

Vickers, Deborah A. and Riley, Alison and Ricaud, Rebecca and Verschuur, Carl and Cooper, Stacey and Nunn, Terry and Webb, Kath and Muff, Joanne and Harris, Frances and Chung, Mark and Humphries, Jane and Langshaw, Alison and Poynter-Smith, Emma and Totten, Catherine and Tapper, Lynne and Ridgwell, Jillian and Mawman, Deborah and de Estibariz, Unai Martinez and O'Driscoll, Martin and George, Nicola and Pinto, Francesca and Hall, Anne and Llewellyn, Carol and Miah, Razun and Al-Malky, Ghada and Kitterick, Pádraig T. (2016) Preliminary assessment of the feasibility of using AB words to assess candidacy in adults. Cochlear Implants International, 17 (sup1). pp. 17-21. ISSN 1467-0100

Access from the University of Nottingham repository:

<http://eprints.nottingham.ac.uk/40152/1/Post-review%20manuscript%20for%20archiving.pdf>

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

Preliminary assessment of the feasibility of using AB words to assess candidacy in adults

Deborah A. Vickers¹, Alison Riley², Rebecca Ricaud³, Carl Verschuur³, Stacey Cooper⁴, Terry Nunn⁴, Kath Webb⁴, Joanne Muff⁵, Frances Harris⁵, Mark Chung⁵, Jane Humphries⁶, Alison Langshaw⁶, Emma Poynter-Smith⁶, Catherine Totten⁷, Lynne Tapper⁷, Jillian Ridgwell⁷, Shahad Howe⁸, Unai Martinez de Estibariz⁸, Deborah Mawman⁸, Nicola George⁹, Francesca Pinto¹⁰, Anne Hall¹⁰, Carol Llewellyn¹¹, Razun Miah¹¹, Ghada Al-Malky¹, Pádraig T. Kitterick¹²

¹UCL Department of Speech, Hearing and Phonetics Science, London. WC1N 1PF, UK

²Midlands Auditory Implant Programme (Adults), Birmingham B15 2TH

³University of Southampton Auditory Implant Service, Southampton SO17 1BJ

⁴St. Thomas' Hearing Implant Centre, London SE1 7EH

⁵Emmeline Centre for Hearing Implants, Cambridge CB2 0QQ

⁶Oxford Hearing Implant Programme, Oxford. OX3 9DU

⁷Yorkshire Cochlear Implant Service, Bradford. BD9 5HU

⁸The Richard Ramsden Centre for Auditory Implants, Manchester. M13 9WL

⁹Cardiff Adult Cochlear Implant Programme, Cardiff, CF4 4XW

¹⁰University College London Hospital Implant Programme, London. WC1X 8DA

¹¹Bridgend Cochlear Implant Programme, Bridgend, CF31 1RQ

¹²National Institute for Health Research Nottingham Hearing Biomedical Research Unit, Nottingham NG1 5DU

Abstract

Adult cochlear implant (CI) candidacy is assessed in part by the use of speech perception measures. In the United Kingdom the current cut-off point to fall within the CI candidacy range is a score of less than 50% on the BKB sentences presented in quiet (presented at 70 dB SPL).

The specific goal of this article was to review the benefit of adding the AB word test to the assessment test battery for candidacy. The AB word test scores showed good sensitivity and specificity when calculated based on both word and phoneme scores. The word score equivalent for 50% correct on the BKB sentences was 18.5% and it was 34.5% when the phoneme score was calculated; these scores are in line with those used in centres in Wales. The goal of the British Cochlear Implant Group (BCIG) service evaluation was to determine if the pre-implant assessment measures are appropriate and set at the correct level for determining candidacy, the future analyses will determine whether the speech perception cutoff point for candidacy should be adjusted and whether other non-speech based measures should be used in the candidacy evaluation.

Introduction

Current practice in the United Kingdom (UK) for assessing candidacy for cochlear implants (CIs) was determined based on research conducted over ten years ago (UK study group 2004). The project involved the collaboration of 13 UK CI centres working with academics from the Medical Research Council Institute of Hearing Research to determine the appropriate limits for candidacy for CIs in adults. All study centres used the same testing apparatus, and the same standard measures for assessing CI candidacy. This included the Bamford, Kowal and Bench (BKB) sentences presented in quiet and without any visual information from the talker's face or lips.

The vocabulary used in the development of the BKB sentences was taken from the utterances of eight to fifteen year old partially hearing children when describing everyday scenarios. The sentences could therefore be used with the majority of implant users pre-and post-implant. The scoring of BKBs was based on the percentage of keywords correct; the recommended cut-off criterion for falling within the CI candidacy range was set at <50% correct. This criterion was adopted by the National Institute for Health and Care Excellence (NICE) in 2009 (NICE, 2009) and is in place today.

