Diagnosis of hearing loss using automated audiometry in an asynchronous telehealth model: a pilot accuracy study

Christopher G. Brennan-Jones a b
Robert H. Eikelboom a b c
De Wet Swanepoel a b c

a) Ear Science Institute Australia, Subiaco, Perth, Western Australia
b) Ear Sciences Centre, School of Surgery, The University of Western Australia
c) Department of Speech-Language Pathology and Audiology, University of Pretoria

CONTACT INFORMATION:
Mr Christopher Brennan-Jones
chris.brennan-jones@earscience.org.au

Ear Science Institute Australia
Suite 1, Level 2, 1 Salvado Road
Subiaco, Western Australia 6008

Abbreviations:
Hz: Hertz
dB: Decibel
WHO: World Health Organisation

ABSTRACT:

Introduction: Standard criteria exist for diagnosing different types of hearing loss, yet audiologists interpret audiograms manually. This pilot study examined the feasibility of standardised interpretations of audiometry in a telehealth model of care. The aim of this study
was to examine diagnostic accuracy of automated audiometry in adults with hearing loss in an asynchronous telehealth model using pre-defined diagnostic protocols.

**Materials and Methods:** We recruited 42 study participants from a public audiology and otolaryngology clinic in Perth, Western Australia. Manual audiometry was performed by an audiologist either before or after automated audiometry. Diagnostic protocols were applied asynchronously for normal hearing, disabling hearing loss, conductive hearing loss and unilateral hearing loss. Sensitivity and specificity analyses were conducted using a two-by-two matrix and Cohen’s kappa was used to measure agreement.

**Results:** The overall sensitivity for the diagnostic criteria was 0.88 (range: 0.86 – 1) and overall specificity was 0.93 (range: 0.86 – 0.97). Overall Kappa (κ) agreement was “substantial” \( \kappa = 0.80 \) [95%CI 0.70, 0.89] and significant at \( p<0.001 \).

**Discussion:** Pre-defined diagnostic protocols applied asynchronously to automated audiometry provide accurate identification of disabling, conductive and unilateral hearing loss. This method has the potential to improve synchronous and asynchronous tele-audiology service delivery.

**KEYWORDS:** automated audiometry, audiometry, hearing loss, teleaudiology, accuracy

**INTRODUCTION**

There is a global shortage of audiological services that is not exclusive to low and middle-income countries.\(^1\) In developed economies, people living in rural areas have been shown to have higher rates of hearing loss with lower uptake of interventions\(^2\, 3\) and as many as 25% of children may require diagnostic audiometric testing to identify hearing loss associated with recurrent ear disease.\(^4\) Telehealth models of care utilising automated audiometers have therefore been proposed as a potential means to increase access to hearing services in underserved regions.\(^5\, 7\)

A number of automated audiometers have recently been clinically validated, including the AMTAS\(^8\, 9\) and KUDUwave\(^10\) and the consensus across studies is that automated audiometry is a suitable alternative to manual audiometry.\(^11\) However, these approaches rely on the availability of an audiologist to interpret the automated audiometry results. Pre-defined
diagnostic protocols that can correctly identify patients with a hearing diagnosis that requires further examination may have the potential to increase the efficiency of audiological services. In synchronous models of tele-audiology, pre-assessment using diagnostic protocols applied to automated audiometry could reduce the number of audiologist-administered audiograms required per session and decrease the time to diagnosis and referral in asynchronous models of care.

Classification systems for audiometry results have been developed in an effort to standardise the reporting of severity, type and configuration of hearing loss.\textsuperscript{12-15} However, these are general descriptors and not specific to a diagnosis.\textsuperscript{16} As such, previous studies have expressed the need for a consensus on diagnosing the site of lesion for a given hearing loss (i.e. is the hearing loss conductive or sensorineural?).\textsuperscript{15} There has been a lack of formal guidance on this issue, however a number of professional bodies have issued relevant diagnostic classification criteria that will be applied in this study.\textsuperscript{14,17}

