

Research Article

A Follow-Up Clinical Trial Evaluating the Consumer-Decides Service Delivery Model

Larry E. Humes,^a Dana L. Kinney,^a Anna K. Main,^a and Sara E. Rogers^a

Objectives: There were 2 main objectives. The primary objective was to replicate a prior clinical trial of a consumer-decides (CD) approach to selecting hearing aids in older adults as a potential model for over-the-counter (OTC) intervention using less front-end screening of participants and a wider range of frequency-gain characteristics in the devices. The 2nd objective, only feasible if participant choices allowed, was to evaluate the efficacy of the CD approach relative to a CD-based placebo device.

Design: The design of this study is a single-site, prospective, double-blind clinical trial. Outcome measures were obtained after a typical 4- to 5-week trial period. An optional follow-up of a 4-week audiology-based (AB) best practices trial was also included for replication and comparison purposes.

Setting: Older adults from the general community were recruited via newspaper and community flyers to participate at a university research clinic.

Participants: Participants were adults, aged 54–78 years, with mild–moderate hearing loss. Forty-one participants enrolled as a volunteer sample; 40 completed the intervention.

Intervention(s): All participants received the same high-end digital mini–behind-the-ear hearing aids fitted bilaterally. CD participants self-selected their own preprogrammed hearing aids via an OTC-type model. One of the 3 devices from which participants could choose was programmed to be a placebo device with no functional gain.

Primary and Secondary Outcome Measures: The primary outcome measure is the 66-item self-report Profile of Hearing Aid Benefit (Cox & Alexander, 1990). The secondary outcome measure is the Connected Speech Test (Cox,

Alexander, & Gilmore, 1987) benefit. Additional measures of hearing aid benefit and usage were also obtained.

Results: Per-protocol analyses based on the data from the 40 (of 41) participants who completed the study were performed. Hearing aid outcomes from this follow-up CD (CD2) cohort were positive and generally the same as for the original CD cohort. CD service delivery model was efficacious relative to CD-based placebo control, with medium effect sizes observed. Approximately half of the CD2 group was likely to purchase hearing aids after the trial, similar to findings for the original CD cohort. Outcomes improved significantly for the 32 CD2 participants who elected to complete the optional 4-week AB trial. For this largely unscreened sample, more individuals with healthy hearing sought amplification, and many of these individuals (35%) chose placebo devices for both ears.

Conclusions: Prior positive outcomes for CD service delivery have been replicated in a less restrictive approach to participant recruitment. The CD approach was again found to be efficacious. Significantly better outcomes were observed after subsequent AB service delivery follow-up, also replicating prior findings. Efficacious OTC models, including those using similar CD approaches to hearing aid self-selection, may increase accessibility and affordability of hearing aids for millions of older adults. Front-end guidance to consumers regarding the best path to intervention, ranging from self-screening of hearing online to a full audiologic assessment, appears to be critical to optimize the success of OTC approaches.

Trial Registration: Clinicaltrials.gov: NCT01788432; <https://clinicaltrials.gov/ct2/show/NCT01788423>

Supplemental Material: <https://doi.org/10.23641/asha.7728479>

^aDepartment of Speech and Hearing Sciences, Indiana University, Bloomington

Correspondence to Larry E. Humes: humes@indiana.edu

Editor-in-Chief: Sumitrajit Dhar

Editor: Ryan McCreery

Received May 17, 2018

Revision received July 24, 2018

Accepted August 26, 2018

https://doi.org/10.1044/2018_AJA-18-0082

Humes et al. (2017) published the first randomized placebo (P)-controlled clinical trial of hearing aids in older adults comparing service delivery models and device purchase prices. Purchase price was found to have no influence on outcomes for either of the service delivery models considered. The two service delivery models, audiology-based (AB) best practices and a consumer-decides (CD) alternative, were both found to be efficacious relative to P devices fit with AB service delivery. In addition, the

Disclosure: The authors have declared that no competing interests existed at the time of publication.

overall hearing aid outcomes observed between the AB and CD groups did not differ significantly, with relatively minor differences in reported satisfaction with hearing aids. To our knowledge, this prior study represents the full extent of the prior literature comparing outcomes across an AB best practices service delivery model and a CD service delivery model. Given the dearth of prior pertinent literature on this topic, replication of the findings of Humes et al. (2017), both in additional participants and with a different set of preprogrammed devices from which the participants could choose, is critical.

This study is a replication of the CD branch of the prior parallel-branch randomized clinical trial with two significant changes. First, the research protocol implemented by Humes et al. (2017) required that prospective participants be screened rigorously before random assignment to one of the study groups. This initial screening included a complete audiometric evaluation with pure-tone and speech audiometry to establish that the prospect's hearing loss was within the targeted range of hearing loss for the devices used in the study. In fact, the targeted range was more restrictive than that of the device alone in that the minimum acceptable hearing loss was adjusted upward from 0 dB HL, according to the manufacturer, to 10–20 dB HL from 250 to 2000 Hz, sloping down to 50 dB HL at 6000 Hz. The maximum acceptable hearing loss for this device, however, followed the manufacturer's suggestions. The primary reason for raising the minimum acceptable hearing loss for participation in the clinical trial was due to the inclusion of a P group who would be receiving the same hearing aids, but programmed and adjusted on ear to have 0-dB insertion gain. Without this restriction on the lower end of the acceptable range of hearing loss for that study, it was conceivable that all or most of the participants could have very slight hearing losses for which no differences could possibly be observed between P and well-fit hearing aids. Although the minimum possible hearing loss was restricted in the study by Humes et al. (2017), the applied minima were typical of the average hearing loss for 70-year-olds from a highly screened (otologically healthy) population (International Standards Organization, 2000).

Pure-tone audiometry, as noted, was only one part of the initial screening of participants. Participants were excluded for several reasons, as noted by Humes et al. (2017), including the presence of medically treatable ear conditions, bilateral flat tympanograms, and known fluctuating or rapid-onset hearing loss. If excessive cerumen was observed during otoscopy, this was removed by the audiologist prior to the audiometric evaluation or, if not possible to remove, the participant was referred to a physician for subsequent treatment prior to enrollment. In addition, a brief screen of cognitive function, the Mini-Mental State Examination—2nd Edition (MMSE-2; Folstein, Folstein, White, & Messer, 2010), was administered, and scores > 25 were required for enrollment in the subsequent clinical trial.

In summary, the participants in the clinical trial of Humes et al. (2017) were highly screened prior to random

assignment to one of the groups in that study. The recently enacted Over-the-Counter Hearing Aid Act of 2017 identifies prospective candidates for over-the-counter (OTC) hearing aids as adults with “perceived mild to moderate hearing loss.” The CD service delivery model of Humes et al. (2017) was one instantiation of many possible service provision models for OTC hearing aids. Although the CD model was found to be efficacious and nearly as effective as the AB model by Humes et al. (2017), the study sample was screened more than would be anticipated for prospective OTC hearing aid purchasers. It is unclear how such screening may have impacted the outcomes, and it is appropriate to examine the outcomes obtained for the CD model from a less thoroughly screened sample—a sample more likely to be representative of the anticipated pool of OTC hearing aid prospects. This represents one of the two main purposes of the present clinical trial. In this trial, we used the same CD service delivery model as in Humes et al. (2017), but with a study sample that simply responded to a newspaper ad for participants and completed a brief screening over the telephone (confirming age, native language of English, and no prior use of hearing aids). The question addressed here is: Do the results from the original highly screened CD participants (original CD [CD1]) in Humes et al. (2017) generalize to a new group of unscreened CD participants (follow-up CD [CD2]), participants more likely to be representative of future OTC hearing aid prospects?

A second significant change from the CD1 service delivery model used by Humes et al. (2017) was to extend the range of devices that could be selected by participants to include a P in this study. This can be considered another assessment of the generalization of the findings of Humes et al. (2017), this time to a broader range of frequency-gain characteristics, including one that would be inappropriate for those enrolled who have hearing loss. This also is a check of the validity of the participants' ability to self-select their own hearing aids. Finally, if a sufficient number of participants select P devices for themselves, this will enable an evaluation of the efficacy of the CD service delivery model. In Humes et al. (2017), all those who received P hearing aids had received AB services. The primary difference between those in the P group and those in the AB group was that the devices for the Ps had been programmed on the participants' ears to have no functional gain. Other interactions with the audiologist were identical for AB and P participants. It is possible that such interactions within the AB model could have impacted various self-report measures obtained as outcomes, as well as the participants' impression as to whether they had received a P device and, in turn, whether they were likely to keep the devices (a query made of all participants in that trial prior to revealing their group assignment). For the CD group in Humes et al. (2017), none of the available choices for hearing aids included a P. Rather, the three choices available to the CD participants were hearing aids preprogrammed to be appropriate for three of the most common audiometric configurations expected among older adults. In this follow-up trial, we retained two of those preprogrammed frequency

responses, the two most frequently selected in Humes et al. (2017), and replaced the third least frequently selected, the one with the greatest gain, with a P device. Thus, the efficacy of the CD service delivery model could now be established by comparing outcomes to Ps who had the identical CD service delivery experience. If the outcomes for the present CD2-CD group were found to be superior to those of the CD2-P group, then this would better establish the efficacy of the CD model. We also anticipated, by eliminating all screening of hearing loss prior to enrollment in this follow-up trial, that there may be several older adults who enrolled with milder hearing loss than in Humes et al. (2017). In this case, a choice of “placebo” may not be too far from the desired target gain for a given participant. As will be seen below, in the end, several participants selected the P devices for themselves. Given that such a selection was entirely determined by the participant and not assigned randomly by the experimenter, we had no idea whether P control would be a reality for this study at the time of study design. As a result, due to self-selection of P devices by several participants, we will be able to compare the outcomes obtained here for non-P and P groups, all within the CD service delivery context, to further evaluate the efficacy of the CD approach, this time with a less screened study sample. This represents the second basic question to be addressed in this follow-up (CD2) clinical trial: Are the results from non-P self-fit devices (CD2-CD) superior to those from P self-fit devices (CD2-P) in this study, and what are the effect sizes observed? Comparison of outcomes between intervention (CD2-CD) and P (CD2-P) groups, with the parallel groups differing only in the treatment (devices) provided, is the standard means of establishing efficacy in clinical trials (Moher et al., 2010). In addition, given that audiology best practices represent the preferred method of hearing aid delivery (American Speech-Language-Hearing Association, 2015; Valente et al., 2006), as in Humes et al. (2017), we also asked how the outcomes for devices delivered with the CD method compared with established best practice service delivery.

