nature biomedical engineering

Article

https://doi.org/10.1038/s41551-022-00947-6

An off-the-shelf otoacoustic-emission probe for hearing screening via a smartphone

Received: 8 November 2021

Accepted: 2 September 2022

Published online: 31 October 2022

Check for updates

Justin Chan[1] ✉, Nada Ali[2], Ali Najafi[3], Anna Meehan[4], Lisa R. Mancl[5], Emily Gallagher[4], Randall Bly[2,4] ✉ and Shyamnath Gollakota[1] ✉

Otoacoustic emissions (OAEs) provide information about the function of the outer hair cells of the cochlea. In high-income countries, infants undergo OAE tests as part of the screening protocols for hearing. However, the cost of the necessary equipment hinders early screening for hearing in low- and middle-income countries, which disproportionately bear the brunt of disabling hearing loss. Here we report the design and clinical testing of a low-cost probe for OAEs. The device, which has a material cost of approximately US$10, uses an off-the-shelf microphone and off-the-shelf earphones connected to a smartphone through a headphone jack. It sends two pure tones through each of the headphone's earbuds and algorithmically detects the distortion-product OAEs generated by the cochlea and recorded via the microphone. In a clinical study involving 201 paediatric ears across three healthcare sites, the device detected hearing loss with 100% sensitivity and 88.9% specificity, comparable to the performance of a commercial device. Low-cost devices for OAE testing may aid the early detection of hearing loss in resource-constrained settings.

It is estimated that 5.3% of the world's population suffers from disabling hearing loss. Also, a disproportionate brunt of this problem falls on low- and middle-income countries (LMICs)[1]. Hearing loss can be especially harmful for neurodevelopment if untreated in early childhood. However, the impact of hearing loss may be mitigated when detected and treated early[2]. It is common practice for high-income countries to adopt guidelines for universal infant hearing screening using otoacoustic emission (OAE) or auditory brainstem response (ABR)[3] testing. In spite of this, the test equipment remains expensive and costs thousands of dollars, which contributes to limited hearing screening in LMICs. In these countries, access to hearing assessment and equipment often requires travel to an urban setting and long wait times[4,5].

In this Article, we present the design and clinical testing of a low-cost OAE probe made from off-the-shelf earphones and microphones, with a material cost of about US$10. OAEs are sounds generated when the outer hair cells move in a healthy cochlea and provide information about their function[6,7]. Unlike conventional audiometry tests[8–11], OAE testing does not require a behavioural response from patients. As a result, it is frequently used for hearing screening in infants as well as young children (before they can participate) and as part of a diagnostic audiologic test battery for differential diagnosis of hearing conditions[7,12,13].

The earphone-based design sends two pure tone signals using each of the earphone's earbuds. When stimulated by two frequencies, the cochlea generates distortion-product OAEs (DPOAEs) due to intermodulation[6]. These emissions occur at frequencies not present in the input stimuli, which we measure using a microphone located at the probe head. Using algorithms run on a smartphone connected to the earphones, our system detects OAEs at various frequencies. We designed real-time algorithms that run on the smartphone to perform calibration, noise detection and automatic pass or refer testing for hearing screening and tested our design in a clinical study with a cohort of paediatric patients. Given the inexpensive cost of the earphones used

[1]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA. [2]Department of Otolaryngology — Head and Neck Surgery, University of Washington, Seattle, WA, USA. [3]Department of Electrical and Computer Engineering, University of Washington, Seattle, WA, USA. [4]Seattle Children's Hospital and Research Institute, Seattle, WA, USA. [5]Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA. ✉e-mail: jucha@cs.washington.edu; randall.bly@seattlechildrens.org; gshyam@cs.washington.edu

https://doi.org/10.1038/s41551-022-00947-6

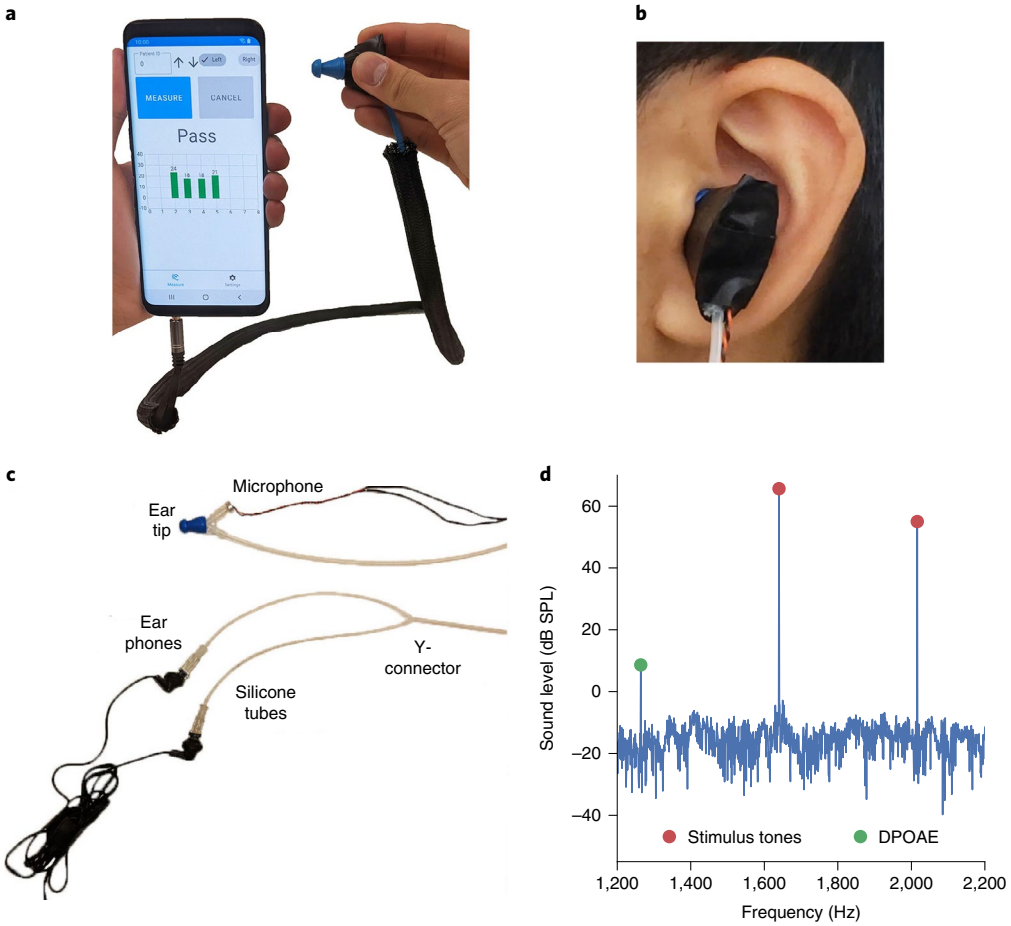

Fig. 1 | Overview of the earphone-based OAE probe system. a–**c**, A pair of earphones send pure tone sound stimuli through silicone tubes into the ear. A microphone positioned directly by the ear tip measures the DPOAEs emitted from the cochlea. The attachment connects to the phone via a 3.5 mm headphone jack (see Supplementary Fig. 7 for more details). **a**, The assembled system includes a nylon sleeve to protect the tubing from wear and tear and a black plastic casing to shield the microphone from damage. **b**, The probe head is lightweight and can rest in an ear without being held in place. **c**, The key components of the system without the sleeve and casing. **d**, The earphones send two stimulus tones, f_1 and f_2, through each of the earphone earbuds. At the same time, the recording from the microphone is averaged over time by the smartphone to reduce noise and detect a DPOAE signal.

and the ubiquity of smartphones, our earphone-based design could be used to increase early access to hearing screening across the world.

Results

Concept and prototype

Given the importance of OAE in infant hearing screening, there has been recent interest in designing better OAE hardware. Recent reports have proposed using a single transducer hardware for both stimulus transmission and recording[14]. High-end personalized headphone products such as Nuraphone (US$350) claim to measure OAEs to determine the listener's sensitivity to different acoustic frequencies to customize the audio signal sent to the listener[15]. Previous work[16,17] created a smartphone interface for the probes from an existing commercial OAE device. Recent devices[18] have used bone conduction to stimulate OAEs through a headband consisting of bone transducers. In addition to not using commodity earphones, none of these previous efforts present data from clinical studies for patients with hearing loss. Here we provide a demonstration that re-purposes earphones to create a low-cost OAE probe. We also provide clinical testing of our earphone-based OAE probe with real-time algorithms running on an attached smartphone to detect DPOAEs.

