Recommended Procedure

Auditory Brainstem Response (ABR) Testing in Newborns

Date: February 2025

Due for review before: July 2029

General foreword

This document presents a recommended procedure by the British Society of Audiology (BSA). This recommended procedure represents, to the best knowledge of the BSA, the evidence-base and consensus on good practice, given the stated methodology and scope of the document at the time of publication. Although care has been taken in preparing this information, the BSA does not and cannot guarantee the interpretation and application of it. The BSA cannot be held responsible for any errors or omissions, and the BSA accepts no liability whatsoever for any loss or damage howsoever arising. This document supersedes any previous recommended procedure by the BSA and stands until superseded or withdrawn by the BSA.

This replaces the previous document dated 2019. It is recognised that occasionally, for technical reasons, clinicians may need to depart from BSA advice. It is the clinician's responsibility to identify and justify any such departure in the clinical report and acknowledge any associated limitation in the precision or scope of results. Their decision should be evidence based. Such departure from BSA guidance should not be routine. If clinicians are unsure how to interpret the guidance or seek clarification, they are encouraged to contact the appropriate BSA Special Interest Group via email: bsa@thebsa.org.uk

This document will be reviewed by the date given on the front cover. The formal BSA process of reviewing its documents can be lengthy which reflects the diligence required in ensuring the content is accurate. It is recognised that rapid technological changes can occur and for this reason, BSA will strive to provide interim updates where this is possible. These will be advertised via the usual BSA communication channels and will be available on the BSA website. However, should any individual or organisation feel that the content requires immediate update, review or revision, they should contact the BSA using the email bsa@thebsa.org.uk. Please add 'BSA document revision request' in the title. You will be asked to complete a short form with your reasons and this will be passed to the Professional Guidance Group for assessment. Comments on this document are welcomed and should be sent to:

British Society of Audiology
Blackburn House,
Redhouse Road,
Seafield,
Bathgate,
EH47 7AQ.
Tel: +44 (0)118 9660622

bsa@thebsa.org.uk
www.thebsa.org.uk

Published by the British Society of Audiology

Authors & Acknowledgments

Produced by: The Electrophysiology Special Interest Group (EPSIG) and the Professional Guidance Group
Key Authors:

Guy Lightfoot	ERA Training & Consultancy Ltd
Inga Ferm	Croydon University Hospital
Jason Smalley	Nottingham University Hospitals NHS Trust

Declarations of interests

ERA Training & Consultancy Ltd offer training courses in ABR testing, training and accreditation in ABR peer review and offer technical advice to centres performing ABR testing.

Jason Smalley is a paid Associate Lecturer for Anglia Ruskin University which offer training courses in Audiology and ABR testing.

With thanks to:

Those who made contributions to this version, including:
Paul Bacon, Rachel Beeby, Steven Bell, Siobhan Brennan, Victoria Evans, John Fitzgerald, Michelle Foster, Constantina Georga, Jackie Harding, Verity Hill, Mike Maslin, Lesley Peplow, Emanuele Perugia, Vivien Thorpe.

Those who contributed to earlier versions, including:
Rob Gardner, Graham Sutton, John Stevens, Rachel Booth, Rachel Feirn, Rhys Meredith, Rob Low, Steve Mason, David Stapells, Clive Elliott, Dave Parker, Mike Vidler and Amanda Hall.

BSA members who provided feedback in the public consultation.

Citation

Please cite this document in the following format:

British Society of Audiology (*Year of publication*), *Full title of document.* [Online]. Available from: *insert web link*. [Accessed *date*]

Shared Decision-Making

It is implied throughout this document that the service user should be involved in shared decision-making when undertaking audiological intervention, receiving subsequent information and understanding how it will impact on the personalisation of care. Individual preferences should be taken into account, and the role of the clinician is to enable a person to make a meaningful and informed choice. Audiological interventions bring a variety of information for both the clinician and the patient which can be used for counselling and decision-making regarding technology and anticipated outcomes.