Method

Participant Recruitment and Selection

Participants were recruited primarily by ads posted in the local newspapers and around the community. All testing of this volunteer sample took place in a university research clinic at Indiana University, Bloomington, Indiana. Those interested in participating contacted the clinical trials coordinator (CTC) by phone for an initial eligibility screen, which consisted of confirming the inclusion criteria noted in the newspaper ad: (a) aged 55–79 years, (b) English as native language, and (c) no prior hearing aid experience. They were also told, as noted in the ad, that enrollment in the study would require payment of \$600 for the purchase of the hearing aids, less \$50 of credit for completion of Visit 1, for a total payment of \$550 due at Visit 1. (This was the reduced price used as one of two purchase prices

in Humes et al., 2017, and is believed to approximate the purchase price anticipated for OTC devices.) Those who successfully completed the phone screen were scheduled for Visit 1. Based on statistical power calculations from the data in Humes et al. (2017), a minimum of 35 participants was the targeted enrollment. A total of 41 were enrolled, one of whom withdrew during Visit 1, leaving 40 total participants, 20 female and 20 male, ranging in age from 54 to 78 years ($M = 69.2$ years, $SD = 5.6$ years). Additional demographics are presented below. The trial commenced on July 12, 2017, and data collection for the main trial ended on January 19, 2018, with the posttrial follow-up sessions (Visits 3 and 4) completed on February 28, 2018. (Supplemental Material S1 and S2 available online with this publication include a complete protocol manual and the various data collection forms used in this clinical trial.)

Visit 1: Outcome Baselines and CD Hearing Aid Selection

Unaided baseline scores were collected for two self-report outcome measures, the 66-item Profile of Hearing Aid Performance (PHAP; Cox & Gilmore, 1990) and the 25-item Hearing Handicap Inventory for the Elderly (HHIE; Ventry & Weinstein, 1982). Both were used in Humes et al. (2017), and the PHAP formed the basis of the primary outcome measure here as in the prior clinical trial. A tablet PC was used to collect and score the responses for both self-report measures.

The CTC met the participant and escorted the participant to a small room where the CD service delivery model was implemented, beginning with the completion of the baseline surveys and then proceeding to hearing aid selection. As in Humes et al. (2017), containers for the ear tips/domes and tubing to be used, together with three bins containing three pairs of hearing aids each with the hearing aid pairs in each bin differing only in color (gray, beige, brown), were located on a table in this room. Each bin was partitioned into three compartments, with each compartment containing a pair of hearing aids labeled X, Y, or Z and programmed in advance to match the National Acoustics Laboratories' Non-Linear Version 2 (NAL-NL2) acoustic gain and output prescriptions (Dillon et al., 2011) in a 2-cm³ coupler for two of the most common patterns of hearing loss among older adults in the United States (Ciletti & Flamme, 2008), corresponding to the X and Y configurations in Humes et al. (2017). The hearing aid manufacturer's Aventa software was used for this programming, and we had verified previously that the NAL-NL2 prescriptions generated by this software matched those using NAL software (Humes et al., 2017). The devices labeled “Z” in this study were programmed assuming 0-dB HL thresholds from 250 through 8000 Hz. Coupler measurements and pilot real-ear measurements on other individuals confirmed that 0-dB insertion gain was observed on average. The default programming options for these hearing aids were identical to those in Humes et al. (2017) regarding the

various features of the hearing aids (microphones, feedback suppression, noise reduction, etc.). Maximum output levels for these devices were established by entering the pure-tone thresholds from the corresponding X, Y, or Z audiogram and using the first-fit option in the Aventa programming software to set the maximum power output (MPO) levels.

The participant watched a brief instructional video overviewing the hearing aid self-selection process then completed each step of the process. A hardcopy of the step-by-step instructions was also provided to the participant in the form of three binders with the contents based on instructions provided for self-fitting hearing aids (Caposecco, Hickson, & Meyer, 2011). The participant first selected an appropriately sized ear tip and tube; then the desired hearing aid color, examining several in a mirror; then the acoustic characteristics desired (X, Y, or Z). A tablet PC was available to provide standardized samples of speech, music, and environmental sounds for listening. Participants could have a significant other in the room with them during the selection process (only one participant opted to do so). As the participant tried the domes, tubes, and hearing aids, those in which he or she was not interested or did not fit properly were placed in a bin labeled *No*. When the participant's selections were finalized, the participant pressed a button to alert the CTC that the selections had been made. The CTC returned to the room, recorded these selections, and noted the contents of the other bin labeled "No." Based on the tallies of items in the "No" bin, 47.5% of the CD2 participants had tried more than one dome; 77.5%, more than one tubing size; and 95%, more than one hearing aid. Typically, two to four hearing aids (out of six available) were in the "No" bin when the participant's selection process had been completed.

For all participants, Visit 1 ended with a meeting with the CTC during which the devices were delivered to the participant for use during the 4- to 5-week trial. Each participant was provided with the devices, tubes and domes selected (as well as an extra set of domes), batteries, cleaning tools, and user guide. Payment was also collected at the end of this session. Participants paid \$550, after applying a \$50 credit for the completion of Visit 1. Those participants who completed Visit 2 were paid \$50 cash at the completion of Visit 2.

If any of the participants had problems with their devices during the trial, they were asked to contact the CTC by phone. The CTC initially instructed the participant to review the user guide for assistance. If the participant did so, the problem persisted, and the problem was occurring for just one device, the CTC instructed the participant to remove both devices and compare them to see if anything differed between the two with regard to the tubing, dome, or battery. If the problem persisted, for either one or both hearing aids, the CTC made an appointment for an unscheduled visit. Unscheduled visits began with the CTC performing a visual inspection of the devices. If no problems were apparent, the CTC had the participant remove the devices and contacted Audiologist A1 or A2. The audiologist first examined the hearing aids and, if nothing was

found, next performed otoscopy, tympanometry, or other audiologic measures to determine the nature of the problem and remedy it accordingly. The CTC received phone calls from a total of nine participants, resulting in 11 calls between Visits 1 and 2, six from non-P and three from P participants. Of these, three non-P and one P participants came in for an unscheduled visit. Most requested to choose a different sized dome; one realized a tube was kinked and asked for a replacement. Three of the non-P unscheduled visits and none of the P unscheduled visits resulted in changes to the domes or tubes. The P unscheduled visit, which occurred 6 days post-Visit 1, was due to a difference noticed in the volume control settings of the hearing aids. The participant noticed that one hearing aid had three volume control beeps (incorrect) and the other hearing aid had four volume control beeps (correct). The participant came in, and the incorrect volume control was reprogrammed.

Visit 2: Hearing Aid Outcomes and Audiologic Assessment

During Visit 2, typically 4–5 weeks after the initial fit of the hearing aids in Visit 1 ($M = 32.1$ days; $SD = 6.4$ days), Audiologist A1 (the one not involved in subsequent Visit 2 audiologic assessment and hearing aid fitting) obtained as-worn aided scores for the Connected Speech Test (CST; Cox, Alexander, & Gilmore, 1987), with each score based on 50 keywords, followed by participant completion of several self-report measures via tablet PC. All CST scores, unaided and aided, were obtained in the sound field within a double-walled sound booth with the participant centered between two loudspeakers, one at 0° azimuth and the other at 180° azimuth. The loudspeakers were each about 1 m from the position corresponding to the center of the participant's head. CST scores were obtained for a speech level of 65 dB SPL and a signal-to-babble ratio (SBR) of +3 dB (the babble supplied with the CST was used).

The primary outcome measure for this trial was the Profile of Hearing Aid Benefit (PHAB). The PHAB score, a relative benefit measure, is the difference between the unaided and aided scores on the PHAB. The unaided PHAB was administered prior to Visit 1 and the aided PHAB following a 4- to 5-week period of hearing aid use (Visit 2), with the difference in ratings on the questionnaire used to compute the PHAB scores. As in Humes et al. (2017), the five communication-related subscales (Familiar Talkers, Ease of Communication, Reverberation, Reduced Cues, Background Noise) were averaged to form a PHABglobal score, and the other two subscales (Aversiveness, Distortion) were averaged to form a PHABavds score. Aided HHIE scores were also obtained with the difference between aided and unaided HHIE scores, HHIE benefit, forming another outcome measure. Next, a brief survey about the difficulty of the CD selection processes and the participants' confidence in their selections was administered.

Next, A1 inspected the participants' ears, hearing aids, tubes, and domes; removed the devices; and then obtained as-worn electroacoustic performance measures for

each hearing aid in the Verifit test box (American National Standards Institute [ANSI], 2009). An inspection checklist was used to note any issues identified during the physical inspection. Following completion of these measures, the hearing aids were reconditioned as needed (domes cleaned, kinked tubing replaced, dead or weak batteries replaced, etc.), and postmaintenance electroacoustic measures (ANSI, 2009), as well as real-ear-to-coupler differences, were obtained on ear using the Verifit system. A1 also did maximum power output measurements and speech mapping measurements at 55, 65, and 75 dB and an unaided real-ear 65-dB response. Finally, Audiologist A1 extracted the data logging information from each hearing aid.

Outcome measures were then obtained by another audiologist, A2. A2 was blind to the real-ear measurement results obtained earlier in this visit and, therefore, to the non-P or P status of the devices. The secondary outcome measure to be used in this clinical trial, as in Humes et al. (2017), was derived from the unaided and aided sound-field CST speech recognition scores. Aided CST scores were again obtained, this time after any required maintenance of the devices resulting from the inspection. The secondary outcome measure was based on the difference between this aided CST score and an unaided CST score, CST benefit. As noted, each CST score was based on one passage pair or 50 keywords and was obtained for a speech level of 65 dB SPL and an SBR of +3 dB.

Each participant was then asked, “Now that you’ve worn these hearing aids for about 4–6 weeks, are you interested in keeping them?” Their responses were recorded as a preliminary indication of the likelihood that the hearing aids would be retained prior to proceeding to the audiologic assessment.