Our design (Fig. 1a–c) consists of a pair of off-the-shelf earphones in which each presents a different tone, f_1 and f_2. These tones cause the cochlea to generate OAEs at $2f_1 - f_2$, which we capture using a microphone. Thus, the primary components for our OAE probe are the two earphone speakers and a microphone. These components have to be integrated such that the probe head is sufficiently lightweight and can rest snugly in a patient's ear without being held in place. Commercial OAE probeheads typically use smaller custom speakers and are able to fit the speaker and microphone elements within the probehead. As we use commodity earphones as speakers, they are larger in size and would be heavy if both the earbuds are placed close to the probehead.

Instead, we place the earphones towards the end of the probe cable, closer to the smartphone, and connect to the probehead via lightweight silicone tubing. This tubing should be long enough so that it can comfortably cover the separation between the smartphone and a subject's ear during a measurement. At the same time, the tubing should not be so long that the sound waves at the probehead are attenuated below the intended sound levels of 65 and 55 dB sound pressure level (SPL) for the two tones.

The earphone speakers are coupled to a pair of 68 cm silicone tubes that are merged with a Y-connector into a single 19 cm silicone tube that connects to the probehead. The tubes merge into a single tube close to the probe head, minimizing the weight at the probe head. The probe head consists of a microphone and a three-dimensional

https://doi.org/10.1038/s41551-022-00947-6

Table 1 | Demographics of the patients in the clinical study

Hearing loss	***n* (%)**
Yes	66 (33)
Sensorineural	38 (19)
Conductive	13 (6)
Mixed	9 (4)
Not known	6 (3)
No	135 (67)
Previous hearing test	***n* (%)**
Behavioural audiometric testing	98 (49)
Auditory brainstem response	14 (7)
Newborn hearing screen	81 (40)
School-based hearing screen	14 (7)
Age (years)	6±6
Sex	***n* (%)**
Male	114 (57)
Female	82 (41)
Not recorded	5 (2)

(3D)-printed enclosure. The probe head is compatible with rubber ear tips that are used in commercial OAE devices[19]. The earphone and microphone are both connected to the smartphone with a 3.5 mm audio jack. The total length of the silicone tubes and the nylon sleeve protecting the earphones is 117 cm.

Figure 1d shows the frequency spectrum in a healthy ear with OAEs captured using our probe. When these emissions exceed a predefined signal-to-noise ratio (SNR) threshold, and optionally an absolute sound-level threshold, we mark the emissions as being present. In our clinical study, we emitted f_2 tones in the 2–5 kHz band, with the ratio f_1/f_2 set to 1.21–1.23. These frequency bands and ratios are commonly used bands for hearing screening[20] and are often available on most commercial screening OAE devices[21,22]. We mark the test as a 'pass' if emissions were detected at three or more of the four test bands and a 'refer' otherwise. As varying ambient and physiologic noise levels can overwhelm the OAEs, our algorithm rejects measurement windows where the noise exceeds a predefined threshold. Additionally, if the noise levels across three or more frequency bands exceed predefined levels, an error is displayed to the user. We designed a probe hardware integrity test to ensure the system did not produce unintended nonlinear acoustic distortions by inserting the probe head into a 2 ml test cavity. During clinical testing, we also perform a real ear test in a healthy ear to confirm that the OAEs can be successfully detected (see Methods for details).

Clinical testing

We conducted a clinical study at Seattle Children's Hospital at the Sandpoint and Bellevue clinics as well as the Center on Human Development and Disability at the University of Washington on a cohort of patients in otolaryngology and craniofacial clinics across three different sites. We tested our devices on 201 ears with patients between 1 week and 20 years of age with a mean age of 6 ± 6 years and a female-to-male ratio of 0.72 (Table 1). Five trained research assistants, including an undergraduate, a resident, a research coordinator, a public health student and a graduate student, performed all testing in a quiet clinic room with the patient awake and sitting upright. The exception was for infants who were tested in a variety of positions depending on what was most convenient for their parents, and we included both awake and asleep infants.

DPOAEs were first measured using a commercial OAE device, followed with our smartphone system using the same probe tip between the two devices. Both devices were calibrated to emit the two tones at 65 and 55 dB SPL respectively, and measurements were obtained for the 2, 3, 4 and 5 kHz bands. The clinical testing was performed using a Samsung Galaxy S9. In software, we performed an in-ear calibration procedure (Methods) that automatically adjusted the sound levels of the stimulus tones based on their recorded sound levels in the ear for 76 of the 201 measured ears (Supplementary Table 1). Otoscopy was performed before each measurement by an otolaryngologist. The hearing status of each patient was assigned after each test. For each tested patient ear, the best available data were interpreted by an otolaryngologist, which included clinical and examination history review, behavioural audiometric testing, newborn hearing screen result and diagnostic ABR. Of the 201 tested ears, 98 had an accompanying behavioural audiogram, and 14 underwent diagnostic ABR. Ears without diagnostic audiometric testing were assigned based on data from clinical history as well as school-based hearing screens ($n = 14$ ears) and newborn screens ($n = 81$ ears). Six ears were assigned hearing status based on clinical assessment, meaning they had no screening or diagnostic hearing tests but had no subjective hearing concerns and no concerns from the otolaryngologist attending who saw and examined the patient. Of the 201 tested ears, 135 had normal hearing, 38 had sensorineural hearing loss (SNHL), 13 had conductive hearing loss and 9 had mixed hearing loss.

We first considered the hearing screening to be a pass if OAEs exceeding a predefined SNR threshold were detected at three or more of the four frequency bands. Figure 2a shows the receiver-operating curve (ROC) for both our earphone-based probe and commercial OAE device by sweeping the SNR threshold from −20 to 40 dB, yielding an area under the curve (AUC) of 0.955 and 0.969, respectively. The best

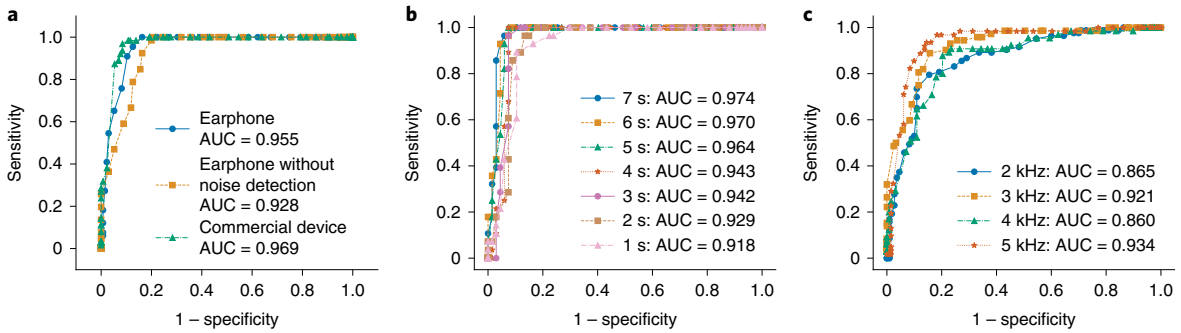

Fig. 2 | Performance of the clinical study. a, ROC curve showing performance of the earphone and commercial DPOAE device for screening performance of hearing loss at different SNR cut-off values. **b**, ROC curves for different signal averaging durations. **c**, ROC curve when comparing the pass and refer performances for individual frequencies to the commercial device for different SNR cut-off values.

https://doi.org/10.1038/s41551-022-00947-6

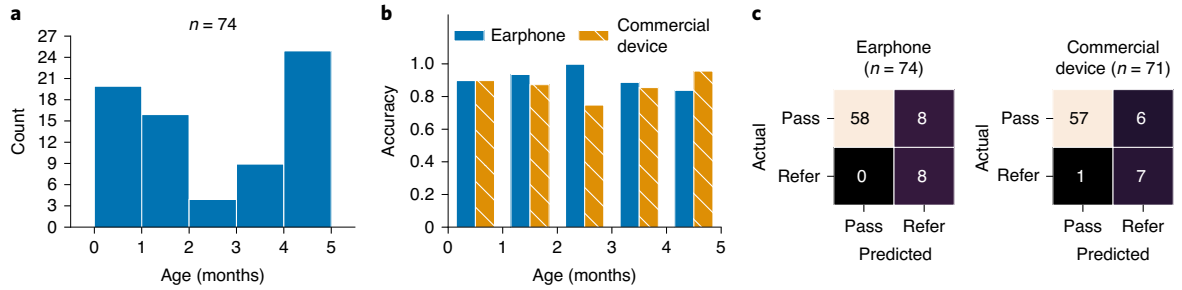

Fig. 3 | Device performance in infant ears under 6 months of age. a, Histogram of ears of patients less than 6 months of age. **b**, Accuracy obtained by earphone and commercial device for different age groups. **c**, Confusion matrix showing performance of the earphone and commercial DPOAE device for screening performance of hearing loss.