Margolis and Saly\textsuperscript{15} have previously focused on improving agreement between audiologists by standardising the classification and configuration of audiograms by creating the AMCLASS software which incorporates 161 audiogram classifications. However, the AMCLASS is currently only available for use with the AMTAS automated audiometry software. This study will therefore utilise simple, freely available, diagnostic guidelines issued by professional bodies that can be applied to any type of manual or automated audiometer and interpreted by telehealth facilitators with minimal training. The primary focus of the diagnostic criteria used in this study was to identify patients suitable for interventions, either medical or audiological. Disabling hearing loss, conductive hearing loss and unilateral or asymmetrical hearing loss are common reasons for referral to specialist medical and audiological professionals. However, using diagnostic criteria can streamline the referral process and limit unnecessary medical and audiological referrals. Patients with a bilateral sensorineural hearing loss will, generally, be considered initially for audiological intervention (e.g. hearing aids) without the need for a medical referral. For patients with a significant conductive hearing loss, both medical intervention (to assess whether any hearing can be restored through surgical or non-surgical intervention) and audiological intervention will typically be required. Patients with a significant unilateral or asymmetrical hearing loss will require a medical and often a radiological referral to exclude acoustic neuroma in additional to an audiological referral.\textsuperscript{17} The aim of the study was to examine the accuracy of standard
diagnostic criteria applied to automated audiometry to identifying a number of key audiometric characteristics which can guide further referral to specialist services.

METHODS

Participants

Study population
We recruited 42 study participants (20 male, 22 female) from a publicly-funded combined audiology and otolaryngology clinic at Sir Charles Gairdner Hospital, Perth, Western Australia. Inclusion criteria were: 18 years or over, no known cognitive disorder, spoke English as a first language, both ears suitable for hearing assessment. Ethics approval was granted by The University of Western Australia.

Participant recruitment and sampling
Patient recruitment was by consecutive series, with all patients attending the clinic offered enrolment in the study, subject to exclusion and inclusion criteria. Recruitment was not based on presenting symptoms (except where they contra-indicated audiological assessment) or results from previous audiometry. No incentives were offered to participants involved in the study.

Test methods

Reference test: Manual audiometry
Manual audiometry is the gold standard assessment of hearing thresholds in adults and children over five years of age\textsuperscript{18} and served as the reference test for this study. Manual audiometry was conducted by a audiologist within a sound-treated room (mean ambient noise level 37 dBA) according to ISO8253-1:2010 using Acoustic Analyser AA30 audiometer (Starkey Hearing Technologies; Minnesota), calibrated to ISO389-1:1998 and TDH-39P (Telephonics; North Carolina) supra-aural headphones and Radioear B-71 bone-conductor (Radioear Corp.; Pennsylvania), calibrated to ISO389-3:1994. Recording the patient’s relevant clinical history, otoscopy and tympanometry using a GSI 38 Auto Tymp (Grason-Stradler; Minnesota) preceded audiometry testing. The reference test was administered by one of five clinical audiologists who were all full members of Audiology Australia.
Index test: Automated audiometry

Automated audiometry was conducted using the KUDUwave (eMoyoDotNet; Pretoria, South Africa) a mobile Type 2B screening, diagnostic and clinical audiometer (IEC 60645-1/2) using the ascending method according to ISO8253-1:2010. The KUDUwave audiometer utilises double attenuation with insert earphones and circumaural earcups, continual noise monitoring, enabling accurate testing down to 0 dB with an ambient noise level of up to 59 dB SPL. Ambient noise levels are monitored in octave bands through an external microphone on each circumaural earcup. The noise monitoring function of the KUDUwave uses low-pass (<125 Hz), even single octave band-pass (125, 250, 500, 1000, 2000, 4000, and 8000 Hz) and high-pass (>8000 Hz) filters to separate the incoming sound. The filters have a stop-band attenuation of 90 dB and pass-band ripple of 0.003 dB. The outputs of these filters are monitored in real-time and the peak value passes to the user interface software (eMOYO) every 100-ms, which is visually represented within the software (Figure 1). Placing insert earphones down to the bony part of the ear canal also reduces the occlusion effect allowing for bone-conduction evaluation with occluded ears using insert earphones.19-21 Air-conduction and bone-conduction thresholds were masked were applicable. See Brennan-Jones et al. for a detailed description of the audiological test parameters.28 The audiometry procedures were automated and recorded on a laptop using the eMOYO (v3.6.7) interface developed by eMoyoDotNet. The diagnostic criteria were calculated post-testing and were not programmed into the eMOYO software. Testing took place in a quiet room that was not sound treated (mean ambient noise level when there was no outpatient clinic in progress was 46 dBA). The researcher gave standard instructions, placed the insert earphones, bone-conductor and headset on the participant and monitored the progress of the test in case of malfunction or patient discomfort.
Hearing thresholds were presented in dB hearing level (dB HL). Participants were tested at air conduction frequencies of 250, 500, 1000, 2000, 4000 and 8000 Hz and bone conduction frequencies of 500, 1000, 2000, 4000 Hz for both the reference standard and index test. The index test had lower maximum testable sound level limits compared to the reference standard (KUDUwave limits for air conduction were 95 dB for 250 Hz, 100 dB for 500, 1000, 2000 and 4000 Hz, and 90 dB at 8000 Hz; sound level limits for bone conduction were 55 dB at 500 Hz and 70 dB at 1000, 2000 and 4000 Hz). In cases where both the index test and the reference standard reached maximum testable limits before a participants hearing threshold was established, a predicted threshold at the index tests maximum level was imputed. The time interval between reference test and the index test being conducted was less than 60 minutes for all participants, as patients proceeded directly to the next test, or after a short break if requested.