Each participant then completed the following measures with A2: (a) otoscopic examination of both ears, (b) a complete audiologic assessment, (c) unaided sound-field CST, (d) immittance measurements, (e) MMSE-2, and (f) a detailed case history. Air-conduction pure-tone audiometry was completed using a calibrated Grason-Stadler Inc. 61 audiometer with Etymotic Research ER-3A insert earphones (ANSI, 2010), at octave intervals from 250 through 8000 Hz plus 1500, 3000, and 6000 Hz. Next, the following measures were completed for each ear and in sequence: (a) speech reception threshold for Central Institute for the Deaf W-1 spondaic words using monitored live voice and 5-dB ascending step size, (b) word recognition scores for recorded Central Institute for the Deaf W-22 monosyllabic words (Department of Veterans Affairs, 2006) presented at 40 dB above speech reception threshold, and (c) bone-conduction hearing thresholds at octave intervals from 250 through 4000 Hz, plus 1500 and 3000 Hz. After the audiologic assessment, an unaided CST score was obtained in the sound field for a 65-dB SPL speech level and +3 dB SBR. Immittance measures followed with 226-Hz tympanometry and automated ipsilateral pure-tone acoustic reflex threshold measurement at 500, 1000, and 2000 Hz completed using a Grason-Stadler Inc. Model 39 immittance device. Next, the MMSE-2 was completed, followed by a

detailed case history (see Humes et al., 2017, for more details).

At the end of Visit 2, with an audiogram now available for the participant, the audiologist reviewed the audiograms for which the device chosen for each ear by the participant had been programmed (X, Y, or Z, with Z being flat 0 dB HL) and compared that to the audiogram just obtained. This provided the audiologist and the participant with some feedback as to how closely their selected hearing aid gain matched that needed for the specific hearing loss measured in each ear. For those who chose Z for one or both ears, it was noted that the programming of Z provided essentially no amplification and was a P hearing aid. They were then asked whether they wished to (a) keep the hearing aids as they were; (b) keep the hearing aids but go through an additional 4-week trial with AB services for individual programming and instruction in the use, care, and maintenance of their devices; or (c) return the study hearing aids for a refund.

Visits 3 and 4: Optional Follow-Up AB Trial

Of the 40 participants, 33 (21 of 26 non-P and 12 of 14 P) opted to complete a second trial period using AB programming protocols, and seven returned their hearing aids for a refund. (An additional P participant returned his hearing aids for refund soon after Visit 3.) The procedures described for AB participants in Humes et al. (2017) were followed for those participating in this additional 4-week trial with the programming and fitting of the devices taking place in Visit 3. The key distinguishing features of the AB approach compared with the CD model were the audiologist’s programming of the hearing aids to the participant’s hearing loss and verifying the match to targets with real-ear measures, together with the completion of a 45- to 60-min hearing aid orientation and counseling session. After 4–5 weeks ($M = 30.5$ days, $SD = 5.1$ days), the procedures described previously for outcomes measurement at Visit 2 were replicated. This second outcomes measurement session is referred to here as Visit 4.

Those participants for whom either Visit 2 or Visit 4 was their final session and who did not wish to purchase the hearing aids at session conclusion received a full refund of the purchase price and were exited from the study. Those who retained their hearing aids at the end of their final visit were fully informed about the make and model of their hearing aids, were provided with information about warranties and options for future follow-up as needed, and were then exited from the clinical trial.

Results and Discussion

The CD2 Study Sample

As noted in the introduction, the original Humes et al. (2017) randomized clinical trial had strict inclusion and exclusion criteria to make sure that all participants were viable candidates for the specific hearing aids used in that study and that there were no contraindications for

the use of hearing aids. In this CD2 clinical trial, there were very few inclusion or exclusion criteria, a scenario more like that envisioned for future OTC hearing aid purchasers. To determine whether this sample of participants was more heterogeneous audiometrically than the participants in the original clinical trial, we examined how many would have been eliminated from the original study for various reasons. In the original study, for example, there were targeted ranges for air-conduction pure-tone thresholds at each frequency, and the maximum asymmetry between thresholds for each ear was 20 dB at a given frequency. Twelve of the 40 CD2 participants, or 30%, would have been excluded from the original clinical trial based on these two audiometric criteria alone. In particular, eight CD2 participants had hearing loss that was too mild for inclusion in the original study, two had too much asymmetry of hearing loss, and two had both too much asymmetry and hearing loss that were too severe. In addition, in the original study, MMSE scores of 26 or greater were required, and three CD2 participants would not have met this criterion, although one of the three would have been eliminated by the audiometric exclusion criteria noted above. Clearly, the elimination of the more rigorous front-end screening of participants from the original clinical trial has achieved the intended goal of allowing a more heterogeneous group of participants into the CD2 clinical trial than in Humes et al. (2017), with the CD2 sample being one that might be considered more representative of the envisioned unscreened population of OTC hearing aid prospects.

Although the CD2 protocol did not require a hearing evaluation prior to participation, it is conceivable that the participants could have obtained one previously on their own. We asked all CD2 participants a series of questions near the end of Visit 2 about prior otoscopic examinations, hearing tests, or hearing screenings, both whether they had these (each defined explicitly and simply) and, if so, how recently. Ninety percent of the CD2 participants had a prior otoscopic examination of their ears, 58% within the past 12 months and most (75%) performed by a primary care physician. Of those 36 with otoscopic examinations, 26 (72%) had also received a hearing test such as that in Visit 2 of this study. Two others who self-reported no otoscopic examination also had a hearing test, resulting in a total of 28 of the 40 (70%) who had a hearing test at some time with 20 of these 28 tests (71%) having taken place over a year prior to enrollment in CD2. Of these 28, most (21, or 75%) had their hearing tested previously by an audiologist, followed by 18% who reported being tested by a physician or nurse. A total of 24 of the 28 (86%) were told they had a hearing loss, but only three of the 24 (12.5%) indicated that they followed through with the recommendation to get hearing aids, and only one of the 24 (4%) followed through with the recommendation to have a hearing retest in a year. Of the three who had indicated that they had tried hearing aids previously, one was incorrectly responding regarding the current study hearing aids. Two of the CD2 participants had previously worn a hearing aid or hearing aids for 6 weeks or less, one about 6 months prior

to enrollment and the other over 20 years prior to enrollment. (We had defined “new users” as those who had never worn hearing aids or had only worn them for a short time [for less than 6 weeks, more than 1 year ago], but clearly one participant failed to meet this definition, having worn hearing aids for a brief period [≤ 6 weeks] 6 months prior to enrollment.) Only nine of the 40 reported that they had received a hearing screening, with most (six, or 67%) indicating that the screen was more than 2 years ago (the longest interval among the response alternatives). Thus, a large portion of this unscreened sample had prior confirmation of hearing loss by a professional but failed to follow up on recommendations for future testing or the pursuit of hearing aids. Again, this is believed to be representative of the very group targeted for OTC devices and service delivery—the estimated 80% of older adults who have significant hearing loss but do not pursue hearing aids (Kochkin, 1993a, 1993b, 1993c, 2000, 2009; Perez & Edmonds, 2012).

Hearing Aids Selected and Demographic Measures

Table 1 shows the hearing aids selected for the right and left ears of each of the 40 CD2 study participants who completed the entire study protocol. As noted previously, because these selections were entirely at the participants’ discretion, it was conceivable that there would be no P participants in this study. As revealed by the summary of the participant selections in Table 1, however, this was clearly not the case. Of the 40 participants, 14 (35%) selected P hearing aids for both ears, and an additional five participants (12.5%) selected a P device for one of their ears. The majority (26, or 65%) selected an X or Y hearing aid for one or both ears. In the prior clinical trial by Humes et al. (2017), 6% of the 51 CD participants selected different hearing aids for each ear, whereas in this study, this was much more common with 22% of the CD2 participants doing so. We suspect that this is due to the elimination of the requirement for bilaterally symmetrical hearing loss in this study. Due to the increased likelihood of asymmetrical and mild hearing loss in the CD2 study than in the original clinical trial, it is possible that selection of a P device (Z) could be a good match for at least one of the ears. As a result, we considered those 14 CD2 participants who selected the P devices (Z) for both ears to be P participants and all others ($n = 26$) to be non-P CD2 participants. Because

Table 1. Hearing aids selected for each ear by the 40 participants in the clinical trial.

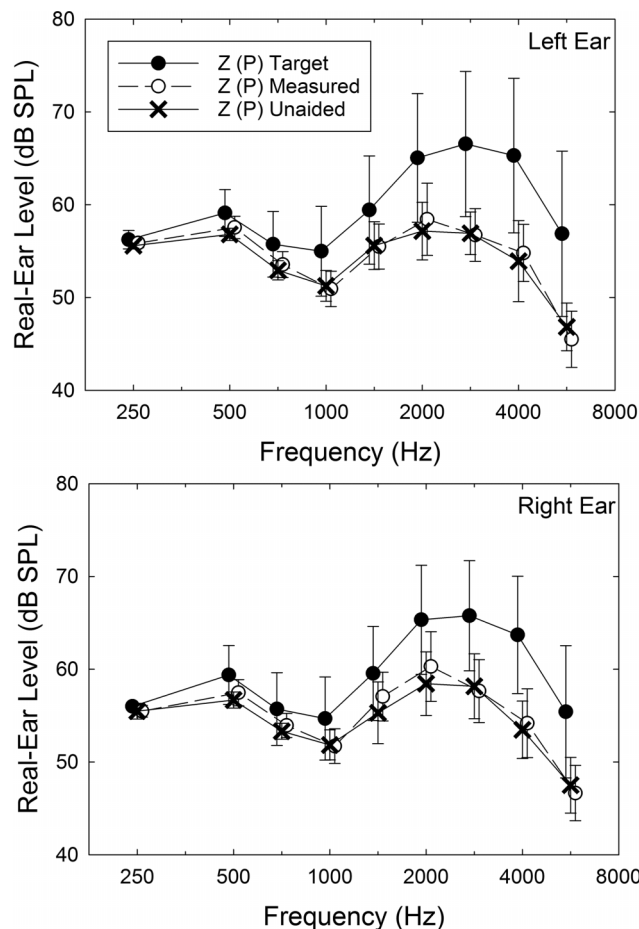
Hearing aid Choice: right	Hearing aid choice: left		
	X	Y	Z (P)
X	7	2	0
Y	2	10	2
Z (P)	2	1	14

Note. Choice Z was programmed to function, on average, as a P with no functional gain. P = placebo.

all 40 participants followed the CD approach to self-selection of the hearing aids, the same CD approach used in Humes et al. (2017), but the subgroups chose devices that were like ($n = 26$; at least one X or Y hearing aid) or unlike ($n = 14$; both Z or P hearing aids) the 51 CD participants in Humes et al. (2017), we have chosen to designate these two subgroups as CD2-CD and CD2-P, respectively.

