36
16 kHz	4.97	0.29

As Friedman's test showed a significant difference between the measurement points for 1 kHz, 1.5 kHz, 3kHz, 4kHz, 6kHz, and 9kHz, data were further subjected to pair-wise comparison using the Wilcoxon Signed-Rank test. The outcome for the Wilcoxon Signed-Rank test is shown in Table 4.1.4.

Table 4.1.4 *The outcome of the Wilcoxon Signed Rank test for pair-wise comparison at each measurement point*

Measurement points	**1 kHz**		**1.5 kHz**		**3 kHz**		**4 kHz**		**6 kHz**		**9 kHz**	
	Z	**p**	**Z**	**P**	**Z**	**P**	**Z**	**p**	**Z**	**P**	**Z**	**P**
Baseline-30 minutes	**-4.01**	**<0.001***	-0.13	0.89	-1.58	0.11	-0.91	0.36	-0.99	0.32	-0.01	0.98
Baseline-1 hour 30 min	-2.39	0.01	-1.10	0.27	-2.31	0.02	-2.34	0.01	-1.42	0.15	-2.11	0.03
Baseline-24 hours	**-4.07**	**<0.001***	-1.60	0.10	**-3.76**	**<0.001***	-2.23	0.02	-2.38	0.01	-2.22	0.02
Baseline- 1 week	-2.17	0.03	-2.45	0.01	-2.36	0.01	-2.74	0.006	**-3.40**	**0.001***	-1.55	0.12
30 minutes-1 hour 30 minutes	-1.51	0.13	-0.75	0.45	-0.41	0.68	-1.24	0.21	-0.36	0.71	-2.06	0.03
30 minutes- 24 hours	-0.21	0.82	-1.14	0.25	-1.89	0.05	-1.13	0.25	-0.92	0.35	-1.79	0.07
30 minutes- 1 week	-1.42	0.15	-1.70	0.08	-0.24	0.80	-1.53	0.12	-1.89	0.05	-1.37	0.16
1 hour 30 minutes- 24 hours	-1.70	0.08	-.630	0.52	-1.70	0.08	-0.10	0.91	-0.67	0.50	-0.10	0.91
1 hour 30 minutes- 1 week	-0.13	0.89	-1.33	0.18	-0.12	0.90	-0.31	0.75	-1.59	0.11	-0.46	0.63
1 day- 1 week	-1.80	0.07	-.875	0.38	-1.94	0.05	-0.46	0.64	-1.08	0.27	-0.64	0.51

Note: *The p-value was considered significant when it was <0.005 [α-corrected p-value for multiple pair-wise comparisons].

There was a significant difference in thresholds at baseline to 30 minutes and baseline to 24 hours ($p<0.005$) for 1 kHz and only at baseline to 24 hours and baseline to 1 week ($p<0.005$) for 3 kHz and 6 kHz, respectively.

Further, the data was subjected to Intraclass Correlation coefficient (ICC) analyses to evaluate the test-retest reliability of thresholds. The average measure ICC for all the frequencies were ≥ 0.85. The outcome of the Intraclass Correlation coefficient test analysis is shown in Table 4.1.5

Table 4.1.5 *The outcome of the Intraclass Correlation coefficient test analysis*

Intraclass Correlation Coefficient								
		Intraclass Correlation	95% Confidence Interval		F Test with True Value 0			
			Lower Bound	Upper Bound	Value	df1	df2	Sig
250 Hz	**Single measures**	0.69	0.63	0.74	12.20	159	636	<0.001
	Average Measures	0.91	0.89	0.93	12.20	159	636	<0.001
500 Hz	**Single measures**	0.70	0.65	0.76	13.17	159	636	<0.001
	Average Measures	0.92	0.90	0.94	13.17	159	636	<0.001
750 Hz	**Single measures**	0.71	0.66	0.76	13.67	159	636	<0.001
	Average Measures	0.92	0.90	0.94	13.67	159	636	<0.001
1 kHz	**Single measures**	0.75	0.70	0.79	16.07	159	636	<0.001
	Average Measures	0.93	0.92	0.95	16.07	159	636	<0.001
1.5 kHz	**Single measures**	0.54	0.46	0.61	6.87	159	636	<0.001
	Average Measures	0.85	0.81	0.88	6.87	159	636	<0.001
2 kHz	**Single measures**	0.78	0.73	0.82	19.11	159	636	<0.001