**Diagnostic criteria**

All participants were examined to identify whether they had normal hearing, a disabling hearing loss, a conductive or sensorineural hearing loss and whether their hearing thresholds were bilateral, unilateral or asymmetrical except where individuals had missing data that
prevented the diagnostic criteria being calculated ($n = 3$). Normal hearing in this study was defined according the WHO criteria of $\geq 26\,\text{dB}$ for hearing loss (i.e. normal hearing is $< 26\,\text{dB}$) in the better ear, with hearing thresholds averaged across 500, 1000, 2000 and 4000 Hz.\(^{14}\) Disabling hearing loss in this study was defined according the WHO criteria of $\geq 41\,\text{dB}$ hearing loss in the better ear, with hearing thresholds averaged across 500, 1000, 2000 and 4000 Hz.\(^{14}\)

Conductive hearing loss was defined in this study as an air-bone gap of $\geq 20\,\text{dB}$ at two or more adjacent frequencies out of 500, 1000, 2000 and 4000 Hz.\(^{17}\) The individual frequency bone-conduction thresholds were subtracted from the individual frequency air-conduction thresholds for both the index and reference test to obtain the air-bone gap.

Unilateral or asymmetrical hearing loss was defined in this study as a $\geq 20\,\text{dB}$ difference between the left and right bone-conduction thresholds at two or more adjacent frequencies out of 500, 1000, 2000 or 4000 Hz.\(^{17}\)

**Blinding**

The audiologist administering the reference standard (manual audiometry) was blinded to the results of the index test (automated audiometry). The researcher administering the index test was not blinded to the results of the reference test as the index test was automated and therefore could not influence the results.

**Statistical methods**

The diagnostic accuracy analysis used air-conduction thresholds for 500, 1000, 2000, 4000 Hz and bone-conduction thresholds of 500, 1000, 2000 and 4000 Hz for both manual and automated audiometry. Sensitivity and specificity along with positive and negative predictive value were calculated for each diagnostic category.

Agreement between automated and manual audiometry across the diagnostic categories was assessed using Cohen’s Kappa statistic. The Landis and Koch (1977)\(^{22}\) recommendations of agreement classification were applied, with $k<0$ indicating no agreement, $k = 0–0.20$ indicating “Slight” agreement, $k = 0.21–0.40$ indicating “Fair” agreement, $k = 0.41–0.60$ indicating “Moderate” agreement, $k = 0.61–0.80$ indicating “Substantial” agreement and $k = 0.81–1.00$ indicating “Almost perfect” agreement.
RESULTS

Participants
The mean age of participants was 49.93 years (SD = 17.35, range of 19.33 to 92.55 years). Hearing loss was not the primary complaint of all participants and many had more than one ear or hearing related disorder at the time of testing. An additional eleven patients who did not complete both hearing assessments (n = 7) or had reliability questioned (i.e. suspected attention or cognitive ability issues) by the clinical audiologist (n = 4) were not included in the analysis.

Diagnostic accuracy of clinical protocols for automated audiometry

Table 1 presents 2x2 tables of the diagnostic data for detecting disabling, conductive and unilateral hearing loss using pre-defined clinical protocols applied to automated audiometry. Three participants had incomplete data that prevented the calculation of diagnostic criteria (one patient for disabling hearing loss, two for unilateral hearing loss). Table 2 presents sensitivity and specificity results; overall sensitivity was 0.88 [95%CI 0.75, 0.95] and overall specificity was 0.93 [95%CI 0.88, 0.97]. The highest sensitivity and specificity was for identifying disabling hearing loss (1.00 and 0.97 respectively), and the lowest was for identifying a normal ear (0.86 and 0.86 respectively). There were generally greater 95%
confidence intervals for sensitivity than for specificity. The positive predictive values (PPV) ranged between 75% and 86% across the diagnostic criteria and negative predictive values (NPV) ranged from 92% and 100% (Table 2).

