Intelligent Earable Systems for Equitable Healthcare

Justin Chan

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee:
Shyam Gollakota, Chair
Shwetak Patel
Emily Gallagher

Program Authorized to Offer Degree:
Paul G. Allen School of Computer Science & Engineering

University of Washington

Abstract

Intelligent Earable Systems for Equitable Healthcare

Justin Chan

Chair of the Supervisory Committee:

Shyam Gollakota
Paul G. Allen School of Computer Science & Engineering

Access to basic medical resources like hearing care is influenced by factors like an individual's birth country. This lack of access is often due to the prohibitively high prices of medical devices which cannot be afforded by most of the world. At the same time, mobile sensors and technologies have advanced substantially over the last two decades and are now ubiquitous. In this dissertation, we develop a unique approach to societally impactful research that involves identifying problems that are important in the medical domain, and formulating solutions which are interesting from a computational standpoint. Through this process, we create a suite of intelligent earable systems for equitable healthcare that breaks the conventional wisdom that expensive medical devices are needed for high quality clinical testing. By working towards adoption in the real-world, these systems are now transforming the field of audiology and having societal impact. The key challenge in designing these systems is in leveraging low-cost, commodity hardware to perform medical diagnostics of the ear at scale while still achieving high-quality clinical performance. We design computational methods spanning applied machine learning, wireless sensing, signal processing and embedded hardware to enable these systems to generalize across different hardware and real-world environments.

The first system can detect ear infections using the speakers and microphones on a smartphone and a paper cone. The second system enables low-cost newborn hearing screening using earphones and wireless

earbuds. The third system is an inexpensive smartphone-based tympanometry system to make screening of middle ear disorders more accessible. Looking forward, this dissertation sets the stage for the mobile systems community which is uniquely positioned to develop wearable and mobile technologies that can alleviate global health inequity and ensure that every human on the planet has access to basic medical tools.

Acknowledgements

First and foremost I would like to thank my advisor Shyam Gollakota. His creativity and boundless energy has been a constant source of inspiration within the lab. Shyam's ability to swiftly identify connections between ideas and articulate a broader vision has been instrumental in guiding and shaping my research direction. I have been privileged to have worked with him on a diverse array of research projects from mobile health to underwater systems and to wireless sensing. It was also a lot of fun co-founding a startup and paving the way for our research to be commercialized. Throughout my PhD journey, Shyam has opened many doors, and pushed me to achieve and explore more than I thought was possible, for which I will always be grateful.

I would like to thank my committee members for their support of my research. Thank you to Shwetak Patel for collaborating on the wireless earbud project and for inspiring me with his many cool mobile health and sensing projects. Thank you to Emily Gallagher for bringing me into the world of global health research, and being a strong advocate for the projects in this thesis. Thank you to Patrick Boyle for supporting this research and providing valuable feedback.

Going on the academic job market is an enormous and demanding effort. I am lucky to have had the support of numerous individuals who have guided me through this journey. Thank you to Shyam Gollakota, Shwetak Patel, Ed Lazowska, Emily Gallagher, Randall Bly, and Xia Zhou for supporting me with letters of recommendation. Thank you to Tim Althoff, Hank Levy, and Ed Lazowska for providing critical feedback on my job talk. Thank you to Arka Majumdar, Daniel Zamora, Edward Wang, Ethan Katz Bassett, Kevin Jamieson, Les Atlas, Les Sessoms, Ludwig Schmidt, Magda Balazinska, Marissa Radensky, Matt Reynolds, Rajalakshmi Nandakumar, Scott Hauck, Stefano Tessaro, Su-In Lee, Tom Anderson, Vikram Iyer, and Yasaman Sefidgar for interview advice and mock interview practice.

I would like to thank my medical collaborators who made much of my research possible. Thank you to Sharat Raju for kicking off my research journey in mobile health with the ear infection project. I deeply appreciate the close working relationship we had recruiting patients, analyzing data, and building up the project together. It was a privilege as well to continue refining the project for wider use, and commercializing it as a startup. Thank you to Jacob Sunshine for working closely with me on the agonal breathing and wearable injector project, and for being supportive during my academic job search. Thank you to Kelly Michaelsen for working together with me on creating and executing on the blood clot testing system. Thank you to Randall Bly for believing in the value of all the ear-related projects, and for making the otherwise challenging prospect of clinical recruitment and analysis across hundreds of patients a much easier task. Thank you to Nada Ali for the support with recruiting and executing on the hearing loss project, and for expanding the testing to the infant population.

I would like to thank my labmates who have taught me so much over the years. I would like to thank Vikram Iyer who, despite being constantly busy, did not hesitate to sacrifice hours of his time to teach me about crucial aspects of research. From teaching me to use software radios and antennas, to helping me break apart smartphone cameras, to explaining to me how I should negotiate faculty job offers and create an equipment list for my future lab, Vikram has been a critical pillar in supporting my work and those around him. I would like to thank Ali Najafi for working with me on the tympanometer and hearing loss project. I learned a lot from him about signal processing and hardware hacking over many hours in front of a computer monitor and at a lab bench. I would like to thank Tuochao Chen for his efficiency, tenacity and attention to detail during the underwater projects. I will never forget the many long and cold nights we spent debugging and evaluating our code outdoors. I would like to thank Antonio Glenn and Malek Itani for their work on the wireless earbud project, and the many idle conversations we had about life. I would like to thank Rajalakshmi Nandakumar and Mehrdad Hessar for their support as senior PhD students, and their camaraderie as we took and taught courses together. Thank you to Anran Wang, Bandhav Veluri, Maruchi Kim, and Solomon Nsumba for being supportive labmates.

Lastly, I would like to express my heartfelt gratitude to my parents, with special acknowledgement to my mother whose relentless dedication and sacrifices have supported my educational and life journey through its ups and downs.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Health inequities often show up in the global health setting, where the country a person is born in has a huge impact on their ability to access even basic medical resources like hearing care [93]. This lack of access is often made challenging due to the prohibitively high prices of medical devices, which cannot be afforded by most of the world. Specifically, ear diagnostic hardware, and more broadly audiology as a field, has not seen significant advances in making medical devices more affordable or accessible. As a result, this has led to the continued use of outdated diagnostic equipment that still costs thousands of dollars. At the same time, mobile computing technologies have advanced substantially over the last two decades since the introduction of the first iPhone in 2007. Today, even budget smartphones and earphones with incredible economies of scale incorporate high-quality microphones and speakers that are ubiquitous around the world.

In this thesis, we present a unique approach to creating societally impactful systems that can help to alleviate health inequities. In this approach we seek to deeply understand a medical domain, identify problems that are intellectually interesting and create computational techniques and systems which can scale and achieve real-world impact. Through this approach, we have created a suite of frugal and accessible earable systems for diagnosing ear disorders in audiology, where for decades the conventional wisdom has been that you need expensive hardware to perform testing. We invent computational and sensing techniques that break this conventional wisdom and are able to transform the way that medical testing can be performed by making medical devices significantly more accessible. While there has been recent interest in using computing and machine learning for health, this dissertation is focused on going beyond building machine

learning models and instead goes towards the design of end-to-end intelligent earable systems for equitable healthcare. These systems are designed to span the stack, from sensor hardware to the software algorithms, work end-to-end in real-time, and generalize to practical environments in the wild. In combination, the suite of earable systems presented in this dissertation are transforming the field of audiology by making medical devices significantly more frugal and accessible, thus reducing health inequities for billions of people around the world.

1.1 Creating intelligent earable systems for equitable healthcare

Designing these intelligent earable systems to scale to the real-world is technically challenging for two key reasons. First, in contrast to traditional medical devices which are designed for a single piece of standardized hardware, the hardware on smart devices like smartphones and earbuds are not designed for medical diagnostics. Furthermore, the sensors and compute on these devices have different characteristics and can also change with time as hardware degrades. Second, medical diagnostics traditionally rely on very expensive and sensitive sensors. For example, medical devices to perform newborn hearing testing in the ear can cost thousands of dollars. To get equitable healthcare, the challenge is to leverage low-cost hardware while still achieving high-quality clinical performance.

In this dissertation, we design computational techniques across signal processing, applied machine learning, wireless sensing as well as embedded systems and hardware to create real-time systems that reliably work in practice. Specifically, our approach to designing these systems is to create methods that can take advantage of the sensing capabilities of the smart devices around us that are inexpensive and ubiquitous due to their economies of scale. We create two types of techniques that can enable these systems to scale: 1) wireless sensing and ML techniques generalize our system across different hardware 2) hardware-software co-design enables our systems to achieve high clinical accuracy at a low-cost.

Building on these techniques, we create intelligent earable systems that can scale to the real-world. Looking forward, the techniques and systems in this dissertation set the stage for the mobile systems community which is uniquely positioned to develop wearable and mobile technologies that can alleviate global health inequity and ensure that every human on the planet has access to basic medical tools

Here, we introduce the following intelligent earable systems for equitable healthcare:

- *Detecting middle ear fluid using smartphones [73].* The presence of middle ear fluid is a key diagnostic marker for two of the most common pediatric ear diseases: acute otitis media (AOM) and otitis media with effusion (OME). We present an accessible solution that uses speakers and microphones within existing smartphones to detect middle ear fluid by assessing eardrum mobility. We conducted a clinical study on 98 patient ears at a pediatric surgical center and achieved results comparable to published performance measures for tympanometry and pneumatic otoscopy. Similar results were obtained when testing across multiple smartphone platforms. Parents of pediatric patients ($n = 25$ ears) demonstrated similar performance to trained clinicians when using the smartphone-based system. These results demonstrate the potential for a smartphone to be a low-barrier and effective screening tool for detecting the presence of middle ear fluid.

- *Newborn hearing screening using low-cost earphones [66; 67].* Otoacoustic emissions (OAEs) are faint sounds generated by the cochlea that provide information about the function of the outer hair cells of the cochlea. In high-income countries, infants undergo OAE tests using sensitive and expensive acoustic hardware 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. We present two low-cost systems for healthcare workers to perform newborn screening:

  - *An off-the-shelf otoacoustic-emission probe for hearing screening via a smartphone [66].* 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.
  - *Wireless earbuds for low-cost hearing screening [67].* We present the first wireless earbud hardware that can perform hearing screening by detecting otoacoustic emissions. We show that by designing wireless earbuds using low-cost acoustic hardware and combining them with wireless sensing algorithms, we can reliably identify otoacoustic emissions and perform hearing screening.

Our clinical evaluations across 251 patient ears and three clinical sites demonstrate that our systems

can detected hearing loss with performance comparable to an $5000 FDA-cleared commercial device.

- *Performing tympanometry using smartphones [71].* Tympanometry is a test used to evaluate the health of the middle ear, which is involved in hearing, and helps in diagnosing middle ear disorders. However, the test currently requires expensive equipment and is limited in availability in resource-constrained settings. We present a low-cost smartphone-based tympanometry system. The device is able to change the air pressure of the ear canal and measure ear drum mobility. Our system, for which the software and hardware are openly available, is tested in 50 ears from children attending an audiology clinic. A panel of pediatric audiologists classified tympanometry measurements from our device and a commercial tympanometer with good agreement. Given the increasing availability of smartphones in developing countries, our system has the potential to make screening of middle ear disorders more accessible.

Below we provide a high-level description for each of these projects.

1.1.1 Detecting middle ear fluid using smartphones

In Chapter 2 we present an accessible system to detect middle ear fluid using the speakers and microphones on existing smartphones. Detecting middle ear fluid requires using either pneumatic otoscopy or tympanometry[114]. Pneumatic otoscopy is used by only 7–33% of primary care providers and is not designed for home screening purposes[143]. Tympanometry necessitates a referral to an audiologist and the use of expensive equipment[115; 116]. For these reasons, in 2016, the American Academy of Otolaryngology called for research into a brief, reliable, and objective method to detect middle ear fluid as well as new in-home strategies to help parents and caregivers monitor fluid after initial physician evaluation[143].

Here, we describe a system that uses the microphone and speaker of existing smartphones to detect middle ear fluid by assessing eardrum mobility as shown in Fig. 1.1. The system uses the smartphone speaker to play soft, 150-ms frequency-modulated continuous wave chirps from 1.8 to 4.4 kHz into the patient's ear canal. The microphone remains active during the chirp, collecting both incident waves from the speaker and reflected waves from the eardrum. Sound reflected from the eardrum will destructively interfere with the incident chirp, causing a dip in sound pressure along a range of frequencies. A normal eardrum resonates well at multiple sound frequencies, creating a broad-spectrum, soft echo; as a result, the

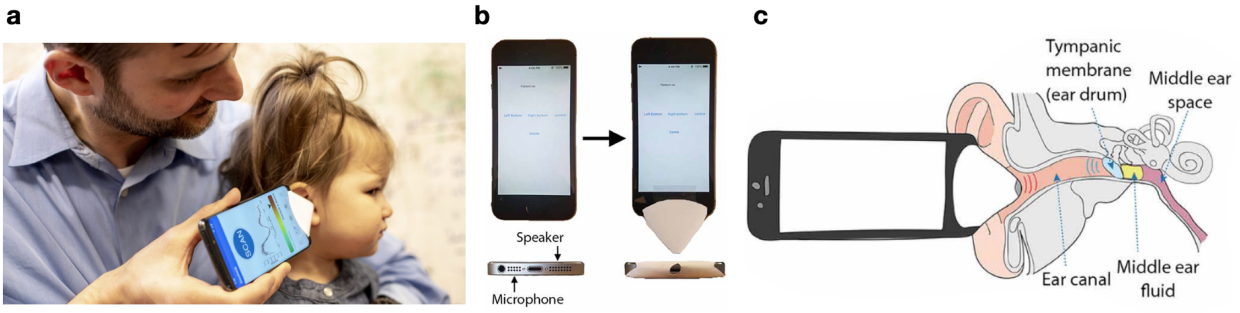

Figure 1.1: Detecting middle ear fluid using smartphones. a, Using a smartphone to detect middle ear fluid. **b,** Process of assembling smartphone funnel. **c,** Proper placement of smartphone and funnel at ear canal entrance.

shape of the resulting acoustic dip is broad and shallow in the frequency domain. In contrast, a fluid or pus-filled middle ear, as found in OME and AOM, restricts the vibrational capacity of the eardrum; sound energy that would have vibrated the eardrum is instead reflected back along the ear canal, creating more destructive interference and resulting in a narrower and deeper acoustic dip. The acoustic dip occurs at the resonant frequency of the ear canal where the quarter-wavelength of the chirp is equal to the length of the canal [157]. Thus, although individual differences in ear canal length affect the location of the dip along the frequency domain, the shape of the dip primarily depends on eardrum mobility.

We conducted a clinical study at Seattle Children's Hospital and evaluated our system on the iPhone 5s on patients between 18 months and 17 years of age using leave-one-out cross-validation to estimate performance on unseen data, we obtained an area under the curve (AUC) of 0.898 for the smartphone-based machine learning algorithm. In comparison, commercial acoustic reflectometry, which requires custom hardware, achieved an AUC of 0.776. Furthermore, we achieved 85% sensitivity and 82% specificity, comparable to published performance measures for tympanometry and pneumatic otoscopy.

Impact. This work demonstrated performance on-par with specialist medical tools, and has been commercialized by startup Wavely Diagnostics. Wavely's first product using this technology is now FDA-listed and is available to select early access participants and healthcare systems.

1.1.2 Newborn hearing screening using low-cost earphones.

In the previous chapter we focused on democratizing access to detecting middle ear fluid, which occurs in the middle ear space. Here, our focus is also related to a common malady related to ears, specifically

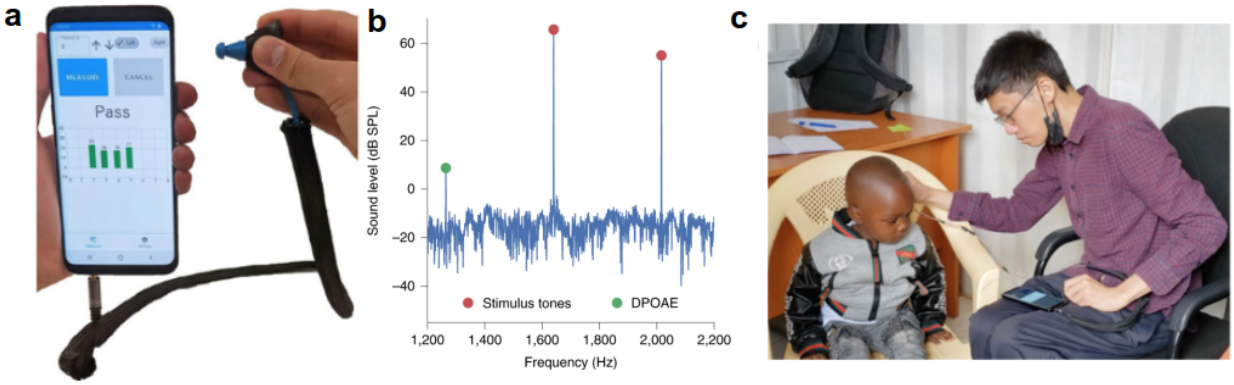

Figure 1.2: Newborn hearing screening using low-cost earphones. **a,** Overview of the earphone-based OAE probe system. **b,** A pair of earphones send two pure tone sound stimuli f_1 and f_2, through silicone tubes into the ear. A microphone positioned directly by the ear tip measures the OAEs emitted from the cochlea. The recording from the microphone is averaged over time by the smartphone to reduce noise and detect an OAE signal. **c,** Patient testing of newborn hearing test using earphone-based system.

hearing loss. While middle ear fluid can block the ear canal and impede hearing, our goal here is to measure the health of the inner ear which is a common area where hearing loss occurs especially in newborns. A key thing to note here however is that both these tests serve complementary functions and can be used in combination to guide a clinician in better understanding if a patient's hearing loss is related to the middle or inner ear space.

It is estimated that 5.3% of the world's population suffers from disabling hearing loss and 80% of people who need hearing care live in low and middle-income countries [11; 125; 165]. 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 [13]. It is common practice for high-income countries to adopt guidelines for universal infant hearing screening using otoacoustic emission (OAE) or auditory brainstem response testing (ABR) [136]. In spite of this, the test equipment remains expensive, and costs thousands of dollars, which contributes to limited hearing screening in low and middle-income countries. In these countries, access to hearing assessment and equipment often requires travel to an urban setting and long wait times [125; 165].

Since the neonatal population cannot provide behavioural response to conventional audiometry tests [46; 47; 48; 172], existing newborn hearing screening technologies instead use the sounds generated by a healthy cochlea called otoacoustic emissions (OAE) [55; 68]. While we think of the ear as a biological organ that

receives sounds like a microphone, a healthy cochlea, the part of the inner ear responsible for converting sound waves into electronic impulses for the brain, also generates sounds. These emissions are created when the cochlea's sensory hair cells vibrate in respond to external sounds [55; 169; 170]. So, we could pick up these faint sounds and use their absence to detect hearing loss.

The challenge is that detecting these faint sounds emitted from the cochlea requires sensitive acoustic hardware and medical devices that are expensive (5000-8000 dollars) [2; 3]. As a result, there is limited to no-hearing screening in low and middle-income countries [91; 158]. Further, in rural and resource-limited settings, getting access to hearing assessment may often require travel to an urban setting and long wait times, significantly limiting the accessibility of hearing care [66; 125; 165].

An off-the-shelf otoacoustic emission device for hearing screening via smartphones. In Chapter 3 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 around $10. Our earphone-based design shown in Fig. 1.2 sends two pure tone signals using each of the earphone's earbuds. When stimulated by two frequencies, the cochlea generates distortion product otoacoustic emissions (DPOAEs) due to intermodulation [68]. 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 otoacoustic emissions at various frequencies. We designed real-time algorithms that run on the smartphone to perform calibration, noise detection, and automatic pass/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 and the ubiquity of smartphones, our earphone-based design could be used to increase early access to hearing screening across the world.

Wireless earbuds for low-cost hearing screening. In Chapter 4 we present OAEbuds, the first wireless earbud design for low-cost hearing screening. Our hardware-software system reliably detects otoacoustic emissions using low-cost acoustic hardware, while being in the form-factor of a wireless earbud as shown in Fig. 1.3. The earbud hardware is designed to work across a wide demographic from new-borns to adults.[1] Our design streams the digital acoustic data via Bluetooth to a nearby smartphone which is then used for processing the signals and displaying the test results.

[1] OAE testing is not limited to just newborns but is also used as part of clinical care in older kids and adults [1].

Figure 1.3: Wireless earbuds for low-cost hearing screening. **a,** OAEbuds in use with an infant. Our low-cost wireless earbud can perform hearing screening by detecting OAEs from the cochlea. **b,** Wireless OAEbuds hardware. **c,** Our algorithms combine frequency modulated chirps with wideband pulses emitted from a low-cost speaker to reliably separate OAEs from in-ear reflections and echoes.

There are two key technical challenges in achieving this design with low-cost acoustic components. First, since speaker hardware components are bulky, it is challenging to incorporate the two-speaker design in recent work [66] into the form-factor of a wireless earbud. Our experiments show that transmitting the dual-tone signals used in [66] on a single low-cost speaker hardware introduces nonlinearities that in turn creates inter-modulation tones that can be confused for OAEs. Second, when an acoustic signal is sent into an ear canal, it first gets reflected and creates echoes not only inside the earbud case but also the ear drum and the walls of the ear canal, before arriving at the cochlea. To accurately identify OAEs, it is important to determine when the reflections and echoes of the input stimuli end and the OAEs begin. We design a two-step protocol that uses wireless sensing techniques to address the above challenges using a single low-cost speaker in our wireless earbuds. Our algorithms combine frequency modulated chirps with wideband pulses emitted from a low-cost speaker to reliably separate otoacoustic emissions from in-ear reflections and echoes.

Impact. This work has been presented to government officials and clinicians at Kenya's Ministry of Health and University of Nairobi, and we are starting deployments at local clinics in Nairobi. It has led to a larger effort called TUNE with the goal of bringing universal newborn hearing screening across Kenya.

1.1.3 Performing tympanometry using smartphones

In the above chapters, we describe the creation of accessible tools to perform screening of disorders in the middle and inner ear space. A referral or failure of these screening tools would often necessitate further test-

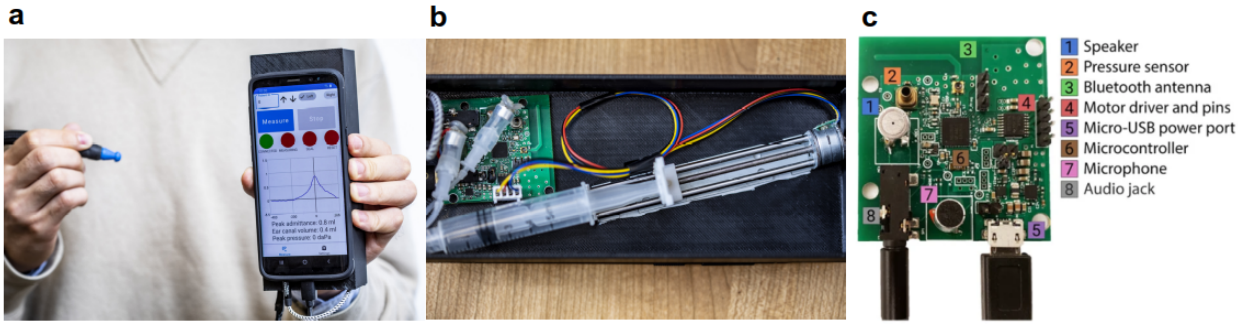

Figure 1.4: Low-cost smartphone-based tympanometer. **a,** All components of the hand-held tympanometer fit into a portable 3D printed enclosure that attaches to the back of a smartphone. **b,** Through precise movements of a plunger and syringe, the stepper motor introduces positive and negative pressure into the ear canal. **c,** Close-up of PCB containing the key acoustic and pressure sensing elements of the tympanometer: a speaker, microphone, motor driver, and pressure sensor. It also contains a microcontroller, Bluetooth antenna and micro-USB port for computing, communication, and power respectively.

ing to get to a precise a medical diagnosis. This chapter goes further by designing a low-cost tympanometer that is able to provide detailed diagnostics of middle ear function, and assist in the diagnosis of a variety of middle ear disorders. Middle ear disorders are one of the most common causes of preventable hearing loss. Unfortunately, while the developing world bears a disproportionate burden of these disorders, it often lacks access to diagnostic tools. Tympanometry is a key test to measure middle ear function, but remains available only on expensive test equipment that costs $2000-$5000. In countries like Kenya where hearing screening programs are less developed, hospitals often have to rely on a very small number of tympanometers donated by non-profit foundations like Hear the World. However, these hospitals still remain strained on resources, and have to serve patients who travel from as far as a nine-hour drive away for hearing screenings. Another problem faced by medical humanitarian groups like Doctors without Borders is that they often only have access to bulky, desktop-bound tympanometers which are inconvenient to transport. Furthermore, these devices require a wall outlet for power, and are not designed to work in the event of electrical failures which occur with greater frequency in developing countries.

Although this core principle behind tympanometry was introduced in the 1950-60s [156], there has not been many major advances in making these devices affordable and accessible for resource-constrained settings. As a result, existing tympanometers remain expensive, ranging from $2000 (handheld Otowave, Amplivox) to $5000 (AutoTymp, Welch Allyn), limiting its availability in resource-constrained scenarios including rural areas and developing countries [130]. At the same time, smartphone technology has ad-

vanced substantially over the last two decades. Today, budget smartphones that cost $40–50 second-hand contain powerful processors, user interfaces and high-resolution displays [31]. In Chapter 5, we introduce open-source algorithms and hardware for a smartphone-based tympanometer that consists of a lightweight and portable attachment to measure middle ear function as shown in Fig. 1.4. We leverage the fact that budget smartphones that cost $40–50 second-hand contain powerful processors to perform computing, user interfaces and high-resolution displays, and are thus able to support many of the operations that existing tympanometers are able to do at a lower price point. It is designed to automatically detect when a seal has been formed with an ear canal, safely vary air pressure, and generate a tympanogram on the smartphone in real-time. Our attachment can be assembled using electronic and passive components with a material cost of $28.

Impact. Due to the prevalence of smartphones in developing countries [14; 15; 23], our tympanometry tool may help increase timely access to otologic healthcare. We open-source our hardware and software making our system free and accessible to use and further adapt.

1.2 Beyond earable systems for equitable healthcare

While this dissertation is focused on creating earable systems for equitable health, my Ph.D. research has also successfully applied our unique approach of societally impactful research more broadly to inventing intelligent mobile and wearable systems that can significantly increase access to medical devices and diagnostics for billions of people. These systems demonstrate the wide applicability and unique potential of mobile systems to reduce health inequities in cardiology, hematology, and anesthesiology. We provide a high-level description for each of these projects below.

1.2.1 Contactless cardiac arrest detection using smart devices

There are many opportunities to develop new systems for clinical monitoring for which there currently do not exist any solutions. In particular, smart devices can be leveraged to enable ambient intelligent systems that can track life-threatening and sensitive conditions such as cardiac arrests. Out-of-hospital cardiac arrest is a leading cause of death worldwide with many victims dying alone in their sleep [129]. Cardiopulmonary resuscitation (CPR) can significantly increase the chances of survival. Yet a significant fraction of cardiac

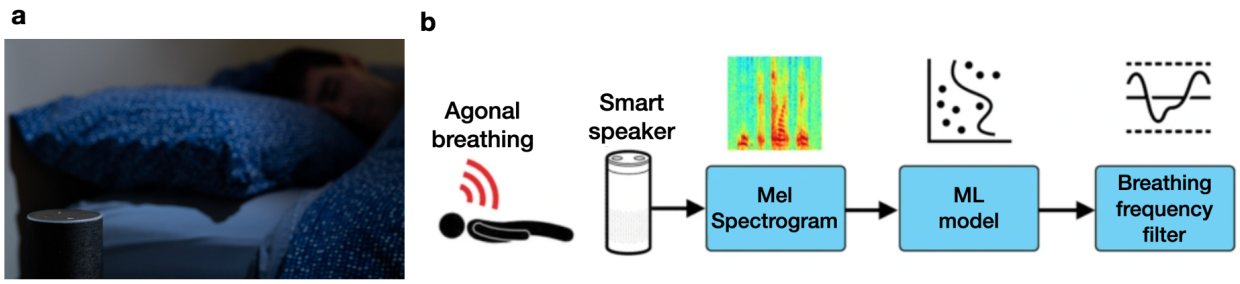

Figure 1.5: Contactless cardiac arrest detection using smart devices [74]. a, Using a smart speaker to detect agonal breathing. **b,** Agonal breathing detection pipeline.

arrest victims have no chance of survival because they experience an unwitnessed event, often in the privacy of their own homes. An under-appreciated diagnostic element of cardiac arrest is the presence of agonal breathing which is a form of disordered breathing that occurs in more than half of cardiac arrests [54; 145] due to severe hypoxia [117; 139].

The widespread adoption of smartphones and smart speakers presents a unique opportunity to identify this audible biomarker and connect unwitnessed cardiac arrest victims to Emergency Medical Services or others who can administer CPR. Using real-world labeled 9-1-1 audio of cardiac arrests, we trained a support vector machine to accurately classify agonal breathing instances in real-time in a bedroom environment (Fig. 1.5). We prototyped our contactless system using commodity smart devices (Amazon Echo and Apple iPhone) and demonstrated its effectiveness at identifying cardiac arrest-associated agonal breathing instances played over the air.

1.2.2 Micro-mechanical blood clot testing using smartphones

Here, we show that beyond speakers and microphones, there are other sensors on smart devices that can be leveraged for medical diagnostics like blood clot testing. For millions of people, medical conditions such as atrial fibrillation increase the risk of morbidity and mortality from blood clotting [89]. These individuals require lifelong administration of anticoagulation drugs such as warfarin, an effective medication but also one of the most common causes of hospitalization due to adverse drug events [63]. Hence, medication effects must be closely monitored via frequent tests of blood clotting times. However, these devices cost hundreds of dollars and patients in developing countries are in the therapeutic range only 40% of the time due to less frequent testing [128; 154; 161].

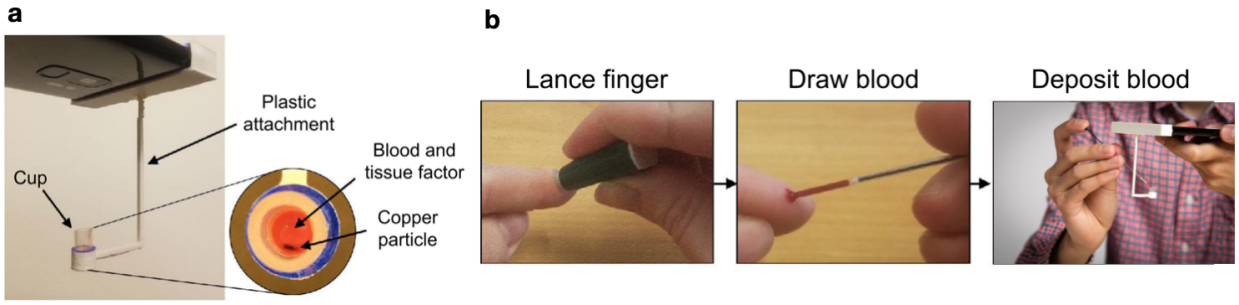

Figure 1.6: Blood clot testing using smartphones [70]. a, Schematic of the system: a plastic attachment containing a cup with 10 µl of whole blood, 20 µl of tissue factor, and a copper particle. The smartphone's vibration motor is coupled to the attachment and vibrates the particle, which is captured with the camera. **b,** A workflow (from left to right) showing how whole blood is added to the cup in our system

We describe a system that uses the vibration motor and camera on smartphones to perform blood clot testing (Fig. 1.6). Our system visually tracks the micro-mechanical movements of a small copper particle in a cup with a single drop of whole blood or plasma and the addition of activators. No additional electronics are required beyond a lightweight plastic attachment that couples the phone's vibrations to the cup. When the mixture is in a fluid state, the copper particle moves freely with the phone's vibration. As the blood clots, the viscous mixture causes the particle to slow to a stationary state. Making coagulation testing accessible in this manner may help improve time within the therapeutic range for users, particularly in rural locations.

