



## Web- and app-based tools for remote hearing assessment: a scoping review

Ibrahim Almufarrij, Harvey Dillon, Piers Dawes, David R. Moore, Wai Yeung, Anna-Pavlina Charalambous, Chryssoula Thodi & Kevin J. Munro

To cite this article: Ibrahim Almufarrij, Harvey Dillon, Piers Dawes, David R. Moore, Wai Yeung, Anna-Pavlina Charalambous, Chryssoula Thodi & Kevin J. Munro (2022): Web- and app-based tools for remote hearing assessment: a scoping review, International Journal of Audiology, DOI: [10.1080/14992027.2022.2075798](https://doi.org/10.1080/14992027.2022.2075798)

To link to this article: <https://doi.org/10.1080/14992027.2022.2075798>



© 2022 The Authors. Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of British Society of Audiology, International Society of Audiology, and Nordic Audiological Society.



View supplementary material [↗](#)



Published online: 09 Jun 2022.



Submit your article to this journal [↗](#)



Article views: 2543



View related articles [↗](#)









View Crossmark data [↗](#)

REVIEW ARTICLE



## Web- and app-based tools for remote hearing assessment: a scoping review

Ibrahim Almufarrij<sup>a,b</sup> , Harvey Dillon<sup>a,c</sup> , Piers Dawes<sup>a,d</sup> , David R. Moore<sup>a,e</sup> , Wai Yeung<sup>f</sup> ,  
Anna-Pavlina Charalambous<sup>g</sup>, Chryssoula Thodi<sup>g</sup> and Kevin J. Munro<sup>a,h</sup> 

<sup>a</sup>Manchester Centre for Audiology and Deafness, School of Health Sciences, University of Manchester, Manchester, UK; <sup>b</sup>Department of Rehabilitation Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia; <sup>c</sup>Department of Linguistics, Macquarie University, Sydney, Australia; <sup>d</sup>School of Health and Rehabilitation Sciences, University of Queensland, Brisbane, Australia; <sup>e</sup>Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; <sup>f</sup>Division of Population Health, Health Services Research & Primary Care, University of Manchester, Manchester, UK; <sup>g</sup>Department of Health Sciences, European University Cyprus, Nicosia, Cyprus; <sup>h</sup>Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

### ABSTRACT

**Objective:** Remote hearing screening and assessment may improve access to, and uptake of, hearing care. This review, the most comprehensive to date, aimed to (i) identify and assess functionality of remote hearing assessment tools on smartphones and online platforms, (ii) determine if assessed tools were also evaluated in peer-reviewed publications and (iii) report accuracy of existing validation data.

**Design:** Protocol was registered in INPLASY and reported according to PRISMA-Extension for Scoping Reviews.

**Study sample:** In total, 187 remote hearing assessment tools (using tones, speech, self-report or a combination) and 101 validation studies met the inclusion criteria. Quality, functionality, bias and applicability of each app were assessed by at least two authors.

**Results:** Assessed tools showed considerable variability in functionality. Twenty-two (12%) tools were peer-reviewed and 14 had acceptable functionality. The validation results and their quality varied greatly, largely depending on the category of the tool.

**Conclusion:** The accuracy and reliability of most tools are unknown. Tone-producing tools provide approximate hearing thresholds but have calibration and background noise issues. Speech and self-report tools are less affected by these issues but mostly do not provide an estimated pure tone audiogram. Predicting audiograms using filtered language-independent materials could be a universal solution.

### ARTICLE HISTORY

Received 6 January 2022

Revised 3 May 2022

Accepted 5 May 2022

### KEYWORDS

App-based; web-based; remote assessment; hearing assessment; online hearing tests



## Introduction


Hearing loss is a leading cause of years lived with disability (World Health Organization 2018), affecting nearly half a billion individuals worldwide (World Health Organization 2020). It causes communication difficulties and is associated with listening effort and fatigue, poor social interactions, anxiety, isolation and loneliness, depression and poor mental health, cognitive decline, increased risk of dementia and reduced quality of life (Cosh et al. 2018; Davis et al. 2007; Ferguson et al. 2017; Loughrey et al. 2018; Maharani et al. 2020; Maharani, Pendleton, and Leroi 2019; McGarrigle et al. 2014). Hearing aids, the most common intervention for hearing loss, can improve the user's quality of life (Ferguson et al. 2017). Nonetheless, it is estimated that only 17% of people who could benefit from hearing aids use them (World Health Organization 2020).

Hearing services are typically based in hospitals or clinics and staffed by registered hearing-health professionals. However, these services are often unavailable due to the shortage of qualified

professionals and resources in many regions around the world, particularly low- and middle-income countries (Mulwafu et al. 2017; World Health Organization 2021). Safety restrictions, such as physical distancing and protracted lockdowns, which are usually imposed during pandemics (e.g. coronavirus disease 2019), prohibit or limit the provision of elective clinic-based health services so as to reduce the risk of infection (Moynihan et al. 2021). Additionally, older adults form the majority requiring hearing services and this population is at high risk for coronavirus-related morbidity and mortality (Grasselli et al. 2020; Wu and McGoogan 2020). Thus, alternative hearing-health delivery models with minimal physical interaction are necessary (Swanepoel and Hall 2020).

As technologies advance and the number of smartphone subscriptions increases (reaching roughly 6.3 billion subscriptions worldwide in 2021; Ericsson 2021), remote health services via smartphone applications (apps) may improve access to and uptake of health care (Agarwal et al. 2016; Wilson et al. 2017). These apps may also meet the demands for hearing services and

**CONTACT** Ibrahim Almufarrij  [ialmufarrij@ksu.edu.sa](mailto:ialmufarrij@ksu.edu.sa)  Department of Rehabilitation Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/14992027.2022.2075798>.

© 2022 The Authors. Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of British Society of Audiology, International Society of Audiology, and Nordic Audiological Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

offer alternative service-delivery routes for those at high risk of coronavirus-related morbidity and mortality (Swanepoel and Hall 2010).

Many hearing-related smartphone apps have been developed and are commercially available in app stores. Bright and Pallawela (2016) conducted one of the earliest reviews identifying English smartphone hearing screening apps, which were evaluated in peer-reviewed publications. They also assessed the methodological quality of the studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2; Whiting et al. 2011). The reviewers identified 30 hearing screening apps, 24 of which had not been evaluated in peer-reviewed studies. The validation studies were of relatively low quality since most had a high risk of bias or/and concerns over applicability.

Since Bright and Pallawela's review, the development of smartphone apps has accelerated for various reasons, including the implementation of coronavirus disease safety measures (Keesara, Jonas, and Schulman 2020). Indeed, between 2019 and 2020, the number of Google Play medical smartphone apps increased by 12% (Appfigures 2021). This provoked the need for an updated review of remote hearing assessment tools to support hearing-health specialists maintain their services. As a quick response to such a need, Irace et al. (2021) updated Bright and Pallawela's review. They identified 44 adult English smartphone hearing assessment apps, seven of which were evaluated in peer-reviewed publications. However, the review lacked a standard critical evaluation of the apps and diagnostic accuracy studies. In addition, numerous web-based tools have also been developed to screen hearing remotely (Leensen and Dreschler 2013) and these have yet to be aggregated and evaluated systematically.

The aims of this review were:

- i. Identify and assess the functionality of commercially available remote hearing assessment tools on app stores and online platforms;
- ii. Systematically search the literature to determine which of the identified tools have been evaluated in peer-reviewed publications and
- iii. Reporting on the accuracy and reliability of these validations.

## Methods

The protocol for this scoping review was registered in the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY; INPLASY2020100073). The review methods were selected according to the Preferred Reporting Items for Systematic Review and Meta-Analyses – Extension for Scoping Reviews (PRISMA-ScR) guidelines. A scoping review was used because this is a broad topic and there are no published guidelines for the conducting and reporting of manuscripts with systematic searches of app stores (Arksey and O'Malley 2005; Grainger et al. 2020).

## Eligibility criteria

For Aim 1 (functionality): The review included any tool intended to screen or assess hearing ability remotely (measured as pure-tone hearing thresholds, ability to understand speech with background noise, or as a self- or caregiver-reported hearing disability or handicap) and was available on online platforms or via smartphone UK app stores. The tools should be self-administered or remotely controlled via a hearing professional. Non-

English tools were translated to English using the Google Translate app and were included when deemed eligible. Tools identifying or assessing other auditory abilities (e.g. temporal resolution) and ear-related disorders (e.g. tinnitus, hyperacusis, auditory processing disorder and balance) were excluded. Additionally, tools that were not available in the UK app stores required special registration numbers (i.e. designed for a particular institution or clinical study) or equipment (other than those routinely used with smartphones) or that did not provide the results of the screen/assessment were excluded.

For Aims 2 and 3 (accuracy and reliability), Diagnostic accuracy studies of any of the tools on human participants (irrespective of their age) were included. The primary outcomes of interest were sensitivity (the proportion of those with hearing difficulties who fail the screen) and specificity (the proportion of adults with no hearing difficulties who pass the screen) measures. Other relevant outcomes (e.g. the correlation between the remote hearing test and reference standard) were also included. Randomised and non-randomised controlled trials were eligible for inclusion. Theses, conference abstracts, clinical guidelines and book chapters were excluded.

## Information sources

For Aim 1 (functionality), Online platforms and application stores were systematically searched in November 2020 to identify relevant tools. The Google Search engine was used to identify web-based tools; only the first 100 hits, sorted by relevance, were searched. Apple App Store and Google Play were searched to identify app-based tools. These platforms were selected because they have the highest share of the global market and are the most commonly searched app stores (StatCounter 2021a). Since changing the user-location settings in app stores can potentially omit or reveal certain tools, this review used the UK/EU as the primary user-location in all stores.

For Aims 2 and 3 (accuracy and reliability), Relevant published, completed (but yet to be published) and ongoing validation studies were identified through a systematic literature search. The following databases were searched on 9 November 2020: EMBASE, EMCare, PubMed, PsycINFO, the Cochrane Library, Global Health and Web of Science. Preprint resources, including MedRxiv and PsyArXiv, were searched. The citations of the identified studies were tracked, and their reference lists were screened to identify additional relevant studies.