Table 2: Diagnostic accuracy of clinical protocols applied to automated audiometry

<table>
<thead>
<tr>
<th>Diagnostic criteria</th>
<th>L95% CI</th>
<th>U95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.86</td>
<td>0.56</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.86</td>
<td>0.66</td>
</tr>
<tr>
<td>PPV</td>
<td>0.75</td>
<td>0.47</td>
</tr>
<tr>
<td>NPV</td>
<td>0.92</td>
<td>0.73</td>
</tr>
<tr>
<td>Disabling HL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1.00</td>
<td>0.40</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.97</td>
<td>0.84</td>
</tr>
<tr>
<td>PPV</td>
<td>0.80</td>
<td>0.30</td>
</tr>
<tr>
<td>NPV</td>
<td>1.00</td>
<td>0.88</td>
</tr>
<tr>
<td>Conductive HL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.86</td>
<td>0.63</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.95</td>
<td>0.86</td>
</tr>
<tr>
<td>PPV</td>
<td>0.86</td>
<td>0.63</td>
</tr>
<tr>
<td>NPV</td>
<td>0.95</td>
<td>0.86</td>
</tr>
<tr>
<td>Unilateral HL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.92</td>
<td>0.60</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.93</td>
<td>0.75</td>
</tr>
<tr>
<td>PPV</td>
<td>0.85</td>
<td>0.54</td>
</tr>
<tr>
<td>NPV</td>
<td>0.96</td>
<td>0.79</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.88</td>
<td>0.75</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.93</td>
<td>0.88</td>
</tr>
<tr>
<td>PPV</td>
<td>0.82</td>
<td>0.69</td>
</tr>
<tr>
<td>NPV</td>
<td>0.96</td>
<td>0.91</td>
</tr>
</tbody>
</table>

HL – hearing loss
PPV – positive predictive value
NPV – negative predictive value
L95%CI – lower 95% confidence interval
U95%CI – upper 95% confidence interval

**Agreement**

Cohen’s Kappa agreement ranged from “substantial” to “almost perfect” across the diagnostic criteria, \(k = 0.80\) [95%CI 0.70, 0.89]; \(p<0.001\).
DISCUSSION
This is the first study to examine the accuracy of pre-defined diagnostic criteria applied to automated audiometry in a telehealth context. The aim of the study was to examine whether standard diagnostic criteria applied to automated audiometry is an accurate method of identifying a number of key audiometric characteristics which can guide further referral to specialist services. The primary use of this method would be to increase access to audiology services in remote areas as part of a tele-audiology program or increase efficiency of audiology practices in urban areas by facilitating patient triaging. The overall agreement ranged from substantial to almost perfect and was highly significant \((p<0.001)\), indicating that pre-defined diagnostic criteria may be an accurate and effective method of identifying hearing disorders using automated audiometry. However, there was a wide range of variance in agreement (0.70 - 0.89) and further research beyond a pilot study with additional participants will enable a more precise range of variance to be calculated. Sensitivity, specificity, positive and negative predictive values appeared to be high across conditions, indicating that automated audiometry may be an accurate diagnostic measure when using pre-defined diagnostic protocols. However, there are no similar published studies that allow us to ascertain whether the values obtained truly represent high sensitivity and specificity.

Screening audiometry in unselected populations perhaps offers the closest comparison and typically has sensitivity values between 50\%\textsuperscript{23} to 75\%, \textsuperscript{24} specificity of 78\%\textsuperscript{23} to 98.5\%\textsuperscript{24}, PPV values ranging from 7.6\%\textsuperscript{23} to 52.9\%\textsuperscript{24} and NPV values of approximately 89-99\%\textsuperscript{23,24,25}. Whilst the results from the present study compare favourably with screening audiometry in terms of accuracy it should be noted that the objectives, target population and disease prevalence in screening audiometry are not directly comparable to diagnostic audiometry in this study.