1.2.3 Closed-loop wearable naloxone injector system

In the systems previously presented, we seek to tackle health inequities that occur in the area of global health or in high-income countries like the United States. Here, we seek to tackle inequities caused by financial incentives. There are many conditions for which there are less financial incentives for medical devices or pharmaceutical companies to create healthcare solutions. For example, opioid overdose deaths disproportionately affect people from low socioeconomic conditions and insecure housing. These deaths remain a major public health problem, with mortality increasing during the COVID-19 pandemic [28; 123; 171].

In this work, we develop a closed-loop, wearable naloxone injector system capable of administering naloxone in the setting of an apnea, specifically prolonged apnea (Fig. 1.7). Our solution comprises a pair of on-body accelerometers that measure respiration and detects apneic motion, a wearable injection system

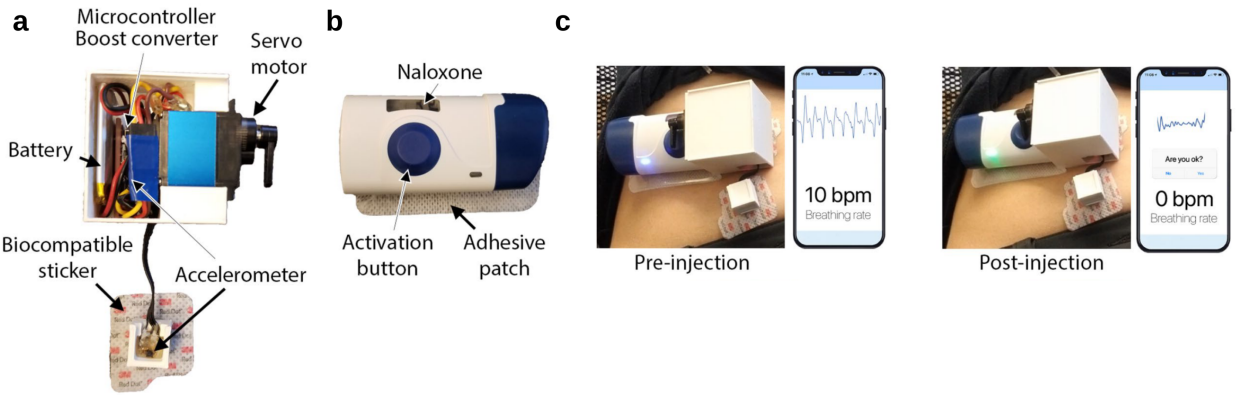

Figure 1.7: Closed-loop wearable naloxone injector system [69]. **a,** Sensor patch consists of two accelerometers to detect respiration and apneas, as well as a servo motor to activate the injector in the event of an overdose. **b,** Wearable injector delivers naloxone subcutaneously when activation button is pressed. **c,** The device as placed on the subject's abdomen prior to and after activation of the injector.

that administers the drug subcutaneously and an actuator to activate the injector system in the presence of prolonged apnea events. We conducted testing on participants in two environments: (1) a supervised injection facility in Vancouver, BC, where people self-inject opioids in the presence of medical personnel, and (2) a hospital environment.

1.3 Organization

The remainder of the dissertation is organized as follows: Chapter 2 presents an accessible system that uses speakers and microphones within existing smartphones to detect middle ear fluid by assessing eardrum mobility. We present the results of a clinical study across multiple smartphone platforms, nonclincian users, and different confounding ear pathologies. Chapter 3 reports the design and clinical testing of a low-cost probe for newborn hearing screening using an off-the-shelf earphone and a microphone connected to a smartphone through a headphone jack. Chapter 4 describes how the low-cost hearing screening test can be integrated into a wireless earbud hardware design using low-cost acoustic hardware and wireless sensing algorithms. Chapter 5 presents an inexpensive tympanometry system to detect middle ear disorders that uses a lightweight and portable attachment to smartphones. Chapter 6 concludes this dissertation, and presents several directions for future work.

Chapter 2

Detecting Middle Ear Fluid Using Smartphones

The presence of middle ear fluid is the key diagnostic marker for the two most common pediatric ear diseases, acute otitis media (AOM) and otitis media with effusion (OME) [114]. AOM, known commonly as an "ear infection," is characterized by the presence of infected fluid in the middle ear and results in symptoms of fever and ear pain. It is a leading cause of pediatric healthcare visits, and although many cases can resolve without antibiotics, complications may include eardrum perforation, mastoiditis, facial nerve palsy, or meningitis [56; 61; 142]. OME is the presence of middle ear fluid without signs of an acute infection and affects up to 80% of children [143; 173]. Although OME has few overt symptoms, making diagnosis more difficult, it is associated with speech delay, sleep disruption, poor school performance, balance issues, and a higher likelihood of developing AOM [143].

Diagnosis of OME or AOM requires detecting middle ear fluid using either pneumatic otoscopy or tympanometry [114]. Pneumatic otoscopy is used by only 7 to 33% of primary care providers and is not designed for home screening purposes [143]. Tympanometry necessitates a referral to an audiologist and the use of expensive equipment [115; 116]. For these reasons, in 2016, the American Academy of Otolaryngology called for research into a brief, reliable, and objective method to detect middle ear fluid as well as new in-home strategies to help parents and caregivers monitor fluid after initial physician evaluation [143].

Here, we describe a system that uses the microphone and speaker of existing smartphones to detect

middle ear fluid by assessing eardrum mobility. The system sends a soft acoustic chirp into the ear canal using the smartphone speaker, detects reflected sound from the eardrum using the smartphone microphone, and uses a logistic regression machine learning model to classify these reflections and predict middle ear fluid status. No additional attachments are required beyond a paper funnel, which acts as a speculum and can be constructed from printer paper, scissors, and tape. Real-time implementation and data processing are performed entirely on the smartphone, compatible with both iPhone and Android devices. The system demonstrated comparable performance across multiple smartphone platforms and when used by parents versus clinicians. Given the ubiquity of smartphones, this system may hold potential as a middle ear screening tool for parents as well as health care providers in resource-limited regions.

Results

2.1 Concept and Prototype

Our system uses the smartphone speaker to play audible, 150 ms frequency-modulated continuous wave chirps from 1.8 to 4.4 kHz into the patient's ear canal. The microphone remains active during the chirp, collecting both incident waves from the speaker and reflected waves from the eardrum. Sound reflected from the eardrum will destructively interfere with the incident chirp, causing a dip in sound pressure along a range of frequencies. A normal eardrum resonates well at multiple sound frequencies, creating a broad-spectrum, soft echo; as a result, the shape of the resulting acoustic dip is broad and shallow in the frequency domain. In contrast, a fluid or pus-filled middle ear, as found in OME and AOM, restricts the vibrational capacity of the eardrum; sound energy that would have vibrated the eardrum is instead reflected back along the ear canal, creating more destructive interference and resulting in a narrower and deeper acoustic dip. The acoustic dip occurs at the resonant frequency of the ear canal where the quarter-wavelength of the chirp is equal to the length of the canal [157]. Thus, although individual differences in ear canal length affect the location of the dip along the frequency domain, the shape of the dip primarily depends on eardrum mobility.

Our system builds upon existing acoustic reflectometry methods in three ways [79; 100; 109; 167; 168]. First, it is a predominantly software-based solution that takes advantage of existing smartphone hardware rather than requiring a separate device. Current acoustic reflectometers require a microphone and speaker in

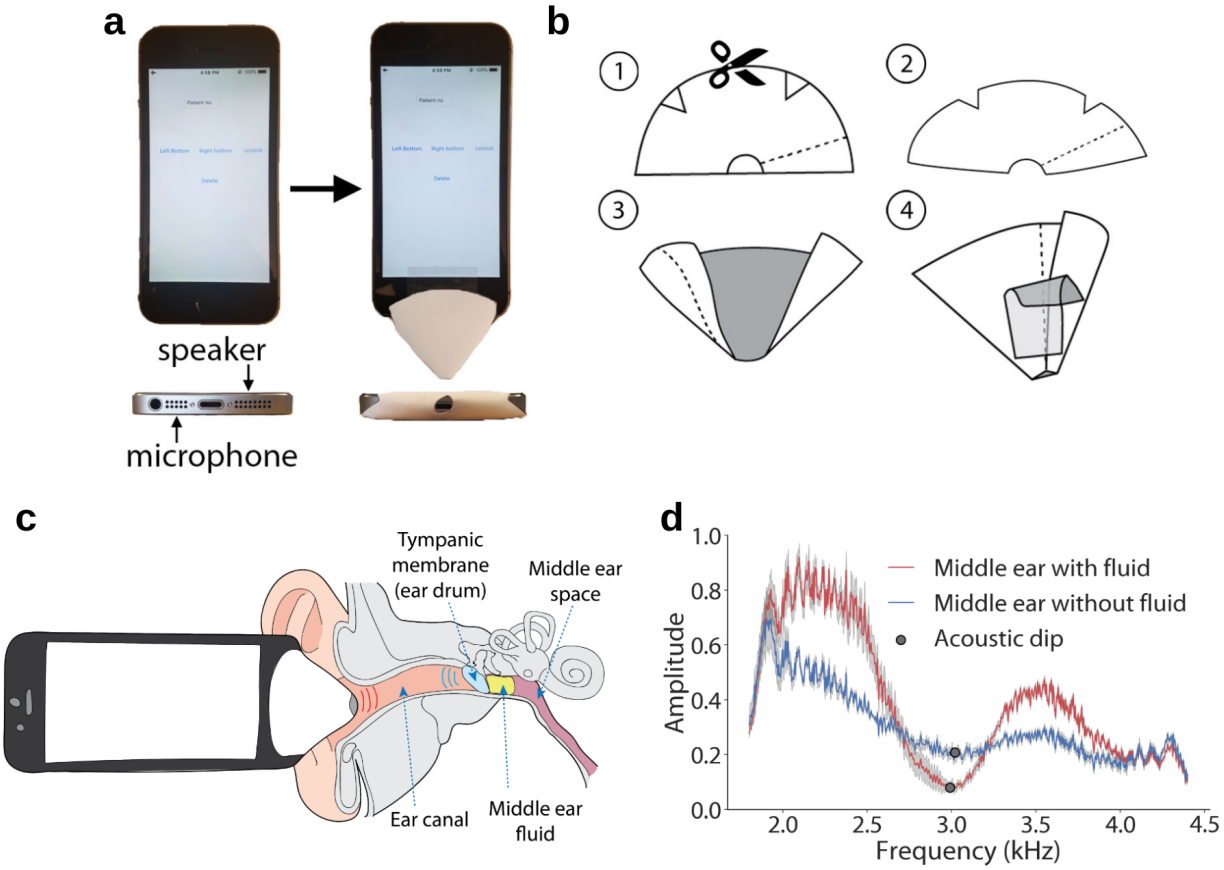

Figure 2.1: Using a smartphone to detect middle ear fluid. **a,** Location of speaker and microphone on the bottom of an iPhone 5s, with and without paper funnel attached. **b,** Process of assembling smartphone funnel. **c,** Proper placement of smartphone and funnel at ear canal entrance. **d,** Raw acoustic waveform obtained when chirps are played into an ear with middle ear fluid (red) and without fluid (blue). The SD (gray) is computed across 10 instances.

close proximity to produce and measure sound waves along the ear canal. Many modern smartphones have a similar configuration, with a co-located speaker and micro-phone on their bottom edge for noise cancellation (Fig. 2.1a). This includes all versions of iPhone, Samsung Galaxy phones after the S5, and other Android phones, including the Google Pixel. Second, we use a paper funnel as a speculum to direct sound into the ear canal. The funnel (Fig. 2.1b) can be assembled using a printed paper template, scissors, and tape. Without this attachment, the resulting waveform can be highly variable because sound could reflect off of different structures of the pinnae. Third, the system uses a logistic regression machine learning model to classify the waveforms received by the microphone. To identify whether a patient has middle ear fluid, we first preprocessed the raw waveform to locate and isolate the acoustic dip (Fig. 2.1 c,d). We then used logistic

regression to determine whether the shape of the dip was more indicative of a normal or fluid-filled ear. A text-based message is presented to the user indicating a result: "suggestive of middle ear fluid" or "middle ear fluid unlikely". On an iPhone 5s and Galaxy S6, data processing and classification took 771.98 ms and 1.2 s, respectively.

2.2 Clinical Testing

We tested system performance for detecting middle ear fluid in two separate cohorts. First, we conducted a clinical study on patients between 18 months and 17 years of age. We used this population to train the algorithm and obtain cross-validated performance measures. Second, we recruited a separate cohort of patients under 18 months of age and evaluated performance using the algorithm trained in the first clinical study.

The first clinical study was conducted at Seattle Children's Hospital surgical centers using a cohort of 98 patient ears between 18 months and 17 years of age from two different subgroups: patients undergoing ear tube placement, a common surgery performed on patients with chronic OME or recurrent AOM (n = 48 ears), and patients undergoing a different surgery, such as a tonsillectomy, with no recent symptoms of AOM or OME and no signs of middle ear fluid by physical examination (n = 50 ears). The median age of recruited patients was 5.0 [interquartile range (IQR), 2.0] years, height was 113.2 (IQR, 19.0) cm, weight was 20.0 (IQR, 9.2) kg, and the female-to-male ratio was 0.6.

A trained clinician performed all patient testing in a private waiting room just before surgery and with the patient awake and held or sitting upright. Soft chirps were played into the ear canal using multiple smartphone models and a new paper funnel for each patient. All patients were also tested in parallel with commercial acoustic reflectometry hardware [79; 107; 115]. After surgery, we prospectively assigned each ear its actual middle ear fluid status. A patient was considered positive for middle ear fluid if, during ear tube placement, an incision into the eardrum (myringotomy) yielded fluid (n = 24) or if the patient had a red, bulging eardrum consistent with AOM (n = 2). A patient was considered negative for middle ear fluid if, during ear tube placement, myringotomy yielded no fluid (n = 24) or if the patient did not receive ear tubes, did not have ear-related symptoms, and was negative for fluid on pneumatic otoscopy performed by the otolaryngologist (n = 48).

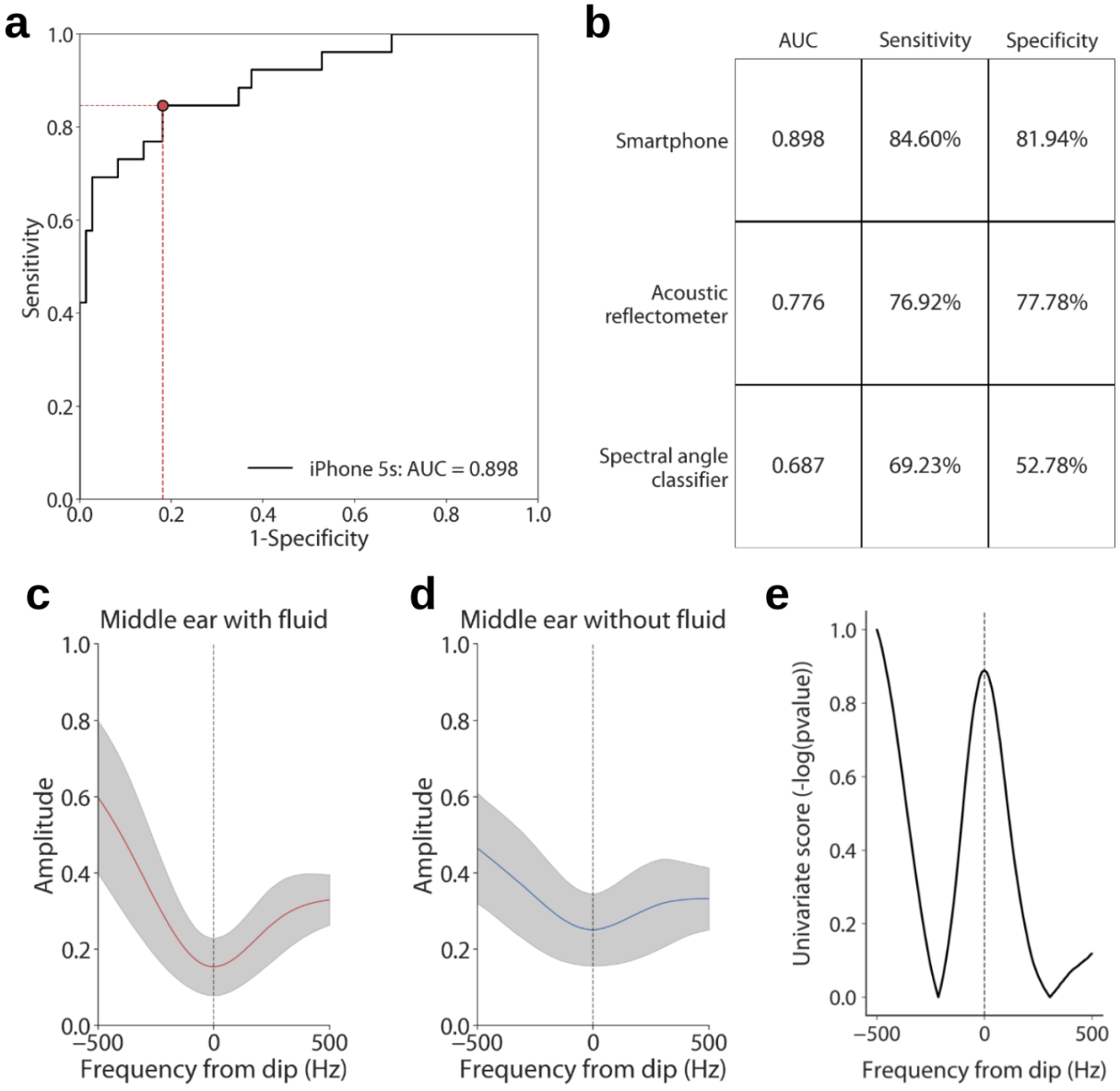

Figure 2.2: Classification of patient ears from clinical testing. **a,** ROC curve for our middle ear fluid detection algorithm, cross-validated on data collected from patients using an iPhone 5s (n = 98), with operating point denoted by the red circle. **b,** Comparison of performance for smartphone-based detection, acoustic reflectometer, and spectral angle-only classification during parallel clinical testing (n = 98). **c, d,** Mean acoustic dip classified by the algorithm as with middle ear fluid (red) and without middle ear fluid (blue). Shaded region represents one SD from the mean. **e,** Feature analysis indicating the weight that the classifier places on each frequency around the acoustic dip.

For classification, we used a logistic regression algorithm on pre-processed microphone acoustic data. The sound intensity (in decibels) of each frequency along the acoustic dip was inputted as a separate feature.

The algorithm was trained with iPhone 5s data collected from patients. Its classification accuracy was evaluated with leave-one-out cross-validation (LOOCV), a rigorous method to validate machine learning models [105]. During each iteration of LOOCV, 97 of 98 patient ears are used to train a model that is then used to output a prediction for the remaining one patient ear. This process is repeated for all 98 ears to estimate the accuracy of a model trained on all 98 ears when tested on unseen data. A receiver-operating characteristic (ROC) curve was generated from the cross-validation step with an area under the curve (AUC) of 0.898 (Fig. 2.2a). The operating point was chosen to have an overall sensitivity and specificity of 84.6% [95% confidence interval (CI), 65.1 to 95.6%] and 81.9% (95% CI, 71.1 to 90.0%), respectively. With k-fold (k = 10) cross- validation [149], we obtained a comparable AUC of 0.906. To address potential bias from training on the same patient's opposite ear, we repeated LOOCV, but during each iteration, we also excluded the contralateral ear from the training set, achieving an AUC of 0.899. We also downsampled the frequency response curve to 100 samples and obtained a similar AUC of 0.888. The fluid type was recorded as either being serous (n = 7), mucoid (n = 11), or purulent (n = 4) for 22 of 36 ears that had middle ear fluid. The algorithm correctly classified 86% (6 of 7) of ears that had serous fluid, 91% (10 of 11) of ears with mucoid fluid, and 100% (4 of 4) of ears with purulent fluid. These estimates predict the real-world clinical performance on unseen data of our final algorithm, which is trained on all 98 ears from the iPhone 5s dataset (Fig. 2.2b). Across male patients, for the iPhone 5s, 15 of 17 positive ears and 31 of 40 negative ears were classified correctly. Across female patients, 7 of 9 positive ears and 25 of 30 negative ears were classified correctly. For two ears, we did not record gender.

Post hoc, we examined how the algorithm classified acoustic waveforms. Figure 2.2c,d, plots the mean sound intensities at each frequency for all ears classified by the model. The algorithm predicted that ears with narrower and deeper acoustic dips were more likely to have middle ear fluid. Similarly, on univariate analysis, sound intensities at the top and bottom of the waveform, which determine the depth of an acoustic dip, were given the most weight by the predictive model (Fig. 2.2e). This result indicates that the algorithm can independently identify an acoustic pattern for middle ear fluid that is consistent with the known acoustic response of the eardrum [79; 107; 115].

The smartphone-based system demonstrated improved clinical performance compared to acoustic reflectometry [18], which uses custom hardware to assess middle ear fluid status. Head-to-head testing (Fig. 2.2b)

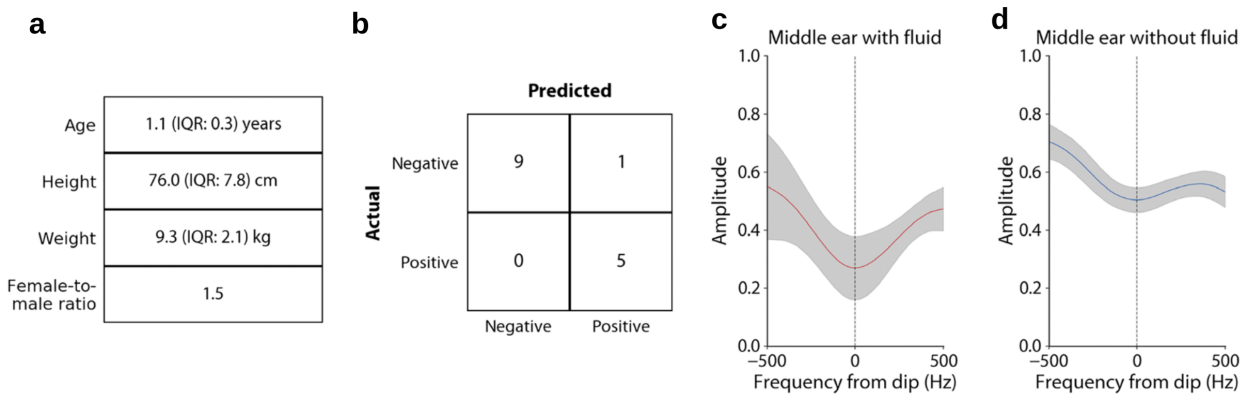

Figure 2.3: Classification of patient ears under 18 months. **a,** Demographic table of patients under 18 months. **b,** Confusion matrix of the algorithm's performance for patients under 18 months. **c, d,** Mean acoustic dip of ears of patients under 18 months (n = 15) classified by the algorithm as with middle ear fluid (red) and without fluid (blue). Shaded region represents one SD from the mean.

across the 98 patient ears demonstrated an AUC of 0.898 for the smartphone-based approach compared to an AUC of 0.776 for commercial acoustic reflectometry (EarCheck Middle Ear Monitor, Innovia Medical). The smartphone algorithm's improved clinical performance may be the result of applying machine learning over the waveform rather than relying on the hand-selected features used by acoustic reflectometers [115]. When classifying patient waveforms obtained from smartphones, we found that using only the spectral angle, as described in previous literature [115], reduces the AUC to 0.687.

We evaluated test-retest reliability in the pediatric patients enrolled in our clinical study. Each ear was tested twice per smartphone; between each attempt, the funnel was fully removed from the ear and reinserted. Of the 66 ears tested twice, 94% of the ears were classified the same between each attempt. When a discrepancy occurred, the algorithm used the positive result to minimize false negatives. Each testing attempt consisted of 10 chirps, and we tested the consistency across these chirps. In the clinical study, 93 of 98 ears showed no difference in classification among all 10 chirps. When doing a majority vote across the first three chirps, 96 of 98 ears showed no difference in classification compared to using a single chirp. Across all 98 ears, there was no difference when considering the classification result to be the majority (more than 5) of the 10 chirps.

Last, using the algorithm trained in the first study, we evaluated the system's performance in a separate cohort of patients under 18 months of age to assess accuracy in a younger population. We again recruited surgical patients at Seattle Children's Hospital, using the same criteria described in the first study with the

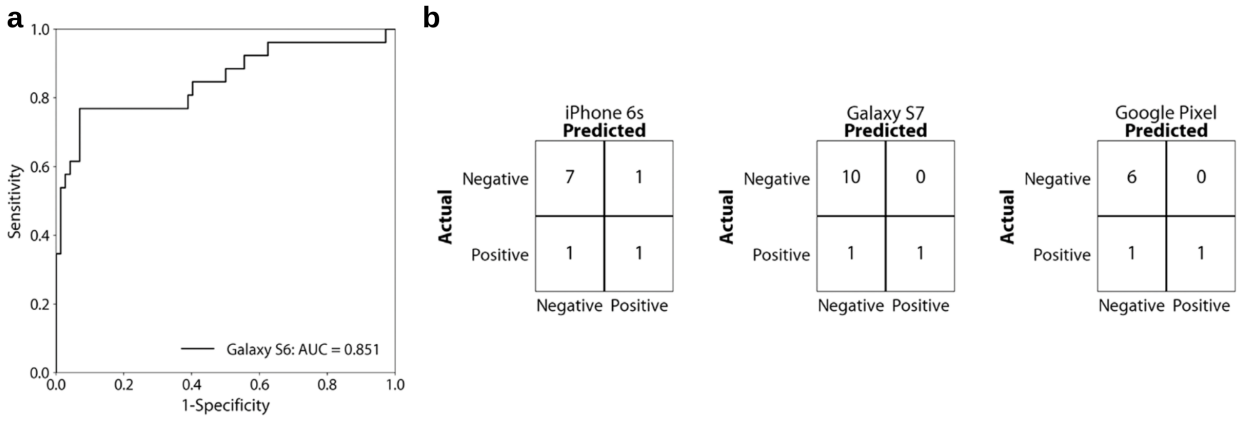

Figure 2.4: Classification performance across other mobile platforms. **a,** ROC curve for our middle ear fluid detection algorithm, cross-validated on data collected from patients using an Samsung Galaxy S6 (n = 98). **b,** Confusion matrices comparing performance on three other smartphones.

exception of age. This cohort included 15 patient ears and had a median age of 1.1 (IQR, 0.3) years, height of 76.0 (IQR, 7.8) cm, weight of 9.3 (IQR, 2.1) kg, and female-to-male ratio of 1.5 (Fig. 2.3a). The lowest age among this cohort was 9 months. All 5 ears that were positive for fluid and 9 of 10 ears that were negative for fluid were classified correctly (Fig. 2.3b). The shape of the acoustic dips paralleled those in the first clinical study: Ears with fluid had a deeper and narrower acoustic dip compared to ears without fluid (Fig. 2.3 c,d). This shows that an algorithm trained on patients over 18 months of age can properly classify patients under 18 months. We also trained and tested our algorithm's performance when using patients under 18 months as the training cohort. When running LOOCV across the 15 patient ears that were under 18 months of age, 5 of 5 ears positive for fluid and 10 of 10 ears negative for fluid were correctly classified. This is similar to the performance of the algorithm that is trained on the 98 patient ears over 18 months of age.

2.3 Performance Across Other Mobile Platforms

All patients in the first cohort (n = 98 ears) were tested in parallel with both the iPhone 5s and the Samsung Galaxy S6. Using LOOCV, we estimated performance of the iPhone 5s–trained system on unseen Galaxy S6 data. Specifically, the entire iPhone 5s dataset was used for training except for one patient ear, which was "held out" for testing. The trained algorithm was then tested on Galaxy S6 data from the held-out ear.

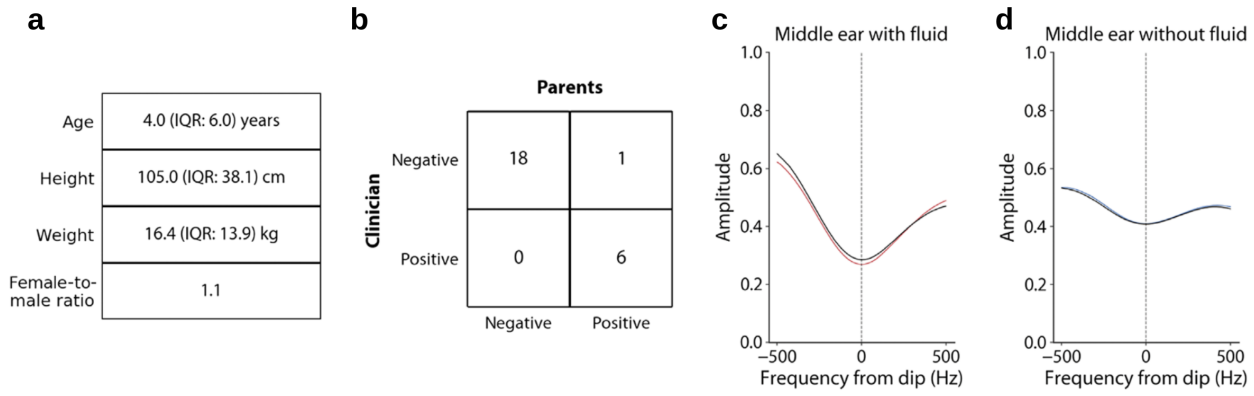

Figure 2.5: Performance testing with trained clinicians versus untrained parents. a, Demographic table of patients that were tested by parents. **b,** Confusion matrix of the algorithm's performance for patient ears (n = 25) tested by parents. **c, d,** Mean acoustic dip of ears tested by parents (black) and clinicians classified by the algorithm as with middle ear fluid (red) and without fluid (blue).

This was repeated for all patient ears in the cohort to generate an AUC of 0.851, as shown in Fig. 5.6a. In the same manner, we also tested a subset of this cohort using an iPhone 6s (n = 10 ears), Samsung Galaxy S7 (n = 12), and Google Pixel (n = 8). The algorithm correctly classified 80% (8 of 10) of iPhone 6s data, 91.7% (11 of 12) of Galaxy S7 data, and 83.3% (7 of 8) of Pixel data (Fig. 5.6b). The low sample size in these subgroups precluded generation of meaningful AUC values.

2.4 Performance Testing with Non-Clinicians

In a clinical setting, we evaluated the system's performance when used by parents. Trained clinicians briefly demonstrated proper technique for testing, and the parent of a pediatric study participant subsequently performed unaided testing on their child. The parent's results were then compared to those of the trained clinician. This cohort included 25 patient ears and had a median age of 4.0 (IQR, 6.0) years, height of 105.0 (IQR, 38.1) cm, weight of 16.4 (IQR, 13.9) kg, and female-to-male ratio of 1.1 (Fig. 2.5a). All 6 ears positive for fluid were classified the same by clinicians and parents, and 18 of 19 ears negative for fluid were classified the same (Fig. 2.5b). In addition, the mean acoustic dip was similar between clinicians (red) and parents (black) (Fig. 2.5c,d).

We tested the usability of funnel construction with a separate cohort of 10 untrained adults. After playing a short instructional video, we first measured the time it took participants to create and mount the

funnel using a paper template, tape, and scissors. The average time was 2.8 (±0.93) min. We then queried participants about the usability of the entire system; they gave an average usability rating of 8.9 (±1.1) on a scale of 1 (unusable) to 10 (extremely usable).

2.5 Effect of confounding ear pathologies

In the above studies, we exclude patients with ear pathologies that affect eardrum mobility. Next, we evaluated the algorithm's performance in the presence of ear pathologies such as cholesteatoma (n = 1), ossicular chain discontinuity (n = 1), acute eardrum inflammation (n = 1), and previous tympanoplasty surgery (n = 3). The algorithm produced false positives for middle ear fluid in all these patients.

Similarly, patients undergoing myringotomy but lacking middle ear fluid may have abnormal middle ear pressure that can affect eardrum mobility. In our cohort, there were seven patients reported by the surgeon as having acutely inflamed eardrums. Only one of these patients presented without fluid on myringotomy. This patient's ear was classified as positive by the algorithm. Thus, in the event that a patient presents with an inflamed eardrum but has not yet developed middle ear fluid, the algorithm would likely test positive and appropriately prompt further evaluation. The other acutely inflamed eardrums (n = 6 ears) had middle ear fluid and were appropriately classified as positive by the algorithm.