Perhaps, the P devices, given that they were preprogrammed to be P on average, were not actually P devices for those who selected them. That is, perhaps, the actual real-ear gain was not zero as planned. Figure 1 shows the real-ear measurements obtained for the left (top) and right (bottom) ears for the P devices during Visit 2, after the devices had been worn for 4–5 weeks in the trial and were reconditioned after visual inspection as needed. Real-ear responses are shown for a 65-dB SPL speech stimulus unaided (X) and aided (open circles), and the two responses are virtually superimposed on one another. That is, the average real-ear response for these devices on these participants

Figure 1. Means, ± 1 SD, for the unaided (X), aided (unfilled circles), and NAL-NL2 prescribed (filled circles) real-ear gain for the ears that chose the Z (placebo) hearing aid in follow-up consumer-decides groups. Top panel shows the data for the left ear and the bottom panel shows the data for the right ear. Speech presentation level was 65 dB SPL.



is at or close to 0-dB insertion gain. Note, moreover, that the filled circles indicate what the target aided response would have been for these same ears when generated from the audiologic information using the NAL-NL2 prescriptive approach. The selected Z (P) devices underamplified the higher frequencies by 7–10 dB, and either Device X or Y would have likely generated an aided response much closer to the target for these same ears. Although the data shown in Figure 1 are for just one of the three input levels, 65 dB, similar results (± 2.5 dB, 250–4000 Hz; ± 5 dB, 6000 Hz) were obtained for each of the other two input levels (55 and 75 dB SPL). Thus, the average real-ear response for the Z (P) devices is consistent with the 0-dB insertion gain desired for the P devices.

Table 2 provides means and standard deviations for a variety of demographic variables and measures of performance for the CD1 group from Humes et al. (2017; CD1, $N = 51$), the entire CD2 group ($N = 40$) from this study, and the CD2-CD ($n = 26$) and CD2-P ($n = 14$) subgroups. Multiple independent-samples t tests were conducted on these data. For the CD1 versus CD2 group comparisons, only unaided CST scores, transformed into rationalized arcsine units (rau; Studebaker, 1985) prior to analyses, differed significantly, $t(89) = -3.82$; $p < .001$, with the CD2 group in this follow-up study showing higher unaided scores. Multiple independent-samples t tests comparing the CD2-CD and CD2-P subgroups on the demographic variables in Table 2 showed no significant ($p > .05$) differences on any measure. In addition, chi-square tests for group differences in gender and level of education were not statistically significant ($p > .05$) for either the CD1 versus CD2 group comparison or the CD2-CD versus CD2-P subgroup comparison. Thus, despite participant self-selection of devices that resulted in the formation of these two subgroups, the two subgroups were well matched on several demographic measures. The same can be said for the CD1 and CD2 group comparisons across studies, despite very few inclusion/exclusion criteria for the present CD2 study, with the lone exception being higher unaided CST scores in the present CD2 study.

Figure 2 shows the air-conduction pure-tone thresholds for the 26 participants in the CD2-CD subgroup (left panels) and the 14 individuals in the CD2-P subgroup (right panels). For each group, thresholds are provided for both the left (top) and right (bottom) ears. Large filled circles in each panel provide the median hearing loss for each subgroup, and the corresponding heavy solid lines show the bounds for hearing thresholds used as inclusion criteria in the prior clinical trial (Humes et al., 2017). The median thresholds for the CD2-CD subgroup are very similar to those for the same group in Humes et al. (2017) study (not shown in Figure 2), whereas the corresponding median thresholds for the P subgroup are better, especially above 2000 Hz. The number of participants in each subgroup with hearing thresholds above or below the bounds used for inclusion in the prior clinical trial (heavy solid lines in each panel) further illustrates the increased audiometric heterogeneity of the CD2 sample.

Table 2. Means and standard deviations for the original CD (CD1; $N = 51$) and follow-up CD (CD2; $N = 40$) groups, as well as the placebo (P; $n = 14$) and consumer-decides (CD; $n = 26$) CD2 subgroups, in the clinical trial for various demographic variables and baseline measures of performance.

Measure	Group CD1		Group CD2		Subgroup P		Subgroup CD	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age (years)	68.2	6.1	69.2	5.6	68.4	5.4	69.7	5.7
Loss duration (years)	8.7	10.6	11.5	13.4	10.1	13.1	12.2	13.7
MMSE	28.6	1.2	28.1	1.7	28.4	2.0	27.9	1.6
hfPTA-R (dB HL)	37.5	8.0	39.5	15.6	33.1	10.8	42.9	16.8
hfPTA-L (dB HL)	39.4	9.0	38.3	12.1	34.6	11.7	40.2	12.1
SRT-R (dB HL)	23.2	9.2	25.1	12.3	20.4	8.4	27.7	13.4
SRT-L (dB HL)	23.3	9.7	24.5	10.1	22.9	8.9	25.4	10.8
WRSQ-R (%)	91.8	6.4	88.1	11.9	90.0	8.1	87.1	13.5
WRSQ-L (%)	89.5	8.7	87.3	10.0	86.9	11.3	87.5	9.5
CST 65 SF (rau)	52.9*	19.4	68.9*	20.2	72.0	20.1	67.2	20.5
HHIE	30.0	17.8	33.4	16.7	30.6	16.5	34.9	16.9

Note. MMSE = Mini-Mental State Examination; hfPTA = high-frequency pure-tone average, 1000, 2000, and 4000 Hz; R = right; L = left; SRT = speech recognition threshold; WRSQ = word recognition score in quiet; CST 65 SF = Connected Speech Test at 65 dB SPL in the sound field; rau = rationalized arcsine unit; HHIE = Hearing Handicap Inventory for the Elderly.

* $p < .01$ for CD1 versus CD2 t test.

Hearing Aid Outcomes

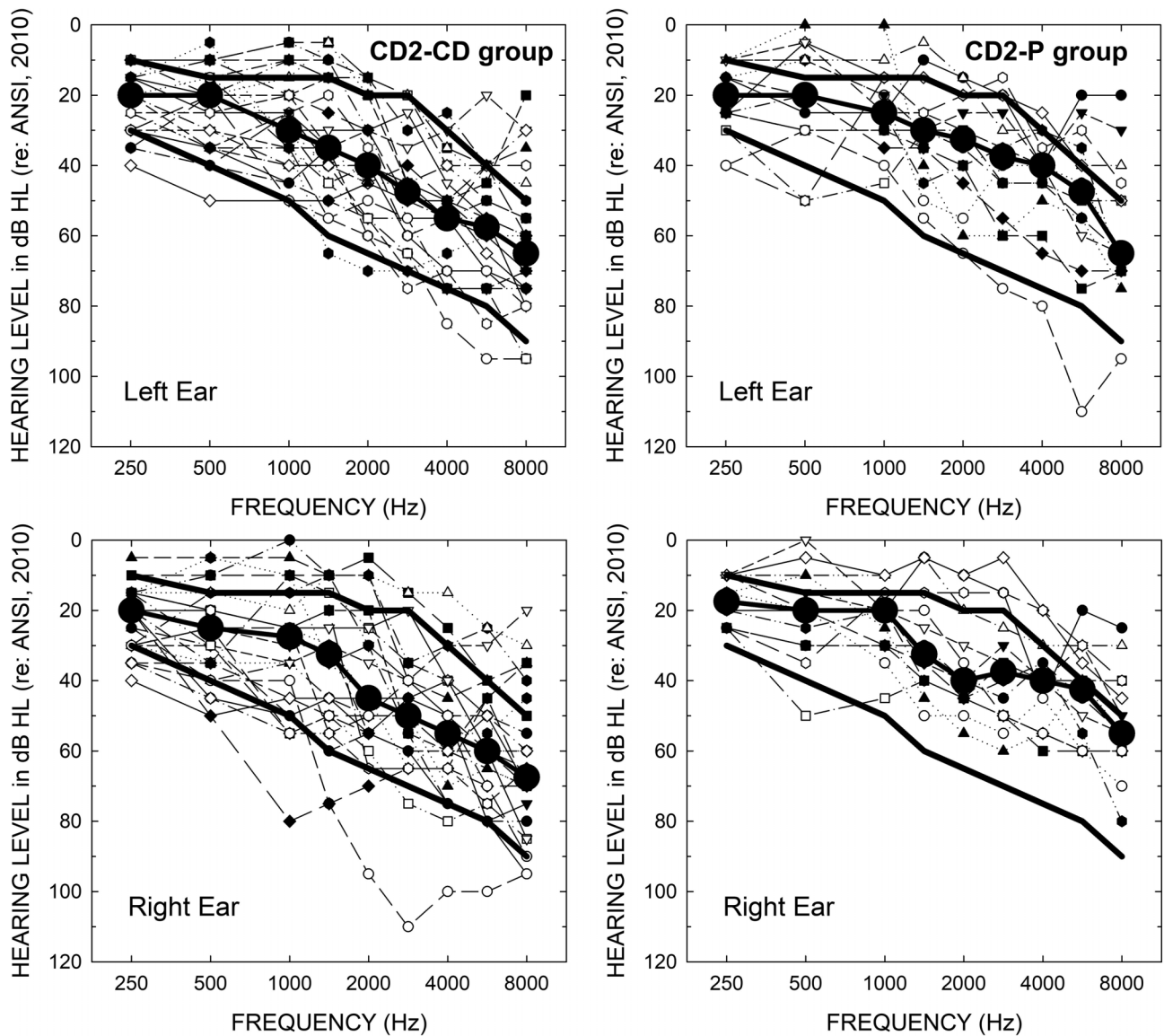
Figure 3 shows the means and standard errors for four outcome measures common to the CD1 (Humes et al., 2017) and CD2 groups. The small arrows in each panel of Figure 3 point to the mean values for the AB best practices group from Humes et al. (2017) for reference. Multiple independent-samples t tests revealed that the only significant difference observed for the four outcome measures in Figure 3 was for the measured CST benefit in rau, $t(89) = 5.66$, $p < .001$, with the CD2 follow-up group showing considerably lower scores than the CD1 group. Although not shown in Figure 3, there were also no differences in daily usage of the hearing aids between the CD1 ($M = 6.1$ hr; $SD = 3.9$ hr) and CD2 ($M = 6.2$ hr; $SD = 4.1$ hr) groups.