Contents

1 Abbreviations 6
2 Introduction 6
2.1 Development of the recommended procedure 6
2.2 Aims 7
2.3 Scope 7
2.4 Staff Training, accreditation and peer review 7
3 Preparation for testing 7
3.1 Test environment 7
3.2 Calibration 8
3.3 Stage A checks 8
3.4 Recording system checks 10
3.5 Choice of electrodes and application 10
3.6 Differences in Electrode Impedance 10
3.7 Electrode location – AC and BC 11
4 Stimulus 11
4.1 Stimulus and stimulus rate: Air and bone conduction 12
4.2 Earphone 14
4.3 Warning – insert earphones 14
4.4 Bone vibrator 14
4.5 Placement of bone vibrator 14
4.6 Pressure to apply to bone vibrator 15
4.7 Effect of age on the bone-conduction and insert stimulus 15
5 Data collection and analysis 16
5.1 Characteristics of the ABR 16
5.2 Artefact rejection (AR) level 17
5.3 Recording filters 18
5.4 Use of digital / display filters on the averaged waveform 18
5.5 Notch filter 19
5.6 Display 19
5.7 Criteria for accepting the presence or absence of a response 19
5.8 Advice on when to stop averaging based on residual noise 25
5.9 Ear-specific recording 26
5.10 Resolving inconclusive results 31
5.11 Definition of ABR threshold for NHSP 31
5.12 Gold standard thresholds 32

5.13 Testing at other frequencies 33
5.14 Reporting thresholds (incl those which are not gold standard) 34
5.15 Baseline drift and the use of blocked-stimulus runs 35
5.16 Post-auricular myogenic (PAM) responses 35
6 Artefacts 36
6.1 Blocking of stimulus artefact 36
6.2 Control (blocked stimulus) recordings during testing 36
7 Records Management 38
Appendix A: Major changes in this recommended procedure 39
Appendix B: 2-channel recording of wave V 40
Appendix C: More detailed advice and examples of ABR waveforms ..... meeting the response criteria CR, RA and Inc 41
C1. Establishing a clear response (CR), satisfying the 3 to 1 ..... signal to noise criterion 41
C2. Estimating residual noise: deciding between response absent ..... (RA) and inconclusive (Inc) 42
C3. The use of blocked-stimulus control runs to assist interpretation 44
Appendix D: ABR dBnHL to dBeHL corrections for babies 46
Appendix E: Objective measures for ABR interpretation in babies 48
E1. Use of response confidence measures (Fsp/Fmp/SNR) in ABR ..... testing in babies 48
E2. Residual noise 50
E3. Using objective confidence measurements as an alternative to ..... waveform replication 51
Appendix F: Baseline drift 54
F1. Possible mechanisms 54
F2. Strategies available to the tester to deal with baseline drift 54
F3. Examples of ‘baseline drift’ and other waveforms with non-flat ..... baselines that represent interpretation difficulties 54
Appendix G: Frequently asked questions 61
References 72

1 Abbreviations

ABR	Auditory brainstem response
AC	Air-conduction
ANSD	Auditory Neuropathy Spectrum Disorder
AR	Artefact Rejection
ASSR	Auditory steady state response
BC	Bone-conduction
ckABR	Click-evoked ABR
CM	Cochlear microphonic
CR	Clear Response
dBeHL	Estimated PTA from electrophysiological thresholds
dBnHL	Decibels Hearing Level
ECochG	Electrocochleography
eSP	e-Screener Plus
FAQ	Frequently Asked Questions
Inc	Inconclusive
NBchirp	Narrowband chirp
NHSP	Newborn Hearing Screening Programme
OAE	Oto-Acoustic Emission
RA	Response Absent
RETFL	Reference Equivalent Threshold Force Level
RETSPL	Reference Equivalent Threshold Sound Pressure Level
RMS	Root mean square
S4H	Smart4Hearing
SNR	Signal to noise ratio
tpABR	Tone pip evoked ABR

2 Introduction

2.1 Development of the recommended procedure

This document has been adapted from the document 'Guidance for Auditory Brainstem Response testing in babies', 2019, (see Appendix A for major changes incorporated into this procedure). The development of this recommended procedure was undertaken by the members of the Electrophysiology Special Interest Group (EPSIG) and has been developed in accordance with BSA Procedure for Processing Documents (2016). Reference is made to other BSA guidance documents. Many will be updated after publication of this guidance so always refer to the latest version.

2.2 Aims

This recommended procedure's aim is to establish how to perform and interpret frequency-specific ABR threshold tests in newborns using air and bone conduction transducers. It also defines criteria with which ear-specific thresholds can be reached to establish a baby's hearing level.

