A systematic review and meta-analysis assessing the effectiveness of alternative listening devices to conventional hearing aids in adults with hearing loss

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: MAIDMENT, D. ... et al., 2018. A systematic review and meta-analysis assessing the effectiveness of alternative listening devices to conventional hearing aids in adults with hearing loss. International Journal of Audiology, 57(10), pp. 721 - 729.

Additional Information:

- This is an Open Access Article. It is published by Taylor and Francis under the Creative Commons Attribution 4.0 Unported Licence (CC BY). Full details of this licence are available at: http://creativecommons.org/licenses/by/4.0/

Metadata Record: https://dspace.lboro.ac.uk/2134/36561

Version: Published

Publisher: © 2018 British Society of Audiology, International Society of Audiology, and Nordic Audiological Society. Published by Taylor and Francis

Rights: This work is made available according to the conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. Full details of this licence are available at: http://creativecommons.org/licenses/by/4.0/

Please cite the published version.
A systematic review and meta-analysis assessing the effectiveness of alternative listening devices to conventional hearing aids in adults with hearing loss

David W. Maidment, Alexander B. Barker, Jun Xia & Melanie A. Ferguson

To cite this article: David W. Maidment, Alexander B. Barker, Jun Xia & Melanie A. Ferguson (2018) A systematic review and meta-analysis assessing the effectiveness of alternative listening devices to conventional hearing aids in adults with hearing loss, International Journal of Audiology, 57:10, 721-729, DOI: 10.1080/14992027.2018.1493546

To link to this article: https://doi.org/10.1080/14992027.2018.1493546

© 2018 British Society of Audiology, International Society of Audiology, and Nordic Audiological Society. Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 02 Nov 2018.

Submit your article to this journal

Article views: 441

View Crossmark data
A systematic review and meta-analysis assessing the effectiveness of alternative listening devices to conventional hearing aids in adults with hearing loss

David W. Maidment, Alexander B. Barker, Jun Xia and Melanie A. Ferguson

Introduction

Acoustic amplification provided by hearing aids is currently the primary clinical management strategy for adults with mild and moderate hearing loss. Hearing aids have been shown to improve hearing-specific health-related quality of life, general health-related quality of life and listening abilities (Ferguson et al. 2017). However, two out of three people who would benefit from using hearing aids do not take them up (Davis et al. 2007). For those who do obtain hearing aids, estimates of non-use vary from 3 to 24% (Ferguson et al. 2017). People with hearing loss report that they are concerned or embarrassed that hearing aids will make them look old and that they will be treated differently (Barker, Leighton, and Ferguson 2017; Heffernan et al. 2015). Indeed, advances in technology have led to a rapid increase in alternative listening devices to conventional hearing aids. Here, we define alternative listening devices as standalone products that provide amplification of sound (e.g. Smartphone hearing aid applications; personal sound amplification products; hearables), as well as assistive listening devices (ALDs) that amplify and transmit sound directly into hearing aids (e.g. Smartphone-connected hearing aids; remote microphone systems). Many alternative listening devices can link wirelessly to Smartphone technologies, allowing users to adjust and personalise their hearing settings (e.g. gain, frequency response) in different listening situations at their own convenience via an application (or app), and without the need to visit a qualified clinician (Taylor 2015).

Existing evidence is mixed in terms of whether alternative listening devices are a suitable management strategy for hearing loss. For example, ‘mid-range’ (US$100–$500) personal sound amplification products (PSAPs), a type of ‘direct-to-consumer’ hearing device, have been shown to provide comparable electroacoustic characteristics (i.e. meet gain and output targets using National Acoustic Laboratories prescriptive procedures) to hearing aids (Callaway and Punch 2008). By comparison, other products defined as ‘low-cost’ (<US$100) may be of limited benefit and potentially damaging to residual hearing due to over-amplification (Callaway and Punch 2008; Chan and McPherson 2015). Smartphone-based ‘hearing aid’ apps have also been shown to provide similar levels of amplification, improved speech-in-noise performance and greater self-reported benefit in comparison to hearing aids (Amlani et al. 2013). Remote microphone systems...
have also been shown to improve hearing outcomes, but may require additional audiological support for optimal use (Boothroyd 2004).