2.6 Benchmark Testing

Figure 2.6 demonstrates benchmark performance of our smartphone-based system across various design and environmental conditions. We identified design and environmental conditions that could affect system accuracy, including background noise, incident angle of the smartphone, alterations to the funnel, and changes in chirp volume. Testing was performed with a 2.5-cm closed-ended plastic tube used to calibrate existing acoustic reflectometers as a positive control [7]. The hard-backed end of the tube reflects sound to produce a narrow and deep acoustic dip that mimics an ear with middle ear effusion.

First, we used four different paper types of varying thickness and consistency to construct the funnel: filler paper, inkjet paper, laserjet paper, and cardstock. Changing paper type did not affect the classification accuracy (Fig. 2.6a). We also tested whether changes to tip opening diameter affected accuracy. The funnel

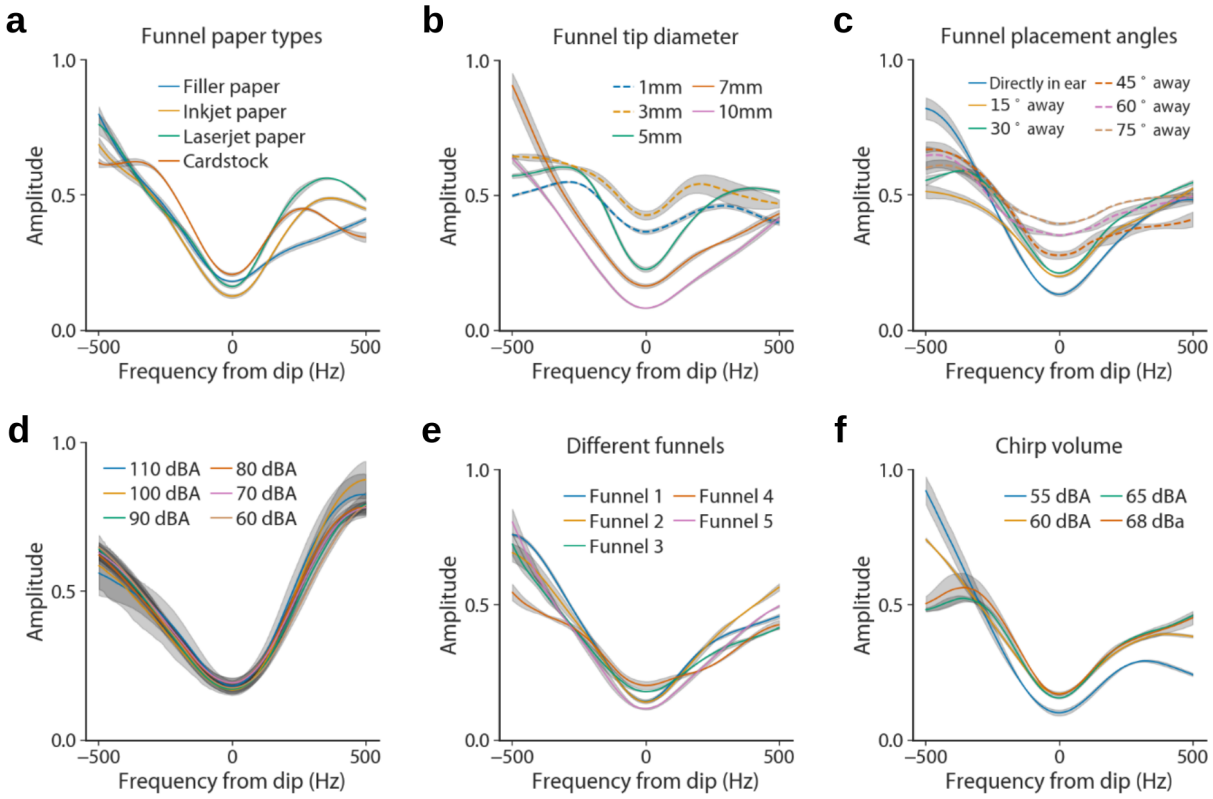

Figure 2.6: Benchmark testing across different scenarios. **a,** Different paper types used to construct the funnel. **b,** Different tip diameters of the funnel. **c,** Different funnel placement angles.**d,** Different background noise (infant crying) intensities. **e,** Funnels created by different individuals. **f,** Different chirp volumes. Solid and dashed lines indicate conditions where the algorithm classifies the waveform correctly and incorrectly respectively. The figure shows the mean for each test and a SD computed across five chirp instances.

is designed to have a tip opening diameter of 7 mm to approximate the diameter of the ear canal [118]. Variations in tip diameter from 5 to 10 mm did not affect performance. However, diameters of 1 and 3 mm produced false negatives (Fig. 2.6b). This suggests that a tip opening diameter between 5 and 10 mm is required.

Second, we varied the incident angle of our smartphone with respect to the calibration tube to examine the system's performance with slight deviations from direct line of sight. The smartphone tolerated up to a 45° offset from line of sight (Fig. 2.6c). Offsets of 60 or 75° produced a less prominent dip and false negatives. This suggests that while a parallel orientation is ideal, the system has some tolerance for non-ideal positioning. To validate the benchmark testing, we evaluated the effect of angle of insertion in an

upright patient (16 months of age) with middle ear fluid confirmed on myringotomy. To accurately assess angle of insertion, we used the built-in smartphone gyroscope to measure smartphone rotation. Initially, the smartphone was placed in line with the axis of the ear canal and began playing and recording chirps. Angular data were recorded for each chirp while the phone was rotated up (positive) and down (negative), up to 30° off axis. All chirps were correctly classified as positive within this range. Different insertion angles are also accounted for during clinical testing, where there was natural variance in measurement angle.

Third, we examined whether changes in background sound affected device accuracy, particularly with a crying child. We used an external speaker to play an audio file of a baby crying, with an average volume from 80 to 110 dBA. We tested our system in the calibration tube when it was placed directly next to the speaker (within 2 cm). In a measurement attempt, where five chirps were played, most of the chirps were correctly classified across tested volume levels during three different measurement attempts (Fig. 2.6d). When the background volume was lower than 80 dBA, all chirps were correctly classified.

Fourth, we tested the effect of deforming the funnel on classification of waveforms when using the calibration tube as a positive control. No deformation and partial deformation (more severe than typical use), were appropriately classified as positive. Full deformation resulted in a false negative. Our clinical study had variance in funnel deformation among different users and ears, and thus, the clinical results account for slight variations during actual use. We also tested the effect of different funnel instances on classification accuracy. Five different untrained users were instructed to construct a funnel and test it on the positive control. All acoustic curves were correctly classified as a positive ear (Fig. 2.6e). We also varied the sound intensity of chirps from 55 to 68 dBA. We found no difference in classification of the positive control (Fig. 2.6f).

Last, we consider the presence of cerumen (ear wax) and its effects on system performance. Our patients had partial cerumen occlusion (range, 0 to 50%) as estimated by the surgeon; our data indicate that this did not impair the algorithm's performance. Because none of our patients had complete cerumen occlusion, we used a positive control calibration tube for further testing. As expected, playing chirps into the tube generated a deep and narrow acoustic dip. Using putty to mimic cerumen, we found that partially occluding wax (60 to 70%) had little effect on the shape or position of the dip. This is consistent with previous observations that acoustic-based techniques are unaffected by less than 50% cerumen occlusion [152]. In

contrast, 100% cerumen occlusion, also known as impaction, occurs in 10% of children [124] and alters the waveform substantially. As the site of impaction moves closer to the entrance of the ear canal, the acoustic dip appears shallower and occurs at a higher frequency due to an effectively shorter canal and corresponding quarter-wavelength. In these cases, chirps can reflect off cerumen, generating a false acoustic dip that does not reflect middle ear status. For example, at a depth of 1 cm, which is the deepest point cerumen would naturally accumulate [58], our tests produced a false acoustic dip located about 1 kHz higher compared to a normal dip from eardrum reflections. At shallower depths of impaction, the false dip was even more right-shifted. Impaction at the entrance produced a waveform similar to calibration chirps played into open air. Given that the mean acoustic dip in our patients was located at 3 kHz (range, 2.4 to 3.7 kHz), a cerumen detection system in future prototypes could include an error display if an acoustic dip is identified outside the normal range or if the waveform resembles an in-air calibration chirp. Such a system must acknowledge that cotton swab insertion or iatrogenic manipulation can result in cerumen impaction deeper than 1 cm. These findings suggest that partial occlusion does not affect results, and full occlusion can be readily identified.

2.7 Discussion

Proper diagnosis of AOM and OME requires an examination of middle ear fluid status [114; 143]. Currently, most assessments of middle ear fluid are made either in primary care clinics and urgent care centers or, in recent years, remotely using smartphone-attached otoscopes [17; 19; 85; 120]. These assessments can be costly and time consuming. Furthermore, they usually rely on visual information without an assessment of eardrum mobility, which can compromise accuracy because middle ear fluid often produces only subtle changes in the eardrum's appearance [98; 138; 155]. Techniques that assess eardrum mobility, such as pneumatic otoscopy and tympanometry, have high sensitivity and specificity (90 and 80% for specialist-performed pneumatic otoscopy) but are infrequently performed outside of a specialist's office [134]. Other methodologies, such as air-coupled ultrasound, short-wave infrared imaging, and optical coherence tomography, hold promise in terms of accuracy, but they require additional specialized and expensive hardware [87; 164; 174]. Thus, there is a need for a middle ear fluid screening technique that does not use costly equipment or attachments, can evaluate eardrum mobility, and requires minimal expertise.

The value of a low cost, smartphone-based screening tool lies in its accessibility and user familiarity.

Ninety-six percent of all parents we queried regarding potential participation in the study consented, and many were interested in learning more about the technology. Although some children were apprehensive before the study, the phone's chirps (which sound like a small bird) had a calming effect, causing many children to respond with smiles or laughs.

Our system has several limitations. As with many screening tools, interpretation of the results requires appropriate clinical context such as symptoms and time course, and positive results should prompt further clinical evaluation for potential misclassifications. This system also does not distinguish between different types of middle ear fluid (purulent, serous, or mucoid). Knowing fluid type could potentially be useful for identifying AOM versus OME, although this distinction is also made on the basis of clinical history and symptoms. Furthermore, as with most middle ear assessment techniques, the system requires that a child not be agitated and remain relatively still for the duration of testing: The algorithm needs a minimum of three chirps for reliable results, which takes 1.2 s; considerable head movement during this duration can cause interchirp inconsistency.

This smartphone-based screening tool relies on an evaluation of eardrum mobility to detect middle ear fluid. This is also true for tympanometry and pneumatic otoscopy, which are the screening techniques recommended by the American Academy of Pediatrics and American Academy of Otolaryngology for middle ear fluid detection. Ear pathologies that affect eardrum mobility, such as cholesteatoma, ossicular chain discontinuity, acute eardrum inflammation, and previous tympanoplasty surgery, can produce false positives for middle ear fluid. Similarly, patients undergoing myringotomy but lacking middle ear fluid may have abnormal middle ear pressure that can affect eardrum mobility. Despite the potential for false positives, the use of eardrum mobility to predict middle ear effusion is well established both in active clinical practice and in previous large-scale pediatric studies [103]. Furthermore, in the context of screening, these false positives would appropriately prompt additional evaluation.

In summary, we present a proof-of-concept screening tool that can be implemented on commodity smartphones to determine the presence of middle ear fluid. Given the ubiquity of smartphones, the system we describe may have clinically relevant applications in developing countries and rural communities where smartphone availability is rapidly growing, in primary care settings as an adjunct to visual otoscopy, or for home screening by parents as a platform to reduce health care costs. Further longitudinal clinical trials are

required to determine the technology's impact in these and other potential scenarios.

Chapter 3

An off-the-shelf otoacoustic emission device for hearing screening via smartphones

It is estimated that 5.3% of the world's population suffers from disabling hearing loss. Further, a disproportionate brunt of this problem falls on low and middle-income countries [11]. 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 [13]. It is common practice for high-income countries to adopt guidelines for universal infant hearing screening using otoacoustic emission (OAE) or auditory brainstem response testing (ABR) [136]. In spite of this, the test equipment remains expensive, and costs thousands of dollars, which contributes to limited hearing screening in low and middle-income countries. In these countries, access to hearing assessment and equipment often requires travel to an urban setting and long wait times [125; 165].

Here 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 around $10. OAEs are sounds generated when the outer hair cells move in a healthy cochlea and provide information about their function [55; 68]. Unlike conventional audiometry tests [46; 47; 48; 172], OAE testing does not require a behavioral 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 [55; 169; 170].

Our 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 otoacoustic emissions (DPOAEs) due to intermodulation [68]. 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 otoacoustic emissions at various frequencies. We designed real-time algorithms that run on the smartphone to perform calibration, noise detection, and automatic pass/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 and the ubiquity of smartphones, our earphone-based design could be used to increase early access to hearing screening across the world.

Results

3.1 Concept and Prototype

Given the importance of OAE in infant hearing screening, there has been recent interest in designing better OAE hardware. Recent reports proposed using a single transducer hardware for both stimulus transmission and recording [96]. High-end personalized headphone products such as Nuraphone ($350) claim to measure OAEs to determine the listener's sensitivity to different acoustic frequencies in order to customize the audio signal sent to the listener [32; 65]. Prior work [95; 97] created a smartphone interface for the probes from an existing commercial OAE device [35]. Recent devices [20; 33] have used bone conduction to stimulate OAEs through a headband consisting of bone transducers. In addition to not using commodity earphones, none of these prior efforts present data from clinical studies for patients with hearing loss. Here, we provide a demonstration that repurposes 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 distortion product otoacoustic emissions.

Our design (Fig. 5.1a-c) consists of a pair of off-the-shelf earphones which each presents a different tone f_1 and f_2. These tones cause the cochlea to generate otoacoustic emissions 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

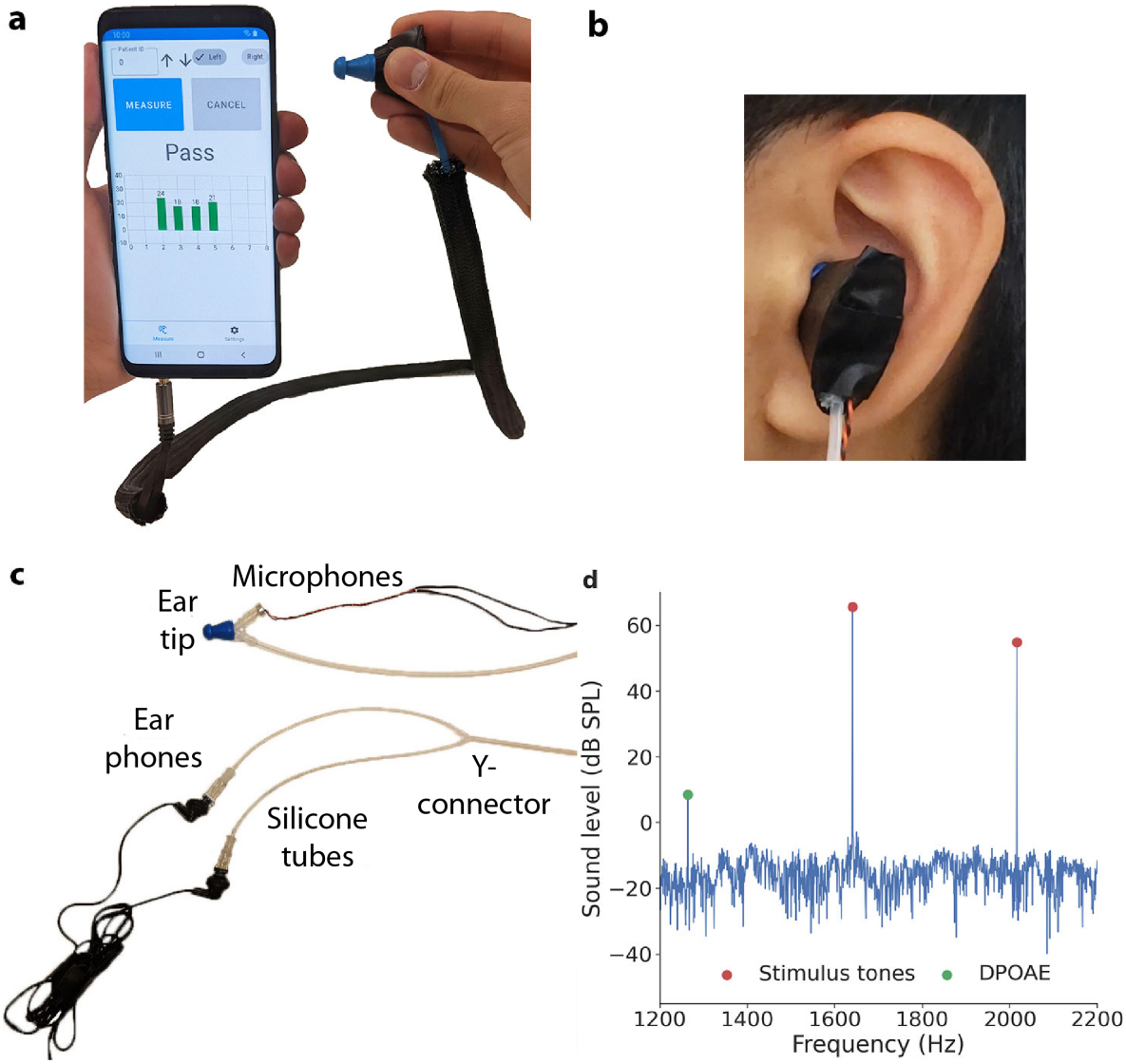

Figure 3.1: Earphone-based OAE probe system overview. **a, b, c** A pair of earphones send pure tone sound stimuli through silicone tubes into the ear. A microphone positioned directly by the ear tip measures the distortion-product OAEs (DPOAEs) emitted from the cochlea. The attachment connects to the phone via a 3.5 mm headphone jack. **a** and **b** show the assembled system which includes a nylon sleeve to protect the tubing from wear and tear, and a black plastic casing to shield the microphone from damage. **c,** shows 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 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. Since 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 65/55 dB 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 3D-printed enclosure. The probe head is compatible with rubber ear tips (Grason & Associates LLC, Rindge, NH) that are used in commercial OAE devices [30]. 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.

Fig. 5.1d shows the frequency spectrum in a healthy ear with otoacoustic emissions 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 $\frac{f_2}{f_1}$ set to 1.21–1.23. These frequency bands and ratios are commonly used bands for hearing screening [8] and are often available on most commercial screening OAE devices [45; 135]. We mark the test as a 'pass' if emissions were detected at three or more of the four test bands, and a 'refer' otherwise. Since varying ambient and physiologic noise levels can overwhelm the otoacoustic emissions, 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 non-linear 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 otoacoustic emissions can be successfully detected.

Table 3.1: Demographics of patients in clinical study.

Hearing loss, *n* (%)	
Yes	66 (33)
Sensorineural	38 (19)
Conductive	13 (6)
Mixed	9 (4)
Not known	6 (3)
No	135 (67)
Prior hearing test, *n* (%)	
Behavioral 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)

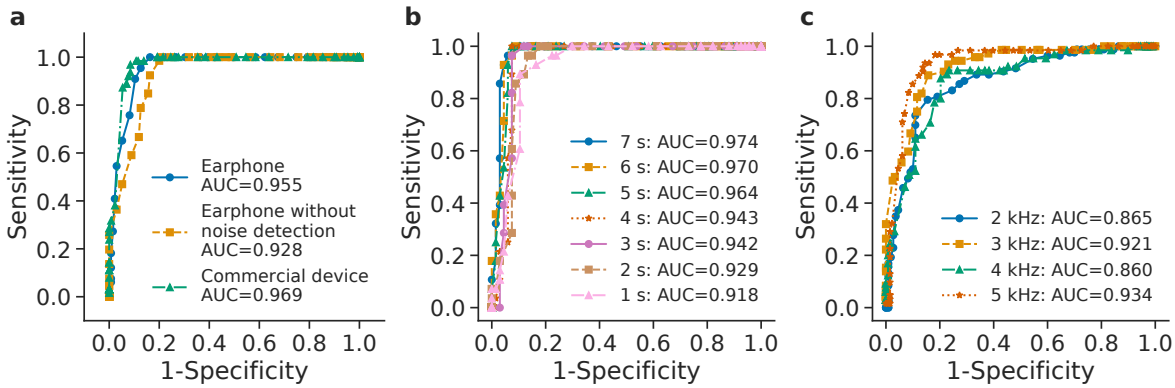

Figure 3.2: Clinical study performance. **a,** ROC curve showing performance of the earphone and commercial DPOAE device for screening performance of hearing loss at different SNR cutoff values. **b,** ROC curves for different signal averaging durations. **c,** ROC curve when comparing the pass/refer performance for individual frequencies to the commercial device for different SNR cutoff values.

3.2 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, craniofacial and hearing 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 (see Table 3.1). Five trained research assistants including a 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.

Distortion-product otoacoustic emissions were first measured using a commercial OAE device (AudX Pro, Bio-logic, 2006), 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 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. Otoscopy was performed prior to 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 was interpreted by an otolaryngologist, which included clinical and examination history review, behavioral audiometric testing, newborn hearing screen result, and diagnostic ABR. Of the 201 tested ears, 98 had an accompanying behavioral audiogram, 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, 13 had conductive hearing loss, 9 had mixed hearing loss.

We first considered the hearing screening to be a pass if otoacoustic emissions exceeding a predefined SNR threshold were detected at three or more of the four frequency bands. Fig. 3.2a 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 operating point for our earphone-based probe was obtained using a measurement time of 6 seconds per frequency band and an SNR threshold of 7 dB for patient ears over 6 months of age, and a measurement

time of 5 seconds 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% CI, 94.7 to 100.0%) and a specificity of 88.9% (95% CI, 82.5 to 93.1%). Of the 15 ears misclassified by the earphone, five of them had a history of middle ear disorders (eustachian tube dysfunction, cholesteatoma), three of them had prior tympanoplasty, and three 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 to 100.0%) and a specificity of 77.8% (95% CI, 70.1 to 84.0%). In comparison, the commercial device had an operating point with a sensitivity of 96.8% (95% CI, 89.1 to 99.1%) and a specificity of 91.5% (95% CI, 85.5 to 95.2%) using a measurement time of 6 seconds per frequency band and the default manufacturer prescribed SNR threshold of 6 dB.

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

Fig. 3.2b shows the device performance as the measurement duration changes from 1 to 7 seconds per frequency, for the 95 ears where the maximum measurement time of 7 seconds per frequency was used. As expected, longer measurement durations slightly improve the AUC. Fig. 3.2c also shows the agreement of our earphone-based probe with the commercial device in identifying the presence of otoacoustic emissions 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.

Prior to 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 226Hz as it was most responsive to changes in probe position, and can be used to infer if the probe is placed securely in the ear canal. We collected this 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. 6 of these 7 ears were correctly classified by our

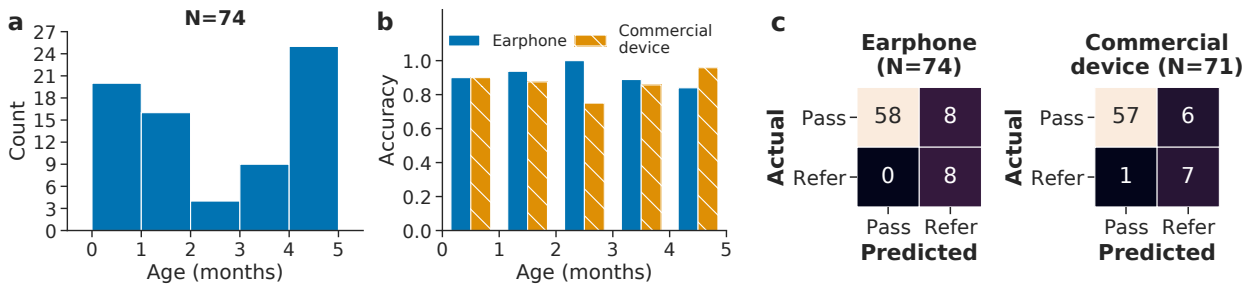

Figure 3.3: Device performance in infant ears under 6 months of age. a, Histogram of patient ears less than 6 months of age. **b,** Accuracy obtained by earphone and commercial device for different age groups. **c,** Confusion matrix showing performance of the earphone and commercial DPOAE device for screening performance of hearing loss.

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.

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 seconds. This is comparable to prior work which described that the coupling time for commercial screening OAE devices is slightly under 10 seconds [135].

3.3 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 one month (Fig. 3.3a). 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, 8 ears failed a newborn hearing screen or ABR test, 2 of these ears had conductive hearing loss based on a 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. 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 seconds per frequency as the

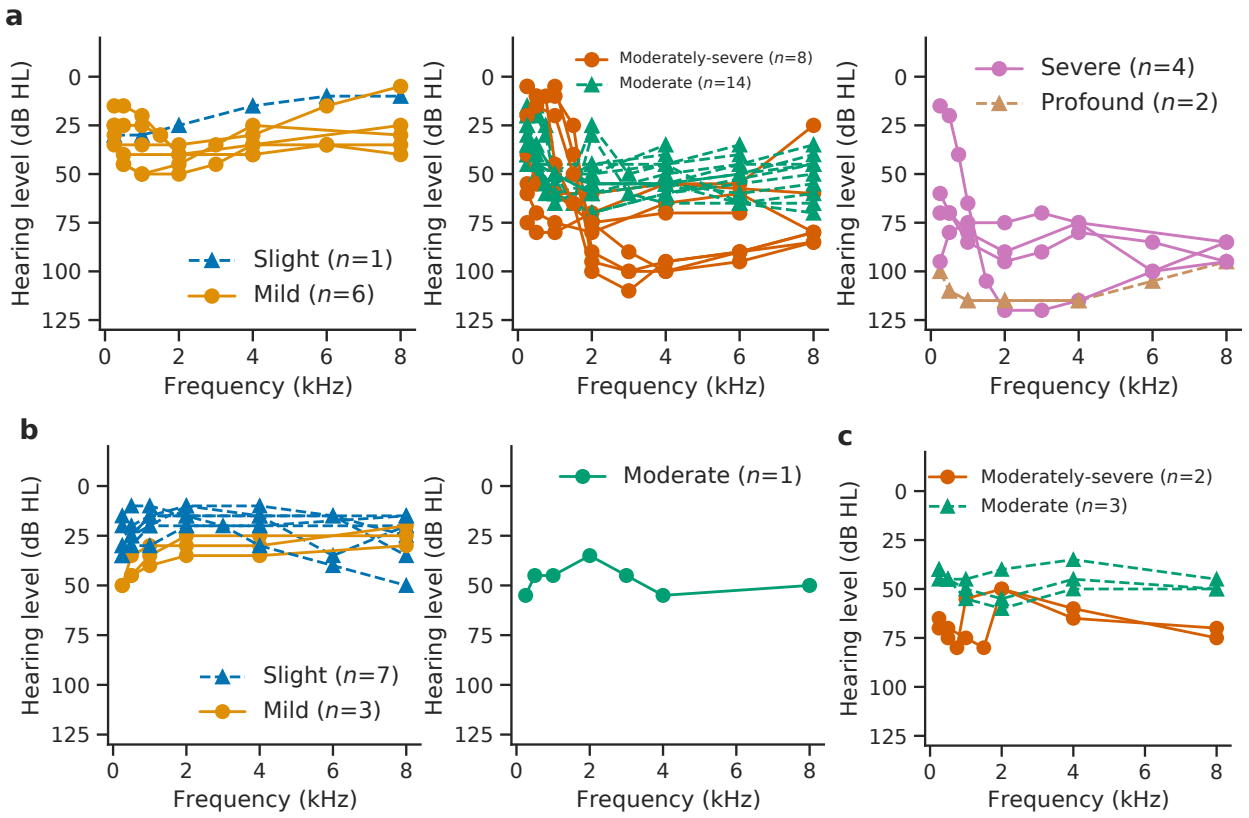

Figure 3.4: Audiograms of patient ears. Ears with **a,** sensorineural, **b,** conductive, and **c,** mixed hearing loss categorized by degree of hearing loss.

infant population was less likely to tolerate long measurements compared to older subjects. The SNR cutoff was also reduced to 5 dB for the earphone device to compensate for the increased noise and movement in this population.

Fig. 3.3b 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%. Fig. 3.3c 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

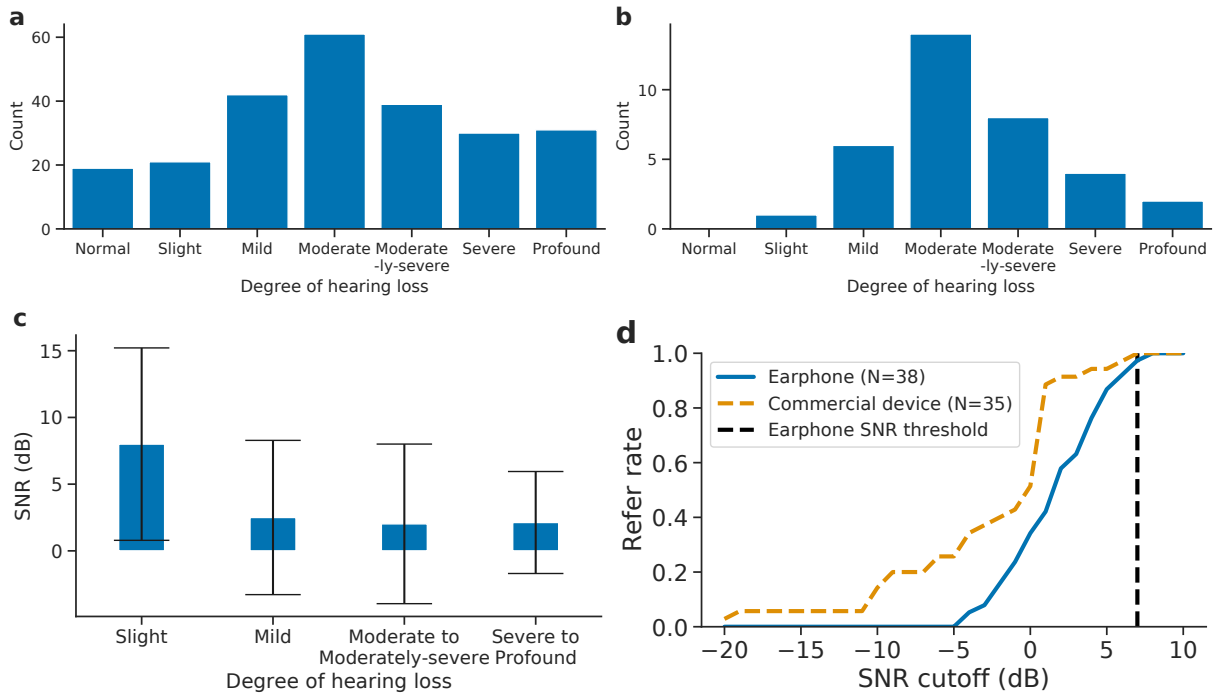

Figure 3.5: Device performance in patient ears with sensorineural hearing loss. a, b Histogram showing degree of hearing loss for ears with SNHL **a,** by audiogram frequency band $(n = 243)$ **b,** and by ear $(n = 35)$. **c,** Comparison of SNR levels obtained on the earphone device categorized by degree of hearing loss. **d,** Refer rate of earphone and commercial DPOAE device for different SNR cutoff values. **(c).** Data are presented as mean of SNR levels measured for ears with different degrees of hearing loss. $n = 4, 24, 88, 24$ 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.

correctly classified 27 of the 31 ears with normal hearing. Both devices correctly classified 5 of the ears with hearing loss.

3.4 Clinical performance on sensorineural hearing loss ears

Fig. 3.4a–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. We perform a subgroup analysis on patient ears with data on different degrees of sensorineural hearing loss (SNHL) from slight to profound. We collected data from a total of 38 patient ears

of which 35 had an accompanying behavioral 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 (Fig. 3.5a). Fig. 3.5b 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. 3.5c). 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 cutoff 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. Fig. 3.5d shows the refer rate obtained by both devices for different SNR cutoff 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, it obtained a refer rate of 97% (34 out of 35 ears). The 1 ear that was misclassified by both devices was classified by the audiologist as having mild to moderate hearing loss.