Although the focus has been placed on measures of benefit, aided performance relative to unaided performance, a similar pattern of findings was observed for mixed-model generalized linear model (GLM) analyses with a repeated-measures factor of aided versus unaided and a between-subjects factor of CD1 versus CD2. That is, the actual aided and unaided scores were analyzed, rather than the differences between those scores as measures of benefit. For the aided and unaided PHAPglob, PHAPavds, CST, and HHIE outcome measures, only the PHAPavds measure failed to show a significant main effect of condition with aided performance superior to unaided, $F(1, 89) = 2.14$, $p > .10$. Thus, the other three outcome measures showed significant benefit. Little or no benefit is expected for the PHAPavds measure (Cox & Gilmore, 1990; Humes et al., 2017). In addition, only the rau-transformed CST scores showed a significant interaction, $F(1, 89) = 32.03$, $p < .001$, between groups (CD1, CD2) and condition (aided, unaided). For all four outcome measures, significant main effects of group were not observed [largest, $F(1, 89) = 3.31$, $p > .05$].

Thus, both the analyses of relative benefit with t tests and the analyses of the actual aided and unaided scores comprising the benefit measures with GLM analyses failed to find significant differences in outcomes between the highly screened CD1 and unscreened CD2 samples, with the exception of the CST outcome measure.

Figure 4 explores the lone difference in hearing aid outcome between the CD1 and CD2 groups in more detail. CST benefit is the difference between the aided and unaided CST scores, and Figure 4 provides the means and standard errors for each of these rau-transformed constituent scores for both groups. As noted previously regarding the demographic measures in Table 2, the CD2 group has significantly higher unaided CST scores compared with the CD1 group. Note, however, that there are no significant differences between these two groups in aided CST scores, as shown in Figure 4. Thus, the diminished CST benefit observed previously for the CD2 group in Figure 3 is due to the differences between these groups in their unaided scores, a measure unaffected by the hearing aids. Given that several CD2 participants had hearing thresholds that were better than allowed in the CD1 group (see Figure 2) and that greater asymmetry was possible in this study, we thought that differences in better ear hearing thresholds between the CD1 and CD2 groups might explain the superior performance of the CD2 group for unaided sound-field CST scores. There were no significant differences ($p > .05$), however, between the better ear pure-tone thresholds of the CD1 and CD2 groups at any frequency, and the correlations between unaided sound-field CST scores and better ear thresholds were moderate (-0.30 to -0.55) and significant ($p < .05$) for both the CD1 and CD2 groups. Given the other similarities demographically (see Table 2) between the CD1 and CD2 samples, it is unclear why their unaided CST scores differ significantly,

Figure 2. Audiograms for the 26 members of the CD2-CD (left) and the 14 members of the CD2-P (right) subgroups. Top panels show data for the left ears and bottom panels show data for the right ears. Large filled black circles show the median audiograms for the CD and P subgroups. The heavy solid lines in each panel show the limits for inclusion in the original clinical trial of Humes et al. (2017). CD = consumer-decides; CD2 = follow-up; P = placebo.



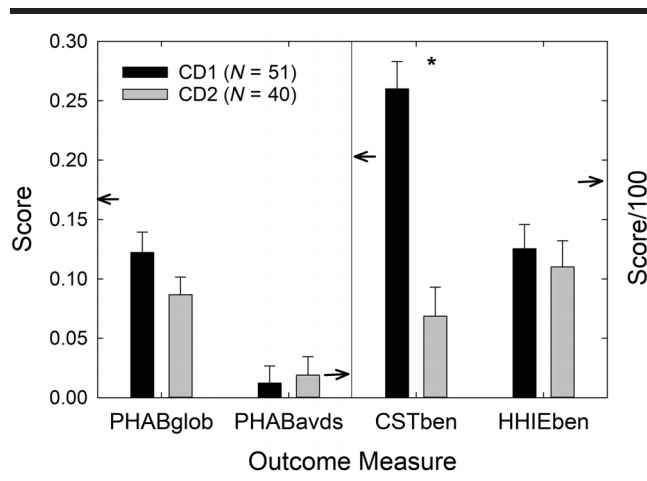
but this is clearly the reason underlying the reduced CST benefit in the CD2 sample (see Figures 3 and 4). This difference, being unaided, has nothing to do with the performance of the hearing aids selected by the participants.

Recall that there were two different aided CST scores obtained in this study and in Humes et al. (2017), an “as-worn” aided CST score and a “postmaintenance” CST score. Figures 3 and 4 report the latter scores, but statistical findings for both the independent-samples *t* tests of benefit and the mixed-model 2×2 factorial GLM analyses of the raw scores were identical with either the as-worn or post-maintenance CST scores serving as the aided measure of speech recognition. When the as-worn and postmaintenance-aided

CST scores were compared, transformed as-worn CST scores were 10 rau lower than transformed postmaintenance scores, $F(1, 89) = 57.79, p < .001$, as would be expected, but there were no main effects of group (CD1, CD2), $F(1, 89) = 0.61, p > .10$, and no interactions with group, $F(1, 89) = 0.04, p > .10$.

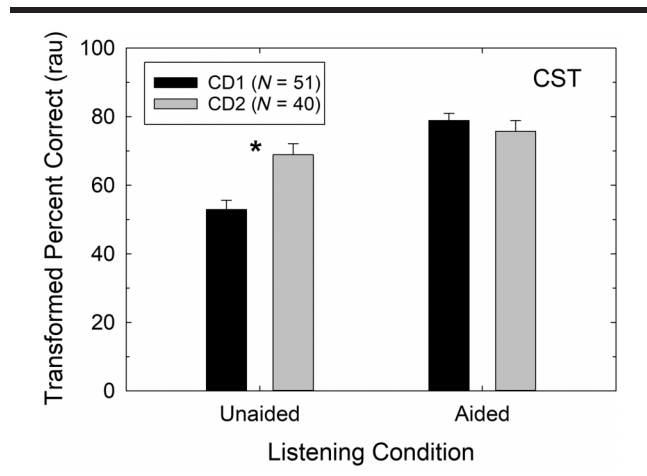
For three of the four outcome measures in Figure 3, the CD2 group performed worse than the CD1 group, although this is only a statistically significant difference for CST benefit, as noted. For the CD1 cohort, however, it was not possible to choose a P device for either ear. In the original study, all P devices were fit within the AB best practices service delivery model. At least a portion of the reduced

Figure 3. Mean (+1 SE) for the four main outcome measures common to both the CD1 (black bars) and CD2 (gray bars) cohorts. PHAB scores in the left panel should be referenced to the left ordinate, whereas CST and HHIE benefit scores in the right panel should be referenced to the right ordinate. Note that the right ordinate has been scaled by 1/100 to make it possible to plot all the benefit measures in the same figure. Small arrows pointing to ordinates reference the audiology-based means from Humes et al. (2017). The asterisk above the CST benefit scores indicates that this difference between the CD1 and CD2 cohorts was statistically significant ($p < .05$). CD = consumer-decides; CD1 = original CD; CD2 = follow-up CD; PHAB = Profile of Hearing Aid Benefit; CST = Connected Speech Test; HHIE = Hearing Handicap Inventory for the Elderly.



outcomes observed in CD2 participants relative to CD1 participants in Figure 3 could be due to the inclusion of the 14 Ps within the group of 40 CD2 participants. The hearing aid outcomes for the 40 CD2 participants, after partitioning them into the two subgroups, CD ($n = 26$; left panels) and

Figure 4. Mean (+1 SE) for the unaided and aided CST score, in rau, for the CD1 (black bars) and CD2 (gray bars) cohorts. The asterisk above the unaided CST scores indicates that this difference between the CD1 and CD2 cohorts was statistically significant ($p < .05$). rau = rationalized arcsine unit; CST = Connected Speech Test; CD1 = original consumer-decides; CD2 = follow-up consumer-decides.



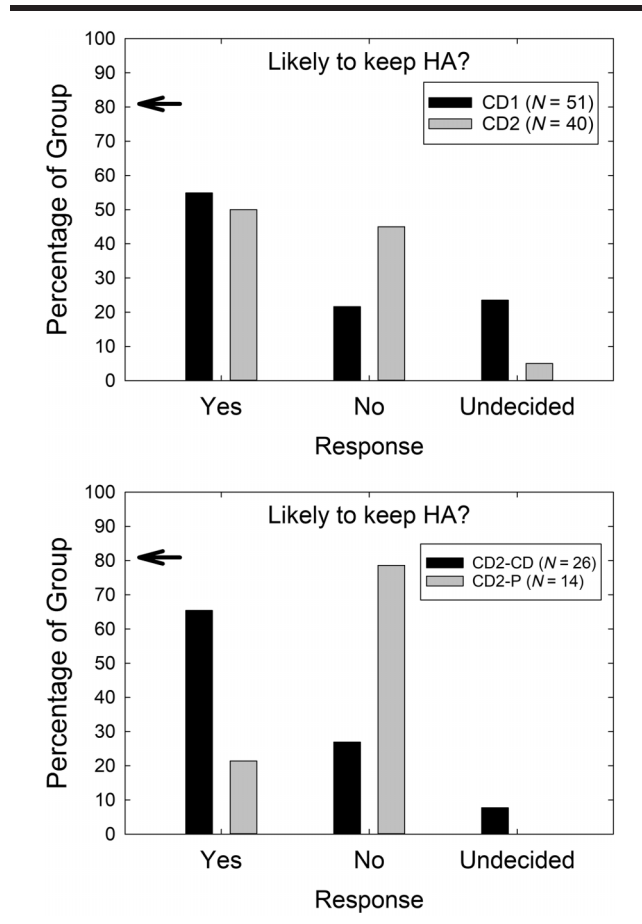
P ($n = 14$; right panels), yielded a closer match between the original ABCD CD group (CD1) and the CD2-CD subgroup in all cases, but a large difference between these groups remains for CST benefit. Thus, the inclusion of P devices and the selection of these devices by several CD2 participants tended to reduce the size of the positive outcomes observed relative to CD1, but not sufficiently to yield statistically significant decreases.

The foregoing analyses address the first research question posed in the Introduction. Specifically, do the results from the original highly screened CD participants (CD1) in Humes et al. (2017) generalize to a new group of unscreened CD participants (CD2) and a wider range of frequency-gain characteristics among the devices? The answer is “yes,” with the most notable exception being the CST benefit measure for which the current unscreened CD2 participants show significantly less benefit (see Figure 3). As noted above, however, this is not a result of diminished aided performance but of significantly higher unaided performance in the CD2 sample. That is, if one chose an equally feasible outcome measure for this study based on aided speech recognition performance, rather than the relative benefit, then the CD1 and CD2 samples would not differ on that aided outcome measure (see Figure 4).