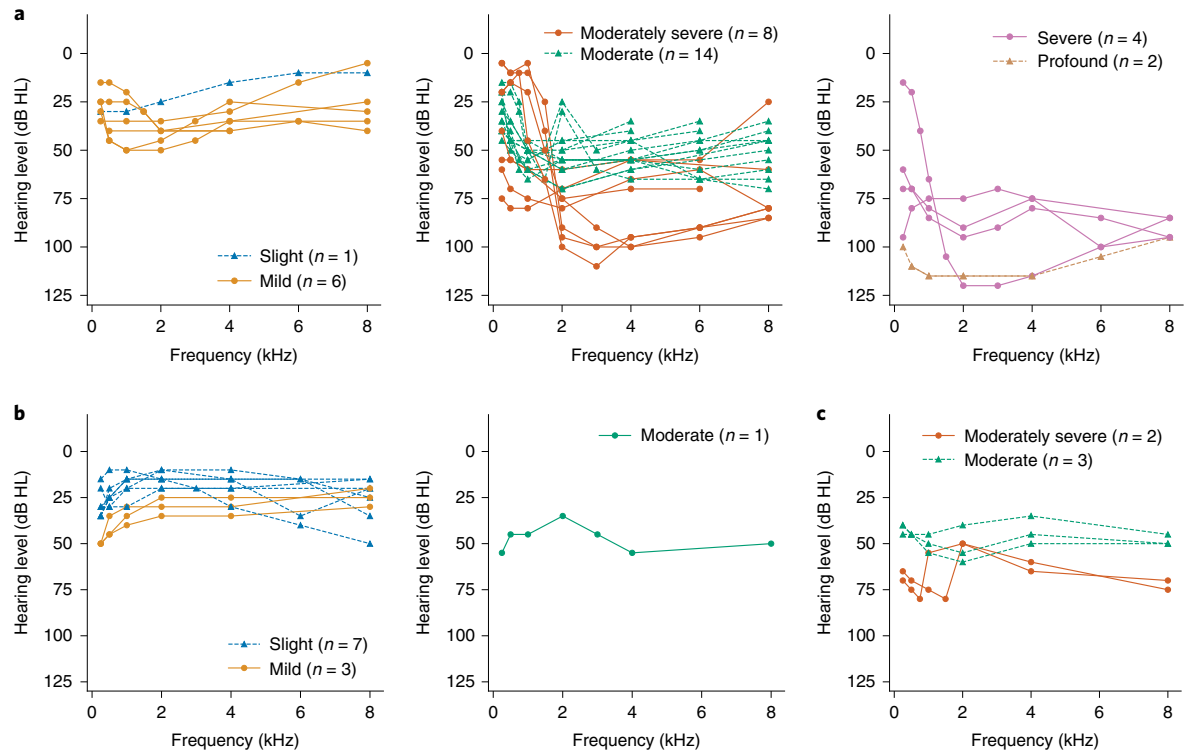

Fig. 4 | Audiograms of patient ears. a–**c**, Ears with sensorineural (**a**), conductive (**b**) and mixed (**c**) hearing loss categorized by degree of hearing loss. Each line refers to audiogram information from a single patient ear. Line colours refer to degree of hearing loss for an ear. Audiograms are plotted in units of dB hearing level (HL).

operating point for our earphone-based probe was obtained using a measurement time of 6 s per frequency band and an SNR threshold of 7 dB for patient ears over 6 months of age, and a measurement time of 5 s per frequency band and an SNR threshold of 5 dB for patient ears at or less than 6 months of age, yielding a sensitivity of 100.0% (95% confidence interval (CI), 94.7–100.0%) and a specificity of 88.9% (95% CI, 82.5–93.1%) (Supplementary Table 2). Of the 15 ears misclassified by the earphone, 5 of them had a history of middle ear disorders (Eustachian tube dysfunction, cholesteatoma), 3 of them had prior tympanoplasty and 3 of them had middle ear fluid. Disabling our noise detection algorithm, which rejects measurement windows with high noise, the best SNR threshold was 7 dB and yielded a sensitivity of 100.0% (95% CI, 94.5–100.0%) and a specificity of 77.8% (95% CI, 70.1–84.0%). In comparison, the commercial device had an operating point with a sensitivity of 96.8% (95% CI, 89.1–99.1%) and a specificity of 91.5% (95% CI, 85.5–95.2%) using a measurement time of 6 s per frequency band and the default manufacturer-prescribed SNR threshold of 6 dB.

To reduce the likelihood of classifying noise artefacts as OAEs, we added another threshold requiring the OAE levels to be at or above −10 dB SPL. After applying this criterion to the above SNR threshold, the operating point for our device had a sensitivity of 100.0% (95% CI, 94.5–100.0%) and a specificity of 88.1% (95% CI, 81.6–92.6%). When disabling the noise detection heuristic, the specificity of our device reduces to 76.3% (95% CI, 65.5–82.7%). In comparison, the commercial device had a sensitivity of 98.4% (95% CI, 91.5–99.7%) and a specificity of 89.2% (95% CI, 82.7–93.5%) with this additional threshold.

Figure 2b shows the device performance as the measurement duration changes from 1 to 7 s per frequency for the 95 ears where the maximum measurement time of 7 s per frequency was used. As expected, longer measurement durations slightly improve the AUC. Figure 2c also shows the agreement of our earphone-based probe with the commercial device in identifying the presence of OAEs at each of the four frequency bands for SNR threshold values in the range of −20 to 40 dB. The AUC values for each of the curves range from 0.860 to 0.934.

https://doi.org/10.1038/s41551-022-00947-6

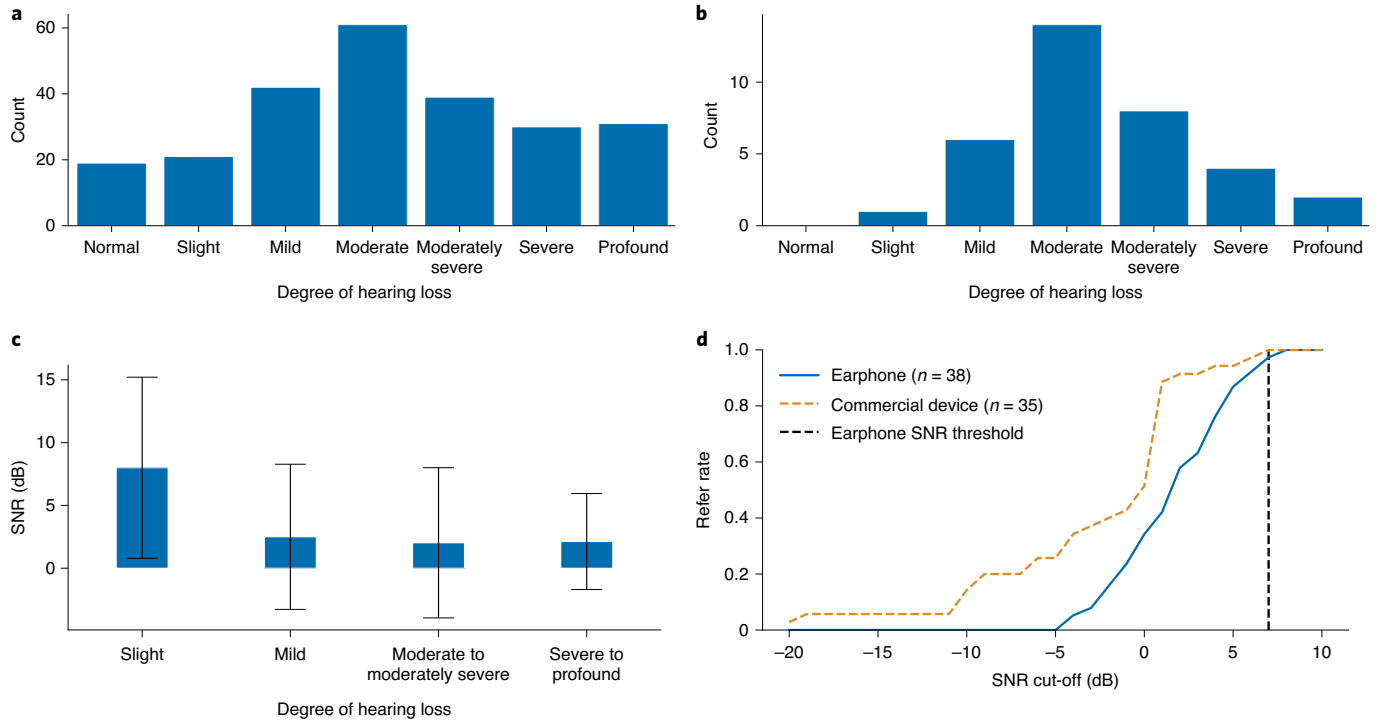

Fig. 5 | Device performance in patient ears with SNHL. a,**b**, Histogram showing degree of hearing loss for ears with SNHL by audiogram frequency band ($n = 243$) (**a**) and by ear ($n = 35$) (**b**). **c**, Comparison of SNR levels obtained on the earphone device categorized by degree of hearing loss. Data are presented as mean of SNR levels measured for ears with different degrees of hearing loss. $n = 4, 24, 88$ and 24 values were used to derive the mean SNR level for the slight, mild, moderate to moderately severe and severe to profound hearing loss conditions. Error bars denote s.d. from the mean. **d**, Refer rate of earphone and commercial DPOAE device for different SNR cut-off values.