Limitations of this study include the low number of participants with the target hearing conditions, particularly for disabling hearing loss, which contributed to the wide confidence intervals. The current results lack the statistical power to provide conclusive evidence to support adopting pre-defined diagnostic protocols. However, it demonstrates proof of concept and shows that this method warrants further investigation as it has significant potential for improving audiology service delivery. Sensitivity and specificity for identifying normal hearing (the largest clinical group) was 86\%. We would therefore expect future studies to have a similar level of accuracy, albeit with more precise confidence intervals.
Role of diagnostic protocols in a telehealth framework

With over 360 million people around the world with disabling hearing loss and a shortage of hearing health specialists, new models of diagnosis and care are required to provide equitable access to healthcare. Automated audiometry is becoming increasingly sophisticated and has recently been validated for use without a sound treated environment, and in a clinically heterogeneous population, meaning it has the potential to overcome some of these obstacles. However, the results of a hearing test still need to be interpreted to determine the next step in the clinical pathway. Diagnostic protocols such as those applied in this study have the potential to address this, for synchronous, asynchronous and hybrid tele-audiology services. An example of a possible clinical pathway using diagnostic protocols for automated audiometry is shown in Figure 2. Applying these protocols to automated audiology could enable local health workers to triage patients with ear and hearing disorders according to their type and level of hearing loss. In particular, the identification of

Figure 2: Flowchart of the potential patient journey through a service incorporating pre-defined diagnostic criteria applied to automated audiometry.
participants with normal hearing levels will directly reduce the workload of audiologists and enable more patients to be seen or allow more time spent on complex cases, whether synchronous or asynchronous methods are used.

**Clinical relevance & further research**

This study presents a method of identifying patients with hearing disorders that can be applied to automated audiometry within telehealth practices. Assistants currently working to facilitate both synchronous and asynchronous tele-audiology services would be able to apply, conduct and interpret these results from pre-defined diagnostic protocols using automated audiometry, reducing the time and administrative burden placed on clinical audiologists. Classification systems do have their limitations.\(^{16}\) Classifying hearing loss by severity does not provide any medical information (conductive or sensorineural hearing loss) or an understanding of the patient’s perceived experience of their hearing loss (i.e. low-frequency versus high-frequency hearing loss) or quality of life.

The diagnostic accuracy of audiometry will often be strengthened by interpretation by an experienced audiologist. However, clinical protocols that adhere to these diagnostic criteria could allow basic treatment and referral decisions to be made based on results of automated audiometry conducted in a primary care or tele-audiology setting where experienced audiologists may not be available. In addition to complementing tele-audiology models of care, the diagnosis of these conditions could be an effective method of triaging appropriate referral to audiologists and otolaryngologists in secondary and tertiary care.

It is expected that there will be disparities in clinical guidelines locally, nationally and internationally which would affect the definitions used for the diagnostic protocols. However, this study serves to provide a proof of concept and numerous iterations of pre-defined diagnostic protocols are possible for identifying target conditions. Current automated audiometers can accommodate different diagnostic criteria, but the use of alternative guidelines will vary the sensitivity and specificity. It is therefore suggested that a formal analysis of alternative diagnostic definitions should be conducted before implementation.

Further research should also consider patient perceptions of automated audiometry testing as part of a telehealth service. Patient non-acceptance of telehealth applications in audiology is decreasing, but has ranged from 9-30% in previous studies.\(^{29, 30}\) Quality assessment of
automated audiograms is also an area for future research, in order to identify patients for whom automated testing may not be suitable.

CONCLUSION

The current study demonstrates significant potential for the use of diagnostic protocols applied to automated audiometry to complement current telehealth models of care in audiology. Pilot data demonstrate that pre-defined diagnostic protocols applied to automated audiometry asynchronously are sensitive, specific and have substantial to almost perfect agreement for identifying disabling, conductive and unilateral hearing. We have demonstrated proof of concept for this method to be used in a telehealth model of care to improve synchronous and asynchronous tele-audiology service delivery. However, further studies utilising a greater number of participants with the target conditions are now required.

ACKNOWLEDGMENTS

We would like to acknowledge all the participants in this study for the generous use of their time. We would also like to thank all the medical, nursing and support staff from the Sir Charles Gairdner Hospital, Department of Otolaryngology, and audiologists Amie Grandidge, Jonathan Constantine, Susan Tegg-Quinn, Varsha Matthews and Joanne Sahdeo.

REFERENCES