Headphone Evaluation for App-Based Automated Mobile Hearing Screening

Adam W. Pickens[1]   Lakshmi Dakuri Robertson[1]   Matthew Lee Smith[2,3]   Qi Zheng[4]   Sejun Song[5]

[1] Texas A&M Ergonomics Center, Department of Environmental and Occupational Health, Texas A&M Health Science Center (TAMHSC) School of Public Health, College Station, TX, United States
[2] Center for Population Health and Aging, Department of Environmental and Occupational Health, Texas A&M Health Science Center (TAMHSC) School of Public Health, College Station, TX, United States
[3] Department of Health Promotion & Behavior, College of Public Health, University of Georgia, Athens, GA, United States
[4] Department of Epidemiology and Biostatistics, Texas A&M Health Science Center (TAMHSC) School of Public Health, College Station, TX, United States
[5] University of Missouri-Kansas City (UMKC) School of Computing and Engineering, Kansas City, MO, United States

Address for correspondence Adam W. Pickens, PhD, MPH, Texas A&M Health Science Center School of Public Health, 1266 TAMU, College Station, TX 77843-1266, United States (e-mail: pickens@sph.tamhsc.edu).

Abstract

Introduction With the need for hearing screenings increasing across multiple populations, a need for automated options has been identified. This research seeks to evaluate the hardware requirements for automated hearing screenings using a mobile application.

Objective   Evaluation of headphone hardware for use with an app-based mobile screening application.

Methods   For the purposes of this study, hEAR, a Bekesy-based mobile application designed by the research team, was compared with pure tone audiometric tests administered by an audiologist. Both hEAR and the audiologist's test used 7 frequencies (125 Hz, 250 Hz, 500 Hz, 1,000 Hz, 2000 Hz, 4,000 Hz and 8,000 Hz) adopting four different sets of commercially available headphones. The frequencies were regarded as the independent variable, whereas the sound pressure level (in decibels) was the dependent variable. Thirty participants from a university in Texas were recruited and randomly assigned to one of two groups, whose only difference was the order in which the tests were performed. Data were analyzed using a generalized estimating equation model at $\alpha = 0.05$.

Results   Findings showed that, when used to collect data with the mobile app, both the Pioneer HDJ-2000 (Pioneer, Bunkyo, Tokyo, Japan) ($p > 0.05$) and the Sennheiser HD280 Pro (Sennheiser, Wedemark, Hanover, Germany) ($p > 0.05$) headphones presented results that were not statistically different from the audiologist's data across all test frequencies. Analyses indicated that both headphones had decreased detection probability at 4kHz and 8kHz, but the differences were not statistically significant.

Conclusion   Data indicate that a mobile application, when paired with appropriate headphones, is capable of reproducing audiologist-quality data.

Keywords
► hearing tests
► mobile applications
► automation

Introduction

Hearing impairment is one the most prevalent debilitating illnesses affecting the worldwide populations.[1] The two primary populations impacted by hearing loss are those in occupational settings and older adults (that is, those aged 65 years and older). It is estimated that approximately one-third of Americans between the ages of 65 and 74 have some degree of hearing loss, which increases to almost half of the population aged 75 years and older.[2] Additionally, the aging of the baby boomers is having a dramatic effect on the elderly population in the U.S. Data indicate that from 2012–2050 the number of persons over age 65 is expected to double.[3] With prevalence of hearing loss estimated at 40–80% of individuals between the ages of 65–85 years,[4] early identification and proper management are often cited as the most important factors in minimizing the effects of age-induced hearing loss in the patients' quality of life.[5] These trends among the aging population will continue to place additional burden on the population of audiologists, which is projected to have a shortage of qualified professionals available in coming years.[6]

In addition to the strain that the aging population is placing on the community of qualified audiologists, the rates of hearing loss among employees of governmental agencies in the U.S. and Europe have led these groups to identify noise exposure as a high-priority issue.[7,8] In the U.S., the National Institute for Occupational Safety and Health (NIOSH) has designated occupational noise as a high priority research area in their National Occupational Research Agenda (NORA).[8] Because occupationally-induced hearing loss is frequently targeted as a major concern for employees,[9] the overall shortfall of qualified audiologists has created the demand for valid hearing screening options for workers in general (regardless of the industry, be it rural or urban, public or private) and for the aging populations.[10]

To address this need for alternatives of assessment methodologies that can accurately screen the at-risk populations, the research team developed the hEAR mobile screening application with the aim of producing quality automated screening results while increasing the ease of access. This is in line with current research, in which mobile applications and mobile software-based audiometers are becoming more commonplace.[11–14] Our previous work indicated the capacity of the hEAR mobile application to replicate audiologist-collected screening data, but with a strong dependence upon headphone reproduction capacity.[15] This study examined the use of commercially available headphones when combined with a full-spectrum mobile application and their capability of re-producing audiologist-quality hearing screenings.