Before each measurement, we sent a 226 Hz tone into the ear and recorded the sound level at the microphone that records the acoustic signals including the reflections from the eardrum. We selected 226 Hz as it was most responsive to changes in probe position and can be used to infer whether the probe is placed securely in the ear canal (Methods and Supplementary Fig. 1). We collected these data for 74 ears, of which 67 had intact eardrums and 7 ears had either perforation of the tympanic membrane or a patent ear tube. Six of these seven ears were correctly classified by our earphone device. The mean sound levels of the tone recorded at the microphone were 59 ± 2 and 54 ± 4 dB SPL for the ears with and without intact eardrums (Supplementary Fig. 2).

Finally, we report how long it took to couple our probe head to the ear. To compute this, we used the start time as just before the clinician placed the probe into the ear and the end time as when the clinician positioned the probe head into the ear and was satisfied with the fit. Of the 41 ears where this measurement was performed, the mean time was 10 ± 4.0 s (Supplementary Fig. 3). This is comparable to previous work that described that the coupling time for commercial screening OAE devices is slightly under 10 s (ref. [21]).

Clinical performance in infant ears

We perform a subgroup analysis on infant patients under the age of 6 months. In our clinical study, we recruited a total of 74 infant ears, 20 of which were from newborns less than 1 month (Fig. 3a). For 71 of these 74 ears, we tested using both the earphone and commercial device. The age of patients ranged from 1 week to 5 months of age, with a mean age of 3 ± 1 months. The hearing status of each ear was assigned based on data from either a newborn hearing screen or an ABR test. Of the 74 tested ears, 66 had normal hearing, whereas 8 ears failed a newborn hearing screen or ABR test; 2 of these ears had conductive hearing loss based on an ABR test.

In this population, we used ear tips with a diameter of 3 mm to accommodate the smaller ear canal sizes of the infant subjects (Supplementary Fig. 4). Aside from this, no other modifications were made to the hardware of the earphone device. We reduced the maximum measurement time from 7 to 5 s per frequency as the infant population was less likely to tolerate long measurements compared with older subjects. The SNR cut-off was also reduced to 5 dB for the earphone device to compensate for the increased noise and movement in this population.

Figure 3b shows the screening accuracies obtained by the earphone and commercial device for different age groups. Across these age groups, the earphone obtained accuracies ranging from 84% to 100%, while the commercial device obtained accuracies ranging from 75% to 96%. Figure 3c shows the confusion matrix of device performance for the earphone and commercial DPOAE device in the infant population. Of the 74 infant ears, the earphone correctly classified 58 of the 66 ears with normal hearing and all 8 ears with hearing loss. Of the 71 infant ears tested on the commercial device, it correctly classified 57 of the 63 ears with normal hearing and 7 of the 8 ears with hearing loss. Of the 36 infant ears that were 0 to 2 months of age, the earphone correctly classified 28 of the 31 ears with normal hearing. The commercial device correctly classified 27 of the 31 ears with normal hearing. Both devices correctly classified five of the ears with hearing loss.

Clinical performance on SNHL ears

Figure 4a–c shows the audiograms of patient ears with hearing loss, broken down by degree of hearing loss for patient ears with sensorineural, conductive and mixed hearing loss. The degree of hearing loss for an ear is computed based on the average hearing levels measured across the audiogram, which are then mapped to hearing thresholds (Supplementary Table 3). We perform a subgroup analysis on patient

https://doi.org/10.1038/s41551-022-00947-6

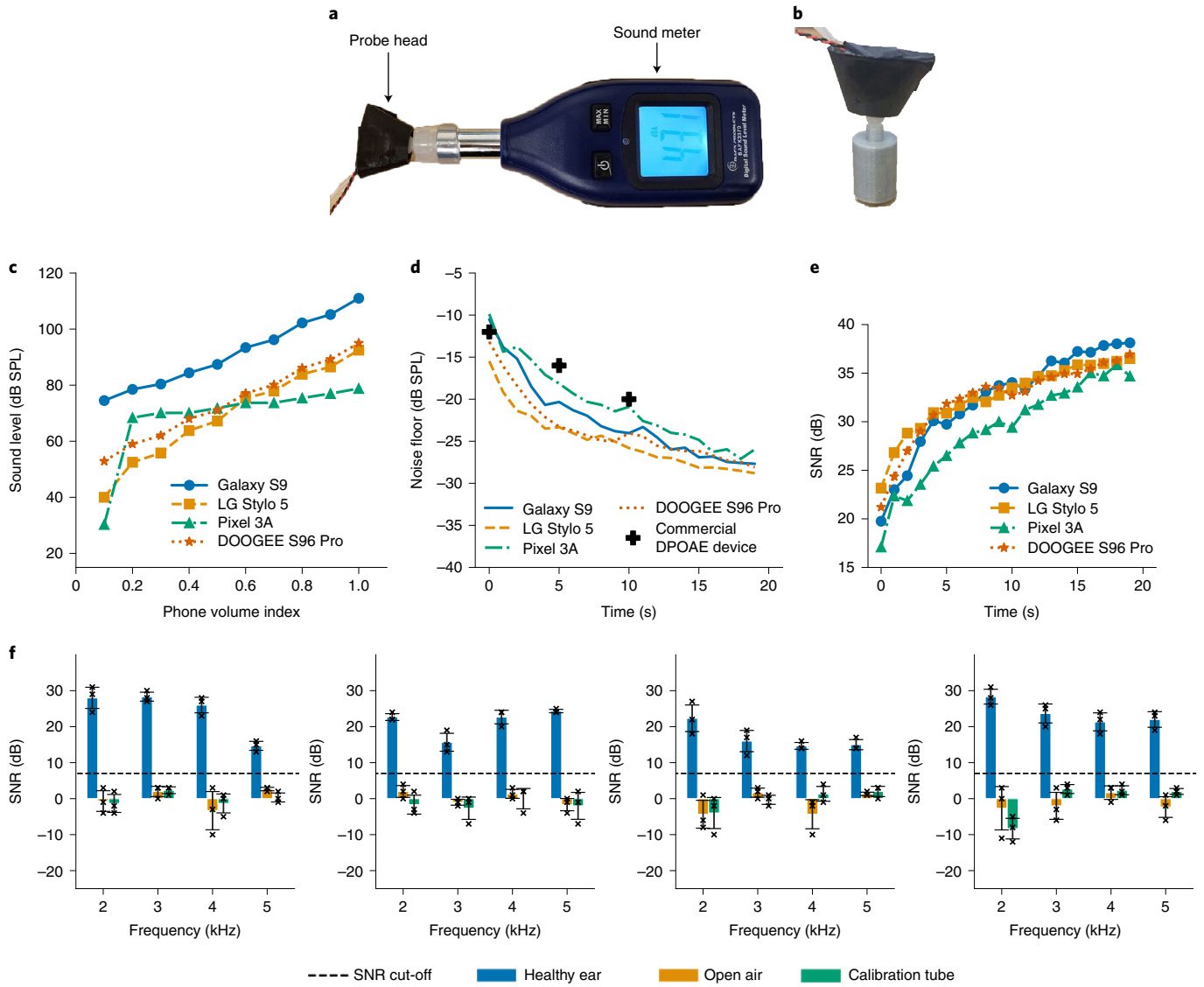

Fig. 6 | Benchmark testing across multiple smartphones. a, The probe head of the device is coupled to a sound level meter that outputs the sound levels emitted by the device in absolute physical units. This set-up is used to calibrate the sound levels of f_1 and f_2 to 65 and 55 dB SPL. **b**, The probe head is coupled to a 3D-printed 2 ml plastic calibration tube to check for system distortions during a DPOAE measurement. **c**, The volume index on smartphones and the absolute sound levels at 1 kHz. **d**, The noise floor for each smartphone decreases over time as additional signal averaging is performed. **e**, The SNR of DPOAEs increases over time for different smartphones. **f**, The SNR at the DPOAE frequency when a measurement is performed in a healthy ear, in open air and in a 2 ml calibration tube for different smartphones specifically (left to right) Samsung Galaxy S9, LG Stylo 5, Pixel 3a XL and DOOGEE S96 Pro. Symbols, mean of three technical measurement replicates. Error bars denote s.d. from the mean.

ears with data on different degrees of SNHL from slight to profound. We collected data from a total of 38 patient ears, of which 35 had an accompanying behavioural audiogram. For 35 of these 38 ears, we tested using both the earphone and commercial device. The age of patients ranged from 10 months to 16 years, with a mean age of 8 ± 4 years. Hearing levels for each frequency band were classified into different degrees of hearing loss based on thresholds (Supplementary Table 3 and Fig. 5a). Figure 5b shows the degree of hearing loss based on the average hearing levels measured across the audiogram.