Replication of findings for the CD service delivery model from Humes et al. (2017) was examined in several other ways as well. For example, in both studies, at the end of the 4- to 5-week trial, the CD participants were asked whether they were likely to keep their hearing aids. This was prior to any explanation of the study design or the devices selected. The responses to this query from the CD2 study participants, as well as the CD1 cohort from Humes et al. (2017), are shown in the top panel of Figure 5. There was a significant difference in responses across studies, $\chi^2(2) = 8.97, p < .05$, with a higher percentage of CD2 participants (45%) than CD1 participants (22%) indicating that they were not likely to keep the hearing aids. The percentages in the lower panel of Figure 5 for the two CD2 subgroups, CD and P, show that it is primarily the P participants who indicated that they were unlikely to keep their hearing aids, 80% so indicating. There were significantly different response distributions across these two subgroups, $\chi^2(2) = 9.99, p < .05$, with the majority of Subgroup P responding “no” and the majority of Subgroup CD responding “yes” to this interim query about the likelihood of keeping their hearing aids. The distributions of responses for the CD1 group (top, black bars) and the CD2-CD group (bottom, black bars) in Figure 5 are very similar and further support replication of findings across studies when Ps are removed from consideration for CD2.

At completion of the trial at Visit 2, each participant had to decide whether to keep the hearing aids or request a refund. Each participant was presented with the opportunity to go through AB service delivery for another 4-week trial at no cost at this time too, deferring their final decision regarding a refund until the completion of that additional optional 4-week trial period. Thirty-three of the 40 CD2 participants kept the study aids at this time and agreed to

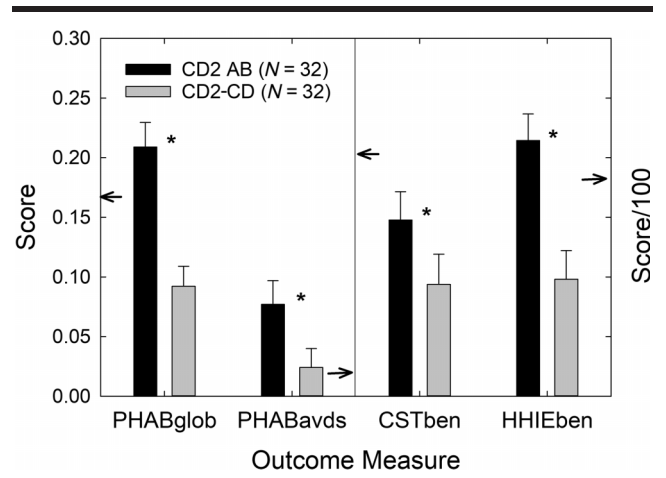
Figure 5. The percentage of each group reporting each of the possible responses (abscissa) in response to the query, “Based on your experience in the trial, are you planning to keep the hearing aids?” The top panel compares the results between the CD1 (black) and CD2 (gray) studies, whereas the bottom panel compares the CD2-CD (black) and CD2-P (gray) subgroups. CD = consumer-decides; CD1 = original CD; CD2 = follow-up CD; P = placebo; HA = hearing aid.



complete the AB service delivery and return for outcome measures 4 weeks later. The remaining seven participants, or 17.5% of the CD2 group, opted to return their hearing aids at the conclusion of Visit 2. In contrast, only one of the 51 CD1 participants opted to return their hearing aids at the end of the trial and forego an optional follow-up trial period following best practices. This difference in rate of returns for credit at the official end of the trial between the CD1 and CD2 samples was statistically significant, $\chi^2(2) = 7.4, p < .05$. Of the seven returns at the end of Visit 2, three were from the P subgroup and four were from the CD subgroup.

One of the 33 CD2-P participants who had planned to return for Visits 3 and 4 changed his mind soon after Visit 3 and opted to return his hearing aids for a refund instead. As a result, outcome data are available from Visit 4, following the optional 4-week follow-up trial with AB service delivery, for 32 of the 40 original CD2 participants, 11 of 14 from subgroup P and 21 of 26 from subgroup CD. Figure 6 shows the performance of these 32 participants at

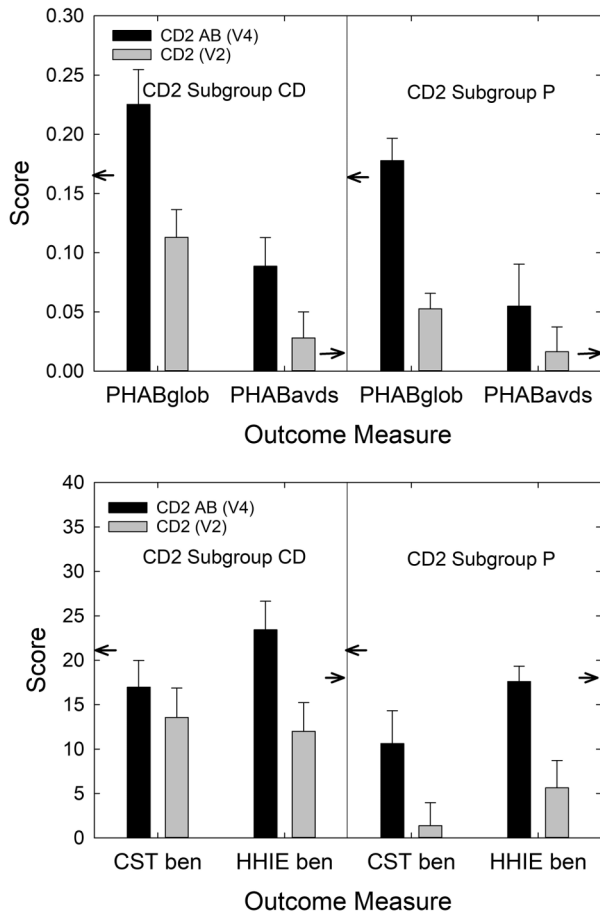
Figure 6. Means (+1 SE) for the four main outcome measures for the 32 CD2 participants who completed the 4- to 5-week clinical trial (black bars) and the additional 4-week AB follow-up trial (gray bars). PHAB scores in the left panel should be referenced to the left ordinate, whereas CST and HHIE benefit scores in the right panel should be referenced to the right ordinate. Note that the right ordinate has been scaled by 1/100 to make it possible to plot all the benefit measures in the same figure. Small arrows pointing to ordinates reference the AB means from Humes et al. (2017). Asterisks above each pair of vertical bars indicate that the two groups differed significantly ($p < .05$) in performance for all four outcome measures. CD = consumer-decides; AB = audiology-based; CD2 = follow-up CD; PHAB = Profile of Hearing Aid Benefit; CST = Connected Speech Test; HHIE = Hearing Handicap Inventory for the Elderly.



Visit 2 after CD service delivery and again at Visit 4 after AB service delivery. The small arrows pointing to various axes in each panel again provide a reference to the AB means from the original clinical trial (Humes et al., 2017). Multiple paired-samples t tests found significant ($p < .05$) differences between AB and CD service delivery for all outcomes shown, with an AB service delivery consistently superior to CD service delivery (see Figure 6).

Figure 7 shows these same data from the 32 CD2 participants for the outcomes obtained at Visit 4 partitioned by subgroup (21 CD, 11 P). A 2×2 mixed-model GLM analysis with a between-subjects factor of subgroup (CD, P) and a repeated-measures factor of service delivery model (CD, AB) was performed for each of the four outcome measures in Figure 7. The analyses found a significant main effect of service delivery model [AB > CD; minimum, $F(1, 30) = 6.1, p < .05$] for all four outcomes; no significant effect of initial subgroup on outcomes [maximum, $F(1, 30) = 2.7, p = .11$], except for CST benefit, $F(1, 30) = 4.7, p < .05$, for which the CD group outperformed the P group; and no significant Service Delivery Model \times Subgroup interaction [maximum, $F(1, 30) = 1.3, p = .26$]. A similar analysis for daily usage found no significant main effects or interactions [maximum, $F(1, 30) = 1.8, p = .19$]. Humes et al. (2017) reported that hearing aid outcomes improved significantly for CD and P participants following subsequent AB service provision, and this CD2 study replicates those findings with an unscreened sample of older adults.

Figure 7. Mean (+1 SE) for the four main outcome measures after the 4- to 5-week clinical trial (gray bars) and after the completion of the optional AB follow-up 4-week trial (black bars). PHAB scores are shown in the top panel, whereas CST and HHIE benefit scores appear in the bottom panel. In each panel, the CD2 data have been partitioned to display the results for the CD2-CD subgroup (left) and the CD2-P subgroup (right). Small arrows pointing to ordinates reference the AB means from Humes et al. (2017). AB = audiology-based; PHAB = Profile of Hearing Aid Benefit; CST = Connected Speech Test; HHIE = Hearing Handicap Inventory for the Elderly; CD = consumer-decides; CD2 = follow-up CD; P = placebo.



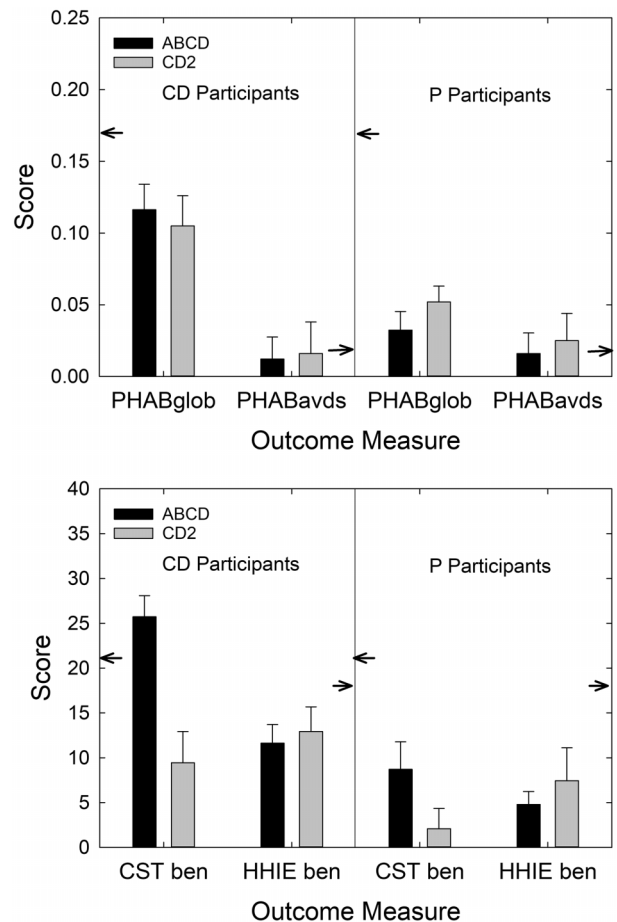
When asked at the end of Visit 4 if the participant was likely to keep their hearing aids, only two of the 32 answered “no,” one from the CD subgroup and one from the P subgroup. Of the 32 participants, 18 (56%) had answered “yes” to this query after the initial 4- to 5-week trial, and this increased to 27 (84%) “yes” responses after AB service delivery, with another three, or 9%, answering “undecided.” Humes et al. (2017) reported similar increases in likely retention of hearing aids for CD and P participants following subsequent AB service delivery.