Table continued to next page

Table continued from previous page

	Average Measures	0.94	0.93	0.95	19.11	159	636	<0.001
3 kHz	**Single measures**	0.81	0.77	0.85	23.35	159	636	<0.001
	Average Measures	0.95	0.94	0.96	23.35	159	636	<0.001
4 kHz	**Single measures**	0.75	0.70	0.80	16.43	159	636	<0.001
	Average Measures	0.93	0.92	0.95	16.43	159	636	<0.001
6 kHz	**Single measures**	0.80	0.76	0.84	22.02	159	636	<0.001
	Average Measures	0.95	0.94	0.96	22.02	159	636	<0.001
8 kHz	**Single measures**	0.83	0.79	0.86	25.89	159	636	<0.001
	Average Measures	0.96	0.95	0.97	25.89	159	636	<0.001
9 kHz	**Single measures**	0.83	0.80	0.87	26.85	159	636	<0.001
	Average Measures	0.96	0.95	0.97	26.85	159	636	<0.001
10 kHz	**Single measures**	0.82	0.78	0.86	24.96	159	636	<0.001
	Average Measures	0.96	0.94	0.96	24.96	159	636	<0.001
11 kHz	**Single measures**	0.84	0.81	0.87	28.87	159	636	<0.001
	Average Measures	0.96	0.95	0.97	28.87	159	636	<0.001
12.5 kHz	**Single measures**	0.84	0.81	0.87	28.54	158	632	<0.001

Table continued to next page

Table continued from previous page

	Average Measures	0.96	0.95	0.97	28.54	158	632	<0.001
14 kHz	**Single measures**	0.90	0.88	0.92	48.07	159	636	<0.001
	Average Measures	0.97	0.97	0.98	48.07	159	636	<0.001
16 kHz	**Single measures**	0.96	0.95	0.97	129.50	159	636	<0.001
	Average Measures	0.99	0.99	0.99	129.50	159	636	<0.001

CHAPTER V

DISCUSSION

The present study aimed to evaluate the test-retest reliability of pure tone audiometry from 250 Hz to 16 kHz across multiple sessions. To achieve the above-stated aim, for each subject, thresholds were obtained using a 5-dB step-size at five-time points across sessions within one week. These time points were 30 minutes, 1 hour 30 minutes, 24 hours and one week after the first baseline measurement. The data used in this study was the secondary data of a project funded under the aegis of AIISH Research Fund, and therefore only unilateral data could be used for examining the test-retest reliability.

5.1 The within-group comparison of thresholds between the measurement points

The mean and median difference between any two measurement points remained within 5 dB HL at every frequency. This possibly points to highly repeatable responses. This finding was in accordance with a study which showed 1-week threshold variability within 5 dB from 8 to 14 kHz for 80–90% and at 16 kHz for 58% of individual ears (Valente et al., 1992). However, the standard deviation increased with the frequency increase, indicating a slightly increased threshold variability at higher frequencies.

The results of the within-group comparison of thresholds between the measurement points showed no significant difference except for frequencies of 1 kHz, 3 kHz, and 6 kHz, which showed a significant difference ($p<0.05$). This significant difference could be attributed to the chance factor, as there was no pattern to these differences.

The reliability measure used in the present study was the Intraclass Correlation Coefficient test. Versiono et al. (2003) recommended that an ICC value of >0.7 be considered to represent excellent test-retest reliability, and 0.4-0.7 and <0.4 be considered moderate and poor test-retest reliability, respectively. Using this reference point, the present study's data had an excellent test-retest reliability for thresholds across the frequencies from 250 Hz to 16 kHz. This finding was in accordance with a study by Beiter and Talley (1976) that evaluated the test-retest reliability of EHFA from 8 to 20 kHz and reported that there was no significant difference between the mean threshold values for the right and left ear at all the individual frequencies tested. They also reported that the average rate of change of threshold is greater from 14 to 20 kHz compared to 8 to 14 kHz. They concluded that test-retest reliability measures indicated accurate repeatability of thresholds even after an extended period of time. However, the test-retest reliability measures were only done for 8 out of 41 subjects due to the unavailability of subjects. Another study by Beahan et al. (2012) also investigated the test-retest reliability of EHFA in children and revealed good test-retest reliability of HFPTA with no significant difference in mean HFPTA thresholds across test and retest conditions for all age groups. Further, at any given frequency, the variation of 10 dB HL or higher was observed in ≤11 (≤6.87%). Schmuziger (2004) reported intrasession threshold changes exceeding 10 dB HL at 16 kHz in 6% of the otologically healthy subjects. They also concluded that measurements could be done in the extended high-frequency range of 8 to 14 kHz but not up to 16 kHz. Therefore, the findings of the present study are in consonance with those reported previously.