2.3 Scope

These guidelines are for the use of ABR in assessing hearing in newborns (up to a corrected age of 12 weeks). For this cohort, frequency-specific information is required[1]. The document should be read in conjunction with the BSA practice guidance "Guidelines for the Early Audiological Assessment and Management of Babies Referred from the Newborn Hearing Screening Programme" (BSA, 2021) which describes the whole process of assessing hearing in neonates including the use of ABR. Reference should also be made to the other NHSP Audiology protocols and guidelines available on the BSA web site http://www.thebsa.org.uk/, and in particular the BSA recommended procedures: Assessment and Management of Auditory Neuropathy Spectrum Disorder (ANSD) in young infants (BSA, 2019a) and Cochlear Microphonic Testing (BSA, 2019b).

This document covers the technical procedure of carrying out an ABR test together with the interpreting and reporting of the results. It does not cover equipment for 'automatic' or screening ABR.

2.4 Staff Training, accreditation and peer review

The minimum requirement for staff conducting and interpreting newborn ABR tests is to attend a specialist ABR training course, supplemented by a period of directly supervised clinical practice, followed by clinical sign-off (accreditation) by a recognised external clinical expert. It is beyond the scope of this guidance to make specific recommendations for this; at the time of writing, arrangements for clinical sign-off are in development. Testers must participate in a scheme of structured external peer review (BSA, 2019c).

3 Preparation for testing

3.1 Test environment

Threshold ABR tests should be performed in a sound-proofed room or environment in which the lowest air and bone conduction stimulus levels that are to be used (typically 15dBnHL) can be clearly heard by a normally hearing adult over any fan or equipment noise.

[1] This may include tone pip ABR, narrowband chirp ABR and ASSR. ABR tests should be used to exclude ANSD prior to performing ASSR testing.

The environment must also have suitably low levels of electrical interference (e.g. 50Hz) that the signal base line is not adversely affected. Test rooms should therefore not be sited close to transformer, plant or high power equipment and careful selection of the local test area or room may be necessary in order to achieve satisfactory environmental conditions.

Where ABR testing is performed outside the designated clinic area - for example on the ward or in the operating theatre - levels of acoustical and electrical interference must be sufficiently low so as not to influence the results of the test.

3.2 Calibration

The transducers must be objectively calibrated at least annually by a calibration authority which can show clear traceability and use methods given in IEC 60645-3 (IEC, 2020). It must not be used beyond its "calibration next due" date.
ISO 389-6 (2007) provides reference equivalent threshold sound pressure levels (RETSPL) for click and tone pip stimuli used for AC ABR for certain types of transducer. It also provides a standard for reference equivalent threshold force levels (RETFL) for use for click BC ABR for the B71 type bone vibrator. However, there are no RETFLs for tone pip BC ABR. A procedure for calibration is given in IEC 60645-3. In July 2005 a provisional set of reference levels for AC and BC ABR was agreed for NHSP; including reference threshold values for narrow band CE-chirps, ("BSA Recommended Stimulus Reference Levels for ABR Systems, 2012" 2012), with the ISO389-6 values used where appropriate. These values should be used to calibrate equipment used for hearing assessment.

It is important to note that the stimulus rate affects the psycho-acoustic threshold but not the ABR threshold (Lightfoot et al. 2007). When performing a 'stage A' listening check therefore, using the recommended stimulus rate of close to 50 per second will make the stimuli sound about 3dB too loud. This should be taken into account or the listening check performed at 20 per second.

The equipment should not change the physical intensity of the stimulus when the stimulus repetition rate is changed since this would introduce errors. Detecting if it does is difficult subjectively and further information on this is available in the equipment-specific ABR parameter document (ERA Training and Consultancy, 2023).

3.3 Stage A checks

The level of stimulus output should be checked at the start of a session and monitored by listening to the earphone at critical points during the test, particularly if unexpected results occur.

Check list for daily and monthly function check of auditory brainstem response systems.

(Based on B700/701, BS EN ISO 8253-1:1998).