No search restrictions were imposed with regard to the participant's age, publication date, status or language.

## Search strategy

The review team developed the search protocol in consultation with a medical information specialist. The search strategy consisted of controlled terms (e.g. Medical Subject Headings) and free text words, where appropriate. An iterative process was conducted to test the proposed strategies. The final search strategies are reported in [Supplement Table 1](#).

## Data management

Retrieved records from all databases were exported to reference management software (i.e. EndNote). The same software was used to remove duplicates automatically. The surviving e-records

were then exported to an Excel spreadsheet for eligibility screening.

### **Selection process**

For Aim 1 (functionality), the tool's names (and descriptions, when needed) were independently screened by two authors to assess eligibility. Eligible tools were downloaded and installed on Huawei Mate 10 Pro (Android OS version 10) or iPhone 6 (iOS 12.4.8) for further screening.

For Aims 2 and 3 (accuracy and reliability), the titles and abstracts of the retrieved records were screened against the inclusion and exclusion criteria by two independent researchers. Studies that passed the initial screening stage were inspected by two independent researchers. Reasons for exclusion were documented. Disagreements, which accounted for 3%, were resolved by discussion.

### **Data collection process and data items**

The data extraction process was conducted by the first author (IA), and a proportion (30%) was verified by an independent author. All discrepancies were resolved by discussion. The data were extracted from the eligible tools using a pre-designed spreadsheet. These data included the tool's name, the developer company, the function, the cost, the dates of release and last update, the number of downloads and the overall consumer rating. A different spreadsheet was used to extract the data from the diagnostic accuracy studies, including the publication details, study design, setting, demographics, funding, accuracy data (e.g. sensitivity, specificity and intraclass correlation coefficients) and any other relevant data. Numerical data from plots were extracted using a web-based extraction tool (e.g. WebPlotDigitizer). The design of the diagnostic accuracy studies was categorised as case study, cross-sectional, case-control or comparative diagnostic accuracy studies (Chassé and Fergusson 2019).

### **Critical appraisal**

For Aim 1 (functionality), the quality and functionality of each app were assessed by two independent authors (either IA and WY, IA and APC or WY and APC.) using the Mobile Application Rating Scale (MARS), designed to classify and evaluate the quality of smartphone health-focused applications. This tool consists of 23 items related to five quality subscales: engagement (e.g. entertaining, customisable and interactive), functionality (e.g. easy to learn and navigate), aesthetics (e.g. visual appeal and stylistic consistency), information quality (e.g. descriptions, instructions and feedback quality) and overall satisfaction (Stoyanov et al. 2015). The rating score of each item ranges from 1 (inadequate) to 5 (excellent). This assessment tool was selected because it is simple and has excellent internal consistency and reliability (Stoyanov et al. 2015).

For Aims 2 and 3 (accuracy and reliability), the methodological quality of the diagnostic accuracy studies was assessed using the revised version of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (Whiting et al. 2011), which was developed to assess the risk of bias and concerns about applicability to the intended use. This tool consists of seven domains, four of which are related to the risk of bias: patient selection (e.g. recruited participants from a representative

sample pool), index test (e.g. used a pre-defined pass and refer criteria), reference standard (e.g. used pure-tone audiometry as a reference standard test) and flow and timing (e.g. acceptable time interval between the index and reference standard tests). The remaining domains are related to the applicability to our intended use: patient selection (e.g. the sample was not over-represented by one hearing loss-severity group), index test (e.g. the test was not administered with calibrated equipment or in controlled acoustic environments) and reference standards (e.g. the gold standard pure-tone audiometry was administered as it would be in audiology clinics). Each domain was judged as high (major risk of bias and concerns about applicability), low (no or minimal risk of bias and concerns about concerns) or unclear (no or insufficient details). Researchers were assigned to an equal number of studies and, using a pseudo-random pairing, two researchers independently assessed the risk of bias and applicability concerns for each of the studies. Major discrepancies (i.e. more than one category difference [high vs low risk]) were resolved by discussion or consulting other members of the review team. Minor discrepancies (i.e. one category difference [unclear versus low risk or unclear versus high risk]), when they existed, were resolved by labelling the risk as "unclear".

### **Data synthesis and missing data**

The data from the eligible studies were aggregated narratively. Missing data, when needed, were acquired from either the developers or the corresponding authors.

## **Results**

### **Search and selection of tools and studies**

Figure 1 displays the selection process of tools assessed for functionality. The search identified 357 records. After removal of duplicates, the names, details and functionality of the remaining tools ( $n=242$ ) were assessed for eligibility. Of these, 167 tools were included, the remainder discarded because they did not match the inclusion criteria (e.g. tools that only determine the highest audible frequency). Twenty additional tools identified from different sources were also included. In total, 187 eligible remote hearing assessment tools were included. All versions of eligible tools (iOS and Android) were included in this review.

Figure 2 displays the PRISMA flow diagrams for the validation studies when assessing accuracy and reliability. Searching the research databases and preprint servers identified 3657 records, of which 67 were deemed eligible for inclusion (see Figure 2 for further details). Conducting comprehensive reference screening and citation tracking of the eligible articles led to the identification of an additional 5208 records. After screening, 34 additional studies met the inclusion criteria. Of these, three were translated into English (from Russian, Chinese and Thai) and included in this review. Ultimately, 101 studies were included in the review. The number of studies from each database is reported in Supplement Table 2.

### **Remote hearing assessment tools**

The tools varied considerably in multiple domains, including their output formats, supported languages, hearing screen/assessment methods and compatibility with devices and operation systems. For the purposes of review, the tools were categorised based on the type of hearing assessment method used: tone,



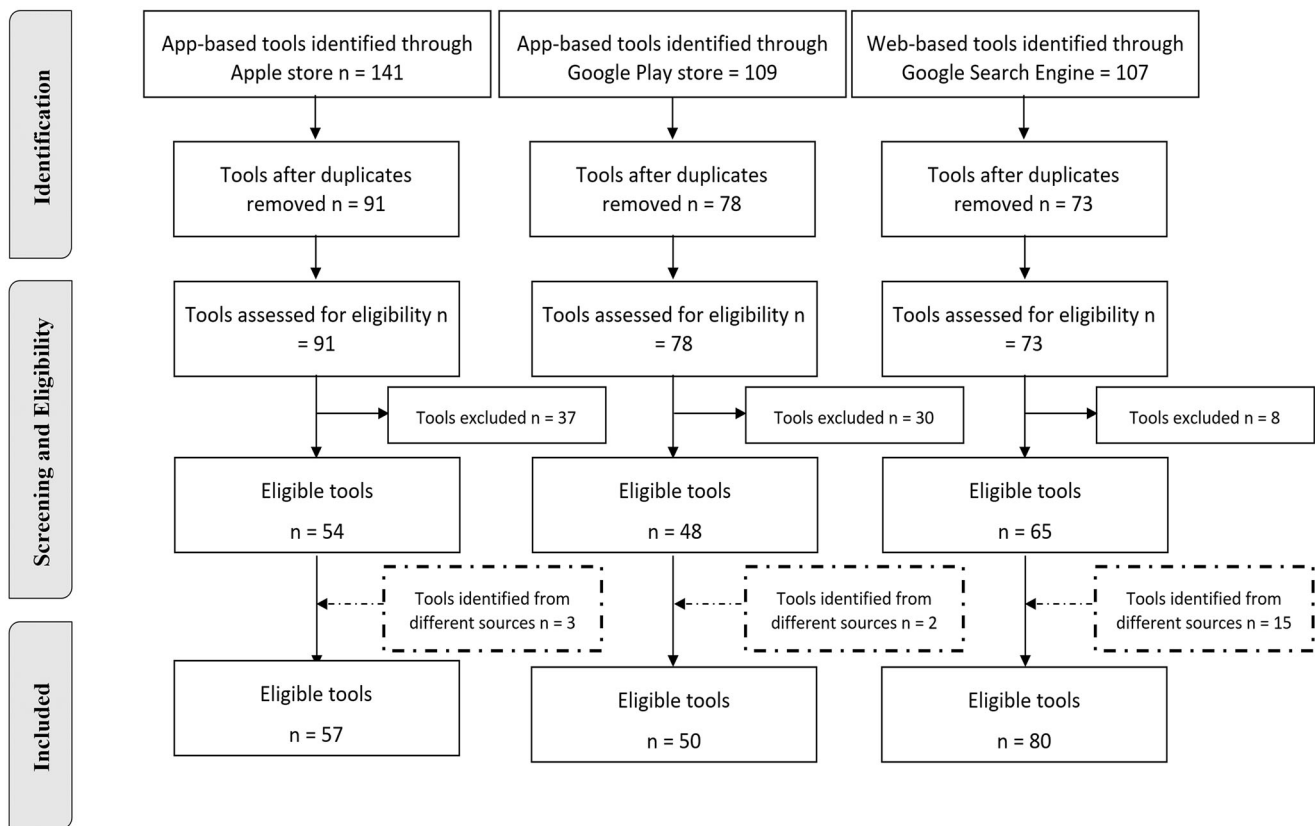


Figure 1. The selection process of the hearing assessment tools for assessing functionality.

speech, self-report or mixed-method hearing assessment tools. Table 1 provides a summary of these four categories.

### Tone hearing assessment tools

Tone hearing assessment tools (as the only assessment method within the app) were the most common category ( $n=92$ ). These tools measure the lowest intensity level of tone (e.g. pure and warble) that is audible, either in a quiet environment or background noise. Table 2 lists these tools and the platforms on which they run, with full details reported in Supplement Table 3. All but two of the tools were self-administered, the two exceptions being controlled by a tester. Almost half of the tone tools operated on iOS systems, and the remaining operated either on Android or as web-based tools.

The output format varied greatly, with qualitative, quantitative and graphical outputs. Audiograms were the most common output format (77%;  $n=70$ ), followed by the other formats, including scores, traffic lights, qualitative descriptions and hearing thresholds in units such as dB HL or dB SPL. Almost half of the tone hearing assessment tools (46%;  $n=42$ ) provided more than one output format to their users.