Efficacy of the CD Model

As noted in the Introduction, a secondary research question addressed in this follow-up study is as follows: Are the results from non-P self-fit devices (CD2-CD) superior

to those from P self-fit devices (CD2-P) in this study? Comparing the two CD2 subgroups to one another allows one to establish the efficacy of the CD service delivery model with a P control within the same CD service delivery framework. For the primary outcome measure, PHABglobal, CD2-CD subgroup scores were significantly greater than CD2-P subgroup scores, $t(35.3) = 2.23, p < .05$. There were no significant differences in the performance of the CD2-CD and CD2-P subgroups on the other three outcome measures in Figure 8, although the difference in CST benefit approached significance ($p = .085$). Using mixed-model 2×2 factorial GLM analyses to examine the aided and

Figure 8. Mean (+1 SE) for the four main outcome measures common to both the ABCD (black bars) and CD2 (gray bars) studies. The data from the ABCD study are either for the CD group (left) or the group from Humes et al. (2017). PHAB scores are shown in the top panel, whereas CST and HHIE benefit scores appear in the bottom panel. In each panel, the CD2 data have been partitioned to display the results for the CD2-CD subgroup (left) and the CD2-P subgroup (right). Small arrows pointing to ordinates reference the AB means from Humes et al. (2017). The asterisk above the CST benefit scores indicates that this difference between the ABCD and CD2 cohorts was statistically significant ($p < .05$). CD = consumer-decides; CD2 = follow-up CD; P = placebo; PHAB = Profile of Hearing Aid Benefit; CST = Connected Speech Test; HHIE = Hearing Handicap Inventory for the Elderly.



unaided constituents of each of the four relative benefit measures, no significant main effects of CD2 subgroup were observed and no interactions with subgroup ($p > .10$). For all but PHAPavds aided and unaided scores ($p > .10$), significant main effects of amplification were observed ($p < .05$). Because the number of participants in the P group was determined entirely by participant choice, the P subgroup was relatively small ($n = 14$), and this likely impacted the failure to observe statistically significant differences in outcomes between the CD and P subgroups. Although not shown in Figure 8, hearing aid usage also did not differ significantly ($p = .40$) between these two CD2 subgroups, with a mean daily usage of 5.7 hr ($SD = 4.3$ hr) for subgroup CD and 6.9 hr ($SD = 3.8$ hr) for subgroup P. The results from non-P self-fit devices (CD2-CD) do not differ statistically from those from P self-fit devices (CD2-P) in this study of unscreened older adults.

Cohen's d (Cohen, 1988) is a common metric of effect size with suggested interpretations as follows: 0.8 is a large effect, 0.5 is a medium effect, and 0.2 is a small effect. For the primary outcome measure, PHABglobal, Cohen's d was 0.65, a medium-to-large effect size and one similar to that for the CD versus P comparison ($d = 0.76$) from Humes et al. (2017). For the secondary outcome measure, CST benefit, Cohen's d was 0.53, a medium effect size and one that is smaller than that ($d = 0.92$) observed in Humes et al. (2017). The effect size for HHIE benefit, 0.54 in Humes et al. (2017), yielded a similar medium effect size here ($d = 0.40$). As in Humes et al. (2017), Cohen's d showed no effect for the PHABavds, with a d value of -0.08 observed previously and -0.09 observed here for CD versus P comparisons. Finally, usage showed no effect in Humes et al. (2017) with $d = 0.04$, whereas in this study, a small negative effect size was observed ($d = -0.28$) indicating that the P participants tended to wear their devices longer each day than the CD participants. Contrary to the foregoing statistical analyses, the observed medium effect sizes support the efficacy of the CD selection method, this time with relative comparisons to P devices within the same CD service delivery model. The observed effect sizes also replicate those of Humes et al. (2017) for the CD versus P comparisons.

General Discussion and Conclusions

This CD2 follow-up clinical trial yielded results for these 40 older adults that were in general agreement with those obtained from another sample (CD1) of 51 older adults in Humes et al. (2017). The primary exception to this broad summary statement pertains to the secondary outcome measure, CST benefit. The benefit measured in this study for CD2 participants was significantly lower than that measured in the CD participants of Humes et al. (2017). Closer examination of the CST benefit measures revealed that the aided performance of the CD2 participants was high and very similar to that observed previously by Humes et al. (2017) for CD1 participants, but the unaided scores for CD2 participants were significantly higher than those

from the CD1 sample. Given the close correspondence in the demographics of the two study samples, CD1 and CD2 (see Table 2), it is unclear why the unaided scores of the CD2 cohort were much higher. The other three outcome measures examined here, including the primary outcome measure (PHABglobal), did not differ between the CD1 group and the current CD2 sample. Thus, in large part, this study replicated the findings of the original Humes et al. (2017) study for the CD service delivery model on a more heterogeneous unscreened sample.

Regarding the efficacy of the CD service delivery model, results were mixed. Statistical comparison of outcomes between the CD2-CD subgroup and the CD2-P subgroup revealed few significant differences. This could be due, in part, to the relatively small sample size of self-selected Ps ($n = 14$). P assignment was by participant choice, rather than random assignment by the investigators, and only 14 of the 40 CD2 participants selected P devices for both ears. The prevalence of medium effect sizes (Cohen's d) between the outcomes for the CD2-CD and CD2-P subgroups, however, supports the efficacy of the CD self-selection procedure.

As noted, *placebo* was defined here as those CD2 participants who selected both devices to be P (Z). However, in addition to the 14 who did so, another five participants selected the P device for either the right or left ear (see Table 1). We did not include these participants in the P subgroup because the CD2 study lacked a priori limitations on the amount of asymmetry that could be manifested by a participant. For those with healthy hearing in one ear, the choice of the P (Z) hearing aid could be entirely appropriate. In addition, given two hearing aids, assumed symmetrical hearing loss, and one hearing aid with higher gain (X or Y) than the other (Z), sound-field performance on the CST, as well as performance in many typical everyday listening situations, would likely be determined by the ear with the functional hearing aid, rather than the ear with the P device. That is, given X/Z or Y/Z hearing aid pairs and the fairly symmetrical hearing loss expected in older adults, sound-field speech recognition performance should be determined by the ear with the lower aided thresholds: the ear with the X or Y hearing aid. In that sense, such mixed X/Z or Y/Z pairs would be considered to function more like corresponding X/X or Y/Y pairs than Z/Z pairs. For this reason, mixed X/Z and Y/Z pairs were grouped with the other non-P pairs (X/X or Y/Y) for the CD2 subgroup comparisons in this report.

The CD2 study, in addition to making it possible to select a P device from among the device alternatives, also eliminated the front-end screening of participants. As noted in the introduction, this was done to better emulate what is envisioned as a common point of entry into the OTC hearing aid market. There was more variation in hearing loss among the CD2 study participants than in the earlier CD group, as was demonstrated by the audiograms in Figure 2. There was also greater variation in MMSE scores among the CD2 cohort than among the CD1 cohort. Nonetheless, despite general loosening of the

inclusion and exclusion criteria for the CD2 study, the demographics of this sample of 40 older adults were very similar to those from the CD1 sample of 51 older adults (see Table 2). Importantly, because the investigators did not perform an audiologic evaluation to determine study eligibility, we added several study questions to the CD2 study about prior ear exams, hearing tests, and hearing screenings by others. Many of the CD2 participants had received such evaluations, but few followed up on the ensuing recommendations to seek hearing aids or to have an annual reevaluation of their hearing. In that sense, this sample is exactly the type targeted by the OTC pathway, representative of the 80% with significant hearing difficulty who do not typically pursue hearing aids (Kochkin, 1993a, 1993b, 1993c, 2000, 2009; Perez & Edmonds, 2012).

Across both the CD1 and CD2 studies, about 50%–60% of the CD participants indicated at the end of the trial that they were likely to keep their hearing aids prior to divulging any information about the nature of the hearing aids worn. If we assume a 50% success rate among participants who are representative of the 80% who fail to access hearing aids, this would represent an additional 40% or a total of 60% of those older adults with significant hearing loss. This would triple the number of such individuals purchasing and keeping hearing aids, a tremendous impact on the uptake of hearing aids by older adults.

As importantly, in both studies, the vast majority of participants, once fully debriefed, opted to complete an additional trial period with AB best practices service provision. That is, at least in the context of these two clinical trials, the CD experience did not have a negative impact on their pursuit of improved hearing. In addition, in both studies, AB service delivery significantly improved outcomes beyond the positive outcomes obtained from the initial CD trial. This also was manifested in the improved uptake or reduced returns for refund following AB service delivery. This suggests that a particularly viable approach to OTC hearing aid provision may be to offer a pathway to OTC devices within a framework that readily allows for subsequent follow-up with AB service provision as needed and when elected by the OTC hearing aid wearer. It should be noted, however, that the follow-up AB services were provided to CD1 and CD2 study participants at no additional cost. It is unclear how many would have completed the additional 4-week trial with AB services if payment was required for such services.