Some manufacturers offer their own equipment-specific version of Stage A check

1. Clean equipment and examine for damage or wear. Check headphones, bone conductor, insert earphones and leads for signs of damage.
2. * Switch on & adjust according to handbook.
3. * Earphone serial numbers or marking tally with equipment.
4. * Check battery state, if appropriate.
5. * Electrode impedance test correct with dummy load.
6. *# Threshold levels of stimuli to be used are subjectively correct for:
    i. Air conduction (For all transducers used)
    ii. Bone conduction
7. * High level (max 80dBnHL) listening test with stimuli to be used satisfactory by:-
    i. Air conduction (For all transducers used)
    ii. Bone conduction
    iii. Masking (including insert
8. Attenuator sweep subjectively satisfactory.
9. Noise, hum and break-through levels are adequately low.
10. Radiated noise from instrument is acceptable at the patient's position.
11. Headbands are in good condition and tensions subjectively correct.
12. Amplifier: select calibration mode (or loop test mode), run test and check averaged waveform is of expected amplitude and morphology.
13. Connect amplifier inputs together, run test and check that flat waveforms are obtained, the noise floor meets equipment specifications, and that there is no significant correlation between repetitions indicating system artefacts.
14. Check test parameters against the relevant departmental / NHSP protocol
15. *Reset all controls to normal operating positions for commencement of patient testing.

* Tests marked with an asterisk are recommended for checking at the start of a session when the equipment is used; other checks may be performed at monthly intervals. Additionally, it is vital that all checks are conducted prior to and following objective calibration and whenever the user has reason to question the correct function or adjustment of the system.

Threshold levels may be tested at the rate employed in the ABR test but note that in theory, these levels are correct only when a rate of 20/s is used in a subjective listening check. If the stimuli appear too loud, repeat the check at a rate of 20/s.

Ongoing vigilance: Whenever an elevated ABR threshold is recorded, check that the stimulus is being delivered at the expected level; monitor waveforms recorded in babies with normal ABR thresholds and report / investigate any unexpected artefacts.

3.4 Recording system checks

The equipment should be checked at regular intervals (weekly is recommended) for system artefacts.

Using the normal protocol for testing, use a “dummy patient” resistance network, “loop back” box or, if not available, connect the electrodes together. Run the normal protocol twice and check that flat waveforms are obtained with a minimal level of residual system noise (peak-to-peak below 50nV and no significant correlation between repetitions indicating system artefacts).

At regular intervals (monthly is recommended), when testing a baby, take the opportunity to carry out an additional control recording (a blocked-stimulus run) to check that there are no artefacts in the recording system. This should be performed on a baby where clear ABRs have been obtained at discharge levels. Set the stimulus to 30dBnHL and block the sound from reaching the ear (see section 7.2 on how to do this); 30dBnHL is chosen as it may not be possible to completely acoustically block a high stimulus level. Obtain replicated waveforms with the stimulus still being presented (do not reduce the level of the stimulus as this invalidates the check). A pair of waveforms should be obtained which meet the RA criteria defined above.

3.5 Choice of electrodes and application

All local infection, prevention and control procedures must be adhered to. These should cover hygiene, use of equipment and electrodes.
To achieve a good impedance with the skin, after checking for any skin condition that would contraindicate skin abrasion, the skin should be gently and carefully abraded. Appropriate options include abrasive electrode preparation paste and a clean gauze pad, a disposable abrasive pad or a cleaning stick with soft cotton material on the end. Single use disposable electrodes should be used.

3.6 Differences in Electrode Impedance

Differences in electrode impedance can cause significant artefacts and electrical interference. This difference is most easily minimised by ensuring all electrodes have low impedances. The impedance, as measured between each electrode pair should be under 5000 Ohms and similar / balanced across electrodes.
Impedance which is higher than 5000 Ohms may also give an unacceptably large stimulus artefact at high stimulus levels; particularly for bone-conduction ABR. However, in good recording conditions and in a screened room, higher electrode impedances can be tolerated.

The ABR system (dedicated hardware or associated computer) must not be switched on or off (re-booted) with the patient attached. If it is necessary to re-boot the ABR system, first disconnect the patient electrodes and transducers then re-connect them once the system is fully operational. This ensures that no potentially hazardous currents pass through the electrodes and no high-level stimuli are presented to the patient.

3.7 Electrode location – AC and BC

Single or two-channel recording may be used for AC and BC.

For single channel, the electrodes should be located as follows:

- Positive electrode: **high forehead** as near to Cz[2] as possible and midline. The fontanelle should be avoided but the electrode should be placed as close as possible to this otherwise the ABR will be reduced in size. A mid-forehead position is not appropriate.
- Negative electrode: **ipsilateral mastoid**. Sufficient space should be allowed for a bone vibrator to be placed on the mastoid without interfering with the electrode. To allow possible recording of CM, the electrode should be no more than 1cm lower than the meatal level of the ear.
- Common electrode: **contralateral mastoid**.