The use of monosyllabic word tests or speech in noise assessments to assess outcomes post-implantation was uncommon in the UK in 2004, because the post-operative performance of implantees listening to sentences in quiet was rarely, if ever, constrained by ceiling effects. With improvements in patient outcomes over the years, the assessment of adult patients has become more sophisticated and sensitive to changes in post-operative performance levels. These changes in outcome over the years have been attributed to changes in candidacy, surgical practice, clinical intervention and device development (Blamey et al., 2013).

In the UK, CI teams struggle with the current candidacy criteria because they do not permit all individuals whom they currently consider as likely candidates to be implanted. In particular, two issues arise from the use of BKB sentences alone for assessing candidacy. The first issue is that they use highly predictable materials. Some candidates may only have minimal access to speech cues but because they have refined cognitive skills, they are able to work out the content of the sentence in spite of the severe loss of auditory information. While this skill is unlikely to benefit their speech perception in everyday complex listening situations, it may result in too high a test score for them to be eligible for implantation.

The second issue is that BKB sentence scores do not reflect the hearing abilities of some patients at the lower end of the performance scale. Individuals who do not speak English as their primary language may not have the same abilities as a native English speaker to fill in gaps in predictable sentences because their linguistic knowledge is poorer. The same issue arises in pre-lingually deaf adults who may not have a high levels of language sufficient to be able to get a representative score on the BKB sentences. Craddock et al. (2016) suggested that the CUNY audio-visual sentences might be more appropriate for this group.

Doran and Jenkinson (2016) reported on the additional criteria used by the Welsh CI teams. Individuals who have borderline BKB scores (<60%) are evaluated with a second measure, the Arthur Boothroyd (AB) monosyllabic word test. They based their cut-off levels on the 10th percentile of post-operative scores for both BKB sentences (50%) and AB words (15%) and used these as their criteria. The advantages of the AB words are that they are a better measure of speech audibility and not subject to the predictability effects associated with the BKB sentences. However, just like the BKB sentences, the confidence intervals for AB words scores are also extremely wide and dependent upon the number of items used (Thornton and Raffin, 1978) and setting an absolute single value for candidacy may not be appropriate. For example, a patient whose ‘true’ performance level is at 50% on the BKB sentences may actually achieve scores between approximately 40-60%. In Wales, the additional rule is applied for individuals who score greater than 50% on the BKB sentences, if their AB word score is less than 15% then they are considered to be potential candidates for CI.

Currently, the recommendation for scoring the AB words is to use the “word” score; i.e. the proportion of whole words reported correctly. However, a score based on the proportion of phonemes reported correctly (“phoneme” score) may give a finer resolution of performance ability. This approach to scoring may also be more appropriate for those patients with English as a second language or pre-lingually deafened patients in whom the capacity to report whole words may not relate well to audibility.

It is known that there is a strong linear relationship between BKB sentence and the AB word scores (Vickers et al., 2009). The use of both assessments to verify candidacy has great potential for improving the quality and reliability of the candidacy assessments. An additional assessment provides greater clinical confidence in the eligibility of an individual because of this “second look” at performance. Exploratory analyses were therefore conducted to assess whether the 15% AB word score described by Doran and Jenkinson (2016) determines eligibility in a similar way to the 50% BKB sentences score.

Ten UK CI centres have been engaged in collecting clinical data to determine if the BKB sentence test is still the most appropriate pre-operative measure for determining candidacy for CIs in the UK. The ultimate goal of the data collection is to determine if the actuarial model for candidacy for CIs in adults in the UK should be updated. However, the specific goal of this phase of the project was purely an exploratory analysis to evaluate the potential usefulness of the AB words.

Method

Study participant inclusion criteria

There were 134 data sets collated by December 1st 2015. Ninety seven of those had scores for both BKB sentences and AB words in quiet; 37 were missing data on one or both measures. This represented all data sets entered and verified at the point of analysis. Participants had all been seen by a participating adult implant programme and were able to undergo speech perception testing. Each centre registered the data collection with their individual Hospital Research and Development boards as part of a multi-site service evaluation for the British Cochlear Implant Group.

Speech perception test delivery

The BKB sentences and AB words were delivered in sound-treated rooms at the individual hospitals and presented at 70 dBSPL. The BKB sentences were scored by the number of key words correct and the AB word test was scored by both phoneme and word.