Methods

Participants

Thirty participants who were university students, faculty, and staff were enrolled in the study. Twelve of the 30 participants were female, and 18 (60%) were male. While participants' ages ranged from 20–57 years, most were aged 25–32 years ($n = 18$). Participants had no previously diag-nosed hearing loss and were required to limit noise exposure 24 hours prior to all tests. All participant recruitment, consent, data collection, and evaluation methodologies were reviewed and approved by the University Institutional Review Board (IRB) for the Protection of Human Subjects.

Materials and Methods

A Samsung Galaxy Tab 3.0 (Samsung, Seoul, Korea), an Android device, was chosen to test the hEAR mobile application because it is built upon the Android platform. The hEAR application itself was designed based on the best practices for automated screenings suggested by a variety of sources, such as the World Health Organization (WHO), who recommend the Bekesy-style audiometry for self-administered hearing screenings.[16] As is best practice with these recommendations, the test tones initiate at inaudible levels and subjects respond to the attenuator control once they hear the tone.

Four sets of headphones were chosen for this study:

1. Pioneer HDJ 2000 (professional DJ)
2. Bose Quiet Comfort 25 (noise cancelling) (Bose, Framingham, MA, USA)
3. Sony MDR 7506 (affordable option) (Sony, Minato, Tokyo, Japan)
4. Sennheiser HD280 Pro (professional audiologist-quality)

Working within the OSHA and WHO guidelines, the application automatically administered a series of mini-trials based on the OSHA-designated frequencies (125, 250, 500, 1,000, 2,000, 4,000, and 8,000 Hz). Each frequency was administered a minimum of four times bilaterally. Each of these mini-trials were administered randomly to the subject. Each participant underwent at least 28 mini-trials; each individual frequency collection period lasted from 27 to 33 seconds. The entirety of one full-spectrum collection period was of ~ 15 to 20 minutes.

With the potential for false positives/negatives in the user feedback, the application has a series of algorithms to identify values that may be false positives/negatives in the data collection. These series of algorithms are primarily based on the amount of deviations from normal hearing responses (dB) of the general public at each test frequency. In this regard, the application, upon identifying a potential false positive/negative in the data collection, automatically and randomly re-administered the identified frequency later in the test sequence.

The study herein described was the continuation of a previously conducted prospective cohort pilot study.[15] The pilot data indicated the presence of confounders on basis of the hardware used, which potentially impacted in the quality of the data collected. Thus, the purpose of this study was to evaluate headphones with different frequency responses and sound reproduction capacities for accuracy of app-based hearing screening data collection.

Each participant was assigned a participant ID, which was a 7-digit random number. These were generated by a uniform distribution random number generator for data collection/ analysis purposes. Subject trials were randomized and counterbalanced so that half of the subjects initiated data collection procedures in the laboratory (Group 1) and the other half with the audiologist (Group 2). This was done to ensure
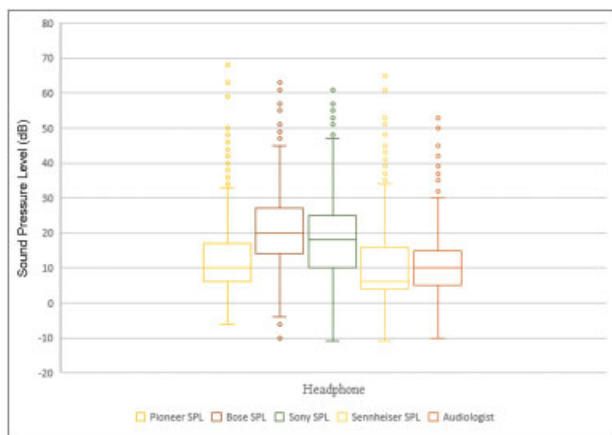
Fig. 1 Boxplots of the headphones and the audiologist's test using summary statistics plotted against the measured sound pressure levels (SPL) on the Y-axis.

unbiased estimates. The scheduling of data collection was performed at an initial meeting with a member of the research team in which subjects completed screening questionnaires. All communication between the researchers and the local audiologist used this identification number to maintain participant protection standards.