3.5 Benchmark testing

We evaluated our OAE probe with four different smartphones released between 2018 to 2020. We calibrated the output sound level to play two tones at 65/55 dB SPL. We performed this calibration by coupling the probe head of our device with a $18 reference sound level meter (Fig. 3.6a). Fig. 3.6c 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 dB SPL 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.

Fig. 3.6d 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 seconds. The measurement was performed three times, and the mean across the measurements is plotted. The noise

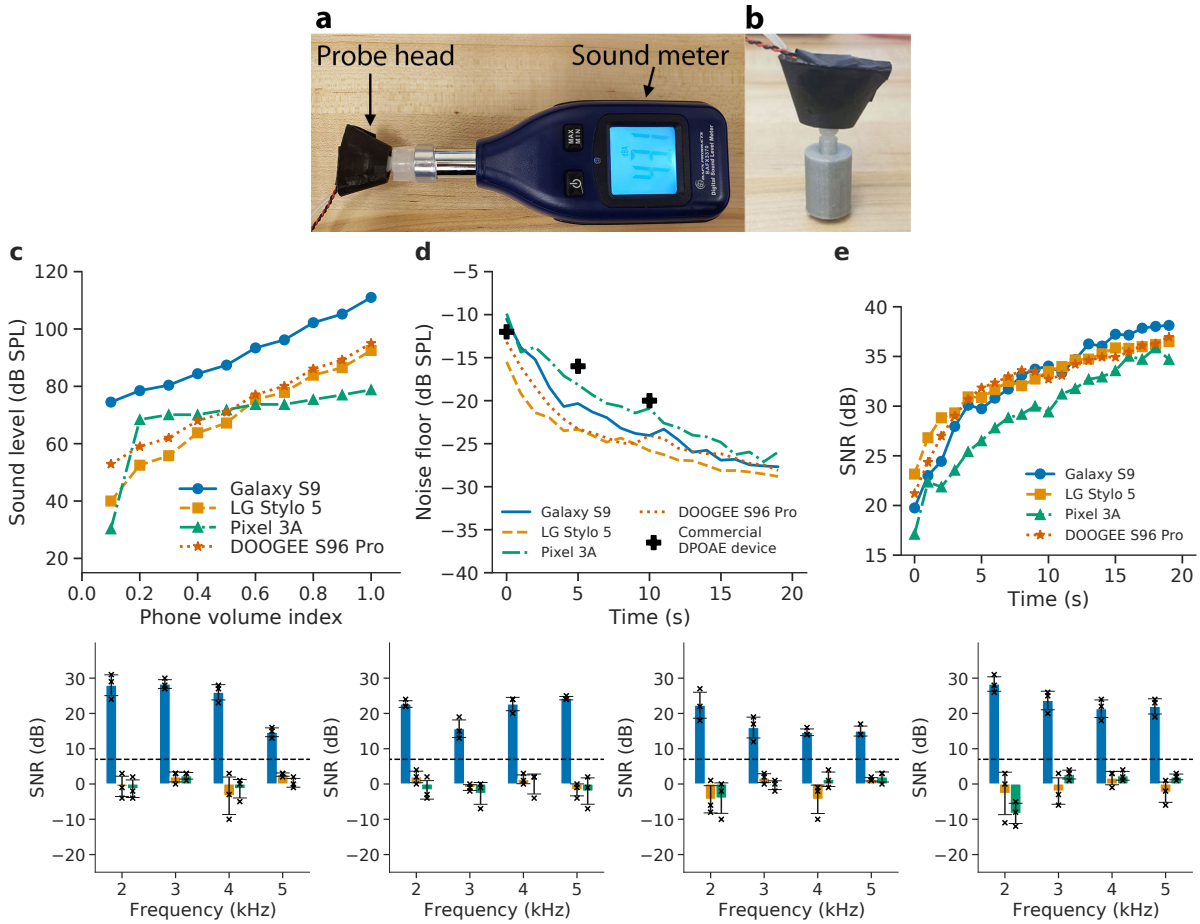

Figure 3.6: Benchmark testing across multiple smartphones. a, The probe head of the device is coupled to a sound level meter (BAFX 3370, Digital Sound Level Meter, $18) which outputs the sound levels emitted by the device in absolute physical units. This setup is used to calibrate the sound levels of f_1 and f_2 to 65/55 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 decrease over time as additional signal averaging is performed. **e,** The SNR of DPOAEs increase 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, DOOGEE S96 Pro. **(f).** Symbols, mean of three technical measurement replicates. Error bars denote s.d. from the mean.

floor of the commercial device decreases from -11 to the minimum reported level of -20 dB SPL after 11 seconds. In comparison, the noise floor of the smartphones ranged from -10 to -15 dB SPL after 1 second of measurement, and reached -20 dB SPL after 6 seconds for three of the phones, and after 11 seconds for the remaining phone. Fig. 3.6e shows the SNR of the recorded OAEs on each of the phones as a function of the averaging time.

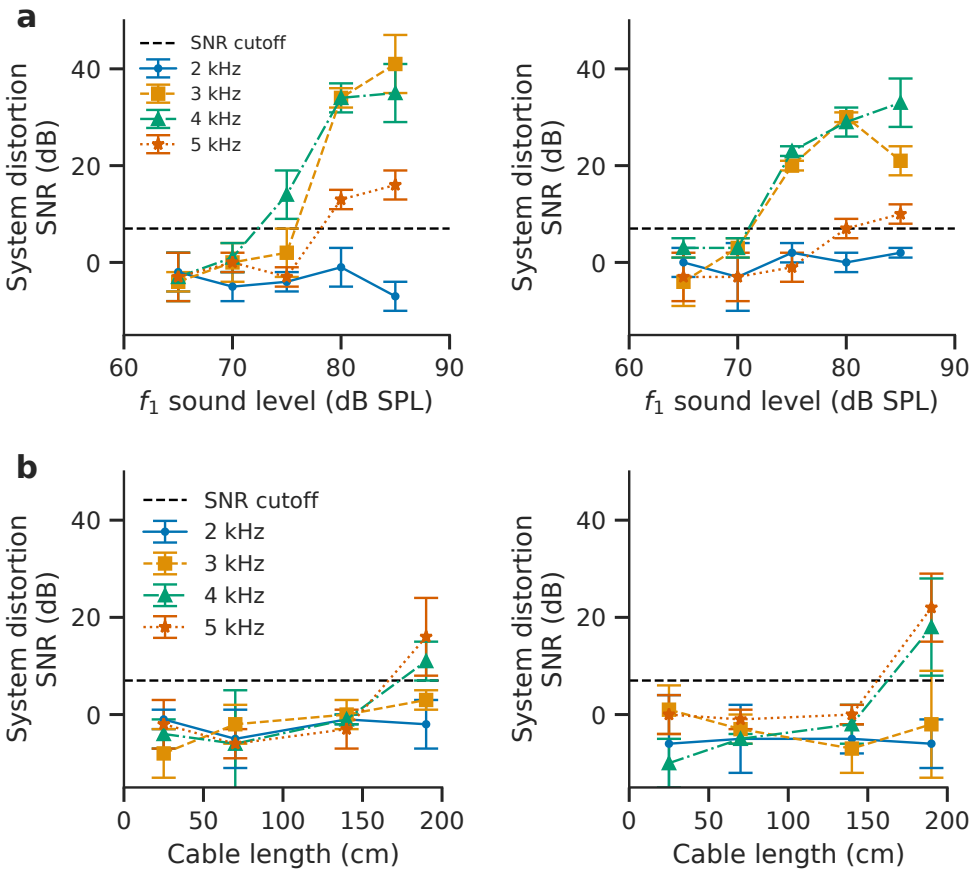

Figure 3.7: Benchmark testing of hardware distortion. The amount of distortion generated by the hardware at the DPOAE frequencies for open air (left) and calibration tube (right) is evaluated for **a,** different stimulus sound levels, and **b,** different cable lengths. **(a – b).** Symbols, mean of three technical measurement replicates. Error bars denote s.d. from the mean.

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. 3.6b). This testing is performed to ensure that the system non-linearities in the $2f_1 - f_2$ frequency would not appear as a false OAE. A 4 second measurement is performed three times for each experimental condition. Fig. 3.6f show the SNR in open air, 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.

Fig. 3.7a shows the effect of sound level on system distortions on the Samsung Galaxy S9. In this test, we initially set f_1/f_2 to 65/55 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 cutoff for the 65/55 and 70/60 dB SPL scenarios. System distortions increase for the 3, 4 kHz band at 75/65 dB SPL, while the distortions in the 2 kHz band remain below the threshold across all measured sound levels.

Finally, Fig. 3.7b shows the effect of cable length on system distortions. We increased the length of the single silicone tube from 25 cm to 190 cm and calibrate f_1/f_2 to 65/55 dB SPL. The system distortions remain below the threshold for all frequencies up till a cable length of 140 cm.

3.6 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 [70]. The assembly of our probe does not require specialized knowledge of electronics, and we estimate the cost of labor for assembly would be less than a dollar at scale — assembly costs of different categories of electronic products typically do not constitute more than 6% of the total cost of device components and assembly [88]. We point out that budget or second-hand smartphones can be purchased in low and middle-income countries for $35-50 due to their economies of scale, which is significantly lower than the upfront cost of commercial OAE devices. Additionally, our system has been built against Android SDK version 29 which has been designed to be compatible with future versions of Android operating systems [43]. While some newer smartphones have eliminated the audio jack interface, audio jack adapters [44] that cost a few dollars can be used to accommodate the wired earphones interface in our design. Finally, the FDA provides guidance for Mobile Medical Applications (MMA) [24; 26] and Software as a Medical Device (SaMD) [12; 16; 22] that regulates the custom software application running on the phone and not the smartphone hardware, thus potentially reducing the associated regulatory costs [70]. Commercially available medical devices that include MMAs and use sensors like microphones and cameras have been cleared or approved by the FDA [25].

The cost of OAE devices is only one factor associated with addressing the complex public health problem of hearing screening in low and middle-income countries. 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 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 (TUNE) 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 [51]. We envision that our earphone-based OAE probe can potentially be combined with existing frugal techniques to detect middle ear fluid on a smartphone [73] and assess eardrum mobility with smartphone tympanometry [72] 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 like technicians and volunteers need to be trained to perform the OAE test. Further, 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 prior work [159; 160] 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 to the sound level at the eardrum, which could result in calibrated sound levels that are higher than intended. While more accurate in-ear calibration procedures [77] 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 [9; 94]. While 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 clinic 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 if the probe is resistant against wear and tear in challenging environments over time. Our earphone-based probe is designed to only measure distortion product OAEs (DPOAEs) at this time; further work is needed to also measure transient-evoked otoacoustic emissions (TEOAEs). TEOAEs typically use 24-bit microphones with stimulus levels of 30 - 90 dB pe SPL [21; 27; 34; 50]. We note that the ADC on iOS devices do support recording at bit depths of 24 and 32 bits [38] and Android devices do support the use of external ADCs to provide 24 bit resolution via their USB Digital Audio interface [52]. 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 [55] and ototoxicity monitoring [146]. In these scenarios, OAEs are often also tested in the 800 Hz to 10 kHz range [55; 146]. Given that higher frequencies undergo more viscothermal losses within small diameter tubing [92], a further investigation into alternative probe designs would likely be required to perform higher frequency tests. Additionally, clinical testing is required to evaluate 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 otoacoustic emission system using off-the-shelf earphones. While hearing loss is one of the more common disorders in low and middle-income countries, early detection can be challenging due to lack of access to affordable hearing screening tools. In comparison to commercial OAE devices which 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.

Chapter 4

Wireless earbuds for low-cost hearing screening

The World Health Organization estimates that 5.3% of the world's population suffers from disabling hearing loss and 80% of people who need hearing care live in low and middle-income countries [11; 125; 165]. Hearing loss is particularly harmful for neuro-development if it is left undetected in early childhood [5; 66; 84]. As a result, high-income countries have guidelines for universal infant hearing screening — the Joint Committee on Infant Hearing, the American Academy of Pediatrics and the Centers for Disease Control and Prevention all recommend universal hearing screening [5; 6; 10; 84] that is now implemented across almost all states, communities and hospitals in the United States [53; 136].

Since the neonatal population cannot provide behavioural response to conventional audiometry tests [46; 47; 48; 172], existing newborn hearing screening technologies instead use the sounds generated by a healthy cochlea called otoacoustic emissions (OAE) [55; 68]. While we think of the ear as a biological organ that receives sounds like a microphone, a healthy cochlea, the part of the inner ear responsible for converting sound waves into electronic impulses for the brain, also generates sounds. These emissions are created when the cochlea's sensory hair cells vibrate in respond to external sounds [55; 169; 170]. So, we could pick up these faint sounds and use their absence to detect hearing loss.

The challenge is that detecting these faint sounds emitted from the cochlea requires sensitive acoustic hardware and medical devices that are expensive (5000-8000 dollars) [2; 3]. As a result, there is limited to

Figure 4.1: OAEbuds in use with an infant. Our low-cost wireless earbud can perform hearing screening by detecting otoacoustic emissions (OAE) from the cochlea.

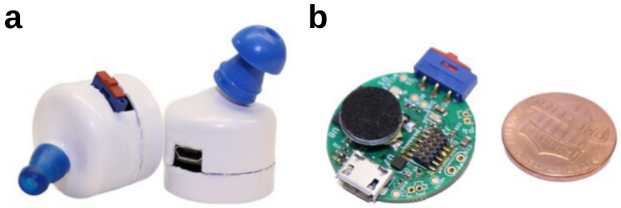

Figure 4.2: OAEbuds hardware. a, The 3D-printed enclosure with pediatric and adult earbud tips. **b,** OAEbud circuit board beside a penny for size comparison.

no-hearing screening in low and middle-income countries [91; 158]. Further, in rural and resource-limited settings, getting access to hearing assessment may often require travel to an urban setting and long wait times, significantly limiting the accessibility of hearing care [66; 125; 165].

We present OAEbuds, the first wireless earbud design for low-cost hearing screening. Our hardware-software system reliably detects otoacoustic emissions using low-cost acoustic hardware, while being in the form-factor of a wireless earbud. The earbud hardware is designed to work across a wide demographic from new-borns to adults.[1] Our design streams the digital acoustic data via Bluetooth to a nearby smartphone which is then used for processing the signals and displaying the test results.

There are two key technical challenges in achieving this design with low-cost acoustic components. First, since speaker hardware components are bulky, it is challenging to incorporate the two-speaker design in recent work [66] into the form-factor of a wireless earbud. Our experiments in §4.1 show that transmitting

[1] OAE testing is not limited to just newborns but is also used as part of clinical care in older kids and adults [1].

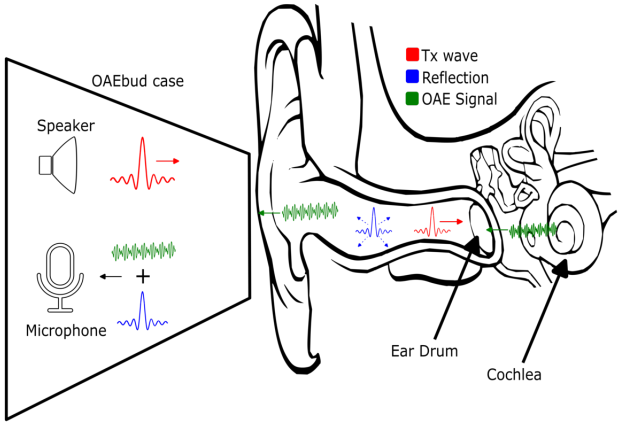

Figure 4.3: In-ear signal propagation. The OAEbud plays a broadband transmit (Tx) pulse to stimulate the cochlea to emit an OAE signal. The signal received by microphone is a superposition of 1) unwanted reflections from the ear canal, eardrum, and within the case and 2) the OAE signal.

the dual-tone signals used in [66] on a single low-cost speaker hardware introduces nonlinearities that in turn creates inter-modulation tones that can be confused for OAEs. Second, when an acoustic signal is sent into an ear canal, it first gets reflected and creates echoes not only inside the earbud case but also the ear drum and the walls of the ear canal, before arriving at the cochlea (see Fig. 4.3). To accurately identify OAEs, it is important to determine when the reflections and echoes of the input stimuli end and the OAEs begin.

We design a two-step protocol that uses wireless sensing techniques to address the above challenges using a single low-cost speaker in our wireless earbuds.

• **Reflection time estimation.** First, we send frequency modulated continuous wave (FMCW) signals as input stimuli. These signals get reflected back from the earbud case, the ear drum and the ear canal which are captured at the microphone. Since OAE signals are faint and non-linear, they do not create linear FMCW reflections. So we can perform FMCW processing to estimate the time-delays corresponding to the reflections and echoes and determine the duration after which their power reduces below a preset threshold (§4.1).

• **OAE signal extraction.** Second, we transmit a train of wideband pulses from the earbud speaker. Since the travelling sound wave traverses more slowly in the cochlea [29], the otoacoustic emissions still arrive delayed in time after the reflections and echoes of the input stimuli. To extract these signals, we first reduce reflections by only considering the signals that arrive after the duration estimated in the previous step. We

then synchronize the responses across multiple wideband pulses, combine them to improve the SNR of OAE signals and detect them using our earbud system (§4.1).

We designed an open-source wireless earbud hardware shown in Fig. 4.2 that is capable of transmitting the above signals from its speaker and wirelessly streaming the microphone audio. We designed our OAEbuds hardware using open source eCAD software, outsourced fabrication and assembly ($28.30 per unit), and 3D printed the enclosures in-house. The earbud is designed to support multiple ear tip sizes that allows it to snugly fit for both new-born infants and adults. The battery in the earbud can be recharged via a USB connection within 3 hours. Our evaluation shows that on a single charge, the earbud can be used to perform up to 91 tests.

We perform a clinical study on 50 ears from 26 pediatric and adult patients across two different healthcare sites. We perform testing with both our earbud device as well as an FDA-cleared medical device that performs OAE detection. For the tested patients, the attending clinicians determined the ground truth for hearing loss using the patient's hearing screen, audiograms, diagnostic auditory brain response, and clinical history. Our evaluation shows that OAEbuds achieves a sensitivity of 100% and specificity of 89.7% in screening for hearing loss. In comparison, the FDA-cleared medical device achieves a sensitivity of 83.3% and specificity of 92.1%. Our techniques also improve the area under the curve (AUC) from 0.847 to 0.950 over existing OAE algorithms. Finally, our system can output a 'pass' or 'refer' result for hearing screening in under 70 seconds.

Contributions. We make the following contributions.

- We design the first wireless earbuds to achieve low-cost hearing screening by detecting otoacoustic emissions.
- We introduce a two-step protocol that combines FMCW signals with wideband pulses to separate reflections and echoes from OAEs while using a single low-cost speaker.
- We perform a clinical study that shows our low-cost wireless earbud detecting hearing loss with accuracies similar to a $8000 FDA-cleared medical device.
- Finally, we make our code and hardware open source to help with adoption across the target settings.

Comparison to prior work. The closest to our work is our probe design described in Chapter 3 that uses

the two speakers in a wired earphone. It transmits a different frequency tone from each speaker and uses an additional microphone that is placed next to the ear to create a smartphone attachment. This work has multiple constraints that limit its adoption in the target use-cases. 1) It uses a wired earphone and external microphone that are connected to a smartphone. Since it uses the smartphone's ADC, DAC and AGC, it requires manual calibration for each smartphone model, which is challenging to generalize. 2) It uses two frequency tones that are transmitted from two different speakers and looks for inter-modulation between the tones to detect OAE. The challenge is that it is difficult to incorporate two speakers pointing into the ear-canal in a wireless earbud form factor and hence the techniques used in Chapter 3 cannot be used for wireless earbuds (see §4.1). 3) The various hardware components are attached using plastic tubing and glue which make it unreliable and difficult to scale and introduces a DIY aspect to the system. In an informal survey of clinicians in an African (anonymized) country conducted by the authors, participants noted that this DIY-aspect could translate to lowered patient confidence in the care received at the clinic. A low-cost yet high-tech device would be required to achieve wider adoption by clinicians. This chapter addresses the above limitations and designs the first wireless earbuds for low-cost hearing screening. Compared to DIY devices, since our wireless earbud is more integrated while being low-cost, it may help broaden the adoption of our hearing screening tool.

4.1 System design

We first describe existing approaches to OAE sensing and their limits. We then present our two-step protocol to estimate the reflection time and extract OAE signals. Finally, we present our low-cost earbud hardware.

Existing OAE approaches

The challenge with reliably detecting OAEs is identifying them in the presence of much stronger in-ear reflections. There are two key prior approaches.

• **DPOAEs.** Distortion-product otoacoustic emissions (DPOAEs) address the reflection problem by using intermodulation. In particular, the cochlea is stimulated by sending two tones f_1 and f_2. Given the nonlinear response of the basilar membrane within the cochlea, it generates a nonlinear intermodulation tone at the

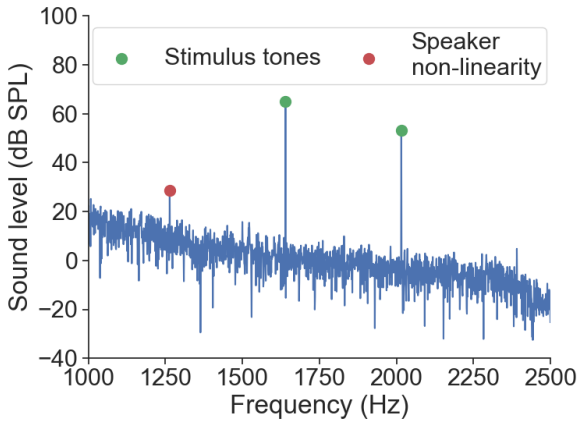

Figure 4.4: Challenge of existing OAE approaches on a single-speaker earbud design. Sending two stimulus tones f_1 and f_2 through a single speaker setup to elicit OAEs creates a hardware non-linearity at $2f_1 - f_2$, which can be stronger than the OAE signal at that frequency.

frequency $2f_1 - f_2$ [66]. Since reflections and echoes do not cause new frequency tones, this dual-tone approach can be used to separate OAEs from in-ear reflections. The challenge with deploying the DPOAE protocol on a single-speaker system is that hardware components introduce nonlinear intermodulation distortion at frequencies which are linear combinations of f_1 and f_2, $k_1 f_1 + k_2 f_2$, where k_1 and k_2 are arbitrary integers. When $k_1 = 2$ and $k_2 = -1$, this matches the DPOAE signals produced by the cochlea. These non-linearities are more prominent for low-cost speaker hardware. Fig. 4.4 shows the amplitude of the intermodulation produced by a single low-cost speaker (Knowles SR-32453-000, $4.42) when sending two tones 1640 and 2016 Hz at 65/55 dB SPL. The figure shows that the unwanted intermodulation component has a sound level of 28 dB SPL; in comparison the typical range of DPOAEs is 5–25 dB SPL [66]. As a result, prior work [66] uses a two-speaker system to separately send the f_1 and f_2 tones. Since wireless earbuds are generally constrained to only a single speaker per bud due to size constraints, it is challenging to use the DPOAE protocol on such a low-cost hardware.

• **TEOAEs.** The transient-evoked otoacoustic emission (TEOAE) protocol extracts the OAEs in the presence of in-ear reflections and echoes using a single speaker. Here a short biphasic click sequence is repeatedly sent, with a polarity and amplitude pattern of $\{1, 1, 1, -3\}$ [106] . The key insight behind this protocol is that the amplitude of the reflections are linearly related to the amplitude of the transmitted clicks. So the responses of all four clicks can be summed to cancel the reflections caused by the eardrum. However

since OAEs are non-linear in nature, they would not be canceled by this addition operation. The challenge with deploying this protocol on a low-cost system, is that 1) the clicks need to be perfectly synchronized and phase-aligned, so that the reflections are cancelled out. Without exact alignment, there will continue to be residual energy caused by imperfect cancellation which will make it difficult to measure the OAEs. 2) Low-cost speaker hardware also introduces non-linearities in polarity and amplitude resulting in imperfect cancellation. Our evaluation in §4.3 shows that this leads to degraded performance.

While other methods for eliciting OAEs have been proposed in the literature [75; 131] they are not used in practice given uncertainty about their reliability.

Our two-stage protocol

Instead of relying on the linearity of the acoustic hardware, we separate the in-ear reflections from OAEs in the time domain. At a high level, we first estimate the time delay at which reflections from both the ear and the enclosure drop below a particular threshold. We then detect the OAEs over the remaining time duration.

Reflection time estimation

The time delay when reflections diminish will differ from one ear to another due to differences in anatomical structures such as ear canal diameter which increases with age. Further, hair and debris can change the reflection profile significantly. At a high level, we send an FMCW signal to estimate when reflections from the case and other parts of the ear diminish beyond a certain threshold. Our algorithm then uses the remainder of the recording after this time estimate to measure OAEs.

Although FMCW signals could be used to estimate the length of an individual ear canal, and convert that to a time delay at which reflections diminish, we find that in practice there is a significant amount of echos caused by reflections from the case, ear drum and ear canal that result in a large delay spread, much larger than the time of flight measurement for a typical ear canal length of 2.5 cm. An analogy to this would be that if one shouts in an empty cave, it can take several seconds for all the echos to diminish due to the significant reflections that occur from the cave walls. Further we note that the speed of sound is slower in the cochlea, meaning the OAE would take a longer time to return compared to if the pulse were only sent

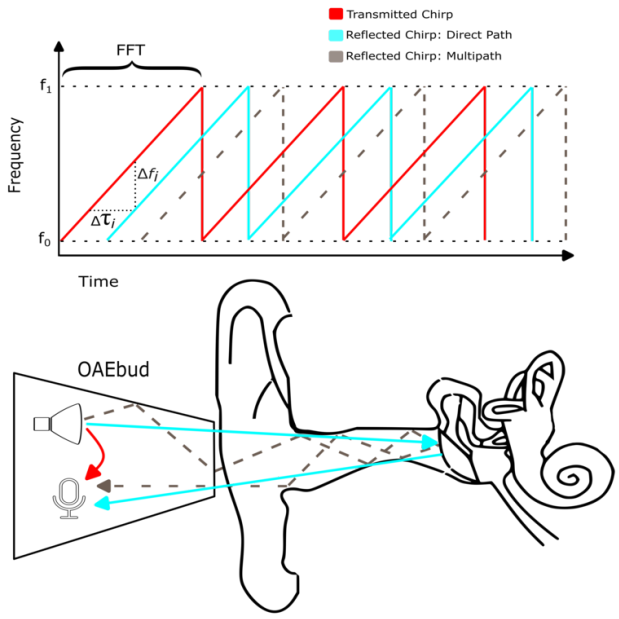

Figure 4.5: FMCW processing to calculate the time of arrival for reflections from the ear canal. The OAEbud transmits a chirp into the ear canal, and record the reflections from the ear and enclosure. It then performs an FFT over the chirp duration to estimate the frequency shift Δf_i and time delay τ_i for the i^{th} reflection from the ear. This estimate is averaged across three chirps.

into the air medium, and this contributes further to a large delay spread [29].

We send a chirp with linearly increasing frequency from f_0 to f_1 where the frequency at a given time t is denoted as $f(t) = f_0 + \frac{Bt}{T}$, where B and T are the bandwidth and duration of the chirp. The phase is computed by integrating $f(t)$ over time, resulting in the function: $\phi(t) = 2\pi(f_0 t + B\frac{t^2}{2T})$. The signal that is then transmitted in the time domain is defined as $x(t) = cos(\phi(t))$, as shown in Fig. 4.5.

Each of the echoes from the ear canal are delayed chirps that arrive at the microphone as the received signal, $y(t)$, that is a combination of all echoes. To estimate the multipath profile of the ear canal, we multiply the receiver signal with the transmitted signal, $x(t)y(t)$. Using the trigonometric identify, $2cos(A)cos(B) = cos(A-B)+cos(A+B)$ and filtering out the high frequency term, $cos(A+B)$, we can translate the time delays of each of the echoes into frequency shifts between the transmitted and received chirps. Note that in contrast to radio signals that have both I and Q components, acoustic signals operate in the real space. So instead of using a downchirp, we multiply the received cosine signal with the transmitted signal and apply a low-pass filter.

In Fig. 4.5, we show a plot of the transmitted FMCW signal, along with several reflections in the fre-

quency domain, each with its own time delay τ_i for the i^{th} reflection. In order to determine individual time delays τ_i when the reflections end, we compare the differences in frequencies between the transmitted and reflected signals. Specifically, a time delay τ_i will result in a frequency shift of Δf_i for the reflected signal from the transmitted signal and can be computed as follows:

$$\tau_i = \frac{\Delta f_i T}{f_1 - f_0}$$

To obtain a precise resolution for our reflection time estimate, we send an FMCW signal that is close to the maximum bandwidth allowable by the sampling rate of our system. As the maximum sampling rate of our OAEbuds hardware is 31250 Hz, we send an FMCW signal with a bandwidth of 5 to 15 kHz so that the upper frequency is close to the Nyquist frequency of 15625 Hz. The time resolution of an FMCW system is, $\frac{1}{2B}$, where B is the bandwidth. This corresponds to a time resolution of 0.05 ms when B=10 kHz. We set the length of our signal based on the maximum number of samples that can be stored in our hardware's memory. In our system we use a 200 ms FMCW signal which corresponds to 6250 samples.

Fig. 4.6 shows the result of this processing in a normal adult ear. We can observe a peak in the zeroth bin that corresponds to the incident chirp, and peaks at subsequent bins that correspond to reflections arriving at increasing time delays. To minimize the interfering effects of reflections in our OAE measurement, we select a time delay that corresponds to the frequency shift where the power level diminishes below a preset power threshold. In our implementation, if the power level of a frequency bin decreases below 55 dB from the power of the incident signal, we use the time delay, t_D, corresponding to that frequency bin. If such a bin cannot be found, a default delay value of 12 ms is used.

We note two key points about our earbud system.

• Performing this estimation is important particularly given the tonotopic geometry of the cochlea (Fig. 4.3) where the high frequency OAEs exit the cochlea first, followed by the low frequencies. The cochlea has a coiled shape where the beginning of the coil responds to high frequency sounds, while the inner most curled part of the coil responds to low frequency sounds. As such, when a stimulus pulse is sent into the cochlea, it is the high frequency OAEs that exit first, and it is these frequencies that would also be most affected by the reflections. By setting a time delay that is too low, there is a risk that the reflections will be confused for the OAEs, whereas setting the delay too high may only result in a measurement over the low frequency OAEs,

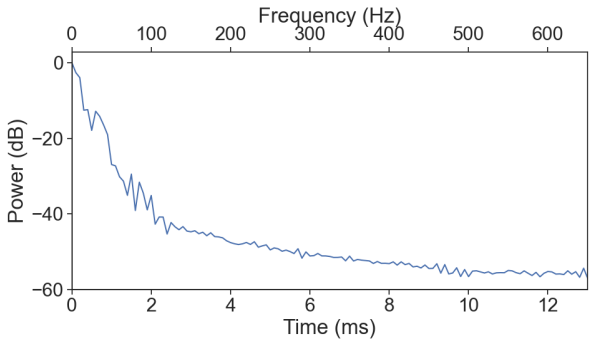

Figure 4.6: Estimating when reflections diminish. By measuring the frequency shifts from reflections of a FMCW signal transmitted into the ear canal, we can estimate the time delay t_D after which reflections diminish to a predetermined power threshold.

but few of the high frequency OAEs. Our algorithm allows us to minimize the power of reflections while increasing the power of the OAE signals.

• Interestingly, the ear canal creates a closed enclosure that can create a large number of strong echoes. As a result, while the length of the ear canal is only around 2.5 cm [176], which translates to an acoustic round trip time of 0.15 ms, given the large number of echoes created within the ear canal, we can have reflections as shown in Fig. 4.6 that arrive even at 5-10 ms. This emphasizes the need for using a system that computes the time-delays for the in-ear reflections which can be much longer than the time it takes to traverse the ear canal. We also note that we compute the above time delay by averaging the values across three continuous FMCW chirps.