To date, no systematic review has evaluated whether alternative listening devices are a clinically effective intervention for people with mild and moderate hearing loss. A systematic review with meta-analysis provides the gold-standard evidence-base to inform future feasibility and effectiveness trials of alternative listening devices. This approach is consistent with the Medical Research Council’s guidelines for evaluating complex healthcare interventions (Medical Research Council 2006). The primary objective of this study, therefore, was to review and synthesise the existing body of evidence to assess the effectiveness of alternative listening devices to conventional hearing aids.

Methods

Prior to commencing the systematic review, the protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO) (registration number, CRD42015029582) and published in a peer-reviewed publication (Maidment et al. 2016). Methods are reported according to the Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) checklist (Moher et al. 2009).

Eligibility criteria

The criteria for inclusion in the review were specified in terms of participants, intervention, comparators, outcomes and study designs (PICOS) as follows.

Participants

Adults (≥18 years) with a mild and moderate hearing loss (average hearing threshold across octave frequencies 0.25–4 kHz ≥20 and ≤70 dB HL (British Society of Audiology 2011). Studies that included both children (<18 years) and adults were not included unless data were reported separately. If the hearing thresholds were not specified, the study author was contacted for further clarification. If hearing threshold data was not reported and could not be obtained, studies were included where the mean average hearing threshold reported fell within the range of either mild (between 20 and 40 dB HL) or moderate hearing loss (between 41 and 70 dB HL). Bilateral and unilateral sensorineural, conductive and mixed hearing losses were included.

Intervention(s)

An alternative listening device to a conventional hearing aid was considered to be non-medical standalone product (e.g. Smartphone app, PSAP, hearable) or an assistive listening device that provides additional functionality to a conventional hearing aid (e.g. remote microphone system; Smartphone-connected hearing aid whereby the Smartphone can be used as a remote microphone and/or allows manipulation of gain and frequency response via an app). An alternative listening device should aim to improve hearing and communication outcomes in people with hearing loss, specifically via the amplification of external sound sources.

Comparators

The comparisons of interest were either passive (e.g. unaided) or active control (e.g. conventional hearing aid). A conventional hearing aid was defined as a device that detects and amplifies sound, delivering an amplified acoustic signal via air conduction to the external auditory canal on the same side that the signals are detected, irrespective of where it is worn (behind-the-ear, in-the-ear or receiver-in-the-canal). Studies evaluating analogue hearing aids were excluded.

Outcomes

As the aim of the review was to assess the clinical-effectiveness of alternative listening devices, studies were restricted to outcomes associated with the consequences of hearing loss. There were no restrictions as to the duration of follow-up. Primary outcomes included one or more of the following: (i) behavioural measures of speech intelligibility (e.g. intelligibility of syllables, words or sentences presented in quiet or in noise); (ii) hearing-specific health-related quality of life (QoL), where participation was the key domain, measured using any self-reported questionnaire (e.g. Hearing Handicap Inventory for the Elderly: Ventry and Weinstein 1982) and (iii) adverse effects of patient, reported as pain, discomfort, tenderness, skin irritation or ear infection as a consequence of device fitting. Secondary outcomes included any of the following self-reported outcomes: (i) General health-related QoL (e.g. Health Utilities Index Mark 3: Furlong et al. 2001); (ii) listening ability (e.g. Abbreviated Profile of Hearing Aid Benefit: Cox and Alexander 1995); (iii) cognition (e.g. working memory); (iv) feasibility (e.g. usability, adherence); (v) adverse effect of noise-induced hearing loss (e.g. due to over-amplification from inappropriate hearing aid fitting).

Study designs

Retrospective or prospective studies, randomised controlled trials, non-randomised controlled trials, before and after studies were included. Articles reporting expert opinions, practice guidelines, case reports, case series, conference abstracts and book chapters were excluded.

Search strategy

An initial literature search was conducted by a medical information specialist (Farhad Shokraneh, University of Nottingham) on 2 April 2016. Searches were last updated on 7 March 2018 to ensure that any newly published studies were included. The following databases were searched: CINAHL (via EBSCO host), Cochrane Library, EMBASE (via Ovid SP), MEDLINE (via Ovid SP), PubMed, Scopus, Citations Indexes of Web of Science, ISRCTN Registry, ClinicalTrials.gov and WHO International Clinical Trials Registry Platform. Supplemental 1 provides full electronic search strategies for all databases. All database searches were completed in one day and with no time, language, document type or publication status limitations. The search terms were collected based on free text and controlled vocabularies (Medical Subject Headings, Excerpta Medica Tree and CINHAL Headings), expert opinion, literature review and checking the test search results.