The primary language of all but four of the tone hearing assessment tools was English. The four exceptions were Italian, Turkish, Polish and Chinese. Around one-third of the English tools ( $n=31$ ) supported more than one language, including Arabic, French, German, Indonesian, Japanese and Spanish. All supported languages are reported in Supplement Table 3.

In terms of cost, 92% of these tools were free to download and use. The remaining tools (all app-based) cost (GBP) £1–£38. In-app purchases (i.e. buying services or advanced features from inside an app) were offered in some of the tools ( $n=23$ ) to access the full features of these tools. On the other hand, all web-based tools were free, although seven required free registration to either perform the test or obtain the results.

Figure 3 shows the MARS ratings (averaged across reviewers) for the engagement, functionality, aesthetics and information quality domains. The higher the score, the better the quality. The median overall score for tone tools was 3.0 (IQR = 0.8). The MARS ratings for each tool and quality domain are reported in Supplement Table 4.

In terms of evaluation, seven (8%) of the tone tools (all app-based: Audcal, Audiogram Mobile, Ear Scale, Hearing Test, Kids Hearing Game [Android and iOS] and Wulira App) were evaluated in 15 peer-reviewed publications. Twenty-five tone tools not readily available (e.g. removed from the app stores or requiring a study/licence number) were also evaluated in peer-reviewed publications. Supplement Table 5 lists these studies and category of the tools, with full details in Supplement Table 6.

The accuracy and quality of the seven tone tools that were evaluated are summarised in Supplement Table 7. Accuracy results were heterogeneous, even for the same tool. In general, all of these tools possessed sensitivity and specificity >70% or correlation coefficients of >0.7 when evaluated on adults using calibrated equipment.

The methodological quality of tone-tool studies are reported in Supplement Table 8. Most studies possessed a low risk of bias in all domains because there were minimal concerns about

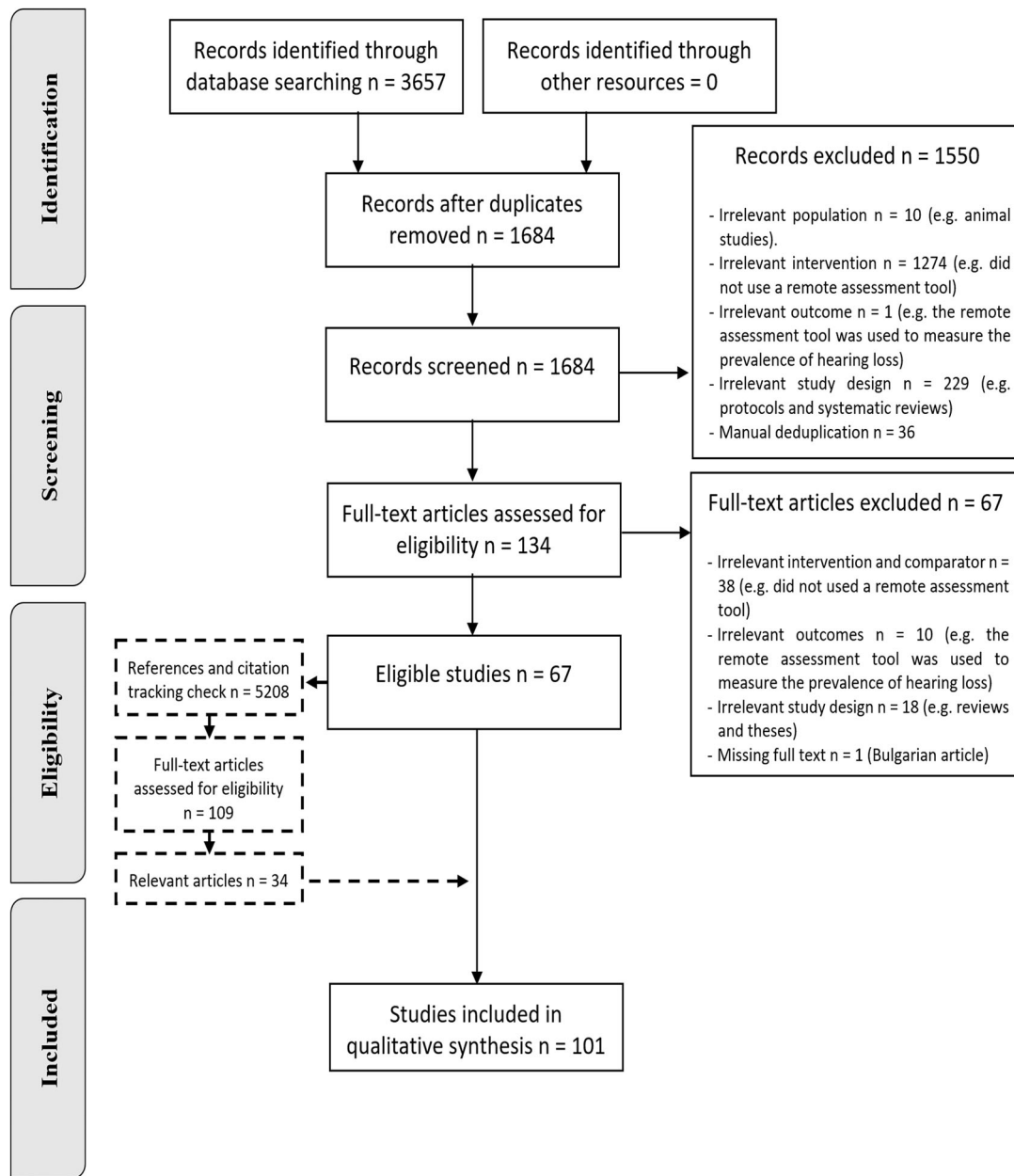


Figure 2. PRISMA flow diagram of the validation studies.

Table 1. Summary of the remote hearing assessment tools included in the review.

Tools category	Operation system			Total
	Android	iOS	Web-based	
Tone	34	44	14	92 (49%)
Speech	12	7	22	41 (22%)
Self-report	0	0	3	3 (2%)
Mixed-method	4	6	41	51 (27%)
Total	50 (27%)	57 (30%)	80 (43%)	187 (100%)

patient selection (e.g. used consecutive or random sampling), index test (e.g. interpreted the index test without knowledge of the findings or reference standard), reference standard (e.g. used pure-tone audiometry as a reference standard) and follow-up timing (e.g. not a long delay between tests). Conversely, most studies possessed high applicability concerns for patient selection (e.g. recruited an unrepresentative sample) and index test (e.g.

performed the index test in an ideal or controlled environment like a sound booth).

### Speech hearing assessment tools

Remotely assessing hearing using speech tasks (either in quiet or noisy settings) was used in 22% of the tools ( $n = 41$ ). Speech tools used various audio materials to assess hearing, including digits, words and sentences. Table 2 lists these tools and the platforms on which they run, with full details reported in Supplement Table 9. All the speech tools were self-administered, and nearly half of them were web-based. The remaining operated on Android or iOS.

The output format differed between tools, with hearing categories (e.g. normal vs abnormal) being the most common output (accounting for 80% [ $n = 33$ ] of the speech tools). Other output formats, including scores, SNRs and qualitative descriptions,

**Table 2.** A list of the tools included in the review.

[illegible]

Table 2. Continued.