The reason for such discrepancies in the findings could be that in previous attempts at studying test-retest reliability of pure tone audiometry, their conclusions were based on findings of repeated measures ANOVA (Beahan et al., 2012) or Wilcoxon signed rank test (Schmuziger, 2004). Unfortunately, these are measures of central tendencies and therefore compare means, medians or ranks and not the performance of the same individual on two or more occasions. Thus, the lack of group difference in mean, median or rank shown by these studies (Beahan et al., 2012; Schmuziger, 2004) does not truly represent the test-retest reliability of pure tone audiometry. The computation of the Intraclass Correlation Coefficient test is a more desirable measure for measuring test-retest reliability as it reflects both degrees of correlation and agreement between measurements (Koo & Li, 2016) and, hence using this statistical analysis is an appropriate method. Therefore the present study is the first attempt, to the best of our knowledge, at evaluating the test-retest reliability of pure tone audiometry using the Intraclass Correlation Coefficient test and findings show an excellent test-retest reliability for thresholds across the frequencies from 250 Hz to 16 kHz.

The previous studies have considered only two sessions for assessing the intrasession test-retest reliability of pure tone audiometry, and the study which tested intersession test-retest reliability used a small sample size (n=8), which might not be appropriate to explain the extent of variability. Several applications require two or more measurements separated in time by days to weeks, hence requiring removal and placement of transducer again. Especially if the test has to be used repeatedly to monitor the hearing status post-treatment or post-noise exposure, the thresholds are compared with the pre-treatment or pre-exposure baseline. In such cases, it is crucial to

know the extent of variation in the thresholds, even without an intervention, to judge whether or not the changes being observed are due to the intervention. Hence considering many sessions will give a clear idea about the variability of thresholds. Therefore, the present study's findings highlight the extent of variability and confirm an excellent test-retest reliability of pure tone audiometry from 250 Hz to 16 kHz. Also, the results of the present study are generalizable since present study had a big data set of 160 subjects.

CHAPTER VI

SUMMARY AND CONCLUSION

Pure tone audiometry is a procedure to obtain frequency-specific hearing thresholds using tonal stimuli (Sliwinska-Kowalska, 2015). While the conventional pure tone audiometry involves frequencies ranging from 250 - 8000 Hz, the extended high-frequency audiometry measures the audiometric thresholds in the frequency range of 8000 – 16000 Hz. EHFA is a suitable method for the early detection of ototoxicity in adults and children when supplemented with an objective test. In ototoxicity, significant threshold shifts are first seen in the extended high-frequency region and then later spread to the conventional frequency region. Similarly, the early effects of the occupational noise and impact noise exposure also occur in the extended high-frequency region, which can be observed even before the otoacoustic emissions show a discernible impact. Thus, extended high-frequency audiometry makes a promising and valuable tool for detecting early hearing damage due to medication and noise trauma.

One factor paramount to the above-mentioned utility is the test-retest reliability. However, test-retest reliability of the thresholds in extended high-frequency audiometry remains less investigated and more uncertain than in conventional audiometry for several reasons. The test-retest reliability is an essential parameter for the clinical utility of any test, especially if the test has to be used repeatedly for monitoring the hearing status. The only known study exploring test-retest reliability of extended high-frequency audiometry showed questionable reliability of thresholds obtained at 14 and 16 kHz; however, it evaluated only the intrasession test-retest reliability. Nonetheless, several applications, as in cases with ototoxicity, require two or more measurements separated in time by days to

weeks, hence requiring removal and placement of transducer again. Despite this, no study has evaluated the inter-session test-retest reliability, especially estimated across multiple sessions. Therefore, the present study aimed to evaluate the inter-session test-retest reliability of pure tone audiometric thresholds measured in the frequency range of 250-16000 Hz for a group of otologically healthy subjects.