OAE signal extraction

After the time delay has been identified, our system needs to reliably measure the faint otoacoustic emissions that are as low as -10 to 30 dB SPL using low-cost microphones that would not have the same sensitivity of the high-end expensive microphones in medical devices. To extract these faint OAE signals, at a high level, we combine the OAE responses across multiple pulses to increase the SNR of OAEs. This can be challenging especially given that the target population of this test is young infants who may move, fidget or otherwise cause noise throughout the measurement. Our measurement should also be able to reliably distinguish between periods of noise caused by the patient and legitimate OAE signals, as an incorrect classification

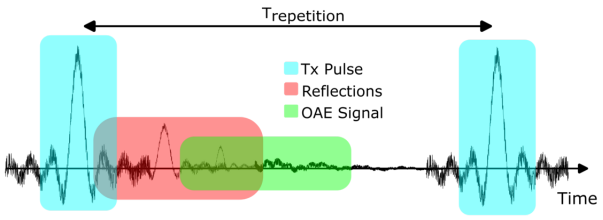

Figure 4.7: OAEbuds pulse transmission scheme. A pulse of 500 μs with a bandwidth from 0–5 kHz is transmitted every 20 ms into the ear to cover the range of frequencies important for hearing screening. The recorded signal consists of reflections of the input stimuli from the ear canal which overlap with the OAE signal.

would result in an inaccurate measurement or an overly lengthy measurement that would result in patient discomfort.

In the rest of this section, we first describe our transmission scheme and then describe the various steps needed to extract the OAE signals.

Pulse transmission scheme. We transmit a sequence of short 500 μs pulses and apply a brick-wall filter with a bandwidth from $f_{start} = 0\ kHz$ to $f_{end} = 5\ kHz$, to cover the full range of frequencies of clinical interest in hearing screening (Fig. 4.7). We use a sampling rate of 15625 Hz, which corresponds to 8 samples to represent the pulse. We multiply the pulse with a hamming window to reduce the effect of ringing. Each of the pulses are separated by a gap of 20 ms, to ensure that the OAEs have enough time to arrive at the microphone. We perform these measurements over the course of 67.5 seconds, which corresponds to approximately 3300 pulses.

Decoding algorithms. We describe the various steps to combine the OAE responses across pulses and extract higher SNR OAE signals.

Step 1. Pulse synchronization. The first step of our algorithm is to establish synchronization with the start of the pulse. To do this, we perform cross correlation of the first one second window with the transmitted pulse, and look for peaks with a minimum peak prominence set to 0.3e8. If we do not find such a peak in this window, we proceed to the next one second window. We note that we may not find peaks within the first one second window if the user initially begins the measurement outside the ear where the amplitude of the reflections will be lower compared to in the ear. Once the start of a pulse can be found, we can add a fixed offset of 20.5 ms to find the start of the next pulse. We also use this windowed approach instead

of performing cross correlation over the entire one minute recording as it allows for real-time computation of the OAEs over the course of the measurement. In other words, our algorithm is able to incrementally compute the OAE result over windows of one second intervals, and provide continuous feedback to the user about the OAE results, and whether the environment is too noisy. This will allow the user to be able to recognize possible problems in probe fit or environmental noise in real-time instead of having to wait until the entire one minute measurement is complete.

Step 2: Noise detection. After we have a set of peaks corresponding to the start of all the pulses in a 1-second window, our next step is to determine which pulses are usable for subsequent processing and identify pulses that have been affected by noise in the environment. To do this, we apply a sliding correlation window over batches of four pulses and calculate the correlation between each of the adjacent pulses. We sum the calculated correlation values and divide it by the sum of the received power within that batch. This produces a normalized value that is invariant to pulse amplitude differences across batches. If this normalized value is above 0.95, we consider that batch to be usable for subsequent measurement, else we discard that batch. This allows us to discard specific OAE signals within the overall measurement that have been corrupted by noise.

Step 3: Combining OAE responses across pulses. For all usable batches in a given window, we look for OAE responses using the time delay t_D computed in the previous section. We use the window size of $t_P - t_D - t_{guard}$ where t_P is the gap between consecutive pulses and t_{guard} is a guard period which we set to 1 ms. In other words, we look for OAEs starting from the time when the reflections have diminished up till the start of the next pulse, minus a small guard period. We then average the power of all odd numbered pulses to compute, P_{odd}^{OAE} and all even numbered pulses for P_{even}^{OAE}. We then compute the signal and noise power, P_{signal} and P_{noise} in the time domain as follows:

$$P_{signal} = \frac{P_{odd}^{OAE} + P_{even}^{OAE}}{2}, P_{noise} = \frac{|P_{odd}^{OAE} - P_{even}^{OAE}|}{2}$$

To obtain the SNRs across 1 to 5 kHz, we convert the above signals to the frequency domain by performing an FFT. We repeat the above process for each frequency band by taking an average across adjacent frequency bins. Specifically we perform an average over the following bands: 750–1250 Hz, 1250–1750 Hz, 1750–2500 Hz, 2500–3500 Hz, and 3500–4500 Hz. The result of this step is a set of SNR measurements for each

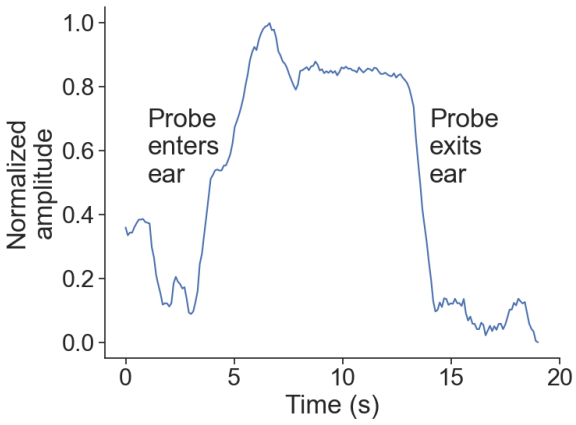

Figure 4.8: Checking if the probe is in the ear. By measuring the amplitude of a chirp at the 200 Hz frequency during the beginning of a measurement, we can detect whether the ear probe is inside or outside the ear.

frequency band.

Hearing testing algorithms. Finally, we describe the algorithms we run while performing the hearing test.

a) Computing the pass/refer result. To calculate the 'pass' or 'refer' screening result, as with existing medical devices, we determine if at least 2/3 of the 5 frequency bands are above a preset threshold (8 dB in our case). These parameters were determined by performing a parameter sweep over different number of frequency bands and SNR thresholds to determine values that optimize our clinical performance. Further, we check that the absolute sound level of the signal component is above a preset value of -10 dB SPL, which is typically regarded as the minimum sound level of an OAE. This ensures that spurious reflections or noise that were not discarded during previous filtering steps are not mistaken as OAEs. Additionally, if the average noise level across the frequency bands exceeds 6 dB SPL, we mark the measurement as noisy.

b) Determining if probe is in the ear. To determine when the measurement can begin, our system performs a check for whether the probe is probably placed in the ear. To do this, we send a sequence of 20 ms chirps from 100 to 5500 Hz and measure the amplitude of the frequency response to determine if the probe has formed a snug fit with the ear. We find that the frequency at 200 Hz is representative of whether the probe is outside or inside the ear (Fig. 4.8). If the average sound level in this frequency range exceeds a predefined threshold for 50 chirps (1 s), we mark the probe tip as being in the ear and begin the measurement.

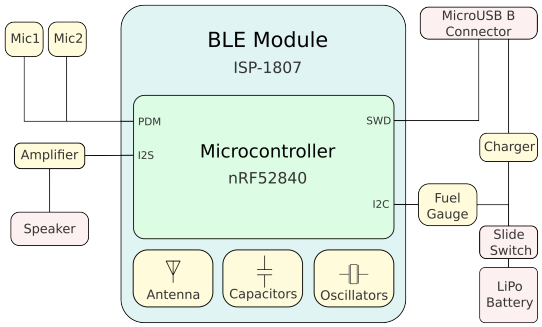

Figure 4.9: OAEbuds hardware design. We include an additional microphone for future research.

Hardware design

We design a custom hardware solution based on the ISP1807 Bluetooth Low Energy (BLE) module, which combines a Nordic nRF52840 microcontroller with a variety of other components such as capacitors, oscillators and an antenna. The device is equipped with a pair of pulse-density modulated (PDM) microphones (TDK Invensense T3903) and a speaker (PUI Audio AS01008MR-3) driven by a digital pulse-code modulation (PCM) input Class D amplifier (Maxim Integrated MAX98357A). The system is powered by a 3.7 V, 100 mAh Lithium Polymer Battery, and a buck converter (Texas Instruments LM3671) is used to bring the system voltage down to 3.3V. A Micro-USB connector is used to program the device over SWD and charge the battery via a charger IC (Analog Devices LTC4124). Battery information, such as cell voltage and state of charge (SOC), is probed using a fuel gauge (Maxim Integrated MAX17048). A high level overview of the system is shown in Fig. 4.9.

The speaker amplifier is interfaced using the controller's Inter-IC Sound (I2S) module. The device is preloaded with a fixed array in the controller's memory that holds the PCM representation of a signal (e.g., pulse). The preloaded waveform must be generated at a sampling frequency matching that of the I2S module clock signals. The sampling frequency is set to 15.625 kHz, the smallest frequency compatible with our amplifier that is larger than twice the pulse bandwidth. When the device starts emitting pulses, the controller supplies a copy of the signal waveform to the I2S module, which transfers the waveform to the amplifier using its direct memory access (DMA). The I2S module is internally double buffered, and it triggers a callback for the controller to supply a fresh audio buffer to output once a previous transfer finishes.

The microphones are interfaced using the controller's PDM module. Specifically, the module's DMA is used to asynchronously convert microphone PDM measurements to PCM values and load the results into memory. The two PDM microphones are connected to the same serial clock line, running at 1 MHz, which the PDM module internally decimates by a factor of 64. This yields an overall sampling frequency of 15.625 kHz. The recording process also uses double buffering and produces 312 channel-interleaved pairs of 16-bit samples (equivalent to 20 ms at 15.625 kHz) at a time. The captured samples are then divided into packets and transmitted over Bluetooth. To maximize throughput, each packet contains 80 two-channel samples (240 bytes total), which, when including the sequence number and overhead bytes, is the largest number of samples we can transmit in a single packet. Additionally, we also use the maximum possible data rate of 2 Mbps. In our design, the earbuds stream the recorded acoustic signals via Bluetooth to a nearby smartphone and the computation to extract the OAEs is performed on the smartphone.

The circuit schematic and PCB layout for the device was designed using KiCAD and was fabricated and assembled by PCBWay. The enclosure was designed in Fusion360 and 3D-printed using a Formlabs Form 3 resin printer. Our enclosure is designed to have a tip diameter and length of 5.4 mm and 7.1 mm respectively, which allows the rubber ear tips to have a snug fit with the enclosure. We note that for all the ear tips in our study, the base diameter is the same, and is able to fit easily on the enclosure. The enclosure also has openings for a switch to power on and off the device, as well as for a micro-USB charging port. The interior of the case is also lined with foam to reduce the effect of acoustic reflections from within the case. Table. 4.1 shows the cost of the individual components in our earbud device estimated using Digikey, Mouser, Alibaba and PCBWay. The above numbers provide a ballpark cost which can be further reduced at higher volumes.

4.2 Clinical study

Our study was approved by the Institutional Review Board and we obtained informed consent for all adults and parental consent was obtained for pediatric patients and patients aged 7 to 17 provided verbal or written assent. We recruited patients from otolaryngology, craniofacial and hearing loss clinics across two clinical sites. We also recruited adults without any known concern for hearing loss ($n = 28$ ears). We tested our

Component	Cost (USD)
BLE Module	10.67
Microphones	2×0.80
Speaker	1.06
Amplifier	1.62
Charger	5.12
Fuel Gauge	1.89
MicroUSB Connector	0.29
Switch	1.52
Battery	1.50
PCB Fabrication & Assembly	1.63
3D Printed Case	1.40
Total	28.30

Table 4.1: Itemized hardware cost. Component prices are estimated for a production lot of 1,000 devices.

device on 50 ears from 26 pediatric and adults patients up to 32 years (mean age: 18 ± 9). Of the 26 total patients, 12 were between the ages of 2 and 17. The remaining 14 participants were between the ages of 22 and 32. Measurements on adult patients were performed in duplicates whenever possible, and a total of 75 measurements are used for subsequent analysis. The female-to-male ratio was 3.2.

To determine the ground truth hearing status of each patient, the best available clinical information was interpreted by the attending physician or clinician. This information includes the patient's newborn hearing screen, audiogram, diagnostic auditory brain response, and clinical and examination history. Our patient population included sensorineural ($n = 5$) and conductive ($n = 1$) hearing loss, as well as auditory neuropathy ($n = 1$) which is a form of hearing loss that affects the auditory nerve's ability to transmit sound to the brain, but which does not affect the cochlea's ability to produce OAEs. We recruited patients with different degrees of hearing loss spanning the full range of degrees from slight to profound (Fig. 4.10(a)). The degree of hearing loss for a given ear is computed by taking the mean hearing level measured from a patient's audiogram in dB HL, and mapping it to the thresholds as defined by the American Speech-Language-Hearing Association [42]. Our dataset also had ears with middle ear infections due to fluid buildup ($n = 3$), as well as ears that recently had ear tubes ($n = 2$). In total, 6 ears were classified as having hearing loss, the remaining 44 ears were classified as having normal hearing or having healthy outer hair cells in the cochlea. For our study, we mark the patient with auditory neuropathy as having healthy outer hair cells in the cochlea, as OAEs are expected in this patient.

During testing, all participants > 6 months were instructed to sit upright for the test. Younger patients were tested in the position that was most comfortable for them and their parents, and included being asleep

in a supine position, or being cradled over the parent's shoulder. All patients were first tested with the commercial OAE device in each ear. Patients recruited from one of the sites were tested with a commercial TEOAE device (Otoport Screener, Otodynamics) that was used regularly at the clinic. This test was performed across the 1, 1.5, 2, 3, 4 kHz bands. We set the device to continue measuring for this full duration even if the test passed early. The remaining patients were measured using a DPOAE device (OAE Hearing Screener, Welch Allyn) that was available for us to use at other locations. This device tested at frequency bands of 2, 3, 4, and 5 kHz. During this portion of the test, we would select a disposable rubber ear tip (Grason & Associates LLC) size based on visual examination of the patient's ear canal. In our study, ear tip sizes 8, 9, 10, 11 an 12 were used 9, 4, 23, 8 and 2 times respectively. For the pediatric population, we used three different ear tip sizes from 8 to 10 mm, while for the adult population four different ear tip sizes from 8 to 12 mm. This suggests that a relatively small number of ear tip sizes can accommodate a large range of ear canal sizes. We note that commercial earbuds such as AirPods contain four different ear tip sizes [40].

After completing the test with the commercial device, we proceeded to test the patient with our wireless earbud device using the same rubber ear tip. For pediatric patients, each ear was tested effectively for 45 to 68 seconds per ear, depending on the compliance of each patient. Adult participants were tested for 68 seconds in each ear, twice. All testing with both children and adults was performed by two computer science graduate students. During testing, we transmitted clicks with a duration of 500 μs and gaps of 20 ms between clicks. The clicks were set to have a bandwidth of 0 to 5 kHz. The clicks were sent at a sound level of 84 dB peSPL (pe = peak-equivalent).

The sampling rate of the speaker and microphone was set to 15625 Hz to allow for streaming the data over Bluetooth. We measure for OAEs at the 1, 1.5, 2, 3, and 4 kHz bands. Specifically, we average the signal and noise responses in the ranges of 750–1250 Hz, 1250–1750 Hz, 1750–2500 Hz, 2500–3500 Hz, and 3500–4500 Hz. On our device, we consider a measurement a passing screen if the SNR of at least two frequency bands exceeds an SNR threshold of 8 dB, and the absolute sound level of those passing frequency bands is greater than -10 dB SPL, which prior work regarded as the minimum power for these OAEs [41].

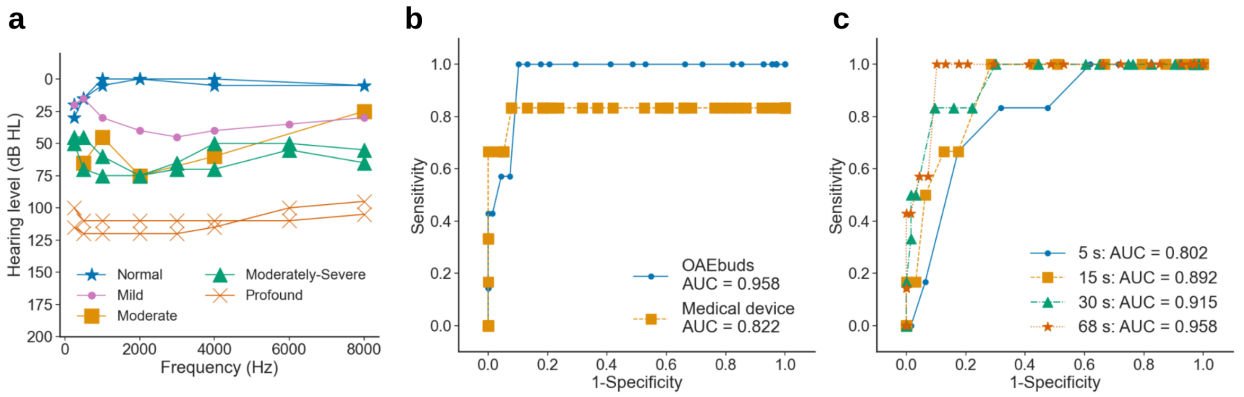

Figure 4.10: Clinical study performance. **a,** Audiograms for ears tested in clinical study with normal hearing and different degrees of hearing loss. **b,** Performance of OAEbuds in comparison with commercial OAE medical device. **c,** Effect of measurement time on clinical performance.

Performance evaluation

To determine our SNR threshold on each frequency band, we generate a receiver-operating curve (Fig. 4.10) to compute the sensitivity and specificity values for SNR thresholds ranging from -20 to 40 dB in increments of 1 dB. We find that the SNR threshold of 8 dB maximizes the sum of sensitivity and specificity, yielding a sensitivity of 100.0% (95% CI, 64.6–100.0%) and specificity of 89.7% (95% CI, 80.2–94.9%). We find that using two frequency bands as the pass criteria yields an AUC of 0.958. Using three or four frequency bands as the pass criteria results in AUCs of 0.950 and 0.884 respectively.

Comparison with medical device. In comparison to our earbuds, Fig. 4.10(b) shows that the medical device had a lower AUC of 0.822 yielding a sensitivity of 83.3% (95% CI, 43.6–97.0%) and specificity of 92.1% (95% CI, 79.2–97.3%). The ground truth for both our device and the medical device is the clinical information that is interpreted by the physician including their clinical and examination history as well as auditory brain response tests. We note that for one ear with hearing loss, the medical device was unable to pass the probe check despite numerous attempts to fit the ear with different sized ear tips, even though the ear tip appeared to fit well visually. For this instance, we marked the medical device as having failed the measurement.

Our device misclassified the hearing loss status for seven ears, three of these ears had middle ear fluid and infection. In these ears, OAEs were not detected, as the fluid acts as a barrier that blocks the OAEs from

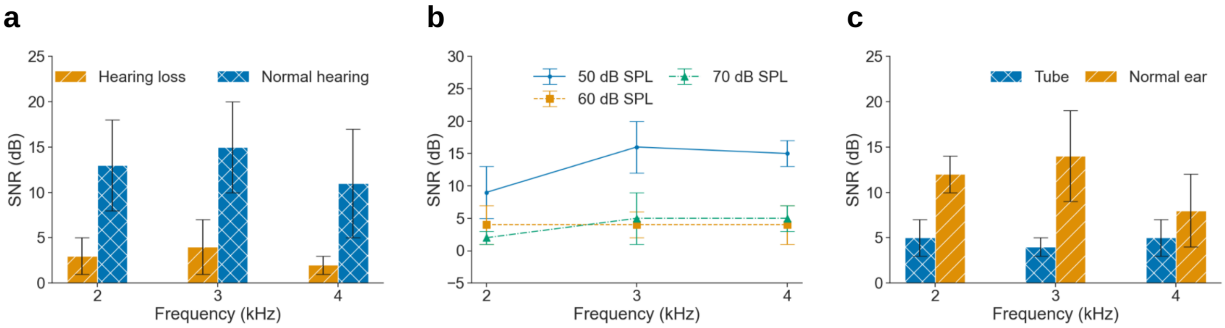

Figure 4.11: Subgroup analysis and benchmark results. a, Subgroup analysis comparing the mean SNR of OAEs measured in patients with hearing loss and normal hearing during the clinical study. **b,** Effect of background noise on system performance for different sound levels. **c,** Probe integrity check in a close-ended tube is used to ensure that the system produces SNRs below the cutoff for healthy hearing when measured outside the ear.

reaching the outer ear, and it is expected that OAEs do not appear in these ears [60]. The commercial device similarly did not detect OAEs in ears with middle ear fluid or infection. One of these ears had hearing loss, but OAEs were detected. We suspect that this is due to reflections or ear tip fit issues, as this was the only ear where the commercial device could not begin a test due to a failure of the initial probe check despite repeated attempts to pass the check. One of these ears recently had ear tubes removed which would have resulted in a hole in the eardrum which would have begun healing, and may have affected the ability to detect OAEs.

Effect of measurement time. To determine the effect of measurement time, we set a limit on the maximum number of clicks used by our algorithm and compute clinical performance when the measurement time is reduced to 15 or 30 seconds. Fig. 4.10(c) shows that we are able to achieve an AUC of 0.892 and 0.915, which is higher than that achieved by the commercial device of 0.822. At these measurement times, our system obtained an optimal sensitivity of 100.0% (95% CI, 61.0–100.0%) and 83.3% (95% CI, 43.6–97.0%) respectively and specificity of 71.4% (95% CI, 59.3–81.1%) and 90.5% (95% CI, 80.7–95.6%) respectively. When reducing the measurement duration to 5 seconds, our sensitivity and specificity, as expected, reduces to 83.3% (95% CI, 43.6–97.0%) and 68.3% (95% CI, 56.0–78.4%) respectively.

Subgroup analysis of hearing status. In Fig. 4.11(a), we show the average SNR obtained in ears with hearing loss as well as with normal hearing. The plot shows that for the ears with hearing loss the average SNR across all frequencies is 3 dB, while it is 11 dB for ears with normal hearing, showing that there is

large separation in SNRs between the two classes of ears. We note that in the hearing loss ears, although the average SNRs are positive, none of the SNRs at any of the frequencies exceeded the 8 dB SNR cutoff. We suspect the SNRs are positive due to residual reflections from within the plastic case of our wireless earbuds. We also note that the SNRs for the OAEs detected by our system are smaller at the lower frequencies. This is in line with existing literature [99] which confirms that transient-evoked OAEs are better at mid-range frequencies than the lower frequencies.

Test-retest evaluation. For the 25 ears where duplicate testing was performed, the screening result for our earbuds matched in all but one ear. In our study, we also tested the commercial device several times if we were not confident in the probe fit in the ear during a given measurement. There were two ears where the commercial device had differing results between tests. Both of these ears were normal hearing ears and it took two and three attempts respectively for these ears in order to obtain a passing screen result.

4.3 Micro-benchmarks

We provide micro-benchmark evaluations including the effect of background noise and a comparison of our wireless sensing techniques with existing OAE algorithms. We also present an evaluation of system level issues including power and run-time analyses.

Effect of background noise. To measure the effect of background noise on our device, we played noise of road traffic from a laptop such that the sound level at a healthy ear varied from 50 to 70 dB SPL and measured the OAEs measured by our earbuds in an ear with normal hearing. These sound levels reflect the typical range of ambient background noise that would occur in a clinical testing facility that is not well insulated from noise, and which might be situated close to a road. Fig. 4.11(b) shows that at a noise of 50 dB SPL, the OAEs can be detected in the ear, and are above the 8 dB threshold at all frequencies. This sound level is the typical ambient noise level in an urban residence [108]. At 60 and 70 dB SPL, the SNRs at all frequencies drop below the 8 dB threshold. These sound levels are similar to conversational speech held at 1 m and a vacuum cleaner at 1 m [49]. These results suggest that testing should be performed in a relatively quiet environment to obtain reliable results.

Probe integrity check. Medical OAE devices use a closed-ended tube between 0.5–2 cc in volume as a

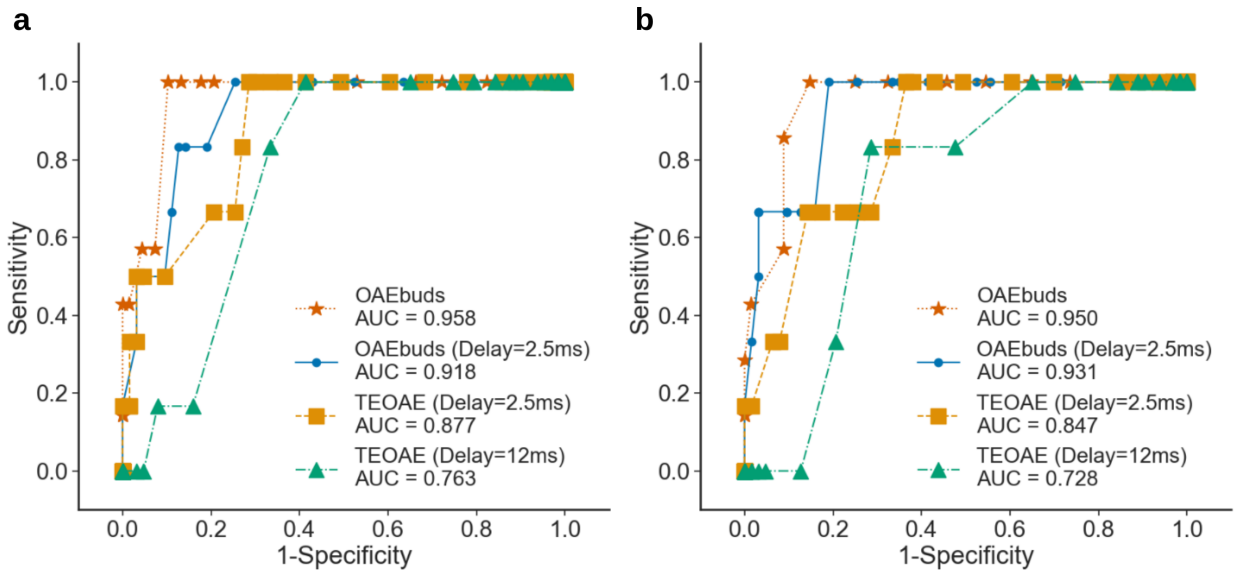

Figure 4.12: Comparison of OAEbuds with prior TEOAE algorithms. Our OAEbuds system achieves better performance compared to prior TEOAE algorithms when **a,** 2 and **b,** 3 frequency bands are required to be above the SNR threshold for the hearing test to pass.

probe integrity check [41]. This range of volumes is selected to mimic the volume of the ear canal for the pediatric and adult population. Similarly, we have our earbuds perform a measurement in a closed ended plastic tube with a volume of 1 cc, and in a healthy ear. Fig. 4.11(c) shows the SNRs obtained in both these scenarios when repeated three times. The plot shows that the SNRs in the tube are below the SNR cutoff of 8 dB for all frequencies across all measurements, and can be used as a probe integrity check to ensure that the device is not incorrectly identifying OAEs.

Comparison with prior OAE algorithms. We evaluate the performance of our OAEbuds system using alternative variations to implementation. We test three different protocols as follows 1) OAEbuds protocol with a fixed delay of 2.5 ms where reflections after a 2.5 ms duration from the pulse are removed, 2) conventional TEOAE protocol with a delay of 2.5 ms and 3) conventional TEOAE protocol with a higher delay of 12 ms. Conventional TEOAE systems transmit a train of pulses with a polarity pattern of $\{1, 1, 1, -3\}$ [39]. The receiver then adds up the response across the four pulses to generate the OAE signals. We select 2.5 ms as the delay as this is what commercial TEOAE devices typically use [50]. Fig. 4.12 shows the ROC curves for these protocol implementations when using either 2 or 3 frequency bands to pass. We note that our implementation of OAEbuds yields a better AUC compared to the conventional protocol regardless of whether

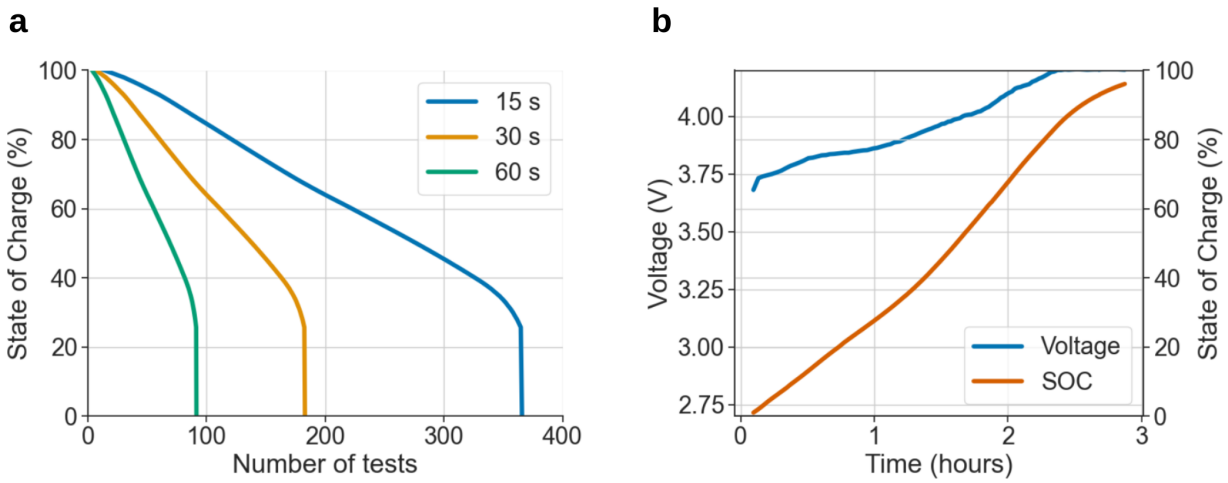

Figure 4.13: Power analysis. **a,** Number of tests that can be performed on a single charge for tests of different durations. **b,** Time required to charge an OAEbud via micro-USB.

2 or 3 frequency bands are needed to pass. We note that when using a delay of 2.5 ms, a significant amount of the signal will be reflections and not OAEs. Because of this, the optimal SNR threshold for these modified protocol is significantly higher at 51 and 48 dB for the 2 and 3 frequency band scenarios respectively. The most likely reason why the TEOAE protocol performs worse than OAEbuds is that it relies on the clicks canceling out each other well. In practice, the cancellation is not perfect and there will be a residual error in the cancellation process, in particular on our low-cost acoustic hardware that has non-linearities. Although the goal of this protocol is to cancel out reflections which are linear, it will also cancel out the linear components of the OAEs themselves, which may also contribute to lowered performance. We note that with TEOAE, using the lower delay of 2.5 ms performs better than the 12 ms delay. With this lower delay, the optimal SNR threshold is 21–22 dB, while it is 4–6 dB when running the TEOAE protocol with a 12 ms delay.

Power analysis. To evaluate how long our earbuds last on a single charge, we first charge the battery of the earbuds to its maximum level and evaluate its performance over multiple tests. Fig. 4.13(a) shows the state of charge on the battery as a function of the number of tests. Each of the lines represents a different duration for a hearing test. To measure the voltages and state of charge, we use an on-device fuel gauge. The fuel gauge uses the proprietary ModelGauge algorithm to continuously track the battery's state of charge (SOC). It simulates the internal nonlinear dynamics of the battery model. By analyzing voltage measurements over

time, it can determine the state of charge much more accurately. The fuel gauge data is interfaced over an I2C bus and it allows us to read out the battery voltage and SOC.

The plots show that on a single charge, even when each test lasts around 60 s, the earbud can support 91 hearing tests. For context, we note that on a typical day, the hearing loss clinic in our institution sees around 20-30 patients. We measure how long it takes to charge the earbuds through a micro-USB cable connected to a wall outlet. Fig. 4.13(b) shows that it takes around 3 hours to charge the earbud which in a practical setting can be done overnight in a clinic.

Runtime analysis. To evaluate the feasibility of running our algorithm on a mobile device like a smartphone, we convert our algorithm to C++ code that can run on the Android smartphone platform and time how long it takes to execute the operations in our system. On a Samsung Galaxy S9, we are able to process windows of 1 s containing 48 clicks in less than a millisecond. This means that we are able to provide a real-time update of the SNR values throughout a OAE measurement. This real-time feedback can help a user determine if a probe fit is snug or if there is too much noise at different points in a measurement, and allows them to make any changes to how the measurement is being performed.