Demographic information

Data were also *collated* ?collected? on home language, aetiology, duration of profound deafness and age at implantation.

Analysis

The scores on these measures were evaluated in two ways. First, the scores for all participants were reviewed to determine if their AB words or BKB sentences fell within the candidacy range proposed by Doran and Jenkinson (2016); i.e. <50% on BKB sentences and/or <15% on AB words. If there were any anomalies the demographic factors were explored to determine if the findings could be explained and highlighting areas to be included in the final actuarial model.

Secondly, comparisons of the AB word and BKB sentence scores were made using Receiver Operating Characteristic (ROC) curve analysis. Patients were first classified as either eligible or ineligible based on their BKB sentences score, where eligible patients were those with a score less than 50% correct, in line with NICE guidance. An ROC curve was then fitted to this binary variable and their corresponding scores on the AB words test. Two analyses were performed: one using the word score and another using the phoneme score. These ROC curves were fit and compared using the pROC package for the R statistical programming environment (Robin et al., 2011). The ‘optimal’ word and phoneme cut-off scores for the AB word test were determined as those scores that simultaneously maximised the sensitivity and specificity of the test. In this context, sensitivity refers to the test’s ability to identify eligible patients and specificity to its ability to identify ineligible patients.

Results

The mean age at implantation was 57 years (median 61 years; range 16 to 88 years) and the mean duration of deafness was 22.5 years (median 15 years; range 0 to 80 years). Eight individuals had English as an additional language (EAL) and five people used British Sign Language as their primary language. Table 1 shows the number of individuals who would pass or fail the BKB sentences (pass cut off level was $\geq 50\%$) or the AB words (pass cut off level was $>15\%$). If the AB words were used as the only candidacy measure fewer people would fall within the candidacy range (78 people) versus 85 people with the BKB sentence criteria. If both assessments were used together an additional one person would be considered for implantation on top of the current BKB cutoff at 50%.

Table 1 here

Table 2 provides the demographic details of the individuals who had different pass/fail criteria based on the two measures.

Table 2 here

The single individual whose scores would fall within candidacy with the use of AB words is shown on the first row of table 2. This individual was pre-lingually deafened and had a conductive component to the hearing loss. This individual was able to detect phonemes but could not identify the words. This individual did remarkably well with the BKB sentences maybe because of the vocabulary level of the materials being appropriate for children. The eight other individuals were those who passed the AB words and failed the BKB sentences, these individuals either did not have English as their primary language or were older than the mean age of the group. Both aspects might make the audibility of phonemes and words easier than understanding the BKB sentences, where gap filling might be necessary to correctly perceive the sentences. Within the dataset forty individuals were over the UK retirement age (65 years old), highlighting the importance of accurately defining the appropriate assessment tools for older adults.

It should also be noted that for some tasks the scores were borderline (between 50 and 60% on BKB sentences (6 people) and between 15 and 33% on AB words (11 people)).

Figure 1 plots the ROC curves for the word and phoneme scores derived from the AB words test. It would be straightforward to choose a very high AB word or phoneme score as a cut-off that would successfully identify all patients with BKB scores <50%; i.e. eligible patients. Similarly, it would be easy to choose a very low word/phoneme score as a cut-off that would successfully identify all patients with BKB scores \geq 50%; i.e. ineligible patients. ROC curves

are used to identify the cut-off score that tries to do both of these tasks simultaneously; i.e. the score that would classify patients as eligible or ineligible as similarly as possible to the 'gold standard' of <50% on the BKB sentence test. The closer the curves are to the top-left corner of the Figure 1, the more similar they are to the BKB 50% cut-off at determining candidacy.

Both word and phoneme scores showed high sensitivity (accuracy with identifying eligible patients) and specificity (accuracy with identifying ineligible patients). The optimal cut-off word score was found to be 18.5% (sensitivity 90.9%, specificity 96.3%) and the optimal phoneme score was 34.5% (sensitivity 100%, specificity 93.0%). A paired comparison of the two curves using DeLong's test (DeLong et al., 1988) revealed no significant difference ($Z=-1.02$, $p=0.31$), suggesting that both word and phoneme scores assessed candidacy in a similar way. **I FOUND THIS RATHER DIFFICULT TO UNDERSTAND! DO YOU THINK KIT MIGHT BE POSSIBLE TO EXPLAIN IT MORE CLEARLY TO THOSE NOT FAMILIAR WITH ROC CURVES AND THEIR USE?!**

Figure 1 here

Discussion

This exploratory analysis was conducted to determine if the data suggests that the addition of the AB words to the candidacy test battery or even the replacement of BKB sentences with AB words would be appropriate. The candidacy criteria recommended by Doran and Jenkinson (2016) were used to determine how the numbers of candidates would change if the AB word criteria were added to the BKB criteria or if the AB words were used alone. The

analysis suggested that fewer people would fall within the candidacy range if only the AB words (<15%) were used. This might suggest that the relationship between the two measures was not appropriately set. Further exploratory analyses based on ROC curve analysis produced an AB word cutoff (18.5%) which was marginally higher but very similar to the 15% suggested. This shows that the BKB and AB tests produce quite consistent data when assessing candidacy across different CI centres and patient catchment areas.