We compare the SNR obtained at different frequencies by the earphone device for ears with different degrees of hearing loss (Fig. 5c). We observe that the mean DPOAE SNR measured by the earphone is 8 dB for frequencies with slight hearing loss, which is above the SNR cut-off of 7 dB used to mark a DPOAE as present. The mean DPOAE SNR measured by the earphone decreases to 3, 2 and 2 dB for mild, moderate to moderate-severe, and severe to profound hearing loss, respectively. Figure 5d shows the refer rate obtained by both devices for different SNR cut-off values. The figure shows that at the predefined SNR threshold of 7 dB, the earphone had a refer rate of 97% (37 out of 38 ears). Using the predefined threshold of 6 dB for the commercial DPOAE device obtained a refer rate of 97% (34 out of 35 ears). The one ear that was misclassified by both devices was classified by the audiologist as having mild to moderate hearing loss.

Benchmark testing

We evaluated our OAE probe with four different smartphones released between 2018 and 2020 (Supplementary Table 4). We calibrated the output sound level to play two tones at 65 and 55 dB SPL. We performed this calibration by coupling the probe head of our device with a US$18 reference sound level meter (Fig. 6a). Figure 6c shows that there is largely a linear relationship between the output audio gain index of the smartphone and the sound level in absolute units of decibel SPL

https://doi.org/10.1038/s41551-022-00947-6

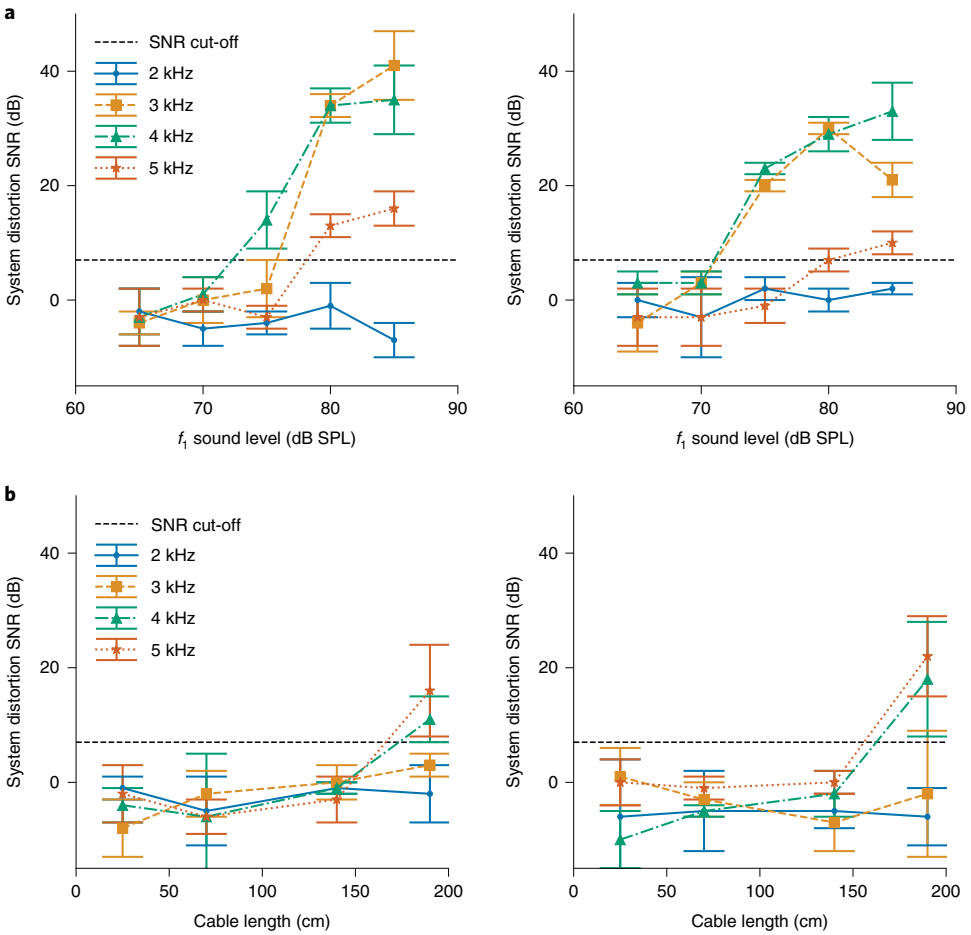

Fig. 7 | Benchmark testing of hardware distortion. a,**b**, The amount of distortion generated by the hardware at the DPOAE frequencies for open air (left) and calibration tube (right) is evaluated for different stimulus sound levels (**a**) and for different cable lengths (**b**). Symbols, mean of three technical measurement replicates. Error bars denote s.d. from the mean.

on the sound meter. To calibrate the sound levels, we find the smallest volume index on the phone that would produce a sound level above 65 dB SPL for a given f_1. We then digitally scale down the amplitude of the f_1 waveform on the phone so that the sound level is 65 dB SPL. The sound level of f_2 is similarly calibrated to 55 dB SPL. In our clinical study, this calibration was performed once a week.

Figure 6d shows the noise floor of our OAE probe and the commercial device in a healthy ear as a function of the averaging duration. We performed a DPOAE measurement at the 2 kHz band for 20 s. The measurement was performed three times, and the mean across the measurements is plotted. The noise floor of the commercial device decreases from −11 dB to the minimum reported level of −20 dB SPL after 11 s. In comparison, the noise floor of the smartphones ranged from −10 to −15 dB SPL after 1 s of measurement and reached −20 dB SPL after 6 s for three of the phones and after 11 s for the remaining phone. Figure 6e shows the SNR of the recorded OAEs on each of the phones as a function of the averaging time.

We perform hardware integrity testing that measures distortions across different smartphones using measurements both in open air and in a 2 ml calibration tube (Fig. 6b). This testing is performed to ensure that the system nonlinearities in the $2f_1 - f_2$ frequency would not appear as a false OAE.

A 4 s measurement is performed three times for each experimental condition. Figure 6f shows the SNR in open air, a calibration tube and a healthy ear for different smartphones. The SNR is low across the tested smartphones in both open air and calibration tube testing.

Figure 7a shows the effect of sound level on system distortions on the Samsung Galaxy S9. In this test, we initially set f_1 and f_2 to 65 and 55 dB SPL and increase both volumes in steps of 5 dB SPL. A measurement is then performed in open air and in a 2 ml calibration tube. We find that system distortions are lower than the SNR cut-off for the 65 and 55 dB SPL and the 70 and 60 dB SPL scenarios. System distortions increase for the 3 and 4 kHz band at 75 and 65 dB SPL, while the distortions in the 2 kHz band remain below the threshold across all measured sound levels. Supplementary Fig. 5 shows an example of system distortions in the 4 kHz band when a sound level of 85 or 75 dB SPL is used in a 2 ml plastic calibration tube and open air. Supplementary Note 1 provides further characterization of system distortion as well as additional benchmark testing on the effect of background noise on device performance.

Finally, Fig. 7b shows the effect of cable length on system distortions. We increased the length of the single silicone tube from 25 to 190 cm and calibrate f_1 and f_2 to 65 and 55 dB SPL. The system distortions remain below the threshold for all frequencies up to a cable length of 140 cm.