Additional information was identified manually through snowballing of the reference lists from included studies, as well as screening of related articles by shortlisted authors to identify any relevant articles that may not have been returned by the initial database searches. Contact with study authors was not necessary to ascertain whether any studies were ongoing.
Study selection

Two investigators (DM, AB) independently screened all identified references to decide eligibility according to the PICO criteria by reading the title and/or abstract. The full text was obtained for articles that appeared to meet eligibility or where there was any uncertainty (i.e. insufficient information to make a clear decision). We did not need to contact study authors for additional information to resolve questions concerning eligibility. Discrepancies were resolved through discussion between investigators.

Data collection process

A standardised data collection form constructed via Covidence (www.covidence.org) was used, which included study details (e.g., sponsorship source, country, setting), author’s contact details (name, institution, email, postal address), study design, population (inclusion/exclusion criteria, baseline characteristics), interventions (and comparators) and outcomes. Prior to starting the review, detailed guidance notes were devised by DM and were piloted by DM and AB to ensure consistency. Data collection was conducted by DM and AB independently, but in duplicate for every included record. Where necessary, study authors were contacted to resolve any uncertainties and to obtain any missing data. If data could not be obtained and were only presented in graphical form, the results were estimated from figures using WebPlotDigitizer (http://arohatgi.info/WebPlotDigitizer/app/). Disagreements about numerical data extracted from figures were discussed and resolved by averaging.

Risk of bias in individual studies

DM and AB independently assessed risk of bias of each included study with the Cochrane risk of bias tool (Higgins and Green 2011), which rates the studies as ‘high risk’, ‘low risk’ or ‘unclear risk’ in the following six domains: sequence generation, allocation concealment, blinding of participants and personnel, blinding outcome assessors, incomplete outcome data, selective outcome reporting and other sources of bias (e.g. influence of funders). As all included studies were non-randomised, the Downs and Black (1998) checklist, which consists of 27 criteria, was used to assess study quality. Criterion 27 was adapted to consider whether or not a power calculation was performed rather than whether there was sufficient power to detect a clinically meaningful change, as there is a lack of consensus regarding clinically meaningful change in hearing loss outcome measurement (Barker et al. 2016). All answers were scored 0 (‘no’) or ‘unable to determine’) or 1 (‘yes’), with the exception of criterion 5 (‘Are the distributions of principal confounders in each group of subjects to be compared clearly described?’), which was scored 0 (‘no’), 1 (‘partially’) or 2 (‘yes’). The total maximum score was 28, with study quality rated as excellent (26–28), good (20–25), fair (15–19) or poor (≤14) (Hooper et al. 2008).

Data synthesis

Included studies were reviewed in order to determine whether their data were suitable for inclusion in the meta-analyses. Meta-analyses were only performed when studies were broadly comparable in terms of study design, interventions and outcomes. For continuous data, where the studies used the same outcome measure, mean differences (MDs) were calculated with 95% confidence intervals (CI). If different outcome measures were employed, effect sizes were calculated as standardised mean differences (SMDs) in which the mean difference between conditions was divided by the pooled standard deviation (between-group) or by the standard deviation of the differences (within-group). Heterogeneity in effect sizes across studies was examined using the I² statistic and its significance was tested using a χ² test. This approach to quantifying heterogeneity provided a value from 0 to 100%, with low (0–40%), medium (41–60%) and high (61–100%) ranges (Higgins and Green 2011). In the absence of meta-analysis, primary and secondary outcomes were assessed at the individual study level through narrative synthesis (Popay et al. 2006).

Deviations from the published protocol

Although we pre-specified that the data would be subjected to both random and fixed effects models, we opted to only use a random-effects meta-analysis. A random effects approach was considered most appropriate based on an assumption that effect sizes would vary across studies, not only because they used different samples of participants (as assumed in a fixed effect approach) but also due to differences in assessment methodologies employed. In addition, we stated previously that we would not pool studies if I² exceeded 60%, suggestive of high heterogeneity (Higgins and Thompson 2002). However, although heterogeneity was high, which was to be expected due to differences across studies, we thought it useful to illustrate the pooled data for comparable studies irrespective of heterogeneity but took this into account when interpreting the data.