At the current 50% BKB cut-off, both word and phoneme scores from the AB word test seem to be highly sensitive and specific. This suggests that they may be suitable replacements/or additions to the test battery for those individuals where there might be concerns over the validity of the BKB sentence test; e.g. EAL, pre-lingually deafened, first-language BSL user.

Current results suggest that it may be productive to use ROC analysis to compare BKB and AB word tests. The future analysis of the BCIG service evaluation data should replicate the current approach with a larger sample of data. Both pre and post-implantation data can be used to ensure that the entire performance range is covered. A larger sample would also allow the analysis to be conducted on only the subset of post-lingually deafened patients with English as a first language to eliminate any concerns about the validity of BKB scores. The cut-off produced by that data could then be applied to the sub-population of patients where BKB scores may not be valid to see whether AB word scores produce a similar proportion of eligible candidates.

This analysis has demonstrated that the AB words may have potential as a measure in the candidacy test battery. There is no evidence at this stage to suggest that it should replace the BKB sentences, but the exploration did identify cases (EAL, older adults, BSL users and

individuals with high levels of cognitive processing) where it could be beneficial as an additional tool for assessment. The absolute values that are appropriate for candidacy and the actuarial equation will re-evaluated in the future analysis.

References

Blamey P, Artieres F, Başkent D, Bergeron F, Beynon A, Burke E, Dillier N, Dowell R, Fraysse B, Gallégo S, Govaerts PJ, Green K, Huber AM, Kleine-Punte A, Maat B, Marx M, Mawman D, Mosnier I, O'Connor AF, O'Leary S, Rousset A, Schauwers, K, Skarzynski H, Skarzynski PH, Sterkers O, Terranti A, Truy E, Van de Heyning P, Venail F, Vincent C, Lazard DS (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiology and Neurotology*, 18 (1), 36-47

Craddock, Cooper, Riley, Wright (2016) Cochlear Implants for Pre-Lingually Profoundly Deaf Adults. **This supplement**

Doran M & Jenkinson L (2016) Mono-syllabic word test as a pre-operative assessment criterion for cochlear implant candidature in adults with acquired hearing loss. **This supplement**

National Institute for Health and Care Excellence Technology Appraisal Guidance (2009). Cochlear implants for children and adults with severe to profound deafness. *NICE technology appraisal guidance [TAG166]*. <http://www.nice.org.uk/ta166> accessed 29/01/2016

- 276 Thornton and Raffin (1978) Speech discrimination scores modelled as a binomial variable.
277 Journal of Speech and Hearing Research. 21, 507-518
278
- 279 UK CI study group (2004) Criteria of Candidacy for Unilateral Cochlear Implantation in
280 Postlingually Deafened Adults I: Theory and Measures of Effectiveness. Ear &
281 Hearing 25, 310–335
282
- 283 Vickers D, Eyles J, Brinton J, Glasberg B, Graham J. (2009) Conversion of scores
284 between Bamford, Kowal and Bench (BKB) sentences and Arthur Boothroyd (AB) words
285 in quiet for cochlear implant patients. *Cochlear Implants International*. 10(3), 142-149
286
- 287 DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L. (1988). Comparing the areas under two or more
288 correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44,
289 837–845.
- 290 Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J-C, Müller, M. (2011). pROC:
291 an open-source package for R and S+ to analyze and compare ROC curves. *BMC*
292 *Bioinformatics* 12, 77.
293

Table 1 Number of individuals who would or would not fall within the implant candidacy range based on BKB and AB word cutoffs. The two shaded boxes indicate the decisions that are the same if either BKB or AB criteria are used. If only AB words were used an additional one individual would fall into criteria but eight individuals would fall outside criteria. If both tests were used together one extra individual would fall within the candidacy range.