In order to meet the requirements of Appendix D of 29 CFR 1910.95, the testing procedures were performed in the laboratories.[17] Ambient sound pressure level (SPL) on each testing cycle, for all 30 subjects, averaged 24 dBA.

Statistical Methods

The statistical analyses were performed using the software SAS Version 13.1 (SAS Institute, Cary, NC, USA). Since the data was non-linear, generalized estimating equations (GEE) were used, which are appropriate for longitudinal data, data with small sample size and data with repeated measures.[18] The outcome variable of interest was the SPL in dB. The SPL measurements were differentiated by both the hEAR mobile application and the pure-tone audiometry test, based on ear side (that is., left and right ears). The preliminary analysis for the pilot data, as well as the current data, showed that the ear side was not significant (t = 0.593, df = 1,478, p = 0.5532).

The means per subject, per headphone, for the respective frequencies were calculated for a preliminary comparison between the SPL measurements between the different chosen headphones and those by the pure-tone audiometry test (►Fig. 1). In ►Figs. 2 and 3, we can observe that the counterbalanced group that the participants were assigned to had only a marginal effect on the overall results.

Each frequency was considered separately, and eight (8) agreement scores were generated for each subject and headphone. Each agreement score was defined as the absolute difference between the SPL response recorded by the hEAR application and the SPL response recorded by the audiologist. The Eq. 1 below was used for calculation of the agreement scores.

$$Ag_i = \sum_{j=1}^{4}\left|LeftApp_{ij} - LeftDoctor_i\right| + \sum_{j=1}^{4}\left|RightApp_{ij} - RightDoctor_i\right| \quad (1)$$

From these 8 scores, an average agreement score was generated for that ID at the particular frequency. The smaller the agreement score, the better it is. Based on the agreement scores, if Yi was a binary random variable, at a threshold tolerance value (Θ) of 8, defined by Eq. (2):

$$Y_i = \begin{cases} 1 \text{ if } Ag_i \leq \theta \\ 0 \text{ if } Ag_i > \theta \end{cases} \quad (2)$$

Eq. 3 is the resulting logistic model.

$$logit\left(P(Ag_i = 1)\right) = \beta_0 + \beta_1 * I_1 + \beta_2 * I_2 + \beta_3 * I_3 . \quad (3)$$

The formula for the probability of success of a headphone is calculated based on the Eq. $1/(1 + E^{\wedge\beta 0})$, whereas the
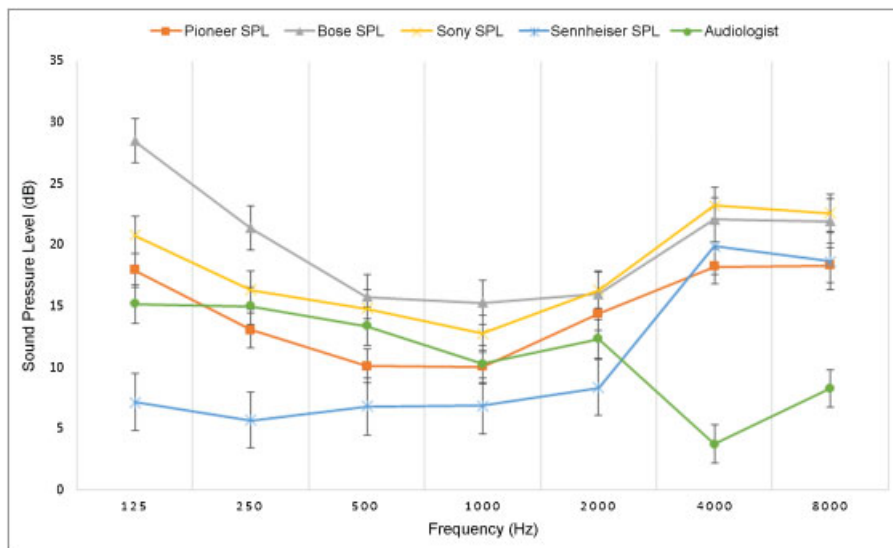

Fig. 2 Sound Pressure Level (SPL) means per headphone for Group 1 calculated using headphone means in SPL (dB) at each measured frequency on the X-axis. The plotted means for headphones show similarities and differences with those measured by the audiologist's test.