Results

A total of 2198 records were identified for screening. Following the removal of 1060 duplicate records, 1138 records were subjected to a three-stage screening process (Figure 1). The full texts of 149 articles that passed the initial title and abstract screen were retrieved. A total of 138 articles were not judged to have met inclusion criteria and were excluded. Eleven studies were included in the review. Supplemental 2 summarises the characteristics of the 11 studies included in the review. Two categories of alternative listening device (i.e. intervention) were evaluated, (i) PSAPs (n = 2); and (ii) remote microphone systems (n = 9). All studies were before-after comparisons, with participants acting as their own control. Comparators included unaided, conventional hearing aids alone or another alternative listening device. Four studies assessed outcomes at a range of follow-up durations, from one month (Sacco et al. 2016) to one year (Chisolm et al. 2007). For studies that were sufficiently similar in terms of interventions and outcomes, meta-analyses were performed. Narrative synthesis (Popay et al. 2006) are reported where meta-analyses were not possible.

PSAPs

Speech intelligibility

Two eligible studies (n = 73 participants) assessed six different PSAPs. Sacco et al. (2016) evaluated the TEO First listening device (Tinteo, Personal Sound Society), classified by the authors as an ‘over-the-counter’ hearing device. Reed et al. (2017) assessed five different PSAPs, which varied in terms of the purchase price as of 4 July 2017.
PSAPs vs. unaided. Summary effects and forest plots are provided in Supplemental 3 and Figure 2 respectively. Overall performance across both studies was better when using a PSAP relative to unaided conditions (Figure 2(A)). Heterogeneity was low ($I^2 = 0\%$). To provide a conservative estimate of effect, performance in Reed et al.'s (2017) study was pooled across all five included PSAPs. However, Reed et al. (2017) also showed that, in comparison to unaided, performance was superior for PSAPs that were priced ≥ US$299.99 (Reed et al. 2017).

PSAPs vs. hearing aids. Only one study (n = 42 participants) compared PSAPs with conventional hearing aids (Reed et al. 2017). This showed an effect favouring hearing aids compared to PSAPs that were priced ≤ US$269.99. However, performance did not differ statistically between hearing aids and PSAPs that were priced ≥ US$299.99.

Hearing-specific health-related QoL

Hearing-specific health-related QoL was only reported by Sacco et al. (2016) (n = 31 participants), who administered the Glasgow Hearing Aid Benefit Profile (GHABP) handicap subscales. Handicap scores reported as percentage change, decreased for the TEO First device compared to unaided handicap for following two items, (i) having a conversation without background noise (-9.6%, p = 0.018), and (ii) having a conversation with several people (-16.2%, p = 0.008).

Listening ability

Listening abilities were only reported by Sacco et al. (2016), whereby statistically significant decreases in GHABP residual disability subscale scores (percentage change) were found for the TEO First device compared to unaided disability for the following four items: (i) watching television (-18.5%, p = 0.011), (ii) having a conversation without background noise (-16.5%, p = 0.002), (iii) having a conversation in noisy background (-17.1%, p = 0.027) and (iv) having a conversation with several people (-20%, p = 0.014).

Feasibility

Sacco et al. (2016) assessed device acceptability using a six-point Likert scale, from zero (‘worst’) to five (‘best’). Mean scores ranged from 1.8 (SD = 1.4) for ‘satisfaction when using the noisy setting’, to 3.2 (SD = 1.6) for ‘ease of use’. The authors conclude that overall acceptability of the TEO First device was low-to-moderate. Mean duration of use was also measured, with participants reporting average daily use of 60 minutes. Feasibility was not reported by Reed et al. (2017).

Adverse effects

Sacco et al. (2016) explicitly reported that ‘no adverse events were observed’ during the course of the study. Adverse effects were not reported by Reed et al. (2017).

Remote microphone systems

Speech intelligibility

Six studies tested speech intelligibility. Summary effects and forest plots are provided in Supplemental 3 and Figure 2 respectively. For all meta-analyses, heterogeneity was high ($I^2 ≥ 80\%$), and statistically significant ($p ≤ 0.024$).

Hearing aids + FM system vs. Unaided. Three studies (Lewis et al. 2004, 2010; Rodemerk and Galster 2015) (n = 61 participants) showed that performance favoured the hearing aids + FM system compared to unaided (Figure 2(B)).