	BKB < 50%	BKB ≥ 50%	
AB ≤ 15% (scored by word)	77	1	78 Would receive CI with only AB criteria
AB > 15% (scored by word)	8	11	19 Would NOT receive CI with only AB criteria
	85 Would receive CI with only BKB criteria	12 Would NOT receive CI with only BKB criteria	

Table 2 Demographic details for nine individuals who had different pass/fail criteria for BKBs and AB words. The shaded row is the one participant who passed BKB sentences but failed the AB words. They hadn't gone forward for a CI and received a bone conduction device. The other failed the BKB sentences but passed the AB words. The individual with Polish as primary language didn't receive an implant due to audiometric thresholds being outside candidacy range.

Primary Language	Aetiology	Age at testing	Duration of Profound Deafness (years)	Onset	% BKBs in quiet	% ABs (phoneme score)	% ABs (word score)
English	Atresia	62	60	Pre-lingual	56	36	0
English	Otosclerosis	60	Unknown	Post-lingual	48	27	16
English	Unknown	70	8	Post-lingual	0	23	17
English	Unknown	78	28	Post-lingual	0	23	17
English	Anaesthesia	78	15	Post-lingual	0	28	17
English	Unknown	69	1	Post-lingual	40	39	17
English	Unknown	76	Unknown	Post-lingual	36	33	20
BSL	Waardenburg type II Syndrome	39	36	Pre-lingual	47	29	20
Polish	Unknown	88	10	Post-lingual	8	45	33

307 Figure legend

308

309 Figure 1 Receiver Operating Characteristic (ROC) curves for the word (solid curve) and
310 phoneme (dashed curve) scores from the AB word test. The symbols mark the score that
311 maximises the sensitivity and specificity of each scoring method; i.e. the cutoff score that
312 identifies eligible and non-eligible candidates as close as possible to the BKB sentence test
313 using the 50% cutoff criterion.

314

315

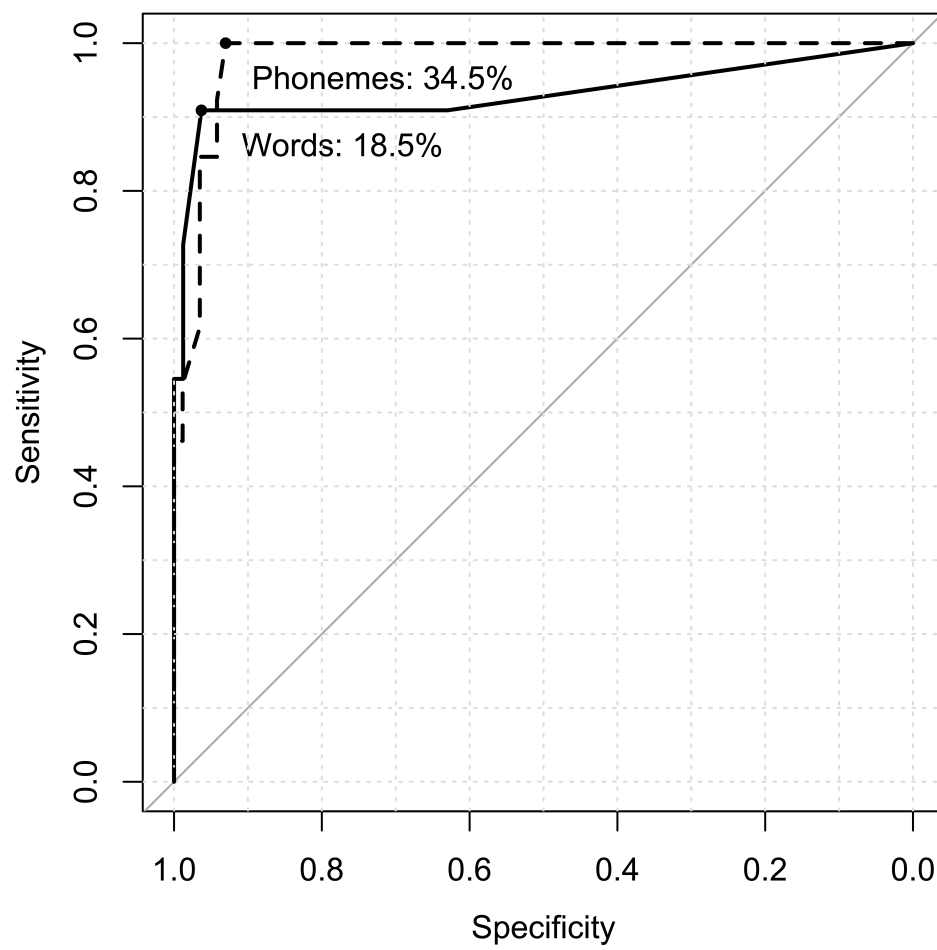


Figure 1