Tool (developer/powered by)	Platform (evaluated in peer-reviewed publication)	Category	Tool (developer/powered by)	Platform (evaluated in peer-reviewed publication)	Category
Easy To Hear Hearing Aid App (IT ForYou)	iPhone	Tone	Q + Hearing Aid (Quadro Devices Private Limited)	iPhone	Tone
EasyHear (UltraDigi Hearing Technology Co., Ltd.)	iPhone	Tone	Resound Online Hearing Test (GN)	Web-based	Mix
EQplus (TOMTOP)	iPad/iPhone	Mix	S hearing (Sense Intelligent Inc.)	iPad/iPhone	Tone
Fennex By Smart Hear (Adam Palmquist)	iPhone	Tone	San Francisco Audiology - Online Hearing Screening (shoebox/WS Audiology)	Web-based	Mix
Healthy Hearing (Healthy Hearing)	Web-based	Self-report	Scrivens Online Hearing Test (sonova)	Web-based	Mix
Hear It Online Hearing Test (hear it)	Web-based	Speech	Shoebox (shoebox/WS Audiology)	Web-based	Mix
Hear ZA (HearX)	Android (E)	Speech	Signia Hearing Test (shoebox/WS Audiology)	Web-based	Mix
Hear ZA (HearX)	iPad/iPhone (E)	Speech	Signia Hearing Test (signia-hearing)	Android	Speech
HearBetter Amplification (Mind:Set Technologies)	iPhone	Tone	Signia Hearing Test (Sivantos Pte. Ltd.)	iPad/iPhone	Speech
HearDigits (HearX)	Web-based	Mix	Simple ear age test (pigpigjumps)	Android	Tone
Hearing Aid App for Android (dectone.pro/)	Android	Tone	SonicCloud Hearing Test for Personalising Sound/SonicCloud Personalised Sound (soniccloud.com/)	Android	Tone
Hearing aid hearing app (Alexander Bredikhin)	iPad/iPhone	Tone	Sonova Hearing Test (sonova)	Web-based	Mix
Hearing Aid Hearing App Amp (HEARING AID. LTD)	iPad/iPhone	Tone	Sound (Starkey Laboratories, Inc.)	iPhone	Tone
Hearing Analyser Lite (TOON, LLC.)	iPhone	Tone	Sound Metre Noise Detector (Ngo Bien)	iPad/iPhone	Tone
Hearing Analyser Portrait (TOON, LLC.)	iPhone	Tone	Sound metre premium (Ngo Bien)	iPad/iPhone	Tone
Hearing Australia - Online Hearing Assessment (NAL)	Web-based	Mix	Sound Relief Online Hearing Test (soundrelief)	Web-based	Mix
Hearing Direct Hearing Test (shoebox/WS Audiology)	Web-based	Mix	Sound Scouts (Sound Scouts HQ Pty Ltd)	iPad/iPhone (E)	Mix
Hearing Exam (Hyunmin Park)	iPad/iPhone	Tone	Sound Scouts Hearing Test (Sound Scouts HQ Pty Ltd)	Android (E)	Mix
Hearing is Believing Online Hearing Test (Robillard Hearing Centres r)	Web-based	Mix	SoundCheck (starkey)	Android	Tone
Hearing Life (AudiKa)	Web-based	Mix	Sounds Good! (Clery Tech LLC)	iPhone	Tone
Hearing Solutions Hearing Test Phillips (Bernafon Hearing Test)	Web-based	Mix	SoundTrack - Online Hearing Test (tetra hearing)	Web-based	Tone
Hearing Test (Android Crazy Studio)	Android	Tone	Specsavers Free Online Hearing Test (sonova)	Web-based	Mix
Hearing Test (dogusmedikal.com.tr)	Android	Tone	Specsavers Hearing Checker (specsavers)	Android	Speech
Hearing Test (e-audiologia.pl)	Android (E)	Tone	Starkey Online Hearing Test (Starkey)	Web-based	Speech
Hearing Test (Nick Maletsky)	Android	Tone	TEL SOUND AMPLIFIER // EASY TO HEAR: HEARING AID APP (it4you.pro/solutions/petralex-caller)	Android	Tone
Hearing Test (thisisappxone)	Android	Speech	Temporary Hearing Loss Test (Marshall Chasin)	iPad/iPhone	Tone
Hearing Test App iOS (Zipdev LLC)	iPhone	Tone	Test My Hearing (justkidsgamesuk)	Android	Tone
Hearing Test Ear Age Test (Yuichi Sakashita)	iPad/iPhone	Tone	Tester Sluchu (github.com/kkrzyzek)	Android	Tone
Hearing Test EN (seniorLABde)	Android	Speech	The national hearing test (Academic Medical Centre Amsterdam)	Web-based (E)	Speech
Hearing Test Online (lr. Stéphane Pigeon)	Web-based	Tone	Tunefork Demo - Hearing Test & Audio Adjustment (tunefork.co.il/)	Android	Tone
Hearing Test Pro (e-audiologia.pl)	Android (E)	Mix	uhear (Unitron Hearing Limited)	iPad/iPhone (E)	Mix
Hearing Test Tone Audiogram (IT ForYou)	iPhone	Tone	U sound (hearing assistant; NEWBRICK S.A.)	iPad/iPhone	Tone
Hearing test, Audiogram (IT ForYou; petralex.pro/en)	Android	Tone	uHear Kiosk (Unitron Hearing Limited)	iPad	Mix
Hearing test-best(Farhan Ghouri)	iPhone	Speech	University of New South Wales Hearing Test (University of New South Wales Hearing Test)	Web-based	Tone
Hearing Tracker (soniccloud)	Web-based	Mix	UrbanDenoyer Player, Volume Booster, Bass Booster (urbanidenoyer.com/)	Android	Tone
HearingPlanet Online Hearing Test (sonova)	Web-based	Mix	uSound for Samsung - Hearing test (usoundforsamsung.com.ar/)	Android	Tone
Hearingscreener (Noyl, LLC)	Web-based	Tone	VHS Hearing Solutions (lr. Stéphane Pigeon)	Web-based	Tone
HearingTest (Leo4google)	Android	Tone	VHS Hearing Solutions (vbshearingsolutions.com/)	Android	Tone
HearingTest4All (schabelDoesIT GbR)	iPad/iPhone	Tone	Whites Hearing North Platte (starkey)	Web-based	Mix
Hearmate (Sabyasachi Ghosh)	iPhone	Tone	Whites Hearing North Platte Test (shoebox/WS Audiology)	Web-based	Mix
HearMate (sapientive.com/)	Android	Tone	Wigan Hearing (Starkey)	Web-based	Speech
hEARo (hearo.pl/)	Android	Tone	WS Hearing Online Hearing Test (sonova)	Web-based	Mix
HearWHO - Check Your Hearing! (HearX)	iPhone	Speech	Wulira App (Joab Technologies Uganda)	Web-based (E)	Tone
HearWHO (HearX)	Android	Speech	听阈测试 (Shenzhen Horn Audio Co.Ltd.)	iPhone	Tone
HearWHO Pro (HearX)	Android	Speech	灯塔听力测试-耳科听力检测挂号 (Beijing shouda hospital management Co., Ltd.)	iPad/iPhone	Tone
HearWHO pro (HearX)	iPad/iPhone	Speech			



were also used by some of the tools. A quarter of the speech tools (24%;  $n = 10$ ) combined two or more output formats.

The primary language of two-thirds of the speech tools was English. The remaining tools were developed in Dutch, German, Norwegian, Swedish and Thai. Some tools ( $n = 11$ ) supported more than one language, including Arabic, Portuguese, French, Turkish, Hungarian, Chinese, Russian, Korean, Japanese, Greek, Vietnamese and Polish. A summary of the supported languages is presented in [Supplement Table 9](#).

In terms of cost, all but one of the speech tools were free. The exception was an app-based tool that cost (GBP) £5. In-app purchases were offered in two of the free tools to access the full features of those tools. Although all web-based tools were free, eight required free registration to either perform the test or obtain the results.

[Figure 3](#) shows the MARS ratings (averaged across reviewers) for all quality domains. The higher the score, the better the quality. The median overall score for speech tools was 3.8 (IQR = 0.5). Higher quality scores were found for functionality, aesthetics and information domains than the engagement domain. The MARS ratings for each tool are reported in [Supplement Table 10](#).

In terms of validation, eight (20%) of the tools (Hear ZA [Android and iOS], Pass Pro Version [Android and iOS], Blamey Saunders, Earcheck, the national hearing test and Kinderhoortest) were evaluated in 14 peer-reviewed publications. Six additional tools not readily available (e.g. unavailable in the app stores) were also evaluated in peer-reviewed publications. [Supplement Table 5](#) summarises the main characteristics of all of these studies, with full details in [Supplement Table 11](#).

The accuracy and quality of the eight speech tools that were evaluated are summarised in [Supplement Table 12](#). Accuracy results were heterogeneous, even for the same tool. Most tools had performance characteristics in excess of 80%. However, the sensitivity and specificity of the Kinderhoortest tool have not yet been measured: the assessors only reported cut-off values for pass and refer criteria.

Most studies possessed low risk of bias and applicability concerns because there were minimal concerns over the patient selection, index test, reference standard and follow-up timing. However, a few studies possessed high applicability concerns for patient selection (e.g. recruited participants with only normal hearing), index test (e.g. performed the index test in an ideal

environment like a sound booth) and reference standard (e.g. used another remote tool as a reference standard). The methodological quality of all studies in this category are reported in in [Supplement Table 13](#).

### Self-report hearing assessment tools

Self-report was the least common remote hearing assessment method, accounting for only 2% ( $n = 3$ ) of the tools. [Table 2](#) lists these tools and the platforms on which they run, with full details reported in [Supplement Table 14](#). All of these were web-based, self-administered tools. The output format of these tools was categorical (e.g. normal or abnormal). In addition to this output format, one of these tools provided a graphical output (i.e. traffic light). The language of these three tools was English, and none of them supported other languages. These three tools were free, but one required users to register before revealing their results.

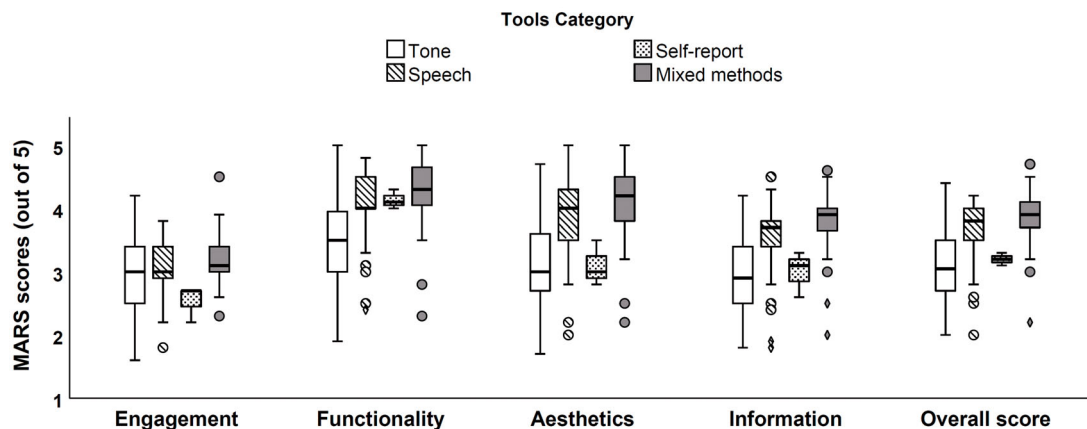
[Figure 3](#) shows the MARS ratings (averaged across reviewers) for all quality domains. The higher the score, the better the quality. The median overall score for self-report tools was 3.2 (IQR = 0.1). Most self-report tools obtained higher scores on the functionality domain than on the domains of aesthetics, engagement and information. The MARS ratings for each self-report tool are reported in [Supplement Table 15](#).

None of the self-report hearing assessment tools was evaluated in peer-reviewed publications.

### Mixed-method hearing assessment tools

Around a quarter ( $n = 51$ ) of the remote hearing assessment tools utilised mixed hearing assessment methods; that is, these tools used two or more of the aforementioned assessment methods. [Table 2](#) lists these tools and the platforms on which they run, with full details reported in [Supplement Table 16](#). These tools were self-administered, and around 80% of them were web-based tools. The remainder operated on iOS or Android.

The output format varied considerably, with hearing categories being the most common output (88%;  $n = 45$ ) followed by the other formats (e.g. audiograms, scores, traffic lights and SNRs). More than 60% of the mixed-method hearing assessment tools ( $n = 31$ ) used more than one output format.