4.4 Related work

Prior work can be broadly divided into three classes.

Health tracking using earphones. Prior work has explored the use of earphones for monitoring physiological signals for cardiovascular sensing [86; 140], blood pressure measurements [64] and respiration [148]. Earphones have also been used for sensing jaw clenching [150], teeth motion and voice detection [110; 141]. Prior work has also explored the use of smartphone attachments for diagnosing middle ear conditions. [73] designed a paper cone that is attached to a smartphone and used a machine learning classifier to detect middle ear fluid behind the ear drum. [102] used a smartphone-connected wired earphone attached to an external microphone to differentiate between middle ear fluid, ear drum ruptures and wax blockage. [72] presented a smartphone attachment to perform tympanometry where the pressure within the ear canal is changed to assess the mobility of the ear drum. All these wired systems however are designed for assessing the state of the middle ear and the ear drum. Hearing screening, in contrast, is primarily focused on the state of the

cochlea. The cochlea is part of the inner-ear that is behind the ear drum and is primarily responsible for converting sound waves into electrical impulses which are then interpreted by the brain.

OAE devices. Prior work [95; 97] created a smartphone interface for the probes from an existing commercial OAE device [35]. In addition to not being wireless, this approach is still constrained by the cost of commercial OAE probes that are expensive. Further, since it is directly connected to a smartphone it requires calibration for each smartphone model which is difficult to generalize. Recent commercial approaches [20; 33] use bone conduction to monitor OAEs through a headband. In addition to not being in the wireless earbud form factor, these have not been demonstrated to be low-cost. [96] proposes to use a single transducer to measure OAEs. However it is primarily focused on the transducer characterization and does not build an end-to-end wireless earbud system. Further, none of these prior efforts have performed clinical studies to evaluate efficacy. Finally, high-end personalized earphones from companies like Nura Sound use OAE measurements to customize music for an adult wearer [32; 65]. Our goal in this work is complementary in that we create a low-cost and open-source earbud system that is designed to achieve hearing loss screening with accuracies similar to medical devices.

Earable platforms. Recent years have seen the introduction of earbud platforms like eSense platform [104; 126], Clearbuds [78] and OpenEarable [151]. Like eSense, the Clearbuds platform does not have speakers or microphones facing the ear canal and is not designed for hearing loss screening. OpenEarable has a rigid over the ear design that is hard to operate across different age groups. We instead design an open-source wireless earbud platform to reliably measure otoacoustic emissions for hearing loss screening using low-cost hardware.

4.5 Limitations and discussion

We describe various limitations of our current system and discuss the regulatory pathway.

Field studies. In practical deployments, OAEbuds will potentially be used by a range of stakeholders including nurses, technicians and volunteers. Our clinical study does show that graduate students with no formal training in audiology were able to select the ear tips and snugly place the earbuds for both infants and adults. However, field studies in low and middle-income countries might be required to ensure that our

design can be used as advertised; this is however not in the scope of this chapter.

Use outside quiet environments. OAEbuds is currently designed and tested to work in relatively quiet environments similar to commercial OAE devices. For the earbud to be used in noisier environments, a future design could incorporate passive noise insulating material and an active noise cancellation system to reduce the effect of background noise on system performance.

Followup care and regulatory costs. Detecting hearing loss is an important first step in addressing this complex public health problem. Other factors include human resources for performing the tests, followup care and regulatory costs. We however note that prior FDA clearances for OAE devices did not require human testing [37], which significantly reduces the cost of clearance.

Software update to commercial earbuds. We develop a custom earbud instead of using existing earbuds for two key reasons: 1) we wanted to achieve a lower cost than existing wireless earbuds and demonstrate that OAEs can be detected using low-cost acoustic components, and 2) commercial wireless earbuds do not provide access to data from the in-ear microphone. We however note that the microphones and speaker used in Apple AirPods and Pixel buds have a higher quality. Given that all the hardware including an in-ear microphone are already present in commercial earbuds, this chapter shows that there is an exciting possibility that using the algorithms presented here, commercial earbuds can potentially enable OAE detection and hearing loss screening using only a software update.

4.6 Conclusion

Over the next decade, the mobile systems community is uniquely positioned to develop wearable and mobile technologies that help alleviate global health inequity. We developed the first wireless earbuds that can detect otoacoustic emissions and perform hearing screening using low-cost acoustic hardware. Our work introduces two components, 1) a low-cost wireless earbud hardware for hearing screening that works across infants and adults, and 2) wireless sensing algorithms to reliably identify otoacoustic emissions in the presence of in-ear reflections and echoes. Our clinical study demonstrates similar sensitivity and specificity to commercial medical devices and shows the potential of our design to enable hearing loss screening in low and middle income countries.

Chapter 5

Performing tympanometry using smartphones

Tympanometry is an objective test of middle ear function and can be used in combination with other tests like otoscopy and pneumatic otoscopy to diagnose middle ear disorders in accordance with clinical guidelines [62; 114]. Accurate screening for middle ear disorders can result in more timely referrals to specialists and could contribute to the reduction of serious complications [122; 147]. During a tympanometry test, the air pressure is changed in the ear canal to evaluate the mobility of the tympanic membrane and ossicular chain [132]. Although this core principle behind tympanometry was introduced in the 1950-60s [156], there has not been many major advances in making these devices affordable and accessible for resource-constrained settings. As a result, existing tympanometers remain expensive, ranging from $2000 (handheld Otowave, Amplivox) to $5000 (AutoTymp, Welch Allyn), limiting its availability in resource-constrained scenarios including rural areas and developing countries [130].

At the same time, smartphone technology has advanced substantially over the last two decades. Today, budget smartphones that cost $40–50 second-hand contain powerful processors, user interfaces and high-resolution displays [31]. Here, we present an inexpensive tympanometry system that uses a lightweight and portable attachment to smartphones. It is designed to automatically detect when a seal has been formed with an ear canal, safely vary air pressure, and generate a tympanogram on the smartphone in real-time. Our attachment can be assembled using electronic and passive components with a material cost of $28.

Due to the prevalence of smartphones in developing countries [14; 15; 23], our tympanometry tool may help increase timely access to otologic healthcare. We open-source our hardware and software making our system free and accessible to use and further adapt.

5.1 Concept and Prototype

We designed our smartphone tympanometer in two different form factors. Fig. 5.1a shows a handheld design that holds all the electronics in a 3D printed plastic enclosure that attaches to the back of a smartphone. A handheld form factor may be suitable for use in mobile health clinics [119; 153; 163] or as part of medical humanitarian trips (e.g., doctors without borders) [76; 166]. The desktop form factor shown in Fig. 5.1b mimics the design of existing commercial tympanometers. All electronic components of the system except for the smartphone are assembled onto a 42 $\times$ 45 mm custom printed circuit board (PCB) (Fig. 5.1c).

The pressure and acoustic sensors are connected to the ear probe through 1 m of lightweight air tight silicone tubes to provide greater mobility during measurement. These silicone tubes are protected from damage using a flexible cable sleeve. The ear probe is designed to be lightweight and can rest securely in a patient's ear during a measurement without any additional applied force. Additionally, the ear probe was designed to be compatible with existing rubber ear tips (Grason & Associates) used with the commercial tympanometers in our clinical study [30]. Our tympanometry platform is easily programmable on the smartphone where parameters such as pressure speed, pressure limits, tone frequency, and volume can be modified in the software app.

The smartphone attachment has multiple low-cost components. It consists of a pressure transducer made from a stepper motor to precisely move the plunger of a 5 mL syringe. Feedback from an onboard pressure sensor is used to accurately change the pressure between -400 and 200 daPa [132]. Fig. 5.1d shows the syringe plunger moving a total distance of 5.3 mm to cover the pressure range. During this pressure sweep, the system sends a 226 Hz audio tone at 85 dB SPL and records the acoustic reflections at a microphone connected to the smartphone by a 3.5 mm audio jack. A frequency of 226 Hz was used in our clinical study, as it is the recommended frequency for patients over 9 months of age [57], though our system can be programmed to emit other frequencies. When the measurement is complete, the digital pressure data is sent to the smartphone using an onboard wireless Bluetooth radio. Fig. 5.3a,b show the acoustic data bandpass

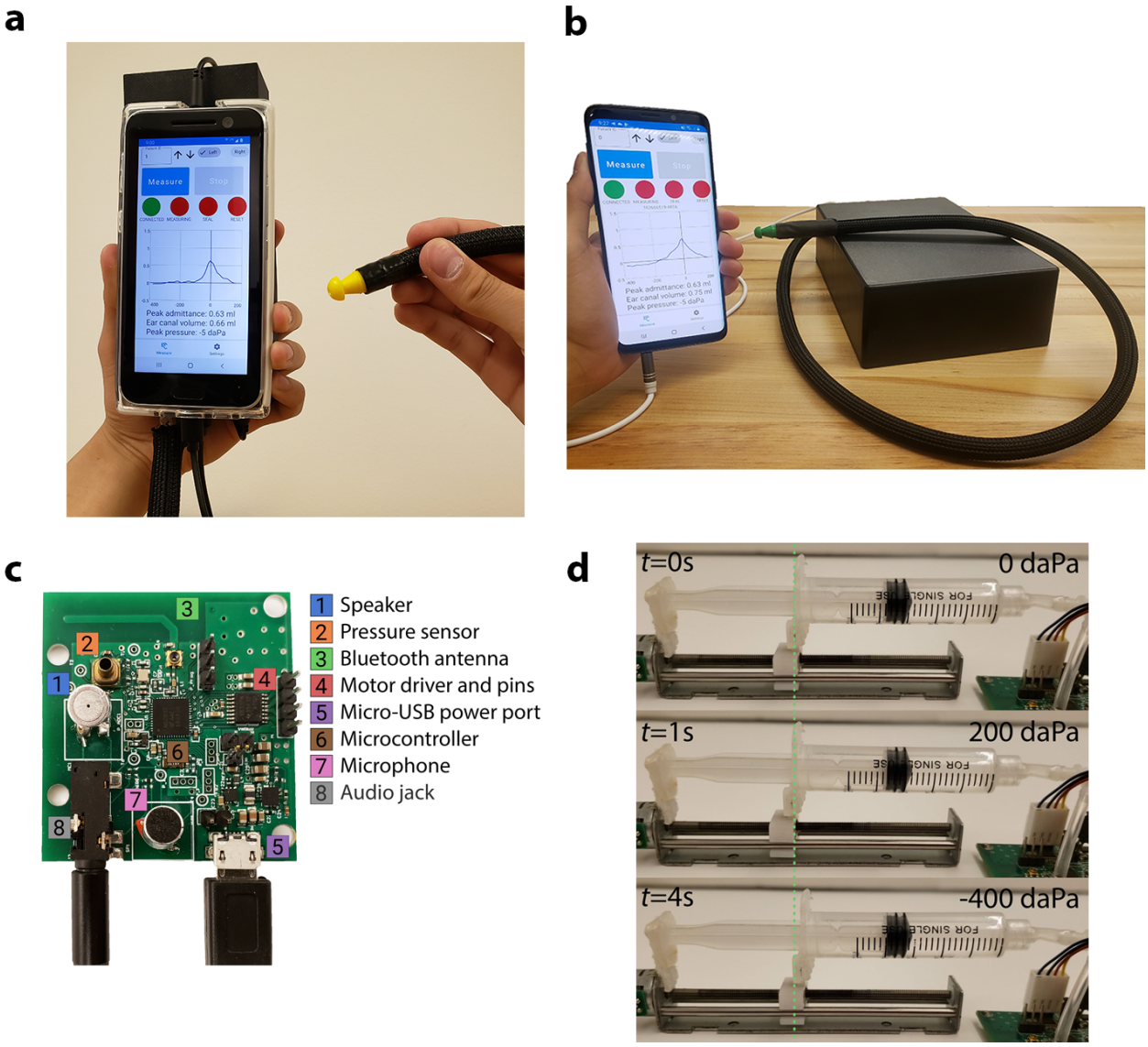

Figure 5.1: Low-cost smartphone-based tympanometer a, All components of the hand-held tympanometer fit into a portable 3D printed enclosure that attaches to the back of a smartphone. **b,** A desktop version of our tympanometer. **c,** Close-up of PCB containing the key acoustic and pressure sensing elements of the tympanometer. The PCB includes a microcontroller, Bluetooth antenna and micro-USB port for computing, communication, and power respectively. **d,** Through precise movements of a plunger and syringe, the stepper motor introduces positive and negative pressure into the ear canal. The green line is a reference marker for the plunger location at the beginning of the measurement. The plunger moves 5.3 mm to perform the pressure sweep from 200 to -400 daPa.

filtered to 220–230 Hz and the pressure data from a single measurement. To ensure safety, our system is designed to terminate the measurement if the syringe of diameter 12.5 mm has moved forward by more than 16 mm. This termination condition was added to avoid large changes in air pressure in the ear. It is a

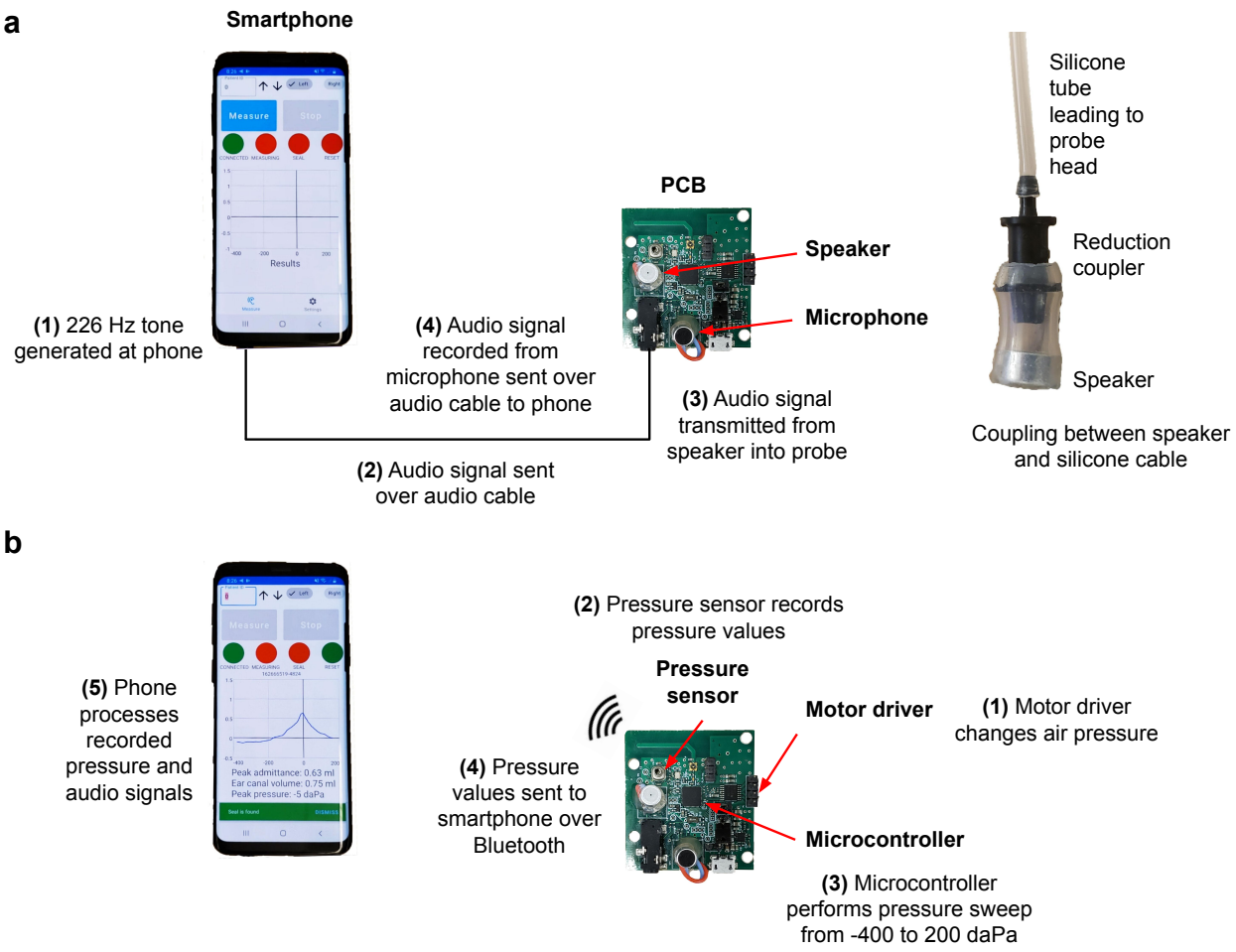

Figure 5.2: Working principle of our smartphone-based tympanometry device. a, Audio signal generation and reception. **b,** Pressure sweeping, sensing and communications.

fail-safe in the event of malfunction with the pressure sensor. We note that during our study there were no failures with the pressure sensor.

Our real-time algorithms run on the microcontroller to detect when a seal is formed within the ear canal and automatically begin measurement, without any user intervention. To detect if there is a seal, the microcontroller moves the syringe plunger back and forth. When the probe tip is outside the ear, only small changes in pressure at an average rate of 2.03 daPa/s are detected (Fig. 5.3c). If the ear probe has an adequate seal, the change in air pressure will exceed a predefined threshold (17 daPa/s). Our algorithm begins the measurement automatically by sweeping the pressure from 200 to -400 daPa and then returning to 0 daPa.

If the probe is dislodged from the ear canal during a measurement, air pressure returns to ambient

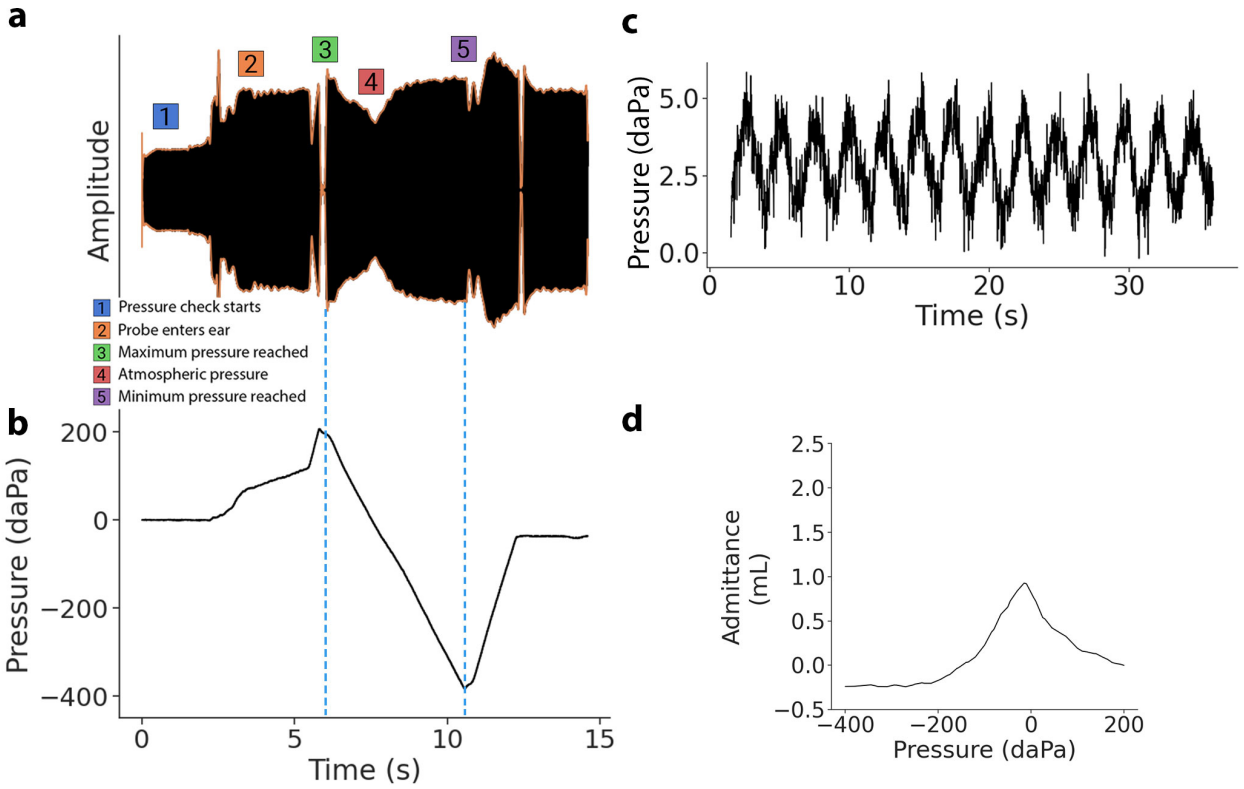

Figure 5.3: Measurement procedure to obtain tympanogram. **a,** Received acoustic signal bandpass filtered to 220–230 Hz. The acoustic signal increases in amplitude when it is inserted into the ear and changes in amplitude in response to air pressure changes. **b,** Air pressure relative to atmospheric pressure as measured by the pressure sensor. The pressure signal spikes when the probe tip enters the ear. The pressure transducer performs a pressure sweep from 200 to -400 daPa, before returning to atmospheric pressure. **c,** When the probe is outside the ear, only small pressure changes are recorded by the microcontroller. **d,** Calibrated tympanogram in units of acoustic admittance magnitude.

pressure. Our algorithms use this to detect that the seal has been lost. Our smartphone algorithm also detects if the eartip is occluded which can occur if it is flush against the wall of the ear canal or obstructed with cerumen. Since pressure changes in a clear and occluded ear canal are similar, our algorithm instead checks for high amplitude acoustic reflections as an indicator for occlusion.

In a healthy patient, an increase in positive and negative pressure tense the eardrum and make it more stiff, which causes an increase in reflected sound energy. The tympanogram is generated on the smartphone by converting the sound energy into calibrated units of acoustic admittance magnitude (Fig. 5.3d). The calibration process is performed by comparing reflections from the ear to reflections generated from cavities of known volumes.

Ears with middle ear fluid may have a stiff eardrum and the amount of reflected sound energy at different pressure levels may not have a pronounced peak. Patients with ear tubes or a tympanic membrane perforation will likely have a consistently attenuated reflection at different pressures that would indicate an unusually large ear canal volume. Middle ear conditions that affect the eardrum, like chronic eustachian tube dysfunction with a retracted eardrum and without fluid, may result in the tympanogram peak occurring at a negative pressure. An overly stiff eardrum, in the case of reduced ossicular mobility such as otosclerosis, may cause a lower than normal peak amplitude and an ossicular discontinuity may cause a higher than normal peak amplitude [132]. The tympanogram output displayed on the smartphone can be interpreted in the same way as it is on commodity tympanometers.

5.2 Hardware design

The electronic components of our hardware (Fig. 5.1a,b) can be divided into three groups: acoustic, pressure, and computing components. Our acoustic hardware consists of a printed circuit board (PCB) (Fig. 5.1c) with a speaker (Knowles SR-32453-000, $4.42) and microphone (DB Unlimited MO064404-1, $0.29) to send and receive a 226 Hz audio tone. The speaker and microphone are connected directly to a smartphone via a 3.5 mm headphone jack. The smartphone sends a tone to the speaker with a sound level of 85 dB SPL at the ear canal, and records audio from the microphone at a sampling rate of 24 kHz. Our pressure hardware consists of a linear stepper motor (DC 2-Phase 4-wire Stepper Motor, $6.16) connected to a syringe and plunger (Frienda syringe, $0.55). The stepper motor is driven by a motor driver (Toshiba TB6612FNG, $0.90). A pressure sensor (Honeywell MPRLS0025PA00001A, $3.97) reads air pressure values in the ear canal. Finally, a microcontroller (Nordic Semiconductor nRF52832, $2.28) continuously monitors the air pressure values and runs the algorithm to check if there is a seal between the ear tip and ear canal, and performs the pressure sweep from 200 to -400 daPa (Fig. 5.1d). The microcontroller sends the recorded pressure values to the smartphone over Bluetooth at the end of the measurement to compute the tympanogram (Fig. 5.2). When the measurement is complete, the plunger of the syringe is automatically moved back to its original position. The PCB and its components can be powered by a smartphone through a micro-USB to USB-C cable. The current draw of the PCB when the motor is in use is 400 mA, which is within the current draw specifications for USB power delivery 2.0 used by USB type-C connectors [36].

Next we describe the set of passive components used to couple the acoustic and pressure components to the ear. There are three main connections. First, the microphone, speaker, syringe and pressure sensor are each directly connected to thin (1 mm ID, 2 mm OD) and flexible silicone tubing. Second, to group the connections, the three silicone tubes are encased in a larger (3 mm ID, 4 mm OD) silicone tube. Finally, the grouped silicone tube is attached to the probe head, which is a plastic adapter that can interface with standard tympanometer ear tips. All the components in the desktop version of our system are enclosed in a plastic enclosure (Hammond Manufacturing 1591XXFSFL, 221.01 $\times$ 150.01 $\times$ 63.50 mm) The total material cost of all electronic and passive components in the system is around $25.

5.3 Smartphone application

A custom smartphone application was created to work with the Android operating system. The application first waits for user input to begin the measurement session. Upon confirmation, the application sends a Bluetooth packet to the microcontroller to begin checking if the ear tip has formed an airtight seal with the ear canal. Whenever a seal has been found or lost, the microcontroller sends a Bluetooth notification to the smartphone, which displays this in the form of a notification to the user. The application is responsible for sending the acoustic tone to the microphone, and for recording the reflected sound energy at the microphone. When the measurement session is complete the application will receive a confirmatory Bluetooth packet from the microcontroller along with a list of timestamped pressure values.

5.4 Algorithm to identify a seal

An air seal between the ear tip and the ear canal is required to change air pressure in the ear canal space. Our seal detection algorithm has two steps. Firstly, we attempt to increase the pressure in the ear canal by moving the plunger of the syringe forward and backward. Specifically, we move the syringe plunger forward by 0.5 mm within an average time of 1.19 s. If the pressure sensor detects a sharp increase in pressure exceeding a threshold of 17 daPa/s, we record that a seal has been formed. When the probe is outside the ear, only a small change in pressure at an average rate of 2.03 daPa/s will be registered (Fig. 5.3c). Even when there is no seal, a small change is still registered as the syringe introduces air pressure more quickly

than air exiting the system. If no seal is detected, we move the syringe backwards by 0.5 mm. These steps are repeated until a seal is registered. We note that the pressure threshold for when a seal has been obtained has to be at least 17 daPa/s. We opted for a pressure threshold of 17 daPa/s as it was practical and yielded accurate measurements.

Relying on pressure changes alone to detect a seal can result in false seals from ear tip occlusions. For example, a false seal would occur if the ear tip were occluded with cerumen. In our second step, we track the amplitude of the reflected audio tone on the smartphone to detect occlusions. When there is an occlusion the amount of sound reflected back to the microphone is much higher than if the ear tip was clear. If the amplitude of the tone exceeds a predefined threshold of 65 dB SPL as recorded on the microphone, we inform the user that the ear tip is occluded. If there is no occlusion, the smartphone sends a Bluetooth notification to the microcontroller, which records that a seal has been formed. This threshold can be calibrated for different smartphones and probes by coupling the probe head to a 5 mL syringe with the plunger set to 0 mL, starting a measurement, and noting the acoustic amplitude value as received on the smartphone.

5.5 Measurement procedure

When a seal is obtained, the measurement begins by increasing the pressure in the ear canal to 200 daPa, performing a pressure sweep from 200 to -400 daPa, and then returning the ear canal back to the atmospheric pressure.

Identifying when to start a measurement. When the probe enters the ear canal and has an airtight seal, the pressure within the tubing system increases sharply. This is due to the pressure applied to the ear tip when it enters the ear canal, which causes air from the ear canal to suddenly enter the tubing system (Fig. 5.3b). If this initial pressure increase is below 200 daPa, the motor moves the syringe forward until the pressure in the ear reaches 200 daPa before starting the measurement. If the initial pressure increases beyond 200 daPa, the system starts the measurement and decreases the pressure to -400 daPa.

The initial pressure increase to 200 daPa, and later pressure increase from -400 daPa back to atmospheric pressure (0 daPa) is performed at a higher speed of 316 $\pm$ 120 daPa/s, compared to the speed of 123 $\pm$ 18 daPa/s during the pressure sweep. This was done to reduce measurement time.

Performing the measurement. During the measurement, the syringe is pulled backwards to sweep the pressure in the ear canal from 200 to -400 daPa. The motor speed is slowed to 123 $\pm$ 18 daPa/s during the measurement phase.

To ensure the plunger does not reach the end of the syringe, the session aborts if the syringe has moved forward by more than 16 mm during a measurement session. This scenario can occur if the ear tip loses its seal with the ear canal multiple times in a single measurement session.

Completing the measurement. After the pressure sweep is completed, the pressure in the ear canal is returned to atmospheric pressure. The user is then instructed to remove the ear tip from the ear canal. After 5 seconds, the syringe plunger is automatically moved back to its original position.

5.6 Algorithm to compute tympanogram

Our algorithm to compute the tympanogram ran in realtime on the smartphone during the clinical study and the tympanogram was displayed on the smartphone. The average runtime was 289 $\pm$ 48 ms across 10 measurements.

Synchronizing pressure and audio streams. In order to generate the tympanogram, pressure values recorded on the microcontroller need to be synchronized with the audio recording stored on the smartphone. This is challenging, as the clocks between the microcontroller and smartphone are not synchronized. The time shifts between the pressure and audio data are estimated using two steps. In the first step, the microcontroller sends a Bluetooth notification to the smartphone when the measurement is at the beginning and end of the pressure sweep (point 3 and 5 in Fig. 5.3a). The smartphone then records a timestamp when it has received these notifications. However, these timestamps do not account for the delay caused by transmitting and receiving the Bluetooth notification. As a result there is some unknown offset between the arrival of the notifications, and the beginning and end of the pressure sweep. In the second step, we estimate when the pressure sweep starts (point 3). To do this, we process the envelope of the audio recording starting from 500 ms before the Bluetooth notification to when the Bluetooth notification arrives. We observe that when there is a change in pressure direction, this causes the acoustic signal to dip for a brief period as shown in Fig. 5.3a. The maximum peak between this dip and the Bluetooth notification corresponds to the point of maximum pressure (point 3). This procedure is similarly repeated to find the point of minimum pressure

(point 5).

Calculating tympanogram. The goal of this step is to convert the synchronized pressure and audio data into a tympanogram. To do this, the audio data is divided into N segments, where N is the number of recorded pressure values between 200 and -400 daPa. For each audio segment, we perform a 24000-point FFT, and track the absolute value of the 226 Hz tone across all segments.

Smoothing tympanogram. The purpose of this step is to smooth the shape of the tympanogram and reduce the effect of noise. This is a two part procedure. In the first part we discretize the admittance values into pressure bins of 5 daPa. Empty buckets are filled with the average admittance values in adjacent buckets. In the second part, we pass the tympanogram through a moving average filter with a window size of 5 samples.

5.7 Calibration procedure

A one-time calibration procedure needs to be performed prior to using the system. As the smartphone does not have the ability to measure the sound level at the microphone in absolute physical units of dB SPL, we perform sound level calibration using a sound level meter which is able to perform this measurement. To do this, we first couple the probe head to a sound level meter (BAFX 3370, Digital Sound Level Meter, $18). We then adjust the volume gain of the speaker through the smartphone UI until the sound level output from the probe head reaches 85 dB SPL. We next calibrate the smartphone so the generated tympanograms can be normalized to absolute units of admittance. This calibration procedure is performed by measuring test cavities with volumes ranging from 0 to 5 mL in increments of 1 mL. For the calibration procedure, we use a 5 mL syringe with a diameter of 12.5 mm as our test cavity. Calibration is performed by fitting the measured values at 200 daPa for all volumes to a cubic curve. The cubic fit produces four coefficients p_1, p_2, p_3, p_4 that are used to normalize an uncalibrated amplitude value x: $p_1x^3 p_2x^2 + p_3x + p_4$. We note that the measured tympanograms exhibit a slope as pressure decreases. We compensate for this slope when computing the root-mean-square error of the tympanograms. To do this, we measure the slope m of the tympanogram measured in the 0 mL configuration between -400 and 200 daPa, and the amplitude value b at -400 daPa. We then created a line of best fit $y = mx + b$ where x is vector of measured pressure values, and subtract this from all subsequent tympanogram measurements to compensate for the slope.