To fulfill the above-stated aim, auditory thresholds in the frequency range from 250Hz to 16 kHz were measured in one ear of 160 otologically healthy subjects in the age range of 18-35 years. For each subject, thresholds were obtained using a 5-dB step-size at five-time points across sessions within one week. These time points were 30 minutes, 1 hour 30 minutes, 24 hours and one week after the first baseline measurement. Friedman's test was administered for the within-group comparison of thresholds among the measurement points. In case of a significant difference on Friedman's test, a Wilcoxon signed-rank test was planned for pair-wise comparison between the measurement points. Additionally, the intra-class correlation coefficient was obtained to categorize the test-retest reliability of thresholds at various frequencies.

The within-group comparison of thresholds between the measurement points showed no significant difference except for frequencies for 1 kHz, 3 kHz, and 6 kHz. This significant difference could be attributed to the chance factor, as there was no pattern to such differences.

The reliability measure used in the present study was the Intraclass Correlation Coefficient test. Versiono et al. (2003) recommended that an ICC value of >0.7 be considered to represent excellent test-retest reliability. Using this reference point, the

present study's data had an excellent test-retest reliability for thresholds across the frequencies from 250 Hz to 16 kHz. This finding was in accordance with a study by Beiter and Talley (1976) that evaluated the test-retest reliability of EHFA from 8 to 20 kHz and reported that there was no significant difference between the mean threshold values for the right and left ear at all the individual frequencies tested. They concluded that test-retest reliability measures indicated accurate repeatability of thresholds even after an extended period of time. Beahan et al. (2012) also investigated the test-retest reliability of EHFA in children. They revealed good test-retest reliability of HFPTA with no significant difference in mean HFPTA thresholds across test and retest conditions for all age groups. Further, at any given frequency, the variation of 10 dB HL or higher was observed in ≤11 (≤6.87%). Schmuziger (2004) reported intrasession threshold changes exceeding 10 dB HL at 16 kHz in 6% of the otologically healthy subjects. They also concluded that measurements could be done in the extended high-frequency range of 8 to 14 kHz but not up to 16 kHz. Therefore, the findings of the present study are in consonance with those reported previously.

The reason for such discrepancies in the findings could be that, in previous attempts at studying test-retest reliability of pure tone audiometry, their conclusions were based on findings of repeated measures ANOVA (Beahan et al., 2012) or Wilcoxon signed rank test (Schmuziger, 2004). Unfortunately, these are measures of central tendencies and therefore compare means, medians or ranks and not the performance of the same individual on two or more occasions. The computation of the Intraclass Correlation Coefficient test is a more desirable measure for measuring test-retest reliability as it reflects both degrees of correlation and agreement between measurements

(Koo & Li, 2016) and, hence using this statistical analysis is an appropriate method. Therefore the present study is the first attempt, to the best of our knowledge, at evaluating the test-retest reliability of pure tone audiometry using the Intraclass Correlation Coefficient test and findings show an excellent test-retest reliability for thresholds across the frequencies from 250 Hz to 16 kHz. Also, the results of the present study are generalizable since the present study had an extensive data set of 160 subjects.

Clinical implications of the study

With its promising test-retest reliability, pure-tone audiometry, including the extended high-frequency audiometry, lends itself to applications requiring multiple points measurements, such as monitoring the damage due to noise exposure and ototoxic drugs and monitoring the outcomes of surgeries. The findings of the present study confirm that such utility may indeed be possible.

Limitations of the study and future directions

The present study measured auditory thresholds in only one ear (right ear) of subjects. Measurements of thresholds for both ears might have given an understanding of whether or not there is a difference in auditory thresholds between two ears across multiple sessions. Additionally, only adult subjects (18-35 years) were taken in the study. Including other age groups like children and older adults would have provided an insight into whether age affects conventional and extended high-frequency audiometric thresholds.

REFERENCES

ANSI/ASA S3.1-1999 (R2013) - Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms.

Beahan, N., Kei, J., Driscoll, C., Charles, B., & Khan, A. (2012). High-frequency pure-tone audiometry in children: A test–retest reliability study relative to ototoxic criteria. *Ear and hearing*, *33*(1), 104-111.

Beiter, R. C., & Talley, J. N. (1976). High-frequency audiometry above 8 000 Hz. *Audiology*, *15*(3), 207-214.