Fig. 3 Sound pressure level (SPL) means per headphone for Group 2 calculated using headphone means in SPL (dB) at each measured frequency on the X-axis. The plotted means for headphones show similarities and differences with t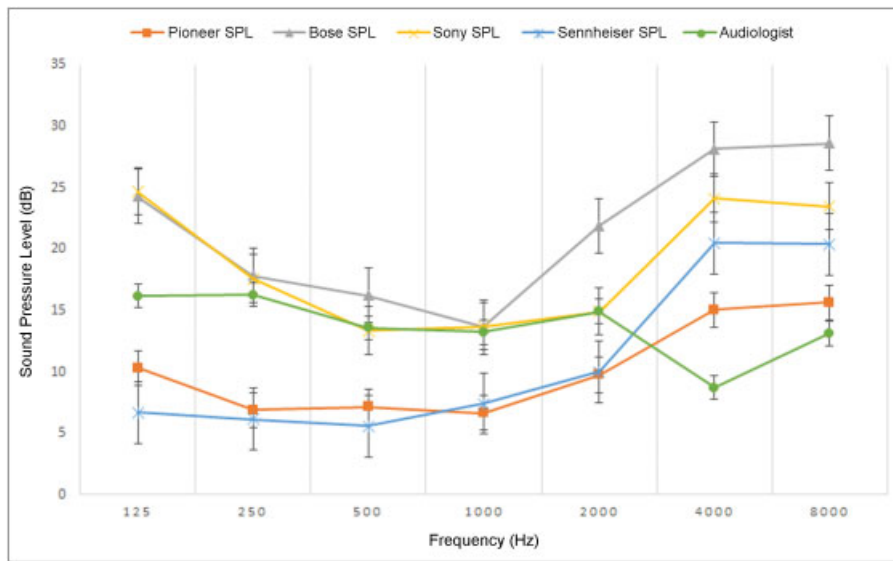hose measured by the audiologist's test. Counterbalanced Groups (Groups 1 and 2) had marginal effect ($p = 0.08$) on the results.

probability of success of the subsequent headphones is calculated based on the Eq. $1/(1 + E^{\wedge \beta 0 + \beta 1})$. Therefore, for headphones at a particular frequency, the higher the probability of success, the better the headphones are when compared with the others.

Results

The results show that neither the group nor the order in which the two tests were conducted (hEAR mobile application and audiologist's test) significantly impacted the probability of success of the headphones ($p = 0.0894$) (►**Table 1**). In general, Group 2 (audiologist test prior to hEAR mobile application testing) had slightly higher probability values ($Z = -1.70$, $p = 0.0894$).

Analysis from the model, produced probability of successfully reproducing test results similar to the audiologist control along with overall statistical significance (p-value). This was performed for every test frequency for each set of headphones. The greater the probability, the more likely the headphones are to reproduce SPL responses similar to the audiologist control when compared with the other sets of headphones.

►**Table 2** indicates that, overall, the Pioneer HDJ 2000 performed the best, or it was statistically similar to the audiologist-administered test, for all frequencies (all $p > 0.05$). Similarly, the Sennheiser HD280 Pro also performed well for all frequencies, and it was not significantly different from the audiologist control (all $p > 0.05$). According to the frequency reproductions of the selected headphones, the accuracy of reproduction decreases after 6,000 Hz, when used under daily,

Table 1 Results of generalized estimating equation model analysis for the counterbalanced headphones and audiologist's test of the test initiation

Analysis of GEE parameter estimates; empirical standard error estimates						
Parameter	Estimate	SE	95% CI		Z	Pr > |Z|
Intercept	10.6054	1.199	8.254	12.957	8.84	< 0.0001
Pioneer Headphones (1)	0.0175	1.462	−2.849	−2.883	0.01	0.9905
Bose Headphones (2)	8.4433	1.434	5.634	11.253	5.89	< 0.0001*
Sony Headphones (3)	6.0960	1.756	2.674	9.517	3.49	0.0005*
Sennheiser Headphones (4)	−1.6569	1.403	−4.407	1.092	−1.18	0.2376
Control (Audiologist's test)	0	0	0	0	.	.
Frequency	0.0008	0.0001	0.0005	0.0011	5.53	< 0.0001
Group	−0.4885	0.2876	−1.0522	0.0752	−1.70*	0.0894*

Abbreviations: CI, confidence interval; GEE, generalized estimating equation; SE, standard error.
GEE Fit Criteria: QIC = 1087.4051; QICu = 1054.0
*Indicates statistical significance