Hearing aids + FM system vs. Hearing aid alone. Two studies (Lewis et al. 2004; Rodemerk and Galster 2015) (n = 51 participants) showed that performance favoured the hearing aids + FM system compared to hearing aids alone (Figure 2(C)).

Hearing aids + FM system vs. FM only. Three studies (Rodemerk and Galster 2015; Lewis et al. 2010; Norrix et al. 2016) (n = 36 participants) showed that while performance favoured FM microphone only compared to hearing aids + FM system, the pooled effect was not significant (Figure 2(D)).

Hearing aids + FM system vs. Hearing aids + 2.4 GHz system. Two studies (Rodemerk and Galster 2015; Thibodeau 2014) (n = 27 participants) showed conflicting results, with one favouring hearing aids + FM (Rodemerk and Galster 2015) and the other favouring hearing aids + 2.4 GHz system (Thibodeau 2014) (Figure 2(E)).

Hearing aids + Bluetooth system vs. Hearing aids alone. Two studies (Kim et al. 2014; Rodemerk and Galster 2015) (n = 46 participants) showed performance favoured the hearing aids + Bluetooth system relative to the when the hearing aid was used alone (Figure 2(F)).

Remote microphone only mode vs. Unaided or hearing aid alone. Only one study (n = 16 participants) compared speech intelligibility across four different remote microphone systems (FM, Bluetooth, 900 MHz wireless, 2.4 GHz wireless) (Rodemerk and Galster 2015). All systems in microphone-only mode significantly improved performance relative to both unaided and hearing aid alone conditions ($p < 0.001$). The magnitude of this effect did not differ statically between systems.
Hearing-specific health-related QoL

There was no robust evidence as to whether remote microphone systems improved self-reported hearing-specific health-related QoL.

Hearing aids + FM system vs. Hearing aids alone. Only one study, using the MarkeTrak VI survey, compared the hearing aids + FM system and hearing aids alone (Chisolm et al. 2007). No statistically significant differences were found (six weeks, $n = 36$, $Z = 0.68$, $p = 0.499$; one year, $n = 30$, $Z = 0.71$, $p = 0.489$).

Hearing aids + Bluetooth accessories vs. Hearing aids alone. Only one study ($n = 12$ participants), using the International Outcome Inventory for Hearing Aids (IOI-HA, Cox and Alexander 2002), compared hearing aids + Bluetooth accessories to hearing aids alone (Smith and Davis 2014). A statistically significant improvement favouring hearing aids + Bluetooth accessories was found for residual participation restrictions ($Z = 2.12$, $p = 0.034$).

Listening ability

There was no robust evidence as to whether remote microphone systems improved self-reported listening abilities.

Hearing aids + FM system vs. Hearing aids alone. Two studies (Chisolm et al. 2007; Lewis et al. 2005) ($n = 59$ participants), using different variants of the Communication Profile for the Hearing Impaired (CPHI, Demorest and Erdman 1987), showed that listening abilities were significantly better ($p \leq 0.03$) for the hearing aids + FM system compared to hearing aids alone for social, work, and home situations.

Hearing aids + Bluetooth accessories vs. Hearing aids alone. Only one study ($n = 12$ participants), using the IOI-HA, GHABP and the Speech, Spatial and Qualities of Hearing Scale (SSQ, Gatehouse and Noble 2004) ($n = 12$ participants), compared listening abilities between hearing aids + Bluetooth accessories and hearing aids alone (Smith and Davis 2014). Statistically

Figure 2. Summary of the random effects meta-analyses for speech intelligibility: (A) PSAPs vs. unaided; (B) Hearing aids + FM system vs. Unaided; (C) Hearing aids + FM system vs. Hearing aids alone; (D) Hearing aids + FM system vs. FM only; (E) Hearing aids + FM system vs. Hearing aids + 2.4 GHz system; (F) Hearing aids + Bluetooth system vs. Hearing aids alone. Black squares = summary effect size of each study for speech intelligibility. Error bars = 95% confidence intervals (CI) for the summary effects. Diamond = overall effect size, lateral points indicate 95% CI for overall effect estimate.
significant improvements favouring hearing aids + Bluetooth accessories were found for residual activity limitations (Z = 2.24, p = 0.025) and residual disability (Z = 2.55, p = 0.011) subscales (IOI-HA, GHABP). No statistically significant differences (p ≤ 0.374) across all sub-scales were found between conditions when listening abilities were measured using the SSQ.