For the calibration procedure, the volume range of 0 to 5 mL is selected to match the recommended

calibration cavity volumes of 0.5, 2, and 5 mL as specified in ANSI S3.39 [4]. Additionally, we note that a meta-analysis [156] of several tympanometry studies show that the 90% range for ear canal volume ranges from 0.3 to 2.2 mL, and falls within the 0 to 5 mL range of calibration volumes. While our system is calibrated for the volume range of 0 to 5 mL, which would result in a plunger displacement of 5.3 to 7.1 mm, our system can accommodate a maximum plunger displacement of 16 mm of the entire reserve.

5.8 Clinical study design

All patients in the study were scheduled for an appointment with an audiologist and were undergoing tympanometry as part of their clinical visit. A licensed audiologist first performed tympanometry on subjects using our smartphone system and then with a commercial tympanometer. The study was performed by two independent audiologists. The first audiologist tested 31 patient ears and used the GSI TympStar Pro as the reference tympanometer. The second audiologist tested 19 patient ears and used the GSI TympStar as the reference tympanometer. Otoscopy was performed prior to the measurement to check for the presence of ear tubes, eardrum abnormalities or cerumen obstruction. A behavioral audiogram was performed on each patient as part of their clinical visit, and any hearing loss, history of ear infection or relevant otologic conditions were recorded. The audiologist performing the test classified the tympanograms measured from the commercial tympanometer into Liden and Jerger classifications [101; 111; 112; 113] using clinical criteria from our institution. All clinical tests were performed using the Samsung Galaxy S9 smartphone.

The study was approved by the Seattle Children's Hospital Institutional Review Board (STUDY00002820). All studies complied with relevant ethical regulations. Parental permission was obtained for participants under the age of 18 years. Children age 7 to 17 provided written assent. Assent was obtained after parental permission was granted. Children age 7 to 12 signed a simple assent form and children age 13 to 17 signed a consent form. Parents co-signed the consent form. Participants 18 years and older signed a consent form. The measured pressure speed was on average 125 ± 19 daPa/s in ears tested in the clinical study. The GSI TympStar and GSI TympStar Pro both had a pressure speed of 200 daPa/s. Recruitment lasted until a sufficient number of measurements were collected to demonstrate proof of concept. Randomization was not applicable, patients were not allocated into different groups. Investigators were not blinded.

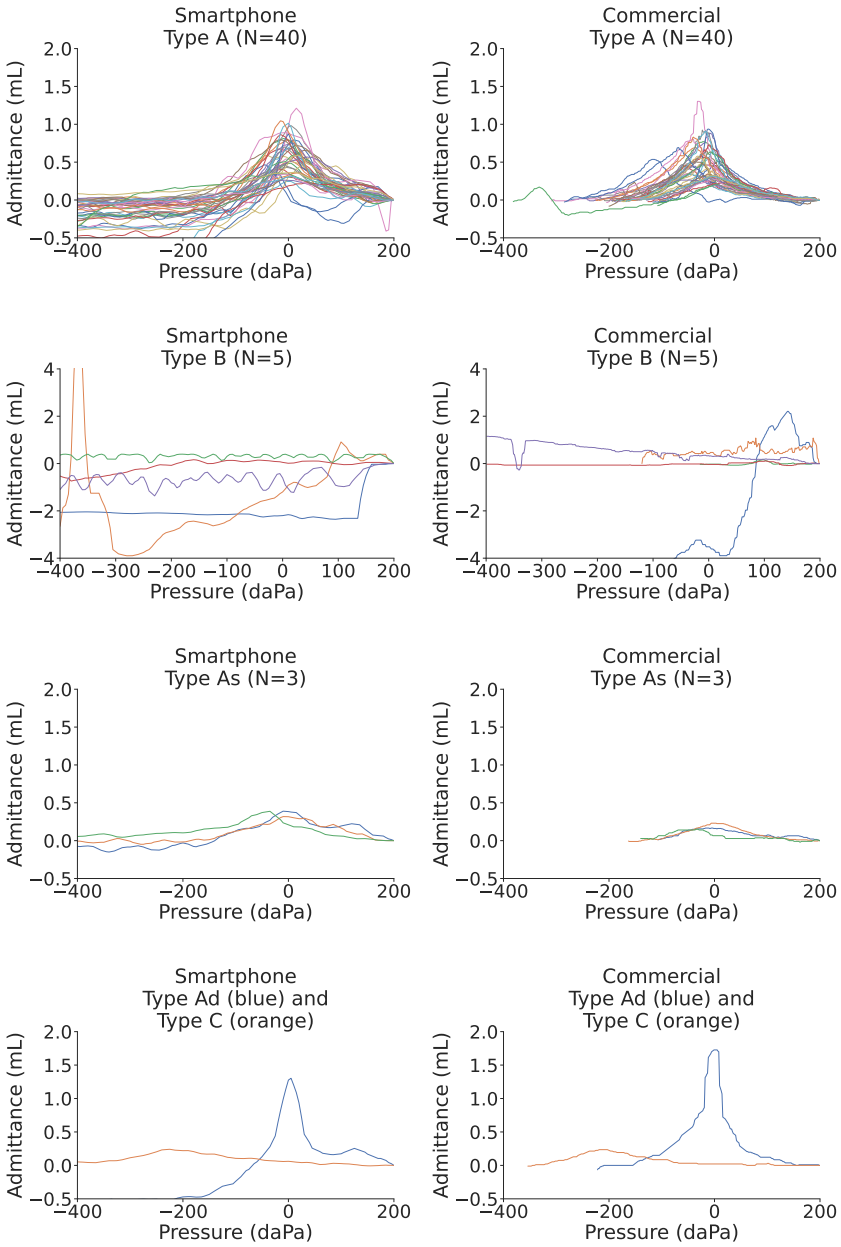

Figure 5.4: Comparison of individual tympanograms from smartphone-based and commercial tympanometer. We show the plots for 40 type A, 5 type B, 3 type As, 1 type Ad, and 1 type C tympanograms measured by the commercial tympanometer, and tympanograms from the same ear as obtained by the smartphone-based tympanometer.

5.9 Clinical testing results

We conducted a clinical study at Seattle Children's Hospital on 50 patient ears. The age range of the patients was 1 to 20 years with a mean age of 9 $\pm$ 5 years, and a female-to-male ratio of 0.52. Of the 50 tested ears, 40 of them were classified as Type A by the audiologist, 5 as Type B, 3 as Type As, 1 as Type Ad, and 1 as Type C. Of the 5 patient ears that were classified as producing Type B tympanograms, 4 had patent ear tubes and 1 had an eardrum perforation. One patient ear with a Type As tympanogram had clinically diagnosed otosclerosis. 27 of the 50 patient ears had a history of hearing loss, and 12 of the 50 ears had a history of ear infections. Fig. 5.4 shows all the tympanograms classified as Type A, B, As, Ad and C by the audiologist, and the corresponding tympanogram obtained from our smartphone system.

Bland-Altman analysis [59] was performed on peak admittance values for the 45 ears without ear tubes (Fig. 5.5b). Since the presence of a patent ear tube would produce a flat tympanogram without a peak, peak admittance would not provide meaningful information about eardrum mobility. Bland-Altman analysis had a bias error (mean of the differences) of -0.02 $\pm$ 0.14 mL. One of 45 measurements were outside the 95% limits of agreement. The mean and standard deviation of absolute error of peak admittance was 0.11 $\pm$ 0.09 mL. The root-mean-square error between the tympanograms measured on the smartphone and commercial tympanometer was on average 0.17 $\pm$ 0.11 mL across all 45 ears without ear tubes. As measurements on the commercial tympanometer were not always recorded to the minimum pressure of -400 daPa, the root-mean-square error is computed for the range of pressure values that were recorded on both the commercial and smartphone device.

Bland-Altman analysis was also performed on ear canal volume measured on the smartphone and commercial tympanometer (Fig. 5.5d). We show the ear canal volume for the 45 ears with an intact eardrum and the 5 ears without an intact eardrum. Measured ear canal volume is dependent on the insertion depth of the probe tip into the ear canal, and may vary across measurements on the same device. Clinically, ear canal volume is primarily used in conjunction with tympanogram shape to screen for intact eardrums, or an occlusion of the probe tip or ear canal. The ear canal volumes measured by the smartphone for the 5 ears without an intact eardrum all exceeded 1.5 mL and appeared as flat tympanograms. Bland-Altman analysis for the full set of 50 ears showed a bias error of 0.15 $\pm$ 0.96 mL, and 2 of 50 measurements fell outside the limits of agreement. The same analysis on the 45 ears with an intact eardrum had a bias error of -0.09 $\pm$

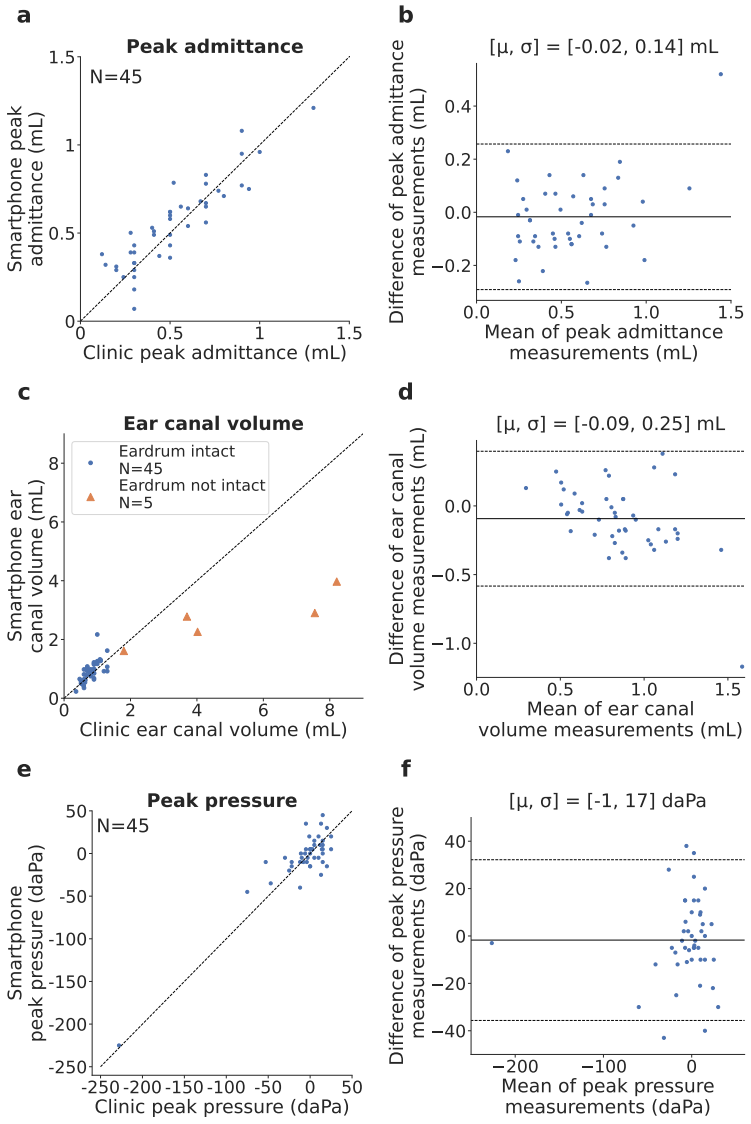

Figure 5.5: Clinical study performance. **a, c, e** Correlation and **b, d, f** Bland-Altman plots compare peak admittance, ear canal volume and peak pressure obtained from the commercial tympanometer and smartphone-based tympanometer. In the Bland-Altman plot μ indicates bias error (mean of the differences) between measurements and σ is the standard deviation (SD) of measurement error. The solid and dotted lines represent the bias error and 95% limits of agreement respectively.

0.25 mL, and 1 of 45 measurements falling outside the limits of agreement. Mean and standard deviation of absolute error of ear canal volume was 0.41 $\pm$ 0.88 mL for the full set of 50 ears, and 0.20 $\pm$ 0.18 mL for the 45 ears. The ear canal volumes for all ears without an intact eardrum ranged from 1.61 to 3.97 mL and were larger than all the ear canal volumes of ears with an intact eardrum.

Finally, Bland-Altman analysis was performed for peak pressure in the 45 ears with an intact eardrum

(Fig. 5.5f). In our clinical study, 1 ear had negative middle ear pressure of -225 daPa. Bland-Altman analysis showed a bias error of -1 $\pm$ 17 daPa and a mean and standard deviation of absolute error of 13 $\pm$ 11 daPa. Four of 45 measurements fell outside the limits of agreement. Prior work [90] comparing the agreement in peak pressure between two commercial tympanometers reported a bias error of -16 $\pm$ 11 daPa.

5.10 Tympanogram classification

At the completion of the clinical study, we presented a total of 100 tympanograms measured on the smartphone and clinic device to a group of five pediatric audiologists and asked them to classify the tympanograms into Liden and Jerger classes (A, As, Ad, B, and C). The 100 tympanograms were anonymized and presented in a randomized order to each audiologist. Since the cutoffs for different tympanogram classes depend on the patient's age, we also provided the age information. The results of otoscopy, any prior ear surgery, and otologic history were also presented. The average percent agreement between both devices was on average 86 $\pm$ 2% across all five audiologists. Additionally, all type B tympanograms measured on the smartphone were correctly classified by the five participating audiologists.

Under the five-class taxonomy, there were ten tympanograms where disagreement was observed and the smartphone classification differed from the ground truth. Eight of the ten disagreements were for tympanograms with a ground truth of Type A or As, but were incorrectly classified as a Type A, As, or Ad. In seven of these cases, the smartphone tympanogram had a peak admittance that was 0.01 to 0.12 mL away from the threshold that would have caused it to be correctly categorized in the same class as the commercial tympanogram. A 0.12 mL peak admittance variation is within the test-retest variability for both the smartphone and commercial device.

5.11 Benchmark testing

We conducted benchmark testing across different variations in system design (Fig. 5.6). Each benchmark test was performed in a single normal and healthy adult ear without history of middle ear disorders. The same ear was used for all tests. The default setup for these benchmark tests matched the setup used in the clinical study and used the Samsung Galaxy S9 and 5 mL syringe.

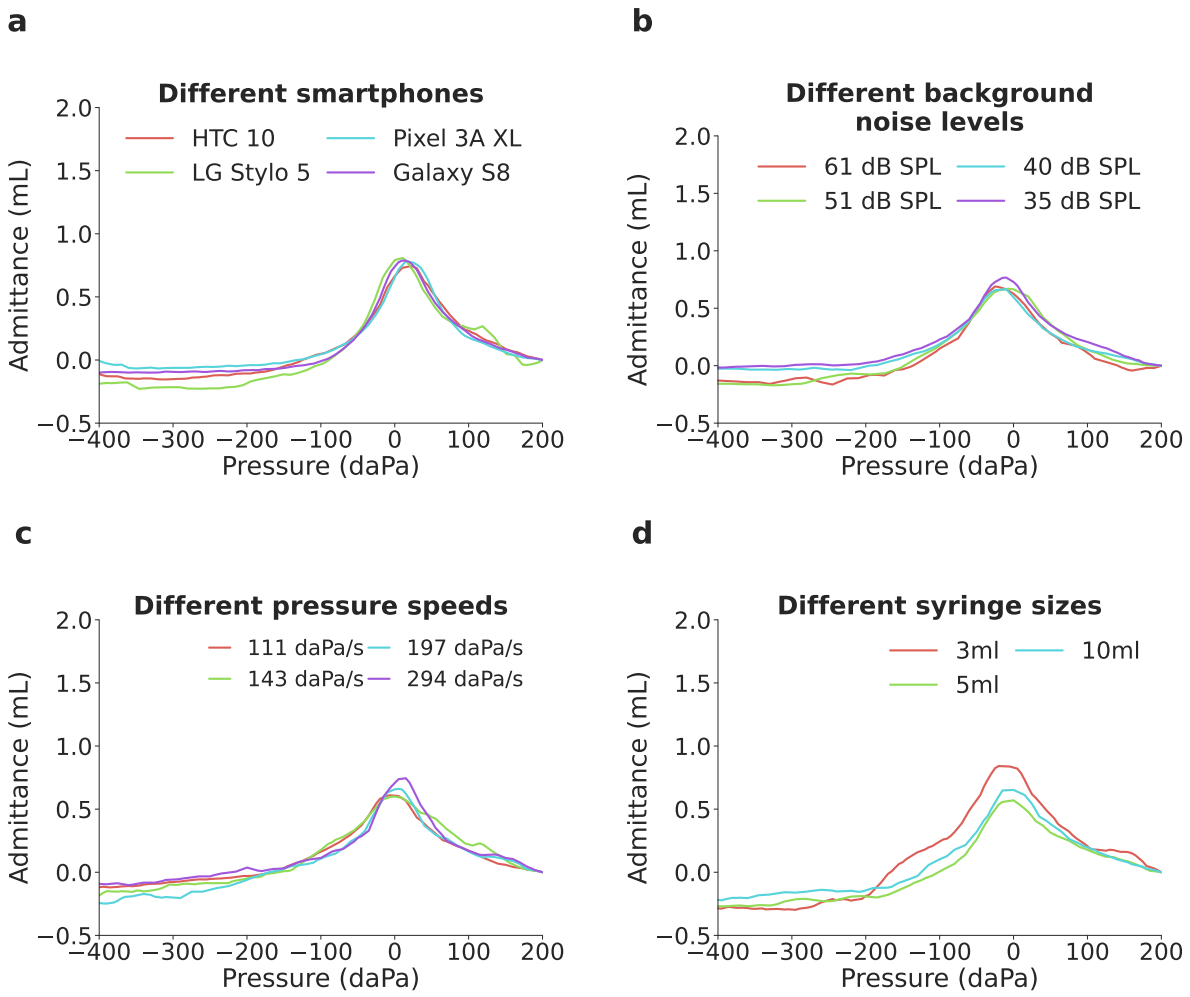

Figure 5.6: Benchmark testing across different measurement scenarios and design parameters. The system was evaluated across **a,** different smartphones including budget models, **b,** different background noise levels **c,** different pressure sweep speeds, **d,** different syringe volumes.

We first evaluated the effect of different smartphones on tympanogram output (Fig. 5.6a). The phones we tested were released during 2016 to 2019 and included smartphones that could be purchased second-hand for $50. We note that identical volume settings on each phone's software will result in slightly different sound levels at the speaker. This may be due to differences in the circuit driving the speaker. Across these phones, the same volume setting resulted in sound levels ranging from 78 to 90 dB SPL. Our calibration procedure was performed for each phone in order to generate comparable tympanograms. We find that across all phones, the peak admittance and peak pressure was in the range of 0.74 to 0.81 mL and 10 to 20 daPa respectively.

Second, we tested the effect of interfering background noise on the tympanogram (Fig. 5.6b). We play an acoustic chirp from 220 to 230 Hz with a duration of 1 s in a continuous loop, using an external speaker (Beats Pill 2.0) positioned 50 cm away from the ear. The sound level of the acoustic chirp was measured at the microphone positioned at the end of the probe cable by coupling a sound meter (Amprobe SM-10) to the silicone tube that would attach to the microphone. The measured sound levels ranged from 35 to 61 dB SPL. The sound levels when measured next to the ear being measured ranged from 43 to 77 dB SPL. The sound levels when measured 1 cm away from the speaker ranged from 75 to 110 dB SPL. The tympanogram is slightly less smooth at the highest sound level tested, but otherwise has no appreciable effect on the overall shape of the tympanogram. Peak admittance ranged from 0.67 to 0.77 mL and peak pressure ranged from -25 to -10 daPa across all sound levels.

Third, we tested how different pressure speeds would affect the tympanogram (Fig. 5.6c). Pressure speed was varied by changing the rate at which the stepper motor pulled the syringe plunger. We tested four different speeds ranging from 111 to 294 daPa/s. These speeds correspond to a pressure sweep duration of 2.04 to 5.45 s. Across different speeds the peak admittance ranged from 0.60 to 0.74 mL and peak pressure ranged from -15 to 5 daPa. Additionally, we tested the effects of different pressure speeds on the commercial tympanometers (GSI TympStar and GSI TympStar Pro) used in the clinical study, and measured a healthy adult ear three times. Both machines had preset measurements speeds of 12.5, 50, and 200 daPa/s. On the GSI TympStar, peak admittance ranged from 0.60 to 0.70 mL while peak pressure ranged from -10 to 5 daPa. On the GSI TympStar Pro, peak admittance ranged from 0.60 to 0.68 mL while peak pressure ranged from -41 to 0 daPa.

Fourth, we show the effect of different syringes volumes on the tympanogram (Fig. 5.6d). We tested three different syringe volumes ranging from 3 to 10 mL, with syringe diameter increasing from 10 to 16 mm. While the stepper motor moved at the same speed for these experiments, the use of larger syringes resulted in higher pressure speeds due to the increased diameter of the plunger. Pressure speeds ranged from 79 to 150 daPa/s across the three syringe volumes. Across all syringes the peak admittance ranged from 0.57 to 0.65 mL and peak pressure ranged from -5 to 0 daPa.

Fifth, we show the admittance values when the ear probe is connected to hard-backed cavities of varying volume. Our cavity is a 5 mL syringe, with the plunger retracted at different volume levels. We can see that

a flat tympanogram is produced at each volume level from 0 to 5 mL. The average root-mean-square error of the calibrated tympanograms was 0.09 ± 0.02 mL across all volumes.

Finally, to evaluate if untrained participants could perform the calibration procedure, we provided two independent participants with written instructions and diagrams for performing the process and compared their calibration results with that of a trained researcher. Specifically, the device was first calibrated to the measurements obtained by the trained researcher, and the obtained cubic coefficients were used to normalize measurements for the two independent participants. No additional calibration was performed for the two participants. Curves from all participants were compared against the ground truth syringe volume. The average root-mean-square error across the pressure range of -400 to 200 daPa, for all volumes from 0 to 5 mL was 0.08 mL for the trained researcher and 0.08, and 0.07 mL for the independent participants. The independent participants were able to read the instructions and perform the entire calibration process in 4:13 and 4:35 minutes.

5.12 Discussion

Existing techniques to measure middle ear function leverage different sensing techniques to probe the mobility of the eardrum and middle ear system. [73] leverages the speakers and microphones on smartphones to assess eardrum mobility using acoustic reflectometry. Unlike prior work, the smartphone-based tympanometer presented here provides similar information to commercially available tympanometry, and can be used to aid clinical diagnosis of a variety of pathologies affecting the middle ear. KUDUwave TMP [144] is a recent commercial device that integrates two pneumatic pumps into circumaural ear cups and ear inserts to perform bilateral tympanometry and audiometry and costs $6950. Our system mimics traditional unilateral tympanometry and is built on a smartphone device in both a handheld and desktop form factor. Visual otoscopy is often performed prior to tympanometry to screen for a bulging or cloudy tympanic membrane or an ear canal occluded with cerumen [137]. Visual otoscopes are also readily available as smartphone attachments [19; 85; 120], but are known to have poor sensitivity and specificity at detecting middle ear disorders [62; 133; 175].

Optical coherence tomography (OCT) uses the reflections of infrared light to try to gain information about the middle ear [81; 82; 127]. While OCT devices are available on the market, they are not widely

adopted. Unlike these tests, our smartphone-based tympanometry is able to provide information about middle ear function similar to commercial tympanometry. The portable nature of our system makes it a potentially viable option for use in developing countries, primary care clinics and mobile health clinics.

Looking forward, our system was tested in a pediatric population with a minimum age of one year with a tone frequency of 226 Hz. Due to known developmental differences in mass and stiffness characteristics of the outer and middle ears of infants age 0 to 9 months [83; 162], tympanometry at 1000 Hz is recommended [80; 121]. Further research is needed to evaluate the utility of our system in young infants. Furthermore, while our current design has been optimized and evaluated for tympanometry, a similar design could potentially be used to perform related tests including acoustic reflexes, which is often performed by commercial tympanometers in ipsilateral or contralateral mode. Related audiological tests that rely on the use of microphones and speakers such as audiometry, and otoacoustic emissions could potentially be integrated into a similar hardware design. Further work would be required to develop these designs and evaluate their clinical performance.

In summary, we present an open-source smartphone-based tympanometry system to measure middle ear function. Although commercial tympanometry is the standard of care to assess middle ear function, the test equipment is expensive and not readily available. Given the prevalence of inexpensive budget smartphones, particularly in developing countries, our frugal system has the potential to be a screening tool for middle ear disorders in resource-constrained environments. Further clinical studies are required to assess the technology's efficacy and acceptability for use in these environments.

Chapter 6

Conclusion

In conclusion, this dissertation develops a unique approach to societally impactful research by deeply understanding problems that are important in the medical domain, and creating computational methods that can scale mobile systems to alleviate health inequities. Specifically, we invent methods spanning applied machine learning, wireless sensing, signal processing, and hardware to create intelligent earable systems for equitable healthcare. Collectively, our suite of systems break the conventional wisdom that expensive medical devices are needed to obtain high-quality clinical testing. Instead, our systems achieve high-quality clinical performance while leveraging low-cost, commodity sensors, generalize across hardware, and can be deployed in real-world environments for impact. This computational toolkit has enabled these systems to transform the field of audiology and pave the way for health equity for billions of people around the world. We have built the following intelligent earable systems for equitable healthcare:

- *Detecting middle ear fluid using smartphone [73].* We develop the first smartphone-based system that uses the onboard speakers and microphones to probe for eardrum mobility and accurately detect middle ear fluid at specialist-level accuracies. Using wireless sensing and machine learning techniques, we enable the system to generalize across multiple unseen smartphones, which allows the system to scale and have realworld impact.

- *Newborn hearing screening using low-cost earphones [66; 67].* We create systems to perform newborn hearing screening by detecting otoacoustic emissions, which are extremely faint sounds produced by the cochlea. We create a low-cost smartphone probe, and a pair of wireless earbuds. Here, we lever-

age hardware-software co-design to enable high clinical accuracies while maintaining a low material cost.

- *Performing tympanometry using smartphones [71].* We introduce an open-source hardware design and real-time algorithms for a low-cost smartphone-based tympanometer that consists of a lightweight and portable attachment to measure middle ear function.

We also demonstrate the broad applicability of our unique approach to societally impactful research and create intelligent mobile systems that have an impact in the fields of cardiology, hematology and anesthesiology:

- *Contactless cardiac arrest detection using smart devices [74].* We invent an ambient sensing system for smart devices to accurately detect audible biomarkers associated with cardiac arrest. We invent data collection and augmentation methods that allow the system to scale to a variety smart devices and unseen environments while maintaining high detection accuracies and a low false alarm rate.

- *Micro-mechanical blood clot testing using smartphones [70].* We create a blood clot testing system using the vibration motor and camera on smartphones to track the micro-mechanical movements of a copper particle and accurately calculate PT/INR time for blood and plasma samples at accuracies comparable to a clinical-grade coagulation analyzer and an at-home point-of-care device.

- *Closed-loop wearable naloxone injector system [69].* We create a wearable device that can detect and reverse the fatal effects of an opioid overdose. The device measures respiration and apneic motion associated with an opioid overdose event using a pair of on-body accelerometers and administers naloxone which is the antidote for reversing an overdose.

6.1 Looking forward

Looking forward, this dissertation is only the beginning of a vision where we can ensure that every human on the planet has access to basic medical diagnostic tools at their fingertips.

Mobile health systems for first-class remote care. While remote health monitoring is one way in which this vision can be achieved, patients still need to go into the clinic to do basic medical tests. The rise in

remote health monitoring adoption during COVID-19 presents an opportunity to develop suites of remote tests that can enable telemedicine appointments to become as effective as in-person visits, and reach the hands of millions of patients. While our work on ear infections is one example of a remote medical test, there are significantly more opportunities to create at-home tests such as tests to distinguish between viral and bacterial infections in biofluids and perform continuous fetal heartbeat monitoring.

Ambient intelligent systems. Going forward, there are many opportunities to expand upon the capabilities enabled by ambient intelligent systems. While my work on cardiac arrest detection is one example of how such ambient systems could work to detect emergent events, they can also be used for long-term and continuous health monitoring especially in the home setting. For example by combining active sonar technology on smart speakers and depth cameras, we may be able to capture the biomarkers of rare seizure events which are difficult to predict. Looking more broadly at audible biomarkers, there are opportunities for smart speakers and voice assistants to track a person's health status and predict events and disorders like COPD relapses, dementia and Alzheimers given that there is a lot more data in the home setting. Such ambient systems can also scale to public spaces like train stations, buses, and offices to contactlessly track the health of everyone in the space in a privacy-preserving manner, create a realtime barometer of a city's health, and track outbreak of diseases.

Building these systems often requires addressing security and privacy issues at all layers of the stack. This is especially pertinent for systems that are designed to continually adapt to the unique physiological patterns of each individual over time. By leveraging hardware and software co-design we can design algorithms to run in real-time on the resource constraints of embedded devices, so that no data is sent to the cloud. At the operating system layer we can explore the use of secure enclaves to provide additional security even if an attacker has physical access to the device. Even at the hardware layer we can create analog components that obfuscates the sensor data in such a way that the identifiable components of the data are erased without compromising the diagnostic accuracy of the system.

Such systems would also require the development of new sensors and algorithms to track different symptoms and diseases, and which can be deployed at scale. For example, mobile and non-invasive diagnostic tools can be created to detect the presence of disease from the gaseous chemical compounds in an individual's breath. The challenge here is in developing accurate systems that can detect even small chemical traces

from mixtures of multiple gaseous compounds, so that diseases can be detected even during early stages.

Small-data intelligent mobile systems. Finally, there are opportunities to develop precision healthcare technologies to reduce inequities due to financial incentives, which often show up in the case of rare diseases. One such disease is apraxia which makes it difficult for a person to speak. Imagine being able to create a real-time translation system running on mobile systems that can allow individuals to communicate seamlessly with the outside world.

Today's speech recognition engines still require individuals with speech impediments to record hundreds to thousands of phrases over several hour-long recording sessions, which is a challenging task to undertake. This is in marked contrast to the short two-minute speech exercises conducted by speech pathologists to measure the full range of speech production mechanisms. In this space, there are opportunities to develop real-time machine learning models that run on mobile devices and are fine-tuned using orders of magnitude less data than current systems, using few-shot or zero-shot based learning methods; thus, enabling practical deployments of personalized AI systems. This is an example of a project where universities can play a unique role in demonstrating the technical path needed to address issues facing underserved populations.

Bibliography

[1] Adult hearing screening. *American Speech Language Hearing Association*. https://www.asha.org/practice-portal/professional-issues/adult-hearing-screening/.

[2] Otoacoustic emissions (OAE) machines. *e3 Diagnostics*. https://www.e3diagnostics.com/products/otoacoustic-emissions---oae.

[3] Welch Allyn OAE hearing screener. https://mdmaxx.com/products/oae-hearing-screener-with-printer?currency=USD&variant=38178860990655.

[4] Specifications For Instruments To Measure Aural Acoustic Impedance And Admittance (Aural Acoustic Immittance). *American National Standards Institute*, 1987. https://standards.globalspec.com/std/14273960/ansi-asa-s3-39.

[5] Joint committee on infant hearing year 2000 position statement: Principles and guidelines for early hearing detection and intervention. *Pediatrics*, 106:798–817, 01 2000.

[6] Universal newborn hearing screening. *American Family Physician*, 75:1349–52, 06 2007.

[7] Earcheck middle ear monitor use & care manual, Innovia Medical, Lenexa, KS, 2008.

[8] Newborn hearing screening protocol. *Iowa's Early Hearing Detection & Intervention Program*, 2008. https://hhs.iowa.gov/sites/default/files/portals/1/userfiles/77/newborn%20hearing%20screening%20protocol-11-08.pdf.

[9] School Nurses Guide to Otoacoustic Emissions (OAEs). https://www.schoolhealth.com/media/pdf/School%20Nurse%20Guide%20to%20OAEs_web%202016.pdf, 2009.

[10] Healthy people. *Centers for Disease Control and Prevention*, 2010. https://www.cdc.gov/nchs/healthy_people/hp2010.htm.

[11] Millions of people in the world have hearing loss that can be treated or prevented. *World Health Organization*, 2013. https://www.who.int/pbd/deafness/news/Millionslivewithhearingloss.pdf.

[12] Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations. *U.S. Food and Drug Administration*, 2014. http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf.