Table 2 Probability and statistical significance (p values) for test headphones

Frequency	Pioneer HDJ 2000	Bose Quiet Comfort 25	Sony MDR 7506	Sennheiser HD280 Pro
125	0.6792 ($p = 0.0754$)	0.2156 ($p = 0.0072^*$)	0.2156 ($p = 0.0013^*$)	0.4591 ($p = 0.0598$)
250	0.7079 ($p = 0.0581$)	0.6126 ($p = 0.4396$)	0.4328 ($p = 0.0158^*$)	0.5788 ($p = 0.1485$)
500	0.6987 ($p = 0.0913$)	0.6035 ($p = 0.4375$)	0.6987 ($p = 1.0000$)	0.6676 ($p = 0.6548$)
1000	0.7372 ($p = 0.0513$)	0.6425 ($p = 0.4043$)	0.6745 ($p = 0.5257$)	0.7372 ($p = 1.0000$)
2000	0.6678 ($p = 0.1549$)	0.4140 ($p = 0.1041$)	0.1791 ($p = 0.0006^*$)	0.5786 ($p = 0.3150$)
4000	0.2630 ($p = 0.0716$)	0.1150 ($p = 0.0367^*$)	0.0387 ($p = 0.0247^*$)	0.1526 ($p = 0.1712$)
8000	0.2500 ($p = 0.646$)	0.2500 ($p = 1.0000$)	0.2500 ($p = 1.0000$)	0.2500 ($p = 1.0000$)
Overall	0.6648 ($p = 0.2897$)	0.4122 ($p = 0.0271^*$)	0.3607 ($p = 0.0003^*$)	0.4954 ($p = 0.0641$)

*Indicates statistical significance.

normal circumstances. Additionally, in ►**Fig. 1** we can see that the boxplots for the Pioneer headphones and the audiologist's test show noticeable similarity with strong overlap in the data. As results in ►**Table 2** indicate, the Bose Quiet Comfort 25 and the Sony MDR 7506 both had statistically significant results overall, and across multiple frequencies when compared with the audiologist control.

Discussion

The results of the study depict interesting outcomes. Both sets of professional headphones (that is, the Pioneer HDJ 2000 and the Sennheiser HD280 Pro) performed well in this test, as both have broad frequency reproduction spectrums. However, the more day-to-day use headphones (such as the Bose Quiet Comfort 25 and the Sony MDR 7506) did not perform as well.

The probability of success indicates that the Pioneer HDJ 2000 are the best headphones for use with the hEAR mobile application. The Sennheiser HD 280 Pro headphones were designed specifically for audiometric testing in noise insulated or soundproof testing rooms to reproduce frequencies in an accurate flat frequency response from 100 to 10,000 Hz in noise-insulted environments. This was not the case in our testing room, where the ambient noise levels were at almost 25 dB, well within the requirement of both OSHA and ANSI. It was also the researchers' experience that even though the Sennheiser HD 280 Pro headphones are marketed as "over the ear," their fit can best be described as "on the ear" since several participants commented that the ear cups did not fully cover their ears during the test. This lack of coverage could potentially have had an effect on the overall results for the Sennheiser headphones.

Although noise cancelling, the Bose Quiet Comfort 25 headphones are specifically designed for high bass music listening or daily use, and also for comfort rather than for

audiometric testing.[19] Therefore, even though these headphones were preferred by most participants due to their comfort, they did not prove as effective for testing hearing. The same was the case with the Sony MDR 7506.

All of the headphones used in this exercise had limitations in accurately reproducing high frequency sounds (above 6,000 Hz). This property could be the reason for the lower probability of success at the frequency of 8,000 Hz.

Overall, the Pioneer HDJ 2000 headphones were the best audiometry screening option for use with the hEAR mobile application. Therefore, they may offer a portable option for full-spectrum audiometric screening. The Sennheiser HD280 Pro headphones, while potentially more capable than the Pioneer HDJ 2000, may be limited to very isolated testing environments. The Bose Quiet Comfort 25 and the Sony MDR 7506 did not have the capability of producing audiologist-quality data when paired with the hEAR application.

Conclusions

It is important to note that there are limitations to this study. The ambient noise in the testing environment was not tested with an octave band analyzer, as is required by most organizations, including OSHA.[17] However, it is not expected that the analysis with an octave band analyzer would have produced results that would have altered the overall interpretation of the data, as the SPL of the overall background noise was just 24 dBA. While there is no way of accurately knowing if there was frequency-specific interference, the research team does not expect that octave band analysis would have significantly changed the results of the study. More evaluation is also needed for a broader test population. However, even with these limitations, the Pioneer HDJ 2000 headphones paired with the self-administered hEAR mobile application were able to reproduce overall and frequency-specific results that were not

significantly different than those of a certified audiologist in a controlled testing environment. These results show a promising trajectory for mobile automated hearing screening options.