Discussion

A key advantage of using off-the-shelf earphones and smartphones is that custom electronics do not need to be manufactured, which lowers development costs[23]. The assembly of our probe does not require specialized knowledge of electronics, and we estimate that the cost of labour for assembly would be less than a dollar at scale (Supplementary

Table 5). Assembly costs of different categories of electronic products typically do not constitute more than 6% of the total cost of device components and assembly[24]. We point out that budget or second-hand smartphones can be purchased in LMICs for US$35–50 due to their economies of scale, which is substantially lower than the upfront cost of commercial OAE devices. Additionally, our system has been built against the Android software development kit version 29, which has been designed to be compatible with future versions of Android operating systems[25]. Although some newer smartphones have eliminated the audio-jack interface, audio-jack adapters[26] that cost a few dollars can be used to accommodate the wired earphones interface in our design. Finally, the US Food and Drug Administration provides guidance for Mobile Medical Applications[27,28] and Software as a Medical Device[29–31] that regulates the custom software application running on the phone and not the smartphone hardware, thus potentially reducing the associated regulatory costs[23]. Commercially available medical devices that include Mobile Medical Applications and that use sensors such as microphones and cameras have been cleared or approved by the US Food and Drug Administration[32].

The cost of OAE devices is only one factor associated with addressing the complex public-health problem of hearing screening in LMICs. There are other factors involved with real-world deployments, including establishing strong multilateral partnerships, support for follow-ups and the cost of regulatory clearance. Financial support is required to fund the manufacturing cost of the OAE probes, as well as support for the screening staff and other healthcare workers who would be administering the test. We envision that initial deployments in the field would be funded by non-governmental organizations and health-insurance funds. Long-term salary support would eventually require partnership and funding from local health ministries.

Towards the goal of adoption, we have open-sourced our hardware and software code to allow anyone to download and recreate the smartphone device. A solution to the complex global health problem that is the diagnosis and management of newborn hearing loss necessitates strong multilateral partnerships. To that end, we direct the readers to our initiative on Toward Universal Newborn and Early Childhood Hearing Screening in Kenya, which has the explicit goal of developing and implementing a hearing-related continuum of care in Kenya inclusive of but not limited to newborn hearing screening (https://tune.cs.washington.edu/). We envision that our earphone-based OAE probe can potentially be combined with existing frugal techniques to detect middle ear fluid on a smartphone[33] and assess eardrum mobility with smartphone tympanometry[34] as part of an audiology toolkit for evaluating middle and inner ear health on smart devices. Over the long term, local non-specialized healthcare workers such as technicians and volunteers need to be trained to perform the OAE test. Furthermore, the results from our device may need to be incorporated into the local medical record system. Finally, a continuum of care will need to be developed for individuals who screen for potentially having hearing loss.

In our study, we tested an in-ear calibration procedure that altered the stimulus levels based on the sound levels recorded by the probe microphone at the entrance of the ear canal. We note that previous work[35,36] has shown that standing waves in the ear canal can cause sound levels at the ear entrance to have a difference of up to 20 dB compared with the sound level at the eardrum, which could result in calibrated sound levels that are higher than intended. Although more accurate in-ear calibration procedures[37], such as Thevenin-equivalent sound calibration, exist, they require carefully engineered probe tips and tend to be complicated. To ensure ease of use during calibration and appropriate stimulus levels during testing, future deployments of our system can be calibrated against different cavities to represent neonate, paediatric and adult ear canals.

Our study has the following limitations: commercial OAE devices are used in practice by nurses, technicians and volunteers[38,39]. Although we used the same ear tips and displayed similar information on a smartphone, subsequent studies are required to determine the reliability and ease of use of our OAE probe and smartphone system by nurses, technicians and volunteers. In our study, although the device has been evaluated by several trained, non-professional and non-clinical testers, it has only been evaluated in controlled clinical environments. Field testing by non-professionals is required to evaluate the long-term durability of our probe design. Such an evaluation is needed to determine whether the probe is resistant against wear and tear in challenging environments over time. Our earphone-based probe is designed to only measure DPOAEs at this time; further work is needed to also measure transient-evoked OAEs (TEOAEs). TEOAEs typically use 24-bit microphones with stimulus levels of 30–90 dB peak equivalent SPL (refs. [40–43]). We note that the analog-to-digital converter on iOS devices do support recording at bit depths of 24 and 32 bits[44] and Android devices do support the use of external analog-to-digital converters to provide 24-bit resolution via their Universal Serial Bus Digital Audio interface[45]. However, more work is required to investigate the feasibility of measuring TEOAEs with our system. Beyond hearing screening, OAEs are also used in conjunction with other tests for the differential diagnosis of hearing conditions[7] and ototoxicity monitoring[46]. In these scenarios, OAEs are often also tested in the 800 Hz to 10 kHz range[7,46]. Given that higher frequencies undergo more viscothermal losses within small-diameter tubings[47], a further investigation into alternative probe designs would probably be required to perform higher-frequency tests. Additionally, clinical testing is required to evaluate the performance of our device at this wider range of frequencies. However, we note that our current design is sufficient to screen for hearing loss, which is a more common test in clinical practice and which is the focus of our system.

In summary, we presented a low-cost OAE system using off-the-shelf earphones. While hearing loss is one of the more common disorders in LMICs, early detection can be challenging owing to lack of access to affordable hearing-screening tools. Compared with commercial OAE devices that cost thousands of dollars, our frugal earphone-based OAE probe has the potential to increase access to hearing screening in resource-constrained environments. Further community deployments are required to determine the technology's impact in these and other potential scenarios.

Methods

Our study was approved by the Seattle Children's Hospital and University of Washington institutional review boards. All studies complied with relevant ethical regulations. Patients were recruited during routine appointments in otolaryngology and craniofacial clinics. Parental permission was obtained for participants under the age of 18 years. Children age 7 to 17 years provided written assent. Assent was obtained after parental permission was granted. Children age 7 to 12 years signed a simple assent form, and children age 13 to 17 years signed a consent form. Parents co-signed the consent form. Participants 18 years and older signed a consent form. We excluded patient ears where the patient was unable to complete the testing. Investigators were not blinded. In the study, the commercial OAE device used for testing was the AudX Pro (Bio-logic, 2006).

Hardware design

We use a pair of inexpensive commodity stereo earphones (Panasonic ErgoFit EP-HJE120, US$4.49) and a small, high-sensitivity electret condenser microphone (PUI Audio POM-2730L-HD-R, US$1.75). The microphone has a 6 mm diameter and a sensitivity of −30 ± 3 dBV. The design of our probe head consists of a small 3D-printed enclosure and a Y-connector (3.175 mm × 3.175 mm × 3.175 mm) that houses the microphone and a silicone tube through which the sound is transmitted.

With this configuration, the microphone is designed to be close to the ear canal. The Y-connector is terminated with a rubber ear tip (Grason & Associates LLC) that can be inserted into an ear canal. The

probe head has a total weight of 3 g. The microphone signal and ground cable, and the earphones are connected to the smartphone via audio adapters (four-contact male audio jack to screw terminal connector, three-contact male audio jack to female four-contact audio jack connector and four-contact male audio jack splitter) that terminate in a 3.5-mm audio jack.

Smartphone user interface

A custom Android app was created to evoke and measure DPOAEs with our earphone hardware attachment. The software allows to select which frequencies to send tones at, as well as the duration of the measurement. The app performs the measurement, signal averaging and noise detection algorithms in real time on the phone. The user interface is shown in Supplementary Fig. 6.

Smartphone computation

We describe the various algorithms that run on the smartphone to measure the SNR of OAEs, detect whether the probe is in the ear and detect noisy measurements.

Computing SNR of OAEs. Our device measures DPOAEs at 2, 3, 4 and 5 kHz bands and uses a sampling rate of 48 kHz. We perform signal averaging to decrease the noise floor of the measurement and increase the SNR of potential OAEs. To do this, the signal over T seconds is split into windows of 1 s and averaged across the T windows in the time domain. The signal is then transformed to the frequency domain by taking the absolute value of an fast Fourier transform of the signal with a window size of 48,000. The fast Fourier transform output is squared to obtain the energy spectral density of the signal $S(f)$. The signal power at f_{DPOAE} is given by $S(f_{\mathrm{DPOAE}})$, and the noise power is defined as the mean of the power at frequencies around f_{DPOAE}, $S(\mathrm{noise}) = \frac{1}{2W}\left(\sum_{i=L}^{L+W} S(f_{\mathrm{DPOAE}} - i) + \sum_{i=L}^{L+W} S(f_{\mathrm{DPOAE}} + i)\right)$. Here W is a noise window size set to 200, L is set to 2 and i is the index of summation. The SNR is computed as $10\log_{10}\left(\frac{S(f_{\mathrm{DPOAE}})}{S(\mathrm{noise})}\right)$.