**Feasibility**

Three different studies assessed the following feasibility domains: (i) satisfaction (Chisolm et al. 2007); (ii) preferences (Thibodeau 2014); and (iii) usability (Smith and Davis 2014).

**Hearing aids + FM system vs. Hearing aids alone.** Using the MarkeTrak VI survey, Chisolm et al. (2007) found statistically significant improvements favouring the hearing aids + FM system for satisfaction in ‘noisy’ listening situations (e.g. restaurant, large group, leisure activities) (six weeks, n = 36, Z = 3.10, p = 0.002; one year, n = 30, Z = 2.27, p = 0.007) and ‘ability to hear soft sounds’ (six weeks, n = 36, Z = 3.31, p = 0.001; one year, n = 30, Z = 3.25, p = 0.001).

**Hearing aids + FM system vs. Hearing aids +2.4 GHz system.** Thibodeau (2014) (n = 10 participants) found that all participants reported that they preferred using hearing aids +2.4 GHz wireless system compared to hearing aids + FM system.

**Hearing aids + Bluetooth accessories vs. Hearing aids alone.** Smith and Davis (2014) (n = 12 participants) observed that the majority of participants (exact data not reported) reported that Bluetooth accessories ‘quite easy’ to use and improved the quality of sound when viewing the TV and using a cell phone.

**Risk of bias assessment**

Using the Cochrane risk of bias tool (Higgins and Green 2011), all studies were judged to be high risk with regard to selection bias (i.e. random sequence generation, allocation concealment) due to the nature of the before-after study design (Table 1). Risk of performance (blinding of participants/personnel) and detection bias (blinding of outcome assessment) was judged to be high for ten and eight studies respectively, as no blinding procedures were reported. Studies were judged to be low risk if blinding was stated, although it should be acknowledged that blinding in before-after studies is not always strictly possible as a consequence of this study design. The risk of attrition bias due to incomplete outcome data was judged to be low for all studies as there was no attrition in nine studies. In the remaining two studies, while attrition ranged from 16.67% (Chisolm et al. 2007) to 25% (Smith and Davis 2014) at the one year and 12 week follow-up respectively, reasons for incomplete outcome data were considered to be clearly reported in each article, increasing confidence that missing data had no undue influence on the results.

With the exception of Reed et al. (2017), risk of reporting bias (selective outcome reporting) was judged to be high for all studies, as numerical values were not sufficiently reported and/or were only provided for statistically significant results. In terms of risk of other bias, with the exception of three studies (Kim et al. 2014; Norrix et al. 2016; Reed et al. 2017), this was judged to be unclear for seven studies because of financial support from the manufacturer of the device(s) being evaluated. For one study (Smith and Davis 2014), an author acted as a consultant for a hearing aid manufacturer. For these reasons, potential vested interest could have posed a threat to validity.

**Quality assessment**

Scores on the Downs and Black (1998) checklist ranged from 13 (Smith and Davis 2014) to 21 (Sacco et al. 2016) out of a possible total of 28, indicative of a poor to good level of quality respectively. In terms of ‘reporting’, with the exception of one study (Sacco et al. 2016), adverse effects as a consequence of the intervention were not reported by any study. Whether participants were representative of the target population from which they were recruited (i.e. ‘external validity’) was also uncertain for the majority of studies, as sufficient detail was often lacking to make a clear judgement. In terms of ‘internal validity’ (e.g. randomisation, blinding), lower quality ratings arose because no studies randomised participants to intervention groups. Furthermore, only one study attempted to blind study participants to the intervention they received (Thibodeau 2014). Similarly, a power calculation was reported for only one study to determine sample size (Rodemerk and Galster 2015).

**Discussion**

In the current review, the scientific literature examining the effectiveness of alternative listening devices in adults with mild to moderate hearing loss was systematically searched. Eleven studies met eligibility for inclusion, two studies evaluated PSAPs and nine assessed remote microphone systems (FM, Bluetooth, wireless). The majority of studies primarily examined behavioural measures of speech intelligibility in noise. Self-reported hearing-specific QoL, listening ability and feasibility (i.e. usability, adherence, acceptability) were also evaluated, but to a lesser extent. There were some outcomes of potential interest that were not measured (i.e. cognition, general health-related QoL, adverse effects). Follow-up ranged from one month (Sacco et al. 2016) to one year (Chisolm et al. 2007), with no long-term follow-up greater than one year. There was considerable heterogeneity, whereby interventions and outcomes varied greatly across studies. The evidence was judged to be of poor to good quality, and subject to bias mainly due to limitations in study design.
Speech intelligibility