**Figure 3.** Boxplot of the MARS ratings for each quality domain and tools category. The central horizontal lines in each box represent the medians. The interquartile ranges are represented by the top and bottom limits of the boxes. The highest and lowest values are represented by the whiskers (exclude outliers). Circles and diamonds represent outliers and extreme outliers, respectively.

The primary language of all mixed-method tools was English, and eight of them supported other languages, including Thai, Chinese, German, Danish, Swedish, Estonian and Hebrew. A summary of the supported languages is reported in [Supplement Table 16](#).

In terms of cost, all but two of the mixed-method tools were free to either download or perform a hearing test. The two app-based exceptions cost (GBP) £1 and £3. In-app purchases were essential in two of the free app-based tools (which had the same developer) to perform hearing tests. All web-based tools were free, similar to the other hearing assessment categories, but 14 of them required registration to either perform the test or obtain the results.

[Figure 3](#) shows the MARS ratings (averaged across reviewers) for all quality domains. The higher the score, the better the quality. The median overall score for mixed-method tools was 3.9 (IQR = 0.4). Most mixed-method tools obtained high scores on all but the engagement domain. The MARS ratings for each mixed-method tool are reported in [Supplement Table 17](#).

In terms of validation, seven (14%) of the tools (Connect Hearing, uHear, Hearing Test Pro e-audiologica.pl, Sound scouts [Android and iOS], NSRT<sup>®</sup> and Medel) were evaluated in 23 peer-reviewed publications. Two tools not readily available (i.e. unavailable in the app stores) were also evaluated in peer-reviewed publications. [Supplement Table 5](#) summarises the main characteristics of all of these studies, with full details reported in [Supplement Table 18](#).

The accuracy and quality of the seven mixed-method tools that were evaluated are summarised in [Supplement Table 19](#). Accuracy results were heterogeneous, even for the same tool. Most tools had performance characteristics in excess of 80%.

Like the other remote hearing assessment tools categories, most studies possessed low risk of bias and applicability concerns because there were limited concerns over the patient selection, index test, reference standard and follow-up timing. However, a few studies possessed high applicability concerns for patient selection (e.g. recruited only participants with hearing loss), index test (e.g. performed using calibrated headphones) and reference standard (e.g. performed in a quiet room instead of a sound booth). The methodological quality of all studies in this category are reported in [Supplement Table 20](#).

## Discussion

This review, the most comprehensive to date, aimed to (i) identify and assess the functionality of remote hearing assessment tools on smartphones and online platforms, (ii) determine if the tools had been evaluated in peer-reviewed publications and (iii) report on the accuracy and reliability of tools with validation data. We identified 187 remote hearing assessment tools from app stores and online platforms (or 167 tools if we treat apps working on multiple platforms as a single app) with considerable variability in functionality and quality. This number of tools is considerably greater than those identified in the reviews by Bright and Pallawela (2016) and Irace et al. (2021). This increase could be attributed to both the passage of time and because this review applied wider inclusion criteria and involved non-English and web-based tools. Additionally, unlike the previous reviews, this was the first to formally assess functionality and report on the accuracy of available remote hearing assessment tools.

Only a small proportion of the tools (12%) were evaluated (and reported in peer-reviewed publications), indicating that the accuracy and reliability of most tools remains unknown. The

validation results and their quality varied greatly, largely depending on the category of the tool. However, speech and mixed-method tools were found in a small number of cases to provide the highest levels of functionality, accuracy and usability among the many tools assessed

## Tone hearing assessment tools

Tone tests are the most common method used in the remote assessment tools. This can be partially attributed to the fact that these tools aim to mimic pure-tone audiometry, the gold standard test for hearing diagnosis.

Tone hearing assessment tools supported 26 languages, mostly from the Americas, Asia and Europe. However, the languages of sub-Saharan African countries, where the need might be greatest because of the scarcity of hearing-health professionals (Mulwafu et al. 2017; World Health Organization 2021), were less supported. This emphasises the need to develop hearing assessment tools to support these languages. Nevertheless, the issue of language barrier can be partially mitigated by using the browser's translation features (e.g. Google Translate) in the web version of tone tools.

The quality score for functionality of most tone tools was within the poor to acceptable range (2–3 out of 5) for all domains. This highlights that many of these tools lack aesthetics, functionality, engagement and entertainment strategies, as well as accurate and comprehensive instructions and feedback. Indeed, four of the seven tools that were evaluated obtained lower scores than the median overall score of the tone tools. The Hearing Test and the Kids Hearing Game were the only evaluated tools to score above the median MARS score for tone tools.

The accuracy of tone tools varied greatly. In general, tone tools provide more accurate results when performed on adults using calibrated equipment in a controlled acoustic environments (i.e. sound booth; Kelly et al. 2018; Masalski and Kręćicki 2013; Pickens et al. 2017). Using uncalibrated headphones with the same device can introduce inaccuracies (Pickens et al. 2018). Indeed, a recent systematic review and meta-analysis of diagnostic accuracy of app-based audiometry found that study population, equipment and test environments can all significantly affect their accuracy (Chen et al. 2021).

In addition, there was a lack of consistency in the accuracy outcomes reported between studies, some reporting Cohen's kappa, intraclass correlation coefficients or sensitivity and specificity (see [Supplement Table 7](#)), preventing a meta-analysis. The lack of consistency issue was also identified and reported in other hearing test reviews (Wasmann et al. 2022). Future studies should follow the Standards for Reporting of Diagnostic Accuracy guidelines (Bossuyt et al. 2015).

The methodological quality of the diagnostic accuracy studies varied considerably, with most having high levels of concern regarding the applicability of patient selection and index tests. These observations reduce our confidence in the generalisability of the findings.

This review identified seven tone tools whose accuracy had been evaluated. Of these, three had functionality scores  $\geq 3$  and sensitivity and specificity  $\geq 70\%$  (the Hearing Test and the Kids Hearing Game [Android and iOS]; see [Supplement Table 7](#)). However, the validation data are not consistent between the studies, even for the same tool, primarily due to issues of smartphone and transducer calibration, which are essential components of the gold standard pure-tone audiometry. Some tools attempted to minimise the effects of these issues by allowing the

users to self-calibrate their own transducers, approximating the reference sound level from a family member with normal hearing (Masalski and Kręćicki 2013). Other tools were pre-calibrated to a set of devices and transducer models (most commonly Apple AirPods); users of such tools can select their transducers before performing the remote hearing test. In principle, it does not seem possible to create accurate tests of absolute hearing threshold for devices with unknown output drive voltage connected to earphones of unknown sensitivity. Those tests found to be reasonably accurate were designed for the Apple iPhone or iPad driving Apple AirPods; thus, all parts of the system had known nominal electroacoustic characteristics.

### **Speech hearing assessment tools**

Speech tools are used less frequently than tone tools. Tools in this category vary in terms of the speech material used, with some utilising nonsense words (e.g. “atta” and “assa”), digits (e.g. 6-4-8), or other words (e.g. “fire” and “car”) and sentences.

Most of the speech tools supported a single language, which is to be expected given the challenge of developing audio materials for multiple languages. Some tools supporting multiple languages used the same audio materials for all languages, while others used different audio materials for each language. The tools that used the same audio materials across languages (English sentences [Absolute Ear Diagnostics] or nonsense words [Signia Hearing Test]) allow users to alter the interface language only. Using English audio materials for non-native English speakers may lead to incorrect classification of the hearing status, as a lack of language proficiency can impact the results (Potgieter et al. 2018). However, tests using generic (nonsense words) or language-tailored audio materials (e.g. Beltone Hearing Test and hearWHO) can minimise or diminish the effects of language proficiency. Similar to the tone tools, most speech tools did not support many African languages. However, unlike tone tools, the browser’s translation features (e.g. Google Translate) cannot assist users in overcoming the issue of language barrier created by non-tailored audio materials.

The quality scores for functionality of most speech tools were within the acceptable to good range (3–4 out of 5) for all but the engagement domain. The scores for the engagement domain were relatively lower than that of the other domains, and this can be ascribed to a lack of entertaining, customisation and interactive features. In general, the functionality scores for the speech tools were greater compared to those for the tone tools, suggesting that more effort is directed towards developing and maintaining speech tools. Indeed, the proportion of speech tools that were last updated in 2019 or later was 10% higher than that of tone tools. The quality scores of all but one of the evaluated tools in peer-reviewed publications were higher than the median score of speech tools, the one exception being Blamey Saunders.

The accuracy of speech tools was generally better than that of tone tools, as they are less prone to calibration issues and noisy test environments (Potgieter et al. 2016). This is because most of the speech tools used speech-in-noise test materials, which can largely bypass the need for a controlled test environment (De Sousa et al. 2021a). In addition, the nature of the outcome measure for the speech-in-noise tests (i.e. SNRs) can also overcome the need for transducer calibration (De Sousa et al. 2021a). However, some studies evaluate the same speech tool using different sound filter settings, and it is unclear which of these settings has been implemented in the currently available tool (e.g. Earcheck; Leensen and Dreschler 2013). There was a lack of

consistency between the studies in terms of the type of accuracy outcome, preventing a meta-analysis (see Supplement Table 12).

The methodological quality of the diagnostic accuracy studies varied slightly, but only a few had high risk of bias and concerns regarding patient selection, index tests and reference standard applicability. This can be partially attributed to the fact that most speech tools have been developed for application in uncontrolled environments using uncalibrated transducers and were tested as such (Leensen and Dreschler 2013; Qi et al. 2018).

This review identified eight speech tools whose accuracy had been evaluated. Of these tools, five had functionality score  $\geq 3.8$  and sensitivity and specificity  $\geq 80\%$  (Hear ZA [Android and iOS], Pass Pro Version [Android and iOS] and Earcheck; see Supplement Table 12). Unlike tone tools, speech tools are less prone to calibration issues and noisy test environments. However, many speech tools are language-dependent, making them suitable only for speakers with adequate knowledge of the vocabulary in that given language.