References

1 World Health Organization (WHO). Deafness and Hearing Loss. 2015 [cited 2016 Jun 05]. Available from: http://www.who.int/mediacentre/factsheets/fs300/en/

2 Oyler AL. Untreated hearing loss in adults - A growing national epidemic. American Speech-Language-Hearing Association (ASHA) 2012 [cited 2017 Jan 31]. Available from: http://www.asha.org/aud/articles/untreated-hearing-loss-in-adults

3 U.S. Department of Health and Human Services (DHHS). Administration on Aging 2010. A profile of older Americans: 2011. [cited 2017 Jun 06]. Available from: http://www.aarp.org/content/dam/aarp/livable-communities/learn/demographics/a-profile-of-older-americans-2011-aarp.pdf

4 Choi JM, Sohn J, Ku Y, Kim D, Lee J. Phoneme-based self hearing assessment on a smartphone. IEEE J Biomed Health Inform 2013; 17(03):526–529

5 Ciorba A, Bianchini C, Pelucchi S, Pastore A. The impact of hearing loss on the quality of life of elderly adults. Clin Interv Aging 2012;7:159–163

6 American Speech and Language Association (ASHA). 2013 Market Trends. [cited 2016 Jul 02]. Avaliable from: http://www.asha.org/Careers/Market-Trends/

7 European Agency for Safety and Health at Work. 2005. The Impact of Noise at Work. [cited 2016 Jul 05]. Available from: https://osha.europa.eu/en/tools-and-publications/publications/factsheets/57

8 National Institute for Occupational Safety and Health (NIOSH). 2012. National Occupational Research Agenda (NORA). [cited 2016 Jul 02]. Available from: http://www.cdc.gov/niosh/nora

9 Kriesky J. Unvonventional drilling for natural gas: The gap between what we know and what we need to know. Pittsburgh, PA: School of Public Health, University of Pittsburgh, Center for Healthy Environments and Communities; 2012

10 Yueh B, Shapiro N, MacLean CH, Shekelle PG. Screening and management of adult hearing loss in primary care: scientific review. JAMA 2003;289(15):1976–1985

11 Ferrari DV, Lopez EA, Lopes AC, Aiello CP, Jokura PR. Results obtained with a low cost software-based audiometer for hearing screening. Int Arch Otorhinolaryngol 2013;17(03):257–264

12 Rourke R, Bromwich M, Chan D. iPad Audiometry in Canada's North: A portable and cost-effective method for hearing screening. Otolaryngol Head Neck Surg 2014;151(1_suppl): P104–P104

13 Van der Aerschot M, Swanepoel W, Mahomed-Asmail F, Myburgh HC, Eikelboom RH. Affordable headphones for accessible screening audiometry: An evaluation of the Sennheiser HD202 II supra-aural headphone. Int J Audiol 2016;55(11):616–622

14 Swanepoel DW. Smartphone-based National Hearing Test Launched in South Africa. Hear J 2017;10(01):14–16

15 Pickens AW, Robertson LD, Smith ML, Zhao H, Mehta R, Song S. Limitations of a mobile hearing test application. Hear J 2017;70 (06):34–37

16 Franks JR. Hearing measurement. Occupational Exposure to Noise: Evaluation, Prevention, and Control. Geneva: World Health Organization (WHO); 1995:183–232

17 Occupational Safety and Health Administration (OSHA). Code of Federal Regulations Title 29 Part 1910: General Industry Subpart 95 Occupational Noise Exposure, Appendix D: Audiometric Test Rooms. [cited 2015 Feb 11]. Available from: https://www.osha.gov/pls/osha-web/owadisp.show_document?p_table=STANDARDS&p_id=9739

18 Pickles A. Generalized estimating equations. Encyclopedia of Biostatistics. New Jersey: John Wiley & Sons, Ltd.; 1998

19 Hertsens T. A better traveler's sanctuary: The Bose Quiet Comfort 25. Inner Fidelity 2014 [cited on 2016 Jun 02]. Available from: https://www.innerfidelity.com/content/better-travelers-sanctuary-bose-quiet-comfort-25