For speech intelligibility performance, data pooled across two studies demonstrated that there was a beneficial effect of PSAPs in improving performance compared to unaided conditions (Reed et al. 2017; Sacco et al. 2016). Findings reported by Reed et al. (2017) further suggest that this effect may be dependent on the cost of the PSAP assessed, whereby PSAPs that were priced ≥ US$299.99 improved speech intelligibility performance relative to unaided and did not differ statistically from hearing aids. A potential explanation for this finding may reside in existing evidence showing that higher priced PSAPs provide comparable electroacoustic characteristics to hearing aids (Callaway and Punch 2008). However, the extent to which these variables (i.e., price and/or electroacoustic characteristics) impact patient-reported outcomes for PSAPs remains to be established.

Similarly, speech intelligibility performance was superior for remote microphone systems used in conjunction with hearing aids (FM, Bluetooth, wireless) relative to both unaided and hearing aids alone (Kim et al. 2014; Lewis et al. 2004, 2010; Norrix et al. 2016; Rodemerk and Galster 2015). While we can be confident in the direction of the effect, due to high heterogeneity, the pooled effect size estimates could change with further evidence. Heterogeneity between studies most likely arose as a consequence of differences in assessment methodologies. In future, there should be greater consistency in the outcome measures used to assess speech intelligibility in adults with hearing loss. Standardised measurement procedures should be employed across studies so that they can be appropriately combined to enable direct comparison of effect sizes.

Self-reported hearing-specific QoL, listening abilities and feasibility

Outcome measures used to assess self-reported hearing-specific QoL, listening abilities and feasibility varied considerably across studies, as did duration of follow-up. This not only limited direct comparison but may also help to explain why the pattern of results was inconsistent across studies. On this basis, there is no robust evidence as to whether alternative listening devices included in this review improve these outcomes, or if the improvements observed are specific to the device, situations specified and/or outcome measures employed in each study. As a consequence, we suggest that the same self-report outcome measures should be consistently applied across studies, which should be appropriately sensitive and tap into the behavioural domains that they aim to reflect (Ferguson et al. 2014, 2017; Heinrich, Henshaw, and Ferguson 2016). There is a clear need for the development of a core outcome set in audiological rehabilitation research (Barker et al. 2015; Ferguson et al. 2017). In addition, longer follow-up durations greater than one-year would improve the certainty in the results, providing a better estimate of potential long-term benefit (Barker et al. 2016; Ferguson et al. 2017).

Study quality and risk of bias

The quality of the evidence included in this review was judged to be poor to good and subject to bias. The design of all studies was classified as ‘observational’ (i.e. before-after comparison), with no studies employing a separate control group. As a result, confidence in the effect size estimates are limited, as the true effects may be different. It cannot be known with certainty whether the effects seen are due to the devices, rather than due to regression to the mean or external factors affecting all participants. Further high-quality evidence is, therefore, required to improve confidence in the effect size estimates.

It should be noted that no studies assessing Smartphone-connected hearing aids or Smartphone ‘hearing aid’ apps met the inclusion criteria during the article screening process. We are aware of at least one non-peer reviewed publication that has evaluated Smartphone hearing aid apps compared to conventional hearing aids (Amlani et al. 2013). This study was not detected in the current review because it was published in an industry-related magazine (i.e. the grey literature). Inclusion of the grey literature could have provided a broader review of the available evidence. However, we opted to exclude databases of the grey literature in our pre-specified search strategy because there is no agreed method of extracting and synthesising evidence obtained from this literature in a clear and transparent way. This would also reduce, though not eliminate, the likelihood of including poor quality studies. Two of the study authors (DM, MF) worked collaboratively with the UK NIHR Horizon Scanning and Intelligence Centre to review new and emerging technologies for hearing loss (NIHR Horizon Scanning Research and Intelligence Centre 2017). Together with the current systematic review, and Cochrane review on hearing aids for mild to moderate hearing loss (Ferguson et al. 2017), all three reviews provide up-to-date high-level evidence of a wide range of listening devices.