One potential concern about the OTC route to improved hearing involves concerns about the potential to miss “red flag” conditions (Food and Drug Administration, 1977) when the consumer is not required to have prior ear or hearing evaluations. Of the eight red flag conditions noted by the Food and Drug Administration, we explicitly asked about or gathered information about six of the eight in the CD2 study; we did not explicitly ask about existing auditory/ear pain or discomfort, nor did we ask about acute or chronic dizziness. Such complaints were noted, however, if volunteered by the participant during the evaluation, and one participant noted occasional pain from behind the

right ear during the audiologic evaluation at Visit 2. Of the remaining six red flag conditions, none of the 40 CD2 participants manifested the problem for four of these conditions (active drainage; air–bone gaps ≥ 15 dB at 500, 1000, and 2000 Hz; sudden or rapidly progressing hearing loss; sudden unilateral hearing loss). One of the 40 CD2 participants had bilateral pinna deformity, and three others had apparent scarring of the tympanic membrane in one ($n = 2$) or both ($n = 1$) ears. The most prevalent red flag problem observed was excessive cerumen, which was reported for 10 CD2 participants (25%) at Visit 2, in one ear for seven participants and in both ears for three others. The participant with pinna deformity was among the seven with excessive cerumen in one ear. Those with excessive cerumen in one or both ears, as observed at Visit 2 (and assumed to be present at Visit 1), did not differ significantly ($p > .05$) from those without excessive cerumen regarding their expressed likelihood to keep their hearing aids at Visit 2, $\chi^2(2) = 0.74, p > .10$, or their decision to return the hearing aids at that visit following debriefing, $\chi^2(1) = 1.4, p > .10$. Thus, the presence of this most prevalent red flag condition did not impact ultimate outcomes regarding purchase of their hearing aids.

The CD2 study largely replicated the findings from the CD1 service delivery group reported in Humes et al. (2017). Positive outcomes were again obtained for CD2 participants self-selecting their own hearing aids despite allowing a more heterogeneous sample to participate and despite including a P device among the hearing aid options in this study. The CD self-selection approach was also found to be efficacious in that medium effect sizes typically were observed for the CD2-CD subgroup relative to the CD2-P subgroup. As in the CD1 cohort, about half of the CD2 participants indicated that they were likely to keep their hearing aids at the completion of the initial 4- to 5-week trial. This percentage increased considerably to about 85% following completion of an additional 4-week trial with AB service provision, also consistent with prior CD1 findings (Humes et al., 2017). Finally, CD2 participants also showed significant improvements in measured hearing aid outcomes following AB service provision, again replicating the findings from the original study in that regard (Humes et al., 2017).

The CD service delivery model shares some attributes with approaches being pursued for self-fitting hearing aids (Convery, Keidser, Dillon, & Hartley, 2011). In fact, the instructions for the self-fitting of the devices used by both the CD1 and CD2 groups were adapted from those developed by Caposecco et al. (2011) for self-fitting hearing aids. Self-fitting hearing aids, however, are devices that may differ from the devices used here (Keidser & Convery, 2016). For example, the devices include signal generators to measure hearing thresholds via in situ audiometry and then generate individualized frequency-gain targets appropriate for the measured hearing loss. Much of the validation of the self-fitting model of service provision to date has focused on evaluation of the instructions provided, the ability of adults to complete the self-fitting procedure, the validity

or reliability of the in situ audiometry, and differences in electroacoustic performance between these devices and conventional hearing aids (Convery, Keidser, Seeto, & McLelland, 2017; Jilla, Johnson, & Danhauer, 2018; Keidser & Convery, 2016; Reed, Betz, Kendig, Korczak, & Lin, 2017; Reed, Betz, Lin, & Mamo, 2017; Wong, 2011). Only recently have outcomes such as those in this study and in Humes et al. (2017) been examined for self-fitting hearing aids. Keidser and Convery (2018) compared the outcomes of 10 experienced hearing aid users using hearing aids fit previously with AB services to outcomes from these same 10 individuals when wearing new self-fitting hearing aids that were self-fitted. No significant differences were observed in HHIE scores or abbreviated PHAB (Cox & Alexander, 1995) scores between the AB and CD fittings. Satisfaction was greater for the conventionally fitted (AB) devices than the self-fitted (CD) devices, an observation consistent with the findings of Humes et al. (2017). Further research is needed to better evaluate consumer-driven fitting options as a means to increase both the affordability and accessibility of hearing health care.

Acknowledgments

This study was supported, in part, by National Institute on Deafness and Other Communication Disorders Research Grant R01 DC011771, awarded to the first author. This trial was supported by National Institute on Deafness and Other Communication Disorders Research Grant R01 DC011771 to Indiana University, awarded to Larry E. Humes, principal investigator. The authors would like to thank Christine Herring for her help in the initial planning stages of this follow-up trial.

References

- American National Standards Institute.** (2009). *Specifications of hearing aid characteristics* (ANSI S.322-2009). New York, NY: Author.
- American National Standards Institute.** (2010). *Specifications for audiometers* (ANSI S.36-2010). New York, NY: Author.
- American Speech-Language-Hearing Association.** (2015). *Hearing aids for adults* [Practice portal]. Rockville, MD: Author. Retrieved from <http://www.asha.org/Practice-Portal/Professional-Issues/Hearing-Aids-for-Adults/>
- Caposecco, A., Hickson, L., & Meyer, C.** (2011). Assembly and insertion of a self-fitting hearing aid: Design of effective instruction materials. *Trends in Amplification, 15*, 184–195.
- Ciletti, L., & Flamme, G. A.** (2008). Prevalence of hearing impairment by gender and audiometric configuration: Results from the National Health and Nutrition Examination Survey (1999–2004) and the Keokuk County Rural Health Study (1994–1998). *Journal of the American Academy of Audiology, 19*, 672–685.
- Cohen, J.** (1988). *Statistical power analysis for the behavioral sciences*. London, United Kingdom: Routledge.
- Convery, E., Keidser, G., Dillon, H., & Hartley, L.** (2011). A self-fitting hearing aid: Need and concept. *Trends in Amplification, 15*(4), 157–166.
- Convery, E., Keidser, G., Seeto, M., & McLelland, M.** (2017). Evaluation of the self-fitting process with a commercially available hearing aid. *Journal of the American Academy of Audiology, 28*(2), 109–118.
- Cox, R. M., & Alexander, G. C.** (1995). The abbreviated Profile of Hearing Aid Benefit. *Ear and Hearing, 16*(2), 176–186.
- Cox, R. M., Alexander, G. C., & Gilmore, C.** (1987). Development of the Connected Speech Test (CST). *Ear and Hearing, 8*(Suppl. 5), 119S–126S.
- Cox, R. M., & Gilmore, G. C.** (1990). Development of the Profile of Hearing Aid Performance (PHAP). *Journal of Speech and Hearing Research, 33*, 343–357.
- Department of Veterans Affairs.** (2006). *Speech recognition and identification materials, Disc 4.0*. Mountain Home, TN: VA Medical Center.
- Dillon, H., Keidser, G., Ching, T. Y. C., Flax, M. R., & Brewer, S.** (2011). The NAL-NL2 prescription procedure. *Phonak Focus, 40*, 1–10.
- Folstein, M. F., Folstein, S. E., White, T., & Messer, M. A.** (2010). *Mini-Mental State Examination—2nd Edition (MMSE-2)*. Lutz, FL: Psychological Assessment Resources.
- Food and Drug Administration.** (1977). Title 21—Food and Drugs, Chapter 1, FDA, Department of HEW, Subchapter H, Medical Devices, Part 801—Hearing Aid Devices; 801.420—Professional and Patient Labeling; 801.421—Conditions for Sale. 42 Fed. Reg. 9286–9296 (31) (proposed Feb. 15, 1977)
- Humes, L. E., Rogers, S. E., Quigley, T. M., Main, A. K., Kinney, D. L., & Herring, C.** (2017). The effects of service-delivery model and purchase price on hearing-aid outcomes in older adults: A randomized double-blind placebo-controlled clinical trial. *American Journal of Audiology, 26*(1), 53–79.
- International Standards Organization.** (2000). *Statistical distribution of hearing thresholds related to age and gender* (ISO 7029:2000). Geneva, Switzerland: Author.
- Jilla, M. A., Johnson, C. E., & Danhauer, J. L.** (2018). Disruptive hearing technologies and mild sensorineural hearing loss II: Current research on affordable technologies and direct-to-consumer models. *Seminars in Hearing, 39*, 146–157.
- Keidser, G., & Convery, E.** (2016). Self-fitting hearing aids: Status quo and future predictions. *Trends in Hearing, 20*, 1–15.
- Keidser, G., & Convery, E.** (2018). Outcomes with a self-fitting hearing aid. *Trends in Hearing, 22*, 1–12.
- Kochkin, S.** (1993a). MarkeTrak III: Why 20 million in U.S. don't use hearing aids for their hearing loss. Part I. *The Hearing Journal, 46*(1), 20–27.
- Kochkin, S.** (1993b). MarkeTrak III: Why 20 million in U.S. don't use hearing aids for their hearing loss. Part II. *The Hearing Journal, 46*(2), 26–31.
- Kochkin, S.** (1993c). MarkeTrak III: Why 20 million in U.S. don't use hearing aids for their hearing loss. Part III. *The Hearing Journal, 46*(4), 36–37.
- Kochkin, S.** (2000). MarkeTrak V: Consumer satisfaction revisited. *The Hearing Journal, 53*(1), 38–55.
- Kochkin, S.** (2009). MarkeTrak VIII: 25-year trends in hearing health market. *The Hearing Review, 16*(11), 12–31.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., . . . Altman, D. G.** (2010). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomized trials. *British Medical Journal, 340*, e869.
- Perez, E., & Edmonds, B. A.** (2012). A systematic review of studies measuring and reporting hearing aid usage in older adults since 1999: A descriptive summary of measurement tools. *PLoS One, 7*(3), e31831.
- Reed, N. S., Betz, J., Kendig, N., Korczak, M., & Lin, F. R.** (2017). Personal sound amplification products vs a conventional hearing

-
- aid for speech understanding in noise. *Journal of the American Medical Association*, 318(1), 89–90.
- Reed, N. S., Betz, J., Lin, F. R., & Mamo, S. K.** (2017). Pilot electroacoustic analyses of a sample of direct-to-consumer amplification products. *Otology & Neurotology*, 38(6), 804–808.
- Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462.
- Valente, M., Abrams, H., Benson, D., Chisolm, T., Citron, D., Hampton, D., . . . Sweetow, R.** (2006). *American Academy of Audiology guidelines for the audiologic management of adult hearing impairment*. Reston, VA: American Academy of Audiology.
- Ventry, I. M., & Weinstein, B. E.** (1982). The Hearing Handicap Inventory for the Elderly: A new tool. *Ear and Hearing*, 3, 128–134.
- Wong, L. L. N.** (2011). Evidence on self-fitting hearing aids. *Trends in Amplification*, 15, 215–225.