[13] Childhood Hearing Loss - Strategies for prevention and care. *World Health Organization*, 2016. https://apps.who.int/iris/bitstream/handle/10665/204632/9789241510325_eng.pdf.

[14] World Development Report, Digital Dividends - Overview. *World Bank Group*, 2016. http://documents1.worldbank.org/curated/en/96162146799469864 4/pdf/102724-WDR-WDR2016Overview-ENGLISH-WebResBox-394840B-OUO-9.pdf.

[15] Global Mobile Consumer Trends. *Deloitte*, 2017. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-global-mobile-consumer-survey-second-edition.pdf.

[16] What are examples of Software as a Medical Device? *U.S. Food and Drug Administration*, 2017. https://www.fda.gov/medical-devices/software-medical-device-samd/what-are-examples-software-medical-device.

[17] Cellscope. https://www.cellscope.com/, 2018.

[18] Earcheck middle ear monitor, Innovia Medical, Lenexa, KS, 2018.

[19] hearScope. https://www.hearxgroup.com/hearscope/, 2018.

[20] Novel pediatric hearing screening device. 2018. http://jhu.technologypublisher.com/technology/40355.

[21] r140 Portable OAE Hearing Screener. *Resonance Audiology*, 2018. http://www.resonance-audiology.com/wp-content/uploads/2018/05/R14O-TECNICAL-SPECIFICATION.pdf.

[22] Software as a Medical Device (SaMD). *U.S. Food and Drug Administration*, 2018. https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd.

[23] Connected Society: The State of Mobile Internet Connectivity. *GSMA*, 2019. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/07/GSMA-State-of-Mobile-Internet-Connectivity-Report-2019.pdf.

[24] Device Software Functions Including Mobile Medical Applications. *U.S. Food and Drug Administration*, 2019. https://www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications.

[25] Examples of Premarket Submissions that Include MMAs Cleared or Approved by the FDA. *U.S. Food and Drug Administration*, 2019. https://www.fda.gov/medical-devices/device-software-functions-inclu,ding-mobile-medical-applications/examples-premarket-submissions-include-mmas-cleared-or-approved-fda.

[26] Policy for Device Software Functions and Mobile Medical Applications. *U.S. Food and Drug Administration*, 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications.

[27] Welch Allyn® OAE Hearing Screener - Directions for Use. *Welch Allyn*, 2019. https://www.henryschein.com/assets/Medical/80022349LITPDF.pdf.

[28] Increase in fatal drug overdoses across the united states driven by synthetic opioids before and during the COVID-19 pandemic. *Centers for Disease Control and Prevention*, 2020. https://emergency.cdc.gov/han/2020/han00438.asp.

[29] OAEs: The difference between TEOAE, DPOAE and SFOAE. *Interacoustics*, 2020. https://www.youtube.com/watch?v=aoIVSyRqWwk.

[30] Generic Eartips from Grason Associates Fitment Guide Eartips. *Grason Associates*, 2021. https://www.emi-canada.com/new-page-5.

[31] HTC 10. *GSMArena*, 2021. https://www.gsmarena.com/htc_10-7884.php.

[32] Nura. 2021. https://www.nuraphone.com/.

[33] Otoglobal Health. 2021. https://otoglobalhealth.wixsite.com/companysite.

[34] OtoRead™ - Instructions for Use. *Interacoustics*, 2021. https://www.interacoustics.com/download/otoread/manuals-otoread/1236-instructions-for-use-otoread.

[35] Path Medical - Audiology | Diagnostic | Screening | Tracking. 2021. https://www.pathme.de/.

[36] Universal Serial Bus Power Delivery Specification Revision 2.0. 2021. https://www.usb.org/document-library/usb-power-delivery.

[37] 510(k) summary of safety and effectiveness: Titan with DPQAE44O and titan with ABRIS440. *U.S. Food and Drug Administration*, 2022. https://www.accessdata.fda.gov/cdrh_docs/pdf10/K103760.pdf.

[38] AVLinearPCMBitDepthKey. *Apple Developers*, 2022. https://developer.apple.com/documentation/avfaudio/avlinearpcmbitdepthkey.

[39] Basic of OAEs. *Otoemissions*, 2022. https://www.otoemissions.org/old/definitions/TEOAE2.html.

[40] Choose your AirPods Pro ear tips and use the Ear Tip Fit Test. *Apple Support*, 2022. https://support.apple.com/en-us/HT210633.

[41] Creating a TEOAE pass/refer protocol. *Interacoustics*, 2022. https://www.interacoustics.com/oae-devices/lyra/support/creating-a-teoae-pass-refer-protocol.

[42] Degree of hearing loss. *American Speech-Language-Hearing Association*, 2022. https://www.asha.org/public/hearing/degree-of-hearing-loss/.

[43] Device compatibility overview. *Android Developers*, 2022. https://developer.android.com/guide/practices/compatibility.

[44] For iPhone Headphone Adapter 3.5mm Jack Aux Cord Dongle Audio Cable Connector US. *eBay*, 2022. https://www.ebay.com/itm/313802818623.

[45] Frequently Asked Questions: EroScan OAE Test System. 2022. https://www.schoolhealth.com/media/pdf/51075_FAQ%20EroScan.pdf.

[46] Hearing test - Audiometry Tone, 2022. https://apps.apple.com/us/app/hearing-test-audiometry-tone/id1368396053.

[47] Hearing Test & Ear Age Test, 2022. https://apps.apple.com/us/app/hearing-test-ear-age-test/id1067630100.

[48] Mimi hearing test, 2022. https://apps.apple.com/us/app/mimi-hearing-test/id932496645.

[49] Table of sound pressure levels. 2022. http://www.sengpielaudio.com/TableOfSoundPressureLevels.htm.

[50] TEOAEs Test Procedures. *Otoemissions*, 2022. https://www.otoemissions.org/index.php/en/basics-of-oaes/teoaes/3-teoaes-test-procedures.

[51] TUNE: Towards Universal Newborn and Early Childhood Hearing Screening in Kenya. 2022. https://tune.cs.washington.edu/.

[52] USB Digital Audio. *Android Developers*, 2022. https://source.android.com/devices/audio/usb.

[53] Why is a hearing screening important for my baby? *Centers for Disease Control and Prevention*, 2023. https://www.cdc.gov/ncbddd/hearingloss/features/baby-hearing-screening.html.

[54] A. M. Schultz, M. A. McCoy, R. Graham. *Strategies to improve cardiac arrest survival: a time to act*. National Academies Press, 2015.

[55] C. Abdala and L. Visser-Dumont. Distortion product otoacoustic emissions: a tool for hearing assessment and scientific study. *The Volta Review*, 103(4):281, 2001.

[56] J. Acuin. Chronic suppurative otitis media: Burden of illness and management options. *World Health Organization*, 2004.

[57] J. Alaerts, H. Luts, and J. Wouters. Evaluation of middle ear function in young children: clinical guidelines for the use of 226-and 1,000-Hz tympanometry. *Otology & Neurotology*, 28(6):727–732, 2007.

[58] L. S. Alvord and B. L. Farmer. Anatomy and orientation of the human external ear. *Journal of the American Academy of Audiology*, 8:383–390, 1997.

[59] J. M. Bland and D. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476):307–310, 1986.

[60] R. T. Boone, C. M. Bower, and P. F. Martin. Failed newborn hearing screens as presentation for otitis media with effusion in the newborn population. *International Journal of Pediatric Otorhinolaryngology*, 69(3):393–397, 2005.

[61] D. Bowyer et al. *Complications of otitis media*. CRC Press, 2015.

[62] K. Bright, C. Greeley, J. Eichwald, C. Loveland, and G. Tanner. American Academy of Audiology childhood hearing screening guidelines. *Reston, VA: American Academy of Audiology Task Force*, 2011.

[63] D. S. Budnitz, M. C. Lovegrove, N. Shehab, and C. L. Richards. Emergency hospitalizations for adverse drug events in older Americans. *The New England Journal of Medicine*, 365(21):2002–2012, 2011.

[64] N. Bui, N. Pham, J. J. Barnitz, Z. Zou, P. Nguyen, H. Truong, T. Kim, N. Farrow, A. Nguyen, J. Xiao, R. Deterding, T. Dinh, and T. Vu. EBP: A wearable system for frequent and comfortable blood pressure monitoring from user's ear. In *MobiCom '19*.

[65] L. J. Campbell and K. D. Slater. Personalization of auditory stimulus, Patent US10154333B2, Dec. 2018.

[66] J. Chan, N. Ali, A. Najafi, A. Meehan, L. Mancl, E. Gallagher, R. Bly, and S. Gollakota. An off-the-shelf otoacoustic-emission probe for hearing screening via a smartphone. *Nature Biomedical Engineering*, 6:1–11, 10 2022.

[67] J. Chan, A. Glenn, M. Itani, L. R. Mancl, E. Gallagher, R. Bly, S. Patel, and S. Gollakota. Wireless earbuds for low-cost hearing screening. In *MobiSys '23*.

[68] J. Chan and S. Gollakota. Inner-ear cochlea testing with earphones. In *MobiSys demo '22*.

[69] J. Chan, V. Iyer, A. Wang, A. Lyness, P. Kooner, J. Sunshine, and S. Gollakota. Closed-loop wearable naloxone injector system. *Scientific Reports*, 11(1):22663, 2021.

[70] J. Chan, K. Michaelsen, J. K. Estergreen, D. E. Sabath, and S. Gollakota. Micro-mechanical blood clot testing using smartphones. *Nature communications*, 13(1):1–12, 2022.

[71] J. Chan, A. Najafi, M. Baker, J. Kinsman, L. R. Mancl, S. Norton, R. Bly, and S. Gollakota. Performing tympanometry using smartphones. *Communications Medicine*, 2(1):57, 2022.

[72] J. Chan, A. Najafi, M. Baker, J. Kinsman, L. R. Mancl, S. Norton, R. Bly, and S. Gollakota. Performing tympanometry using smartphones, Mar. 2022. https://github.com/uw-x/tymp, https://doi.org/10.5281/zenodo.6389337.

[73] J. Chan, S. Raju, R. Nandakumar, R. Bly, and S. Gollakota. Detecting middle ear fluid using smartphones. *Science Translational Medicine*, 11(492), 2019.

[74] J. Chan, T. Rea, S. Gollakota, and J. E. Sunshine. Contactless cardiac arrest detection using smart devices. *NPJ digital medicine*, 2(1):52, 2019.

[75] R. H. Chan and B. McPherson. Test-retest reliability of tone-burst-evoked otoacoustic emissions. *Acta otolaryngologica*, 120(7):825–834, 2000.

[76] E. Chapin and S. Doocy. International short-term medical service trips: guidelines from the literature and perspectives from the field. *World Health & Population*, 12(2):43–53, 2010.

[77] K. K. Charaziak and C. A. Shera. Compensating for ear-canal acoustics when measuring otoacoustic emissions. *The Journal of the Acoustical Society of America*, 141(1):515–531, 2017.

[78] I. Chatterjee, M. Kim, V. Jayaram, S. Gollakota, I. Kemelmacher, S. Patel, and S. M. Seitz. Clearbuds: Wireless binaural earbuds for learning-based speech enhancement. In *MobiSys '22*, 2022.

[79] J. T. Combs and M. K. Combs. Acoustic reflectometry: spectral analysis and the conductive hearing loss of otitis media. *The Pediatric Infectious Disease Journal*, 15(8):683–686, 1996.

[80] L. M. de Resende, J. dos Santos Ferreira, S. A. da Silva Carvalho, I. S. Oliveira, and I. B. Bassi. Tympanometry with 226 and 1000 Hertz tone probes in infants. *Brazilian Journal of otorhinolaryngology*, 78(1):95–102, 2012.

[81] R. Dsouza, J. Won, G. L. Monroy, M. C. Hill, R. G. Porter, M. A. Novak, and S. A. Boppart. In vivo detection of nanometer-scale structural changes of the human tympanic membrane in otitis media. *Scientific Reports*, 8(1):1–13, 2018.

[82] R. Dsouza, J. Won, G. L. Monroy, D. R. Spillman Jr, and S. A. Boppart. Economical and compact briefcase spectral-domain optical coherence tomography system for primary care and point-of-care applications. *Journal of Biomedical Optics*, 23(9):096003, 2018.

[83] M. Emadi, M. Rezaei, M. H. Nahrani, and M. Bolandi. High frequency tympanometry (1,000 Hz) for neonates with normal and abnormal transient evoked otoacoustic emissions. *Journal of Audiology & Otology*, 20(3):153, 2016.

[84] S. Epstein, T. Finitzo, A. Erenberg, and N. Roizen. Joint committee on infant hearing 1994 position statement. *American Speech-Language-Hearing Association*, 5, 01 1991.

[85] N. Erkkola-Anttinen, H. Irjala, M. K. Laine, P. A. Tähtinen, E. Löyttyniemi, and A. Ruohola. Smartphone otoscopy performed by parents. *Telemedicine and e-Health*, 25(6):477–484, 2019.

[86] X. Fan, L. Shangguan, S. Rupavatharam, Y. Zhang, J. Xiong, Y. Ma, and R. Howard. Headfi: Bringing intelligence to all headphones. In *MobiCom '21*.

[87] A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser. Optical coherence tomography-principles and applications. *Reports on Progress in Physics*, 66(2):239, 2003.

[88] J. Funk. Technology change, economic feasibility, and creative destruction: the case of new electronic products and services. *Industrial and Corporate Change*, 27(1):65–82, 2018.

[89] V. Fuster, L. E. Rydén, D. S. Cannom, H. J. Crijns, A. B. Curtis, K. A. Ellenbogen, J. L. Halperin, J.-Y. Le Heuzey, G. N. Kay, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation–executive summary: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients with Atrial Fibrillation) developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *European Heart Journal*, 27(16):1979–2030, 2006.

[90] M. Gaihede and F. Marker. Agreement between two tympanometers: A methodological study of instrument comparison. *Scandinavian Audiology*, 27(2):113–119, 1998.

[91] S. Garg, R. Singh, and D. Khurana. Infant hearing screening in india: Current status and way forward. *International Journal of Preventive Medicine*, 6:113, 11 2015.

[92] S. S. Goodman, D. F. Fitzpatrick, J. C. Ellison, W. Jesteadt, and D. H. Keefe. High-frequency click-evoked otoacoustic emissions and behavioral thresholds in humans. *The Journal of the Acoustical Society of America*, 125(2):1014–1032, 2009.

[93] L. M. Haile, K. Kamenov, P. S. Briant, A. U. Orji, J. D. Steinmetz, A. Abdoli, M. Abdollahi, E. Abu-Gharbieh, A. Afshin, H. Ahmed, et al. Hearing loss prevalence and years lived with disability, 1990–2019: findings from the global burden of disease study 2019. *The Lancet*, 397(10278):996–1009, 2021.

[94] J. W. Hall III. The important role of otoacoustic emissions (OAEs) in preschool hearing screening. *Audiology Online*, 2018. https://www.audiologyonline.com/articles/important-role-otoacoustic-emissions-oaes-22091.

[95] N. Heitmann, P. Kindt, T. Rosner, K. Sikka, A. Chirom, D. Kalyanasundaram, and S. Chakraborty. Sound4All: Towards affordable large-scale hearing screening. In *DTIS '17*, pages 1–6. IEEE, 2017.

[96] N. Heitmann, T. Rosner, and S. Chakraborty. Designing a single speaker-based ultra low-cost otoa-

coustic emission hearing screening probe. In *2020 IEEE Global Humanitarian Technology Conference (GHTC)*, pages 1–8. IEEE.

[97] N. Heitmann, T. Rosner, and S. Chakraborty. Mass-deployable smartphone-based objective hearing screening with otoacoustic emissions. In *International Conference on Multimodal Interaction (ICMI) '21*.

[98] G. D. Hoople. Otitis media with effusion—a challenge to otolaryngology. *The Laryngoscope*, 60(4):315–329, 1950.

[99] D. M. Hussain, M. P. Gorga, S. T. Neely, D. H. Keefe, and J. Peters. Transient evoked otoacoustic emissions in patients with normal hearing and in patients with hearing loss. *Ear and Hearing*, 19(6), 1998.

[100] J. Chianese, A. Hoberman, J. L. Paradise, D. Colborn, K. Kathleen, R. Diana, E. Howard, M. Kurs-Lasky. Spectral gradient acoustic reflectometry compared with tympanometry in diagnosing middle ear effusion in children aged 6 to 24 months. *Archives of pediatrics & adolescent medicine*, 161(9):884–888, 2007.

[101] J. Jerger. Clinical experience with impedance audiometry. *Archives of Otolaryngology*, 92(4):311–324, 1970.

[102] Y. Jin, Y. Gao, X. Guo, J. Wen, Z. Li, and Z. Jin. Earhealth: An earphone-based acoustic otoscope for detection of multiple ear diseases in daily life. In *MobiSys '22*.

[103] P. H. Karma, M. A. Penttilä, M. M. Sipilä, and M. J. Kataja. Otoscopic diagnosis of middle ear effusion in acute and non-acute otitis media. i. the value of different otoscopic findings. *International journal of pediatric otorhinolaryngology*, 17(1):37–49, 1989.

[104] F. Kawsar, C. Min, A. Mathur, and A. Montanari. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing*, 17(3):83–89, 2018.

[105] M. Kearns and D. Ron. Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural Computation*, 11(6):1427–1453, 1999.

[106] D. T. Kemp, P. Bray, L. Alexander, and A. M. Brown. Acoustic emission cochleography–practical aspects. *Scandinavian audiology.*, 25, 1986.

[107] S. Kimball. Acoustic reflectometry: spectral gradient analysis for improved detection of middle ear effusion in children. *The Pediatric Infectious Disease Journal*, 17(6):552–555, 1998.

[108] G. King, M. Roland-Mieszkowski, T. Jason, and D. G. Rainham. Noise levels associated with urban land use. *J Urban Health, 2012*.

[109] R. M. Lampe, M. R. Weir, J. Spier, and M. F. Rhodes. Acoustic reflectometry in the detection of middle ear effusion. *Pediatrics*, 76(1):75–78, 1985.

[110] N. Lezzoum, G. Gagnon, and J. Voix. Voice activity detection system for smart earphones. *IEEE Transactions on Consumer Electronics*, 2014.

[111] G. Liden, E. Harford, and O. Hallen. Automatic tympanometry in clinical practice. *Audiology*, 13(2):126–139, 1974.

[112] G. Lidén, E. Harford, and O. Hallén. Tympanometry for the diagnosis of ossicular disruption. *Archives of Otolaryngology*, 99(1):23–29, 1974.

[113] G. Liden, J. Peterson, and G. Bjorkman. Tympanometry a method for analysis of middle—ear function. *Acta Oto-Laryngologica*, 69(sup263):218–224, 1970.

[114] A. S. Lieberthal, A. E. Carroll, T. Chonmaitree, T. G. Ganiats, A. Hoberman, M. A. Jackson, M. D. Joffe, D. T. Miller, R. M. Rosenfeld, X. D. Sevilla, et al. The diagnosis and management of acute otitis media. *Pediatrics*, 131(3):e964–e999, 2013.

[115] H. Lindén, H. Teppo, and M. Revonta. Spectral gradient acoustic reflectometry in the diagnosis of middle-ear fluid in children. *European Archives of Otorhinolaryngology*, 264(5):477–481, 2007.

[116] J. Lous. Why use tympanometry in general practice: A review. *World Journal of Otorhinolaryngology*, 5(2):53–57, 2015.

[117] T. Lumsden. Observations on the respiratory centres in the cat. *The Journal of Physiology*, 57(3-4):153–160, 1923.

[118] D. D. Mahoney, W. P. Sjursen, W. J. Staab, and M. Leedom. One-size-fits-all uni-ear hearing instrument, Aug. 15 2006. US Patent 7,092,543.

[119] N. C. Malone, M. M. Williams, M. C. S. Fawzi, J. Bennet, C. Hill, J. N. Katz, and N. E. Oriol. Mobile Health Clinics in the United States. *International Journal for Equity in Health*, 19(1):1–9, 2020.

[120] R. Mandavia, T. Lapa, M. Smith, and M. Bhutta. A cross-sectional evaluation of the validity of a smartphone otoscopy device in screening for ear disease in Nepal. *Clinical Otolaryngology*, 43(1):31–38, 2018.

[121] R. H. Margolis, S. Bass-Ringdahl, W. D. Hanks, L. Holte, and D. A. Zapala. Tympanometry in newborn infants—1 kHz norms. *Journal of the American Academy of Audiology*, 14(7):383–392, 2003.

[122] R. H. Margolis and J. W. Heller. Screening tympanometry: criteria for medical referral: original papers. *Audiology*, 26(4):197–208, 1987.

[123] M. Mason, S. B. Welch, P. Arunkumar, L. A. Post, and J. M. Feinglass. Notes from the field: Opioid overdose deaths before, during, and after an 11-week COVID-19 stay-at-home order—Cook county, Illinois, January 1, 2018–October 6, 2020. *Morbidity and Mortality Weekly Report*, 70(10):362, 2021.

[124] D. F. McCarter, A. U. Courtney, and S. M. Pollart. Cerumen impaction. *American Family Physician*, 75(10), 2007.

[125] B. McPherson. Newborn hearing screening in developing countries: Needs & new directions. *The Indian Journal of Medical Research*, 135(2):152, 2012.

[126] C. Min, A. Mathur, and F. Kawsar. Exploring audio and kinetic sensing on earable devices. In *WearSys '18*.

[127] G. L. Monroy, J. Won, R. Dsouza, P. Pande, M. C. Hill, R. G. Porter, M. A. Novak, D. R. Spillman, and S. A. Boppart. Automated classification platform for the identification of otitis media using optical coherence tomography. *NPJ digital medicine*, 2(1):1–11, 2019.

[128] J. C. Mwita, J. M. Francis, A. A. Oyekunle, M. Gaenamong, M. Goepamang, and M. G. Magafu. Quality of anticoagulation with warfarin at a tertiary hospital in botswana. *Clinical and Applied Thrombosis/Hemostasis*, 24(4):596–601, 2018.

[129] A. Myat, K.-J. Song, and T. Rea. Out-of-hospital cardiac arrest: current concepts. *The Lancet*, 391(10124):970–979, 2018.

[130] T. Naidoo. *Audiological Pactice and Service Delivery in South Africa*. PhD thesis, University of the Witwatersrand, 2006.

[131] J. Neumann, S. Uppenkamp, and B. Kollmeier. Chirp evoked otoacoustic emissions. *Hearing research*, 79(1-2):17–25, 1994.

[132] E. Onusko. Tympanometry. *American Family Physician*, 70(9):1713–1720, 2004.

[133] M. Oyewumi, M. G. Brandt, B. Carrillo, A. Atkinson, K. Iglar, V. Forte, and P. Campisi. Objective evaluation of otoscopy skills among family and community medicine, pediatric, and otolaryngology residents. *Journal of Surgical Education*, 73(1):129–135, 2016.

[134] A. Palmu, H. Puhakka, T. Rahko, and A. K. Takala. Diagnostic value of tympanometry in infants in clinical practice. *International journal of pediatric otorhinolaryngology*, 49(3):207–213, 1999.

[135] T. K. Parthasarathy and B. Klostermann. Similarities and differences in distortion-product otoacoustic emissions among four FDA-approved devices. *Journal of the American Academy of Audiology*, 12(8), 2001.

[136] H. Patel and M. Feldman. Universal newborn hearing screening. *Paediatrics & Child Health*, 16(5):301–305, 2011.

[137] S. I. Pelton. Otoscopy for the diagnosis of otitis media. *The Pediatric Infectious Disease Journal*, 17(6):540–543, 1998.

[138] M. E. Pichichero. Diagnostic accuracy, tympanocentesis training performance, and antibiotic selection by pediatric residents in management of otitis media. *Pediatrics*, 110(6):1064–1070, 2002.

[139] C. F. Poets, R. G. Meny, M. R. Chobanian, and R. E. Bonofiglo. Gasping and other cardiorespiratory patterns during sudden infant deaths. *Pediatric Research*, 45(3):350, 1999.

[140] M.-Z. Poh, K. Kim, A. Goessling, N. Swenson, and R. Picard. Cardiovascular monitoring using earphones and a mobile device. *IEEE Pervasive Computing*, 11(4):18–26, 2012.

[141] J. Prakash, Z. Yang, Y.-L. Wei, H. Hassanieh, and R. R. Choudhury. Earsense: Earphones as a teeth activity sensor. In *MobiCom '20*.

[142] A. Qureishi, Y. Lee, K. Belfield, J. P. Birchall, and M. Daniel. Update on otitis media–prevention and treatment. *Infection and drug resistance*, 7:15, 2014.

[143] R. M. Rosenfeld, J. J. Shin, S. R. Schwartz, R. Coggins, L. Gagnon, J. M. Hackell, D. Hoelting, L. L. Hunter, A. W. Kummer, S. C. Payne, D. S. Poe, M. Veling, P. M. Vila, S. A. Walsh, M. D. Corrigan. Clinical practice guideline: otitis media with effusion (update). *Otolaryngology–Head and Neck Surgery*, 154(1_suppl):S1–S41, 2016.

[144] H. Ramatsoma and D. Koekemoer. Validation of a bilateral simultaneous computer-based tympanometer. *American Journal of Audiology*, 29(3):491–503, 2020.

[145] T. D. Rea. Agonal respirations during cardiac arrest. *Current Opinion in Critical Care*, 11(3):188–191, 2005.

[146] K. M. Reavis, G. McMillan, D. Austin, F. Gallun, S. A. Fausti, J. S. Gordon, W. J. Helt, and D. Konrad-Martin. Distortion-product otoacoustic emission test performance for ototoxicity monitoring. *Ear and Hearing*, 32(1):61, 2011.

[147] M. Renko, T. Kontiokari, K. Jounio-Ervasti, H. Rantala, and M. Uhari. Disappearance of middle ear effusion in acute otitis media monitored daily with tympanometry. *Acta Paediatrica*, 95(3):359–363, 2006.

[148] T. Röddiger, D. Wolffram, D. Laubenstein, M. Budde, and M. Beigl. Towards respiration rate monitoring using an in-ear headphone inertial measurement unit. In *EarComp '19*.

[149] S. Rose. Machine learning for prediction in electronic health data. *JAMA Internal Medicine*, Aug 2018.

[150] S. Rupavatharam and M. Gruteser. Towards in-ear inertial jaw clenching detection. In *EarComp '19*.

[151] T. Röddiger, T. King, D. R. Roodt, C. Clarke, and M. Beigl. OpenEarable: Open hardware earable sensing platform. In *EarComp '22*.

[152] S. L. Block, E. Mandel, S. Mclinn, M. E. Pichichero, S. Bernstein, S. Kimball, J. Kozikowski. Spectral gradient acoustic reflectometry for the detection of middle ear effusion by pediatricians and parents. *The Pediatric Infectious Disease Journal*, 17(6):560–564, 1998.

[153] A. Schwitters, P. Lederer, L. Zilversmit, P. S. Gudo, I. Ramiro, L. Cumba, E. Mahagaja, and K. Jobarteh. Barriers to health care in rural Mozambique: a rapid ethnographic assessment of planned mobile health clinics for ART. *Global Health: Science and Practice*, 3(1):109–116, 2015.

[154] J. R. Semakula, J. P. Mouton, A. Jorgensen, C. Hutchinson, S. Allie, L. Semakula, N. French, M. Lamorde, C.-H. Toh, M. Blockman, et al. A cross-sectional evaluation of five warfarin anticoagulation services in Uganda and South Africa. *PLOS ONE*, 15(1):e0227458, 2020.

[155] M. U. Shah, M. Sohal, T. A. Valdez, and C. R. Grindle. iPhone otoscopes: Currently available, but reliable for tele-otoscopy in the hands of parents? *International Journal of Pediatric Otorhinolaryngology*, 106:59–63, 2018.

[156] J. Shanks and J. Shohet. Tympanometry in clinical practice. *Handbook of Clinical Audiology*, 6:159, 2009.

[157] E. Shaw and R. Teranishi. Sound pressure generated in an external-ear replica and real human ears by a nearby point source. *The Journal of the Acoustical Society of America*, 44(1):240–249, 1968.

[158] J. Shinn, A. Jayawardena, A. Patro, M. G. Zuniga, and J. Netterville. Teacher prescreening for hearing loss in the developing world. *Ear, Nose Throat Journal*, 100:014556131988038, 10 2019.

[159] J. Siegel. Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes. *The Journal of the Acoustical Society of America*, 95(5):2589–2597, 1994.

[160] J. Siegel and E. Hirohata. Sound calibration and distortion product otoacoustic emissions at high frequencies. *Hearing Research*, 80(2):146–152, 1994.

[161] D. Singer, A. Hellkamp, J. Piccini, K. Mahaffey, Y. Lokhnygina, G. Pan, J. Halperin, R. Becker, G. Breithardt, G. Hankey, W. Hacke, C. Nessel, M. Patel, R. Califf, and K. Fox. Impact of global geographic region on time in therapeutic range on warfarin anticoagulant therapy: Data from the rocket af clinical trial. *Journal of the American Heart Association*, 2:e000067, 12 2013.

[162] A. S. Sood, C. S. Bons, and G. S. Narang. High frequency tympanometry in neonates with normal otoacoustic emissions: measurements and interpretations. *Indian Journal of Otolaryngology and Head & Neck Surgery*, 65(3):237–243, 2013.

[163] W. Stephanie, C. Hill, M. L. Ricks, J. Bennet, and N. E. Oriol. The scope and impact of mobile health clinics in the United States: a literature review. *International Journal for Equity in Health*, 16(1):1–12, 2017.

[164] R. Stoessel, N. Krohn, K. Pfleiderer, and G. Busse. Air-coupled ultrasound inspection of various materials. *Ultrasonics*, 40(1-8):159–163, 2002.

[165] D. W. Swanepoel, J. L. Clark, D. Koekemoer, J. W. Hall III, M. Krumm, D. V. Ferrari, B. McPherson, B. O. Olusanya, M. Mars, I. Russo, et al. Telehealth in audiology: The need and potential to reach underserved communities. *International Journal of Audiology*, 49(3):195–202, 2010.

[166] K. J. Sykes. Short-term medical service trips: a systematic review of the evidence. *American Journal of Public Health*, 104(7):e38–e48, 2014.

[167] T. Muderris, A. Yazıcı, S. Bercin, G. Yalçıner, E. Sevil, M. Kırıs . Consumer acoustic reflectometry: accuracy in diagnosis of otitis media with effusion in children. *International Journal of Pediatric Otorhinolaryngology*, 77(10):1771–1774, 2013.

[168] D. W. Teele and J. Teele. Detection of middle ear effusion by acoustic reflectometry. *The Journal of Pediatrics*, 104(6):832–838, 1984.

[169] D. C. Thompson, H. McPhillips, R. L. Davis, T. A. Lieu, C. J. Homer, and M. Helfand. Universal newborn hearing screening: summary of evidence. *Jama*, 286(16):2000–2010, 2001.

[170] P. Torre III, K. J. Cruickshanks, D. M. Nondahl, and T. L. Wiley. Distortion product otoacoustic emission response characteristics in older adults. *Ear and hearing*, 24(1):20–29, 2003.

[171] S. E. Wakeman, T. C. Green, and J. Rich. An overdose surge will compound the COVID-19 pandemic if urgent action is not taken. *Nature Medicine*, 26(6):819–820, 2020.

[172] J. J. Walker, L. M. Cleveland, J. L. Davis, and J. S. Seales. Audiometry screening and interpretation. *American Family Physician*, 87(1):41–47, 2013.

[173] I. Williamson. Otitis media with effusion in children. *BMJ Clinical Evidence*, 2015, 2015.

[174] R. H. Wilson, K. P. Nadeau, F. B. Jaworski, B. J. Tromberg, and A. J. Durkin. Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization. *Journal of Biomedical Optics*, 20(3):030901, 2015.

[175] P. Wormald, G. Browning, and K. Robinson. Is otoscopy reliable? a structured teaching method to improve otoscopic accuracy in trainees. *Clinical Otolaryngology & Allied Sciences*, 20(1):63–67, 1995.

[176] J. Zemplenyi, S. Gilman, and D. Dirks. Optical method for measurement of ear canal length. *The Journal of the Acoustical Society of America*, 78(6):2146–2148, 1985.