Algorithm to detect noisy measurements. OAE measurements are sensitive to noise. When noise is introduced into the measurement, it can corrupt the signal and overwhelm any OAE. We implement three algorithms for noise detection, two at the frequency level and one at the test level. In the first algorithm, we look for increases in signal artefact over time. Here, we define an artefact as an increase in the average power in decibels from 0 to 6,000 Hz band, which covers the frequency bands over which we conduct a measurement. Increases in signal artefact generally occur due to environmental noise or motion. For this algorithm, if the artefact in the current segments has risen by more than 4 dB compared with the previous one, we discard the current segment. In the second algorithm, we check that the sound level of the recording has reached a stable sound level. On the smartphone, there is a transient period at the start of measurements where the automatic gain control adjusts and stabilizes the sound level of transmissions. This adjustment period adds unwanted noise to the DPOAE measurements. To detect such a period, we look at the root mean square amplitude of the signal across the first 1,000 samples. If this amplitude value exceeds a predefined threshold of 11,700 (−8.9 dB relative to full scale), that segment is discarded from the signal-averaging calculation. In the last algorithm, we reject measurements where noise levels are high. Specifically, if the noise levels at three or more frequencies is high, the measurement is marked as too noisy, and an error message is displayed to the user.

Determining whether probe is in ear. Our system performs an in-ear check to ensure that the probe is coupled well to the ear canal. To do this, we continuously send 20 ms chirps from 100 to 5,500 Hz into the ear canal and measure the frequency response at the microphone. Supplementary Fig. 1a shows the frequency response of the chirp when the probe is well coupled to the ear, loosely placed in the ear and outside the ear. The figure shows that from 100 to 500 Hz, the sound level of the chirp is low when the probe is outside of the ear and high when the probe is well coupled to the ear. To determine whether a measurement is ready to begin, we measure the sound level at 226 Hz for each chirp and check that the sound level exceeds a threshold of 53 dB SPL and the standard deviation of the sound level across the last 50 chirps (1 s) does not exceed 1.5 dB SPL. This ensures that spurious increases in the sound level due to adjustment of the probe in the ear are not mistaken for a stable probe placement (Supplementary Fig. 1b). The sound level at 226 Hz is shown to the user before a measurement (Supplementary Fig. 6b,c) in the form of a progress bar. If the probe is outside of the ear or loosely placed in the ear canal, the progress bar is partially filled, and the start button is disabled. When the probe is placed securely in the ear canal and the sound level exceeds the predefined sound-level threshold, the progress bar is completely filled, and the start button to begin a measurement is enabled.

Sound-level calibration procedure

A calibration procedure needs to be performed on OAE devices regularly to emit tones f_1 and f_2 at the correct 65 and 55 dB SPL values. For this calibration procedure, we couple the probe head of our device to a reference sound level meter (BAFX 3370, Digital Sound Level Meter, US$18), which outputs the sound level emitted at the probe head in absolute physical units (Fig. 6a). The sound level of a tone can be modified in two ways. It can be altered by adjusting the smartphone's volume gain or by digitally scaling the amplitude of the output waveform. Figure 6c shows that there is a linear relationship between the smartphone's volume gain and the output sound level as recorded on the sound meter. We calibrate the sound levels as follows. Step 1: set the amplitude of the waveform to the maximum digital value. In our implementation each sample has an amplitude range of −32,768 to 32,767. Step 2: play the frequency from the phone and adjust the volume index to the minimum value that would cause the sound level to exceed 65 dB SPL. Step 3: digitally scale down the waveform until the sound levels of f_1 and f_2 reach 65 and 55 dB SPL. This procedure is then repeated for all f_1/f_2 pairs. In our clinical study, we perform this calibration once per week.

Automatic in-ear sound level calibration

In the clinical study on infant ears, we implemented an automatic in-ear calibration procedure to adjust the sound level of the stimuli to account for different ear canal volumes. Infant ears can have a smaller ear canal volume and require lower transmitted sound levels for the f_1 and f_2 sound levels at the microphone to be at 65 and 55 dB SPL. To do this, before a measurement, we automatically adjust the sound levels that need to be transmitted using the linear coefficients generated from the phone volume index to sound level as in Fig. 6c.

The app then sends a 100 ms 1–4 kHz chirp into the ear for synchronization. It then sends eight 200-ms-long tones at the eight f_1 and f_2 frequencies in sequence. The app then measures the sound level of each of the eight tones and calculates the volume index required so that the sound level of f_1 and f_2 are received at the microphone as 65 and 55 dB SPL, respectively.

Probe integrity check

After completion of the sound level calibration procedure, a hardware integrity check is performed to ensure that the device can detect the OAEs and does not generate system distortions that would be mistaken as OAEs (Supplementary Fig. 5). To do this, we first perform a real ear OAE measurement in a healthy ear (for example, the clinician) three times and check that emissions are above the SNR cut-off for the 2, 3, 4 and 5 kHz frequency bands. Next, we perform an OAE measurement in open air and in a 2 ml cavity (Fig. 6b) to check that no OAEs are detected at the desired frequencies. If the SNR at any of the frequency bands

exceeds the cut-off value, the probe should be visually inspected for any debris or wax. In the event that wax has accumulated on the probe head, it should be cleaned thoroughly and carefully with alcohol wipes to avoid pushing the wax into the nozzle of the probe head. If the probe integrity check does not pass even after cleaning, the sound levels of the device should be re-calibrated.

Statistical analysis

Algorithms to perform the OAE measurement are performed on the Android platform. ROC, AUC, sensitivity, specificity, and 95% CI analysis was performed using numpy. Figures were created using matplotlib and seaborn.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data supporting the findings from this study are available within the article and its Supplementary Information. The dataset used to generate the results for this study is available at https://github.com/uw-x/oae and https://zenodo.org/record/7032657 (ref. [48]). Source data are provided with this paper.

Code availability

The custom code used in this study is available at https://github.com/uw-x/oae and https://zenodo.org/record/7032657 (ref. [48]).

References

1. *Millions of People in the World Have Hearing Loss That Can Be Treated or Prevented* (WHO, 2013); https://www.who.int/news/item/27-02-2013-millions-have-hearing-loss-that-can-be-improved-or-prevented
2. *Childhood Hearing Loss—Strategies for Prevention and Care* (WHO, 2016); https://apps.who.int/iris/bitstream/handle/10665/204632/9789241510325_eng.pdf
3. Patel, H. & Feldman, M. Universal newborn hearing screening. *Paediatr. Child Health* **16**, 301–305 (2011).
4. McPherson, B. Newborn hearing screening in developing countries: needs & new directions. *Indian J. Med. Res.* **135**, 152 (2012).
5. Swanepoel, D. W. et al. Telehealth in audiology: the need and potential to reach underserved communities. *Int. J. Audiol.* **49**, 195–202 (2010).
6. Chan, J. & Gollakota, S. Inner-ear cochlea testing with earphones. In *Proc. 20th Annual International Conference on Mobile Systems, Applications and Services* 609–610 (ACM, 2022).
7. Abdala, C. & Visser-Dumont, L. Distortion product otoacoustic emissions: a tool for hearing assessment and scientific study. *Volta. Rev.* **103**, 281 (2001).
8. Walker, J. J., Cleveland, L. M., Davis, J. L. & Seales, J. S. Audiometry screening and interpretation. *Am. Fam. Physician* **87**, 41–47 (2013).
9. *Mimi Hearing Test* (Mimi Hearing Technologies GmbH, 2022); https://apps.apple.com/us/app/mimi-hearing-test/id932496645
10. *Hearing Test—Audiometry, Tone* (IT4YOU Corp., 2022); https://apps.apple.com/us/app/hearing-test-audiometry-tone/id1368396053
11. *Hearing Test & Ear Age Test* (PortTownSoft, 2022); https://apps.apple.com/us/app/hearing-test-ear-age-test/id1067630100
12. Thompson, D. C. et al. Universal newborn hearing screening: summary of evidence. *JAMA* **286**, 2000–2010 (2001).
13. Torre, P. III, Cruickshanks, K. J., Nondahl, D. M. & Wiley, T. L. Distortion product otoacoustic emission response characteristics in older adults. *Ear Hear.* **24**, 20–29 (2003).
14. Heitmann, N., Rosner, T. & Chakraborty, S. Designing a single speaker-based ultra low-cost otoacoustic emission hearing screening probe. In *2020 IEEE Global Humanitarian Technology Conference (GHTC)* 1–8 (IEEE, 2020).
15. Campbell, L. J. & Slater, K. D. Personalization of auditory stimulus. US patent 10154333B2 (2018).
16. Heitmann, N., Rosner, T. & Chakraborty, S. Mass-deployable smartphone-based objective hearing screening with otoacoustic emissions. In *Proc. 2021 International Conference on Multimodal Interaction* 653–661 (ACM, 2021).
17. Heitmann, N. et al. Sound4all: towards affordable large-scale hearing screening. In *2017 12th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS)* 1–6 (IEEE, 2017).
18. Novel pediatric hearing screening device. *Johns Hopkins Technological Ventures* http://jhu.technologypublisher.com/technology/40355 (2018).
19. Generic eartips from Grason Associates fitment guide eartips. *Electro-Medical Instruments* https://www.emi-canada.com/new-page-5 (2021).
20. Newborn hearing screening protocol. *Iowa Department of Public Health* https://idph.iowa.gov/Portals/1/userfiles/77/Newborn%20Hearing%20Screening%20Protocol-11-08.pdf (2008).
21. Parthasarathy, T. K. & Klostermann, B. Similarities and differences in distortion-product otoacoustic emissions among four FDA-approved devices. *J. Am. Acad. Audiol.* **12**, 397–405 (2001).
22. Frequently asked questions: EroScan OAE Test System. *Maico* https://www.schoolhealth.com/media/pdf/51075_FAQ%20EroScan.pdf (2022).
23. Chan, J., Michaelsen, K., Estergreen, J. K., Sabath, D. E. & Gollakota, S. Micro-mechanical blood clot testing using smartphones. *Nat. Commun.* **13**, 1–12 (2022).
24. Funk, J. Technology change, economic feasibility, and creative destruction: the case of new electronic products and services. *Ind. Corp. Change* **27**, 65–82 (2018).
25. Device compatibility overview. *Android Developers* https://developer.android.com/guide/practices/compatibility (2022).
26. For iPhone headphone adapter 3.5 mm jack aux cord dongle audio cable connector US. *eBay* https://www.ebay.com/itm/313802818623 (2022).
27. Device software functions including mobile medical applications. *US FDA* https://www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications (2019).
28. Policy for device software functions and mobile medical applications. *US FDA* https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications (2019).
29. Software as a medical device (SaMD). *US FDA* https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd (2018).
30. *Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations* (2014); http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf
31. What are examples of software as a medical device? *US FDA* https://www.fda.gov/medical-devices/software-medical-device-samd/what-are-examples-software-medical-device (2017).
32. Examples of premarket submissions that include MMAs cleared or approved by the FDA. *US FDA* https://www.fda.gov/medical-devices/device-software-functions-including-mobile-medical-applications/examples-premarket-submissions-include-mmas-cleared-or-approved-fda (2019).