### **Self-report hearing assessment tools**

Self-report measurement is the least common method used in remote hearing assessment tools. Although the process of developing self-report tools is relatively straightforward, the number of tools is considerably lower than for tones or speech. This could be because of the limited and inconsistent correlations between self-report measures and pure-tone audiometry in identifying hearing loss (Brennan-Jones et al. 2016; Newton et al. 2001; Tsimpida et al. 2020). Despite being easy to translate into multiple languages, all self-report tools support only the English language.

The quality scores for functionality of most self-report tools were within the poor to good range (2–4 out of 5). The scores for the engagement domain were relatively low compared to those for the other domains. This could be partially attributed to the limited customisation and interactive features of self-report tools (e.g. the House of Hearing Online Hearing Test tool).

No self-report tools have been evaluated in peer-reviewed publications, although many self-report tests are available off-line and have been well validated (see e.g. Humes and Weinstein 2021). It is also unclear whether the developers of self-report tools used standard hearing-related questionnaires (with known psychometric properties) or developed new questionnaires. Transferring standard questionnaires to online use is a relatively straightforward process.

Overall, self-report measures are easy to administer and can assist in understanding the impact of hearing loss on people’s lives. In addition, self-report measures are also strong predictors of hearing help-seeking and hearing aid adoption (Meyer and Hickson 2012; Meyer et al. 2014). However, on their own, they may have limited and inconsistent capability to identify hearing loss. Thus, they are useful for complementing other remote hearing assessment methods.

### **Mixed-method hearing assessment tools**

Mixed-method tools are the second most common category among remote hearing assessment tools. Most mixed-method tools complement tone or speech assessment methods with self-report measures. This, as mentioned in the previous section, can be partially ascribed to the strong ability of self-report measures in predicting hearing help-seeking and hearing aid adoption.



Some mixed-method tools run the hearing tests independently and provide their respective outputs separately (e.g. Hearing Test Pro and Eartone Hearing Test), helping users comprehensively understand their hearing problems. Other mixed-method tools perform several tests at once and provide a single set of results (e.g. HearX and ReSound Online Hearing Test). For such tools, it is unclear how the findings of these tests are factored into a single set of outputs. One mixed-method tool (Hearing Australia) explicitly contrasted the two outcomes in the report to the user when they produced diverging results.

More than half of the mixed-method tools were powered by companies and organisations (e.g. HearX, Starkey, Sonova, WS Audiology and the National Acoustic Laboratories). These companies developed hearing assessment tools that could be implemented on websites to help business owners attract potential hearing aid users. The development, maintenance and dissemination of these tools is a positive sign of sustainable momentum towards routine remote hearing aid assessment and generating increased uptake.

Mixed-method tools supported a combined total of 21 languages, but most African languages were not supported. This observation is consistent with tone and speech multi-language tools.

The quality score for functionality of most mixed-method tools is higher than the other categories and range from acceptable to good scores (3–4 out of 5). This can be ascribed partially to the development of these tools by research centres, universities and large hearing-related companies. Although the mixed-method tools possess high scores on functionality, aesthetics and information, their engagement scores are relatively low, highlighting that entertainment, retention, interaction and customisation appear not to be priorities for developers. These are prevailing weaknesses of many health-related tools, stressing the need to draw more attention to these elements (Creber et al. 2016; Sarkar et al. 2016). The quality scores of all but two of the tools in peer-reviewed publications are higher than the median quality score of mixed-method tools. The two exceptions were Hearing Test Pro and Connect Hearing.

The accuracy of the mixed-method tools varies considerably and largely depends on the methods used for hearing assessment. Tools that use tone or speech methods exhibit similar accuracy outcomes to those of tone and speech tools, respectively. However, some studies evaluated two different screening approaches for the same tool (e.g. the uHear “original” and “modified Handzel” screening approaches), and it is not clear which one is implemented in the currently available version. There is a lack of consistency between the studies in terms of the type of accuracy outcome, preventing a meta-analysis (see Supplement Table 19).

The methodological quality of the validation studies varied greatly, but most of them have low risk-of-bias ratings. However, some studies exhibit high concerns regarding patient selection and index tests. These concerns reduce our confidence in the generalisability of the findings.

This review identified seven mixed-method tools whose accuracy had been evaluated. Of these, six had functionality  $\geq 3.5$  and sensitivity and specificity  $\geq 80\%$  (uHear, Hearing Test Pro, Sound scouts [Android and iOS], NSRT<sup>®</sup> and Medel; see Supplement Table 19).

### General limitations of hearing assessment tools

This review identified numerous tools that were deemed ineligible for inclusion because they made claims that could not be

substantiated against the design of the tool (e.g. tools that measure the highest audible frequency, such as the Hearing Age Test, do not provide a reliable assessment of functional hearing). This means that the results of these tools are potentially misleading.

About half of the identified app-based tools were labelled as medical or health and fitness tools, which may falsely imply to users that they have been clinically validated and accredited by professional bodies (e.g. US Food and Drug Administration and UK Medicines and Healthcare Products Regulatory Agency).

Although dual sensory impairments (i.e. hearing and vision) are quite common among older adults (Dawes et al. 2014; Saunders and Echt 2007), many remote hearing assessment tools were developed with small, non-customisable buttons. Users of such tools can partially mitigate this issue by increasing the font and display sizes of their smartphones, tablets or computers. However, such issues may instantly deter users from completing the hearing test. Involving patients and the public in developing hearing-health tools could greatly assist in improving the usability of such tools (Vaissou et al. 2021). Using the Web Accessibility Initiative guidelines (<http://www.w3.org/wai>) while developing hearing-related tools will also help enhance the usability for older people and those with sensory, physical and cognitive impairments.

Many of the hearing assessment tools were available free of charge and developed by individuals or small businesses. This raises the question of whether users' data will be used to generate revenues by selling them to third parties. Indeed, many of the free web-based tools required users to disclose personal data, including name, sex, date of birth and address, either to perform the test or obtain their results. This issue was highlighted in previous studies of apps for hearing assessment (Irace et al. 2021), as well as vision (Yeung et al. 2019), cognitive impairment (Charalambous et al. 2020) and many other health areas. Hearing test providers should adhere to and meet the data security laws and guidelines (Swanepoel et al. 2019).

### Future directions

While this review identified numerous tools with various output formats, only two of them (Sound Scouts Android and iOS) provided feedback on the type of hearing loss detected (auditory processing disorder, conductive loss and sensorineural loss). This was achieved by comparing thresholds for speech in quiet (affected by both sensorineural and conductive loss), speech in noise (assumed to be more affected by sensorineural loss than by conductive loss) and tone in noise (assumed to be affected only by sensorineural loss; Dillon et al. 2018). Feedback could help the users decide whether they need to seek professional medical evaluation (e.g. for conductive hearing loss). Research on combining two hearing assessment methods has shown that conductive and sensorineural hearing loss can be distinguished from each other by sequentially performing pure-tone audiometry and digit-in-noise tests (De Sousa et al. 2020; Dillon et al. 2018). In addition, performing different types of digit-in-noise tests (e.g. anti-phasic) has shown promising results in distinguishing between unilateral, bilateral, conductive and sensorineural hearing loss (De Sousa et al. 2021b). The review team encourage developers of remote hearing assessment tools to incorporate these approaches into their tools.

While tone tools can provide users with their approximated hearing thresholds to help them fit their own hearing aids (when needed), such tools are prone to calibration and background noise issues. Speech tools, on the other hand, are less affected by

these issues, but most do not provide users with an approximated audiogram, limiting the usability of their outputs. Research on predicting hearing thresholds from speech-in-noise tests has shown promising results (Blamey, Blamey, and Saunders 2015; Garrison and Bochner 2017). Fitting hearing aids using these predicted values was shown to yield favourable hearing aid outcomes (Blamey, Blamey, and Saunders 2015). Enhancing the accuracy of the predicted hearing thresholds using high- and low-pass filters with language-independent stimuli (e.g. nonsense words) could be a way forward for universal remote hearing aid assessment and fitting.

### Limitations of study

This review was limited to web- and app-based tools available in the Google Play and Apple app stores, and by Google search. Other platforms, including Bing and Yahoo, and the Huawei and Galaxy app stores were not searched. However, the market shares of the app stores and search engine used were higher than 98% and 86%, respectively (StatCounter 2021a, 2021b). Thus, the potential impact of excluding other app stores and search engines is extremely low.

Hearing assessment tools that require a licence number and equipment (other than headphones or earphones routinely used with smartphones, e.g. ShoeBox and HearTest) or are available only in certain countries or to certain stores (e.g. hearingScreening and Jacoti hearing centre) were not assessed for quality because they were not readily available to the public in the UK. In theory, we could ask a non-UK-based reviewer to download, install and assess the functionality of the tools that are not available in the UK app store, but this would require the use of another smartphone with different operating systems versions, and could result in inconsistent quality assessment scores. Thus, we encourage tool developers to remove country restrictions to help reduce the global burden of hearing loss. Although tools that require calibrated equipment (e.g. ShoeBox) can provide more accurate results, shipping this equipment to users can be costly and potentially inconvenient.

It would be good to categorise children's and adults' remote hearing assessment tools independently. However, the age suitability data, available in all app stores, were quite general. That is, most Android and iOS apps were labelled as PEGI 3 (suitable for all age groups) and as 4+ years old, respectively. In addition, almost all web-based tools lacked details of suitability for adults or children.

The user's (star) ratings for each app were extracted from app stores and reported in Supplement Table 3, 9 and 16. Analysing and discussing those data, however, was not an aim of this review due to the limited number of raters and the resulting potential rater biases for many of the apps.

This review did not extract the extent to which remote hearing assessment tools monitored environmental noise due to the variable sensitivity of microphones in Android and Windows devices, which leads to inaccurate measurement of background noise levels (Murphy and King 2016).

Remote hearing assessment and intervention are rapidly growing areas, and it is likely that some of the tools studied have now been updated or removed from stores. Other new tools may have been released. However, many of the identified quality issues and test challenges pertaining to remote hearing assessment tools may persist in new or updated tools if they were not considered when developing these tools.