Review limitations

A potential limitation of this review, as well as the field more generally, is that there is no consensus in terms of audiometric descriptors across different countries and organisations. In this review, we used the audiometric descriptors for mild to moderate hearing loss, based on pure-tone air-conduction thresholds established by the British Society of Audiology (2011). An average hearing threshold in the better hearing ear across octave frequencies 0.25–4 kHz that are ≥20 and ≤40 dB HL is defined as ‘mild’, and ≥41 and ≤70 dB HL as ‘moderate’. Other definitions adopt different frequency ranges and intensity cut-offs, such as the World Health Organisation definitions, whereby average thresholds across 0.5–4 kHz between 26 and 40 dB HL is defined as ‘mild’, and 41 to 60 dB HL as ‘moderate’ hearing loss (Mathers, Smith, and Concha 2000). In addition, individual hearing threshold data were seldom reported for all included studies and could not be made available by the study authors for logistical reasons. Although unlikely, included studies could have included some participants with more severe degrees of hearing loss. Nevertheless, in accordance with our published protocol (Maidment et al. 2016), studies were verified as eligible for inclusion because the mean average hearing threshold always fell within the pre-specified range.

Research recommendations

On the basis of this review, further high-quality evidence, namely randomised controlled trials (RCTs), is needed to evaluate whether alternative listening devices are a clinically- and cost-effective intervention for adults living with hearing loss. This is in-line with a research recommendation for hearing loss assessment and management specified by the UK National Institute for Clinical Excellence (2018). Such research has also been identified as a high-priority need in the US (Humes et al. 2017; National Academies of Science 2016), given that alternative
listening devices could enable new service delivery models (e.g. direct-to-consumer, over-the-counter). Moreover, there have been recent legislative changes in the United States, with the introduction of the Over-the-Counter (OTC) Hearing Aid Act of 2017, which aim to improve accessibility and affordability of hearing-healthcare for adults. High-quality evidence, therefore, is needed as a priority in this area. In a recently published randomised, double-blind, placebo-controlled clinical trial (Humes et al. 2017), hearing aids fitted by an audiologist (i.e. ‘audiology best practices’) and hearing aids fitted using an OTC model resulted in similar effect sizes for measures of speech recognition and hearing aid benefit. Nevertheless, satisfaction and percentage likely to purchase hearing aids post-trial were lower for the OTC model, potentially attributable to the lack of audiological interaction during the provision of the hearing aids (Humes et al. 2017). It has been proposed that adults living with hearing loss may require optional assistance to successfully use alternative listening devices that do not require a hearing healthcare professional in terms of device fitting and/or fine-tuning (Keidser and Convery 2016). In support, the results of our mixed-methods usability study, completed following this systematic review, suggest that people living with hearing loss would like greater instruction to use and adjust alternative listening devices themselves (Maidment and Ferguson 2017, Forthcoming). Remotely-delivered information is one means of providing such assistance, and has been shown to successfully supplement the provision of hearing aids, resulting in improved outcomes (Ferguson et al. 2016; Kramer et al. 2005; Thorén et al. 2014). Therefore, we recommend that this concept be incorporated in the design of future effectiveness trials.

Conclusions
In summary, the evidence included in this review suggests that alternative listening devices improve behavioural measures of speech intelligibility relative to unaided and/or conventional hearing aids. There is no robust evidence as to whether alternative listening devices improve self-reported outcomes. Furthermore, the evidence was judged to be poor to good quality and subject to bias due to limitations in study design. On this basis, we argue that high-quality studies (i.e. RCTs) investigating the clinical- and cost-effectiveness of alternative listening devices are needed in this area. Such evidence, though currently unavailable, is necessary to guide healthcare commissioners and policymakers when considering new service delivery pathways to benefit adults living with hearing loss. Moreover, given that this field is likely to continue to develop in new and unexpected ways, we envisage that the current systematic review will require updating, and it is our intention to do so in two to three years. These rapid developments reflect the innovative nature of the field, which not only has the potential to transform hearing healthcare service delivery in the future but also increase the likelihood that people will seek and use amplification to successfully manage their hearing loss.

Acknowledgements
The authors would like to thank Farhad Shokraneh (Medical Information Specialist, University of Nottingham) for devising the search strategy.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This research was funded by the NIHR Nottingham Biomedical Research Centre and carried out at the NIHR Nottingham Clinical Research Facilities. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

ORCID
David W. Maidment http://orcid.org/0000-0002-8416-519X
Melanie A. Ferguson http://orcid.org/0000-0002-8096-869X

References