33. Chan, J., Raju, S., Nandakumar, R., Bly, R. & Gollakota, S. Detecting middle ear fluid using smartphones. *Sci. Transl. Med.* **11**, eaav1102 (2019).
34. Chan, J. et al. Performing tympanometry using smartphones. *Commun. Med.* **2**, 57 (2022).
35. Siegel, J. & Hirohata, E. Sound calibration and distortion product otoacoustic emissions at high frequencies. *Hear. Res.* **80**, 146–152 (1994).
36. Siegel, J. Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes. *J. Acoust. Soc. Am.* **95**, 2589–2597 (1994).
37. Charaziak, K. K. & Shera, C. A. Compensating for ear-canal acoustics when measuring otoacoustic emissions. *J. Acoust. Soc. Am.* **141**, 515–531 (2017).
38. Maico Diagnostics. *School Nurses Guide to Otoacoustic Emissions (OAEs)* (2009); https://www.schoolhealth.com/media/pdf/School%20Nurse%20Guide%20to%20OAEs_web%202016.pdf
39. Hall, J. W. III *The Important Role of Otoacoustic Emissions (OAEs) in Preschool Hearing Screening* (2018); https://www.audiologyonline.com/articles/important-role-otoacoustic-emissions-oaes-22091
40. Technical specifications—Lyra. *Interacoustics* https://www.interacoustics.com/download/lyra/3111-datasheet-for-lyra (2019).
41. OtoRead™ – instructions for use. *Interacoustics* https://www.interacoustics.com/download/otoread/manuals-otoread/1236-instructions-for-use-otoread (2021).
42. Welch Allyn® OAE hearing screener—directions for use. *Welch Allyn* https://www.henryschein.com/assets/Medical/80022349LITPDF.pdf (2019).
43. r140 Portable OAE hearing screener. *Resonance* http://www.resonance-audiology.com/wp-content/uploads/2018/05/R14O-TECNICAL-SPECIFICATION.pdf (2018).
44. AVLinearPCMBitDepthKey. *Apple Developer Documentation* https://developer.apple.com/documentation/avfaudio/avlinearpcmbitdepthkey (2022).
45. USB Digital Audio. *Android Open Source Project* https://source.android.com/devices/audio/usb (2022).
46. Reavis, K. M. et al. Distortion-product otoacoustic emission test performance for ototoxicity monitoring. *Ear Hear.* **32**, 61 (2011).
47. Goodman, S. S., Fitzpatrick, D. F., Ellison, J. C., Jesteadt, W. & Keefe, D. H. High-frequency click-evoked otoacoustic emissions and behavioral thresholds in humans. *J. Acoust. Soc. Am.* **125**, 1014–1032 (2009).
48. Chan, J. et al. A low-cost smartphone-based otoacoustic emission device for hearing loss. *Zenodo* https://doi.org/10.5281/zenodo.7032657 (2022).

Acknowledgements

We thank our participants and their families at Seattle Children's Hospital. We thank K. Sie, D. Horn and H. Ou at the Seattle Children's Hospital and Research Institute for permitting us to recruit patients from their clinics. We thank S. Norton for her critical and important feedback on the manuscript and methods. N.A. discloses support for the research described in this study from the National Institute on Deafness and Other Communication Disorders award T32DC000018. R.B. and S.G. disclose support for the research described in this study from the Washington Research Foundation, Seattle Children's Research Institute, Research Integration Hub. S.G. discloses support for the research described in this study from the Moore Inventor Fellow award #10617.

Author contributions

All authors designed the experiments and interpreted the results. J.C. and S.G. wrote the manuscript, and N.A., A.M., L.R.M., E.G. and R.B. edited the manuscript. J.C., N.A. and A.M. conducted the experiments and performed the analysis, under technical supervision by S.G. and J.C.; A.N. and S.G. designed the algorithms. J.C. and S.G. conceptualized the study.

Competing interests

S.G., J.C. and R.B. are co-founders of Wavely Diagnostics, Inc. S.G. is a co-founder of Jeeva Wireless, Inc. and Sound Life Sciences. R.B. is a co-founder of EigenHealth, Inc. R.B. is a consultant and stockholder of Spiway, LLC. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41551-022-00947-6.

Correspondence and requests for materials should be addressed to Justin Chan, Randall Bly or Shyamnath Gollakota.

Peer review information *Nature Biomedical Engineering* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Justin Chan, Shyamnath Gollakota

Last updated by author(s): Aug 28, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Custom Android code (v. 1.0) was used to collect data, and is available at https://github.com/uw-x/oae and https://zenodo.org/record/7032657.
Data analysis	Matlab R2021a

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data supporting the findings from this study are available within the article and its supplementary information. Source data are provided with this paper. The dataset used to generate the results for this study is available at https://github.com/uw-x/oae and https://zenodo.org/record/7032657.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were chosen on the basis of whether sufficiently representative data were collected.
Data exclusions	We excluded patient ears when the patient was unable to complete testing. This criterion was established prior to data collection.
Replication	The findings were replicated.
Randomization	Randomization of the subjects was not relevant, as the participants were not allocated into groups.
Blinding	The nvestigators were not blinded.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	We tested our devices on 201 ears, with patients between 1 week and 20 years of age. Of the 201 tested ears, 98 had an accompanying behavioural audiogram, and 14 underwent a diagnostic auditory brainstem response test. Ears without diagnostic audiometric testing were assigned on the basis of data from their clinical history as well as on school-based hearing screens (n = 14 ears) and newborn screens (n = 81 ears). Six ears were assigned hearing status on the basis of clinical assessment, meaning that they had no screening or diagnostic hearing tests but had no subjective hearing concerns and no concerns from the attending otolaryngologist who saw and examined the patient. Of the 201 tested ears, 135 had normal hearing, 38 had sensorineural hearing loss, 13 had conductive hearing loss, and 9 had mixed hearing loss.
Recruitment	We conducted a clinical study at Seattle Children's Hospital and the Center on Human Development and Disability at the University of Washington on a cohort of patients in otolaryngology, craniofacial and hearing clinics. Parental permission was obtained for participants under the age of 18 years. Children aged 7 to 17 provided written assent. Assent was obtained after parental permission was granted. Children aged 7 to 12 signed a simple assent form, and children aged 13 to 17 signed a consent form. Parents co-signed the consent form. Participants 18 years and older signed a consent form.
Ethics oversight	Seattle Children's Institutional Review Board.

Note that full information on the approval of the study protocol must also be provided in the manuscript.