### Conclusion

In conclusion, 187 tools using tones, speech, self-report or a combination, were identified in the review, the largest and most comprehensive to date. The tools varied in accuracy, quality and output/feedback. Twenty-two (12%) tools have been formally evaluated and around half of these have acceptable functionality; therefore, the accuracy and functionality of the majority of tools is unknown. Speech and mixed-method tools were found in a small number of cases to provide the highest levels of functionality, accuracy and usability among the many tools assessed. Tone tools can provide users with approximated hearing thresholds but are prone to calibration and background noise issues. Speech tools are less affected by these issues, but most do not provide users with an approximated audiogram. Predicting audiograms using filtered language-independent materials (e.g. nonsense words) could be a universal solution if accompanied by an automated method of scoring users' on-screen responses.

### Acknowledgments

We thank our colleague Michael Stone for providing insightful comments on an earlier draft of the manuscript.

### Disclosure statement

All but two of the review team members have no conflicts of interest to disclose. David R Moore is a scientific advisor to and shareholder of the HearX group. Harvey Dillon performs occasional ad-hoc work for Sound Scouts Pty Ltd.

### Funding

IA, HD, PD, DRM and KM are supported by the NIHR Manchester Biomedical Research Centre. IA is also supported by the Deanship of Scientific Research at the College of Applied Medical Sciences Research Center at King Saud University. PD, WY, CT and APC are supported by the SENSECog project which has received funding from the European Unions Horizon 2020 research and innovation program under grant no. 668648.

### ORCID

Ibrahim AlmuFarrij  <http://orcid.org/0000-0003-4043-7234>  
 Harvey Dillon  <http://orcid.org/0000-0002-7036-7890>  
 Piers Dawes  <http://orcid.org/0000-0003-3180-9884>  
 David R. Moore  <http://orcid.org/0000-0002-1567-1945>  
 Wai Yeung  <http://orcid.org/0000-0001-8456-0635>  
 Kevin J. Munro  <http://orcid.org/0000-0001-6543-9098>

### References

- Agarwal, Smisha, Amnesty E. LeFevre, Jaime Lee, Kelly L'Engle, Garrett Mehl, Chaitali Sinha, and Alain Labrique. 2016. "Guidelines for Reporting of Health Interventions Using Mobile Phones: Mobile Health (mHealth) Evidence Reporting and Assessment (mERA) Checklist." *BMJ (Clinical Research Edition)* 352: i1174. doi:10.1136/bmj.i1174.
- Appfigures. 2021. "Number of mHealth Apps Available in the Google Play Store From 1st Quarter 2015 to 1st Quarter 2021. Statista." Statista Inc. Accessed 17 December 2021. <https://www-statista-com.manchester.idm.oclc.org/statistics/779919/health-apps-available-google-play-worldwide/>



- Arksey, H., and L. O'Malley. 2005. "Scoping Studies: Towards a Methodological Framework." *International Journal of Social Research Methodology* 8 (1): 19–32. doi:10.1080/1364557032000119616.
- Blamey, P. J., J. K. Blamey, and E. Saunders. 2015. "Effectiveness of a Teleaudiology Approach to Hearing Aid Fitting." *Journal of Telemedicine and Telecare* 21 (8): 474–478. doi:10.1177/1357633X15611568.
- Bossuyt, P. M., J. B. Reitsma, D. E. Bruns, C. A. Gatsonis, P. P. Glasziou, L. Irwig, J. G. Lijmer, et al. 2015. "STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies." *BMJ (Clinical Research Edition)* 351: h5527. doi:10.1136/bmj.h5527.
- Brennan-Jones, C. G., D. S. Taljaard, S. E. Brennan-Jones, R. J. Bennett, D. Swanepoel, and R. H. Eikelboom. 2016. "Self-Reported Hearing Loss and Manual Audiometry: A Rural versus Urban Comparison." *Australian Journal of Rural Health* 24 (2): 130–135. doi:10.1111/ajr.12227.
- Bright, T., and D. Pallawela. 2016. "Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review." *JMIR Rehabilitation and Assistive Technologies* 3 (2): e13. doi:10.2196/rehab.6074.
- Charalambous, A. P., A. Pye, W. K. Yeung, I. Leroi, M. Neil, C. Thodi, and P. Dawes. 2020. "Tools for App- and Web-Based Self-Testing of Cognitive Impairment: Systematic Search and Evaluation." *Journal of Medical Internet Research* 22 (1): e14551. doi:10.2196/14551.
- Chassé, M., and D. A. Fergusson. 2019. "Diagnostic Accuracy Studies." *Seminars in Nuclear Medicine* 49 (2): 87–93. doi:10.1053/j.semnuclmed.2018.11.005.
- Chen, C. H., H. H. Lin, M. C. Wang, Y. C. Chu, C. Y. Chang, C. Y. Huang, and Y. F. Cheng. 2021. "Diagnostic Accuracy of Smartphone-Based Audiometry for Hearing Loss Detection: Meta-Analysis." *JMIR mHealth and uHealth* 9 (9): e28378. doi:10.2196/28378.
- Cosh, S., T. Hanno, C. Helmer, G. Bertelsen, C. Delcourt, and H. Schirmer. 2018. "The Association Amongst Visual, Hearing, and Dual Sensory Loss with Depression and Anxiety over 6 Years: The Tromsø Study." *International Journal of Geriatric Psychiatry* 33 (4): 598–605. doi:10.1002/gps.4827.
- Creber, R. M. M., M. S. Maurer, M. Reading, G. Hiraldo, K. T. Hickey, and S. Iribarren. 2016. "Review and Analysis of Existing Mobile Phone Apps to Support Heart Failure Symptom Monitoring and Self-Care Management Using the Mobile Application Rating Scale (MARS)." *JMIR mHealth and uHealth* 4 (2): e74. doi:10.2196/mhealth.5882.
- Davis, A., P. Smith, M. Ferguson, D. Stephens, and I. Gianopoulos. 2007. "Acceptability, Benefit and Costs of Early Screening for Hearing Disability: A Study of Potential Screening Tests and Models." *Health Technology Assessment* 11 (42): 1–294. doi:10.3310/hta11420.
- Dawes, P., C. Dickinson, R. Emsley, P. N. Bishop, K. J. Cruickshanks, M. Edmondson-Jones, A. McCormack, et al. 2014. "Vision Impairment and Dual Sensory Problems in Middle Age." *Ophthalmic & Physiological Optics: The Journal of the British College of Ophthalmic Opticians (Optometrists)* 34 (4): 479–488. doi:10.1111/opo.12138.
- De Sousa, K. C., D. R. Moore, C. Smits, and D. Swanepoel. 2021a. "Digital Technology for Remote Hearing Assessment—Current Status and Future Directions for Consumers." *Sustainability* 13 (18): 10124. doi:10.3390/su131810124.
- De Sousa, K. C., C. Smits, D. R. Moore, H. C. Myburgh, and D. Swanepoel. 2020. "Pure-Tone Audiometry Without Bone-Conduction Thresholds: using the Digits-in-Noise Test to Detect Conductive Hearing Loss." *International Journal of Audiology* 59 (10): 801–808. doi:10.1080/14992027.2020.1783585.
- De Sousa, K. C., C. Smits, D. R. Moore, H. C. Myburgh, and D. Swanepoel. 2021b. "Diotic and Antiphasic Digits-in-Noise Testing as a Hearing Screening and Triage Tool to Classify Type of Hearing Loss." *Ear and Hearing* 43 (3): 1037–1048. doi:10.1097/AUD.0000000000001160.
- Dillon, H., C. Mee, J. C. Moreno, and J. Seymour. 2018. "Hearing Tests Are Just Child's Play: The Sound Scouts Game for Children Entering School." *International Journal of Audiology* 57 (7): 529–537. doi:10.1080/14992027.2018.1463464.
- Ericsson. 2021. Ericsson Mobility Report. Stockholm: Ericsson. Accessed 16 December 2021. <https://www.ericsson.com/4ad7e9/assets/local/reports-papers/mobilityreport/documents/2021/ericsson-mobility-report-november-2021.pdf>
- Ferguson, M. A., P. T. Kitterick, L. Y. Chong, M. Edmondson-Jones, F. Barker, and D. J. Hoare. 2017. "Hearing Aids for Mild to Moderate Hearing Loss in Adults." *The Cochrane Database of Systematic Reviews* 9 (9): CD012023. doi:10.1002/14651858.CD012023.pub2.
- Garrison, W., and J. Bochner. 2017. "An Application for Screening Gradual-Onset Age-Related Hearing Loss." *Health* 09 (04): 715–726. doi:10.4236/health.2017.94051.
- Grainger, R., H. Devan, B. Sangelaji, and J. Hay-Smith. 2020. "Issues in Reporting of Systematic Review Methods in Health App-Focused Reviews: A Scoping Review." *Health Informatics Journal* 26 (4): 2930–2945. doi:10.1177/1460458220952917.
- Grasselli, G., A. Zangrillo, A. Zanella, M. Antonelli, L. Cabrini, A. Castelli, and D. Cereda, et al. 2020. "Baseline Characteristics and Outcomes of 1591 Patients Infected with SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy." *JAMA* 323 (16): 1574–1581. doi:10.1001/jama.2020.5394.
- Humes, L. E., and B. E. Weinstein. 2021. "The Need for a Universal Hearing Metric-Is Pure-Tone Average the Answer? ." *JAMA Otolaryngology-Head & Neck Surgery* 147 (7): 588–589. doi:10.1001/jamaoto.2021.0417.
- Irace, A. L., R. K. Sharma, N. S. Reed, and J. S. Golub. 2021. "Smartphone-Based Applications to Detect Hearing Loss: A Review of Current Technology ." *Journal of the American Geriatrics Society* 69 (2): 307–316. doi:10.1111/jgs.16985.
- Keesara, S., A. Jonas, and K. Schulman. 2020. "Covid-19 and Health Care's Digital Revolution." *The New England Journal of Medicine* 382 (23): e82. doi:10.1056/NEJMp2005835.
- Kelly, E. A., M. E. Stadler, S. Nelson, C. L. Runge, and D. R. Friedland. 2018. "Tablet-Based Screening for Hearing Loss: Feasibility of Testing in Nonspecialty Locations." *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology* 39 (4): 410–416. doi:10.1097/MAO.0000000000001752.
- Leensen, M. C., and W. A. Dreschler. 2013. "Speech-in-Noise Screening Tests by Internet, Part 3: test Sensitivity for Uncontrolled Parameters in Domestic Usage." *International Journal of Audiology* 52 (10): 658–669. doi:10.3109/14992027.2013.803610.
- Loughrey, D. G., M. E. Kelly, G. A. Kelley, S. Brennan, and B. A. Lawlor. 2018. "Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and Dementia: A Systematic Review and Meta-Analysis." *JAMA Otolaryngology-Head & Neck Surgery* 144 (2): 115–126. doi:10.1001/jamaoto.2017.2513.
- Maharani, A., P. Dawes, J. Nazroo, G. Tampubolon, and N. Pendleton. 2020. "Associations between Self-Reported Sensory Impairment and Risk of Cognitive Decline and Impairment in the Health and Retirement Study Cohort." *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 75 (6): 1230–1242. doi:10.1093/geronb/gbz043.
- Maharani, A., N. Pendleton, and I. Leroi. 2019. "Hearing Impairment, Loneliness, Social Isolation, and Cognitive Function: Longitudinal Analysis Using English Longitudinal Study on Ageing." *The American Journal of Geriatric Psychiatry* 27 (12): 1348–1356. doi:10.1016/j.jagp.2019.07.010.
- Masalski, M., and T. Kręcki. 2013. "Self-Test Web-Based Pure-Tone Audiometry: Validity Evaluation and Measurement Error Analysis." *Journal of Medical Internet Research* 15 (4): e71. doi:10.2196/jmir.2222.
- McGarrigle, R., K. J. Munro, P. Dawes, A. J. Stewart, D. R. Moore, J. G. Barry, and S. Amitay. 2014. "Listening Effort and Fatigue: What Exactly Are We Measuring? A British Society of Audiology Cognition in Hearing Special Interest Group 'White Paper'." *International Journal of Audiology* 53 (7): 433–440. doi:10.3109/14992027.2014.890296.
- Meyer, C., and L. Hickson. 2012. "What Factors Influence Help-Seeking for Hearing Impairment and Hearing Aid Adoption in Older Adults?" *International Journal of Audiology* 51 (2): 66–74. doi:10.3109/14992027.2011.611178.
- Meyer, C., L. Hickson, K. Lovelock, M. Lampert, and A. Khan. 2014. "An Investigation of Factors That Influence Help-Seeking for Hearing Impairment in Older Adults." *International Journal of Audiology* 53 (sup1): S3–S17. doi:10.3109/14992027.2013.839888.
- Moynihan, R., S. Sanders, Z. A. Michaleff, A. M. Scott, J. Clark, E. J. To, M. Jones, et al. 2021. "Impact of COVID-19 Pandemic on Utilisation of Healthcare Services: A Systematic Review." *BMJ Open* 11 (3): e045343. doi:10.1136/bmjopen-2020-045343.
- Mulwafu, W., R. Ensink, H. Kuper, and J. Fagan. 2017. "Survey of ENT Services in Sub-Saharan Africa: Little Progress Between 2009 and 2015." *Global Health Action* 10 (1): 1289736. doi:10.1080/16549716.2017.1289736.
- Murphy, E., and E. A. King. 2016. "Testing the Accuracy of Smartphones and Sound Level Meter Applications for Measuring Environmental Noise." *Applied Acoustics* 106: 16–22. doi:10.1016/j.apacoust.2015.12.012.
- Newton, V. E., I. Macharia, P. Mugwe, B. Ototo, and S. W. Kan. 2001. "Evaluation of the Use of a Questionnaire to Detect Hearing Loss in Kenyan Pre-School Children." *International Journal of Pediatric Otorhinolaryngology* 57 (3): 229–234. doi:10.1016/s0165-5876(00)00453-5.
- Pickens, A. W., L. D. Robertson, M. L. Smith, H. Zhao, R. Mehta, and S. Song. 2017. "Limitations of a Mobile Hearing Test Application." *The Hearing Journal* 70 (6): 34,36,37.
- Pickens, A. W., L. D. Robertson, M. L. Smith, Q. Zheng, and S. Song. 2018. "Headphone Evaluation for App-Based Automated Mobile Hearing