This configuration should result in wave V being plotted upwards on the display. If this is not the case, then check that all the electrodes are in the correct sockets of the pre-amplifier.

For two-channel recording, the electrodes should be located as follows:

- Positive electrode: **high forehead** as near to Cz as possible and midline. The fontanelle should be avoided but the electrode should be placed as close as possible to this otherwise the ABR will be reduced in size. A mid-forehead position is not appropriate.
- Negative electrode: Channel 1: **ipsilateral mastoid**. Sufficient space should be allowed for a bone vibrator to be placed on the mastoid without interfering with the electrode. To allow possible recording of CM, the electrode should be no more than 1cm lower than the meatal level of the ear. Channel 2: **contralateral mastoid**.
- Common electrode: **forehead**, lateral to the positive electrode by at least 4cm.

See section 5.9 for further details on the use and benefits of two-channel recording.

4 Stimulus

The recommended values (or ranges) for stimulus parameters are summarised in Table 1. These ensure optimum recording of the III/V-SN_{10} complex which is crucial to paediatric threshold testing. Appendix D gives the dBnHL to dBeHL corrections for stimuli used in babies ≤12 weeks corrected age, together with the recommended maximum stimulus levels for inserts and supra-aural earphones.

[2] Cz is the standard position used in adults. It is defined in the 10-20 electrode system for electroencephalography. For these purposes it can be taken as the point along the midline of scalp half way between the bridge of the nose (nasion) and the start of the skull at the rear of the head (inion).

4.1 Stimulus and stimulus rate: Air and bone conduction

The stimulus should be of alternating polarity to minimise the stimulus artefact.

An electrical pulse of 100µs should be used for click ABR (ckABR) and a 2 cycle rise/fall and 1 cycle plateau for tone pip ABR (tpABR), the reference stimuli described in IEC 60645-3 (IEC, 2020). The envelope for the rise, plateau and fall phases of the tone pip can be Blackman[3] or linear (Blackman is preferred). Some equipment specifies a Blackman envelope by stating the total number of cycles (this should be 5). Narrow band (pip-like) chirps (NBchirp) can also be used provided these conform to IEC 60645-3 and have a defined nHL to eHL correction. The term tpABR implies frequency-specific ABR and should be assumed to include NBchirp ABR.

The frequencies used for frequency-specific testing should be 0.5, 1, 2 or 4kHz. The primary stimulus must be 4kHz.
In Table 1, the slower stimulus rates for 0.5kHz and 1kHz tpABR allow for longer window lengths to be used so that the full SN_{10} part of the waveform is recorded. A range of recommended values is given in this guidance to fit with those available on commonly used equipment. Lower repetition rates will not give invalid results but will be less time efficient. Equipment-specific parameters for NHSP referrals are available (ERA Training and Consultancy, 2023); it is particularly important to use these if chirps are used since a compromise is required between stimulus rate, stimulus duration and "blocking" which is needed at both the start and end of the recording epoch[4]. Rates such as 35.1/s, 49.1/s etc. (with no common relationship with mains frequency) should be chosen to minimise any mains artefact.

Stimulus levels should be recorded in dBnHL. The "nHL" can be taken to imply the use of either ISO, 2007 or BSA-recommended calibration values (BSA, 2019d). For testing carried out following NHSP referrals, the advice in the BSA early assessment guidelines should be adhered to (BSA, 2021).

[3] A Blackman envelope does not strictly have a plateau. If the equipment has the option of entering the total number of cycles this is preferable for a Blackman envelope. Where it does not, we recommend entering 2 cycles rise/fall and 1 cycle plateau for the Blackman envelope.

[4] A chirp stimulus begins before (and extends to just after) the "zero" point on the recording timebase and evokes a response which does not appreciably change its latency with frequency. For this reason, a 20ms window is used for all frequencies of NB chirps. In order to use a fast rate, we allow the next stimulus to start just before the end of the previous timebase and since this could produce a stimulus artefact, "blocking" periods are needed at both the start and end of the timebase. These periods are affected by the chirp frequency, timebase and rate so it is vital that recommended values are adopted without modification. For more details of chirps see Elberling and Don 2010.

	Click, NBchirp & 2kHz / 4kHz tone pip	0.5kHz / 1kHz tone pip
Electrode location (single channel)	Positive: High forehead (as close to vertex as possible but avoiding fontanelle) Negative: Ipsilateral mastoid Common: Contralateral mastoid	