Békésy, G. V. (1947). A new audiometer. *Acta Oto-Laryngologica*, *35*(5-6), 411-422.

Bell-Lehmkuhler, B., Meinke, D. K., Sedey, A., &Tuell, C. (2009). Reliability of audiometric thresholds obtained with insert earphones when used by certified audiometric technicians. *Noise and Health*, *11*(42), 59.

Borchgrevink, H. M., Hallmo, P., & Mair, I. W. (1996). Extended high-frequency hearing loss from noise exposure. *Scientific basis of noise-induced hearing loss. New York: Thieme*, 299-312.

Hirsh, I. J. (1952). *The measurement of hearing.* McGraw-Hill.

Campbell, K. C., &Durrant, J. (1993). Audiologic monitoring for ototoxicity. *Otolaryngologic Clinics of North America*, *26*(5), 903-914.

Carhart, R., &Jerger, J. F. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of speech and hearing disorders*, *24*(4), 330-345.

Dancer, J., Ventry, I. M., & Hill, W. (1976). Effects of stimulus presentation and instructions on pure-tone thresholds and false-alarm responses. *Journal of Speech and Hearing Disorders*, *41*(3), 315-324.

Davies R. A. (2016). Audiometry and other hearing tests. Handbook of clinical neurology, 137, 157–176.

Dhrruvakumar, S., Shambhu, T., & Konadath, S. (2021). Assessment of Hidden Hearing Loss in Individuals Exposed to Occupational Noise Using Cochlear, Neural, Temporal Functions and Quality of Life Measures. *Indian Journal of Otolaryngology and Head and Neck Surgery*.

Downs, M., & Northern, J. L. (1971). High-frequency hearing. *ASHA*, *13*, 541.

Fausti, S. A., Henry, J. A., Hayden, D., Phillips, D. S., & Frey, R. H. (1998). Intrasubject reliability of high-frequency (9-14 kHz) thresholds: tested separately vs. following conventional-frequency testing. *Journal of the American Academy of Audiology*, *9*(2).

Fausti, S. A., Henry, J. A., Schaffer, H. I., Olson, D. J., Frey, R. H., &Bagby, G. C. (1993). High-frequency monitoring for early detection of cisplatin ototoxicity. *Archives of Otolaryngology–Head & Neck Surgery*, *119*(6), 661-666.

Fausti, S. A., Henry, J. A., Schaffer, H. I., Olson, D. J., Frey, R. H., & McDonald, W. J. (1992). High-frequency audiometric monitoring for early detection of aminoglycoside ototoxicity. *Journal of Infectious Diseases*, *165*(6), 1026-1032.

Guidelines for Manual Pure-Tone Threshold Audiometry. (2005).

Hamernik, R. P., Patterson, J. H., Turrentine, G. A., &Ahroon, W. A. (1989). The quantitative relation between sensory cell loss and hearing thresholds. *Hearing research*, *38*(3), 199-211.

Hughson, W. A. I. T. E. R., & Westlake, H. (1944). Manual for program outline for rehabilitation of aural casualties both military and civilian. *Trans Am AcadOphthalmolOtolaryngol*, *48*(Suppl), 1-15.

Hunter, L. L., Margolis, R. H., Rykken, J. R., Le, C. T., Daly, K. A., &Giebink, G. S. (1996). High frequency hearing loss associated with otitis media. *Ear and hearing*, *17*(1), 1-11.

Isaacson, J., & Vora, N. M. (2003). Differential diagnosis and treatment of hearing loss. *American family physician*, *68*(6), 1125-1132.

Ishak, W. S., Zhao, F., Stephens, D., Culling, J., Bai, Z., & Meyer-Bisch, C. (2011). Test-retest reliability and validity of Audioscan and Békésy compared with pure tone audiometry. *Audiological Medicine*, *9*(1), 40-46.

Knight, K. R., Kraemer, D. F., Winter, C., &Neuwelt, E. A. (2007). Early changes in auditory function as a result of platinum chemotherapy: use of extended high-frequency audiometry and evoked distortion product otoacoustic emissions. *Journal of Clinical Oncology*, *25*(10), 1190-1195.

König, E. (1962). On the use of hearing-aid type earphones in clinical audiometry. *Acta Oto-Laryngologica*, *55*(1-6), 331-341.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, *15*(2), 155-163.

Kutz, Mullin, & Campbell. (2018). *Audiology Pure-Tone Testing: Overview, Indications, Contraindications.*

Larson, V. D., Cooper, W. A., Talbott, R. E., Schwartz, D. M., Ahlstrom, C., & De Chicchis, A. R. (1988). Reference threshold sound-pressure levels for the TDH-50 and ER-3A earphones. *The Journal of the Acoustical Society of America*, *84*(1), 46-51.

Laukli, E., & Mair, I. W. S. (1985). High-frequency audiometry normative studies and preliminary experiences. *Scandinavian Audiology*, *14*(3), 151-158.

Maclennan-Smith, F., Swanepoel, D. W., & Hall III, J. W. (2013). Validity of diagnostic pure-tone audiometry without a sound-treated environment in older adults. *International journal of audiology*, *52*(2), 66-73.

Margolis, R. H., Rykken, J. R., Hunter, L. L., &Giebink, G. S. (1993). Effects of otitis media on extended high-frequency hearing in children. *Annals of Otology, Rhinology & Laryngology*, *102*(1), 1-5.

Mehrparvar, A. H., Mirmohammadi, S. J., Davari, M. H., Mostaghaci, M., Mollasadeghi, A., Bahaloo, M., & Hashemi, S. H. (2014). Conventional audiometry, extended high-frequency audiometry, and DPOAE for early diagnosis of NIHL. *Iranian Red Crescent Medical Journal*, *16*(1).

Musiek, F. E., Shinn, J., Chermak, G. D., & Bamiou, D. E. (2017). Perspectives on the pure-tone audiogram. *Journal of the American Academy of Audiology*, *28*(07), 655-671.

Prosen, C. A., Moody, D. B., Stebbins, W. C., Smith, D. W., Sommers, M. S., Brown, J. N., ... & Hawkins Jr, J. E. (1990). Apical hair cells and hearing. *Hearing research*, *44*(2-3), 179-193.

Recommended Procedure Pure-Tone Air-Conduction and Bone-Conduction Threshold Audiometry with and without Masking, (2018).

Schlauch, R. S., Nelson, P (2015). Puretone Evaluation. *Handbook of Clinical Audiology* (pp. 30). Wolters Kluwer.

Schmuziger, N., Probst, R., &Smurzynski, J. (2004). Test-retest reliability of pure-tone thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 earphones. *Ear and hearing*, *25*(2), 127-132.

Shim, H. J., Kim, S. K., Park, C. H., Lee, S. H., Yoon, S. W., Ki, A. R., ... & Yeo, S. G. (2009). Hearing abilities at ultra-high frequency in patients with tinnitus. *Clinical and experimental otorhinolaryngology*, *2*(4), 169-174.

Sideris, I., &Glattke, T. J. (2006). A comparison of two methods of hearing screening in the preschool population. *Journal of Communication Disorders*, *39*(6), 391-401.

Sliwinska-kowalska, M. (2015). Hearing. *Handbook of Clinical Neurology*, *131*, 341–363.

Somma, G., Pietroiusti, A., Magrini, A., Coppeta, L., Ancona, C., Gardi, S., ... & Bergamaschi, A. (2008). Extended high-frequency audiometry and noise induced hearing loss in cement workers. *American Journal of Industrial Medicine*, *51*(6), 452-462.

Stuart, A., Stenstromb, R., Tompkins, C., &Vandenhoff, S. (1991). Test-retest variability in audiometric threshold with supraaural and insert earphones among children and adults. *Audiology*, *30*(2), 82-90.

Valente, M., Valente, L. M., & Goebel, J. (1992). High-frequency thresholds: circumaural earphone versus insert earphone. *Journal of the American Academy of Audiology*, *3*(6).

VENTRY, I. M., CHAIKLIN, J. B., & BOYLE, W. F. (1961). Collapse of the ear canal during audiometry. *Archives of Otolaryngology*, *73*(6), 727-731.

Vielsmeier, V., Lehner, A., Strutz, J., Steffens, T., Kreuzer, P. M., Schecklmann, M., ... & Kleinjung, T. (2015). The relevance of the high frequency audiometry in tinnitus patients with normal hearing in conventional pure-tone audiometry. *BioMed research international*, *2015*.

Wilber, L. A. (1979). Pure tone audiometry. *Hearing Assessment*, 1-38.