- Screening." *International Archives of Otorhinolaryngology* 22 (4): 358–363. doi:10.1055/s-0037-1607438.
- Potgieter, J. M., D. Swanepoel, H. C. Myburgh, T. C. Hopper, and C. Smits. 2016. "Development and Validation of a Smartphone-Based Digits-in-Noise Hearing Test in South African English." *International Journal of Audiology* 55 (7): 405–411. doi:10.3109/14992027.2016.1172269.
- Potgieter, J.-M., D. Swanepoel, H. Myburgh, and C. Smits. 2018. "The South African English Smartphone Digits-in-Noise Hearing Test: Effect of Age, Hearing Loss, and Speaking Competence." *Ear & Hearing* 39 (4): 656–663. doi:10.1097/AUD.0000000000000522.
- Qi, B. E., T. B. Zhang, X. X. Fu, and G. P. Li. 2018. "Establishment of the Characterization of an Adult Digits-in-Noise Test Based on Internet." *Journal of Clinical Otorhinolaryngology, Head, and Neck Surgery* 32 (3): 202–205. doi:10.13201/j.issn.1001-1781.2018.03.011.
- Sarkar, U., G. I. Gourley, C. R. Lyles, L. Tieu, C. Clarity, L. Newmark, K. Singh, and D. W. Bates. 2016. "Usability of Commercially Available Mobile Applications for Diverse Patients." *Journal of General Internal Medicine* 31 (12): 1417–1426. doi:10.1007/s11606-016-3771-6.
- Saunders, G. H., and K. V. Echt. 2007. "An Overview of Dual Sensory Impairment in Older Adults: perspectives for Rehabilitation." *Trends in Amplification* 11 (4): 243–258. doi:10.1177/1084713807308365.
- StatCounter. 2021a. "Mobile Operating Systems' Market Share Worldwide From January 2012 to June 2021." Statista. Accessed 17 December 2021. <https://www-statista-com.manchester.idm.oclc.org/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/>.
- StatCounter. 2021b. "Worldwide Desktop Market Share of Leading Search Engines From January 2010 to September 2021." Statista. Accessed 17 December 2021. <https://www-statista-com.manchester.idm.oclc.org/statistics/216573/worldwide-market-share-of-search-engines/>.
- Stoyanov, S. R., L. Hides, D. J. Kavanagh, O. Zelenko, D. Tjondronegoro, and M. Mani. 2015. "Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps." *JMIR mHealth and uHealth* 3 (1): e27–e27. doi:10.2196/mhealth.3422.
- Swanepoel, D., K. C. De Sousa, C. Smits, and D. R. Moore. 2019. "Mobile Applications to Detect Hearing Impairment: opportunities and Challenges." *Bulletin of the World Health Organization* 97 (10): 717–718. doi:10.2471/BLT.18.227728.
- Swanepoel, D., and J. W. Hall. 2010. "A Systematic Review of Telehealth Applications in Audiology." *Telemedicine Journal and e-Health : The Official Journal of the American Telemedicine Association* 16 (2): 181–200.
- Swanepoel, D., and J. W. Hall. 2020. "Making Audiology Work During COVID-19 and Beyond." *The Hearing Journal* 73 (6): 20,22,23,24–24. doi:10.1097/01.HJ.0000669852.90548.75.
- Tsimpida, D., E. Kontopantelis, D. Ashcroft, and M. Panagioti. 2020. "Comparison of Self-Reported Measures of Hearing with an Objective Audiometric Measure in Adults in the English Longitudinal Study of Ageing." *JAMA Network Open* 3 (8): e2015009–e2015009. doi:10.1001/jamanetworkopen.2020.15009.
- Vaission, G., T. Provencher, M. Dugas, M.-È. Trottier, S. Chipenda Dansokho, H. Colquhoun, A. Fagerlin, et al. 2021. "User Involvement in the Design and Development of Patient Decision Aids and Other Personal Health Tools: A Systematic Review." *Medical Decision Making* 41 (3): 261–274. doi:10.1177/0272989X20984134.
- Wasmann, J.-W., L. Pragt, R. Eikelboom, and D. Swanepoel. 2022. "Digital Approaches to Automated and Machine Learning Assessments of Hearing: Scoping Review." *Journal of Medical Internet Research* 24 (2): e32581. doi:10.2196/32581.
- Whiting, P. F., A. W. Rutjes, M. E. Westwood, S. Mallett, J. J. Deeks, J. B. Reitsma, M. M. Leeflang, J. A. Sterne, and P. M. Bossuyt. 2011. "QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies." *Annals of Internal Medicine* 155 (8): 529–536. doi:10.7326/0003-4819-155-8-201110180-00009.
- Wilson, B. S., D. L. Tucci, M. H. Merson, and G. M. O'Donoghue. 2017. "Global Hearing Health Care: new Findings and Perspectives." *Lancet (London, England)* 390 (10111): 2503–2515. doi:10.1016/S0140-6736(17)31073-5.
- World Health Organization. 2018. *Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000-2016*. Geneva: World Health Organization. Accessed 27 May 2021. <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys>
- World Health Organization. 2020. *Deafness and Hearing Loss*. Geneva: World Health Organization. Accessed 06 October 2020. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>
- World Health Organization. 2021. *World Report on Hearing*. Geneva: World Health Organization. Accessed 21 April 2022. <https://www.who.int/publications/i/item/world-report-on-hearing>.
- Wu, Z., and J. M. McGoogan. 2020. "Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention." *JAMA* 323 (13): 1239–1242. doi:10.1001/jama.2020.2648.
- Yeung, W. K., P. Dawes, A. Pye, M. Neil, T. Aslam, C. Dickinson, and I. Leroi. 2019. "eHealth Tools for the Self-Testing of Visual Acuity: A Scoping Review." *NPJ Digital Medicine* 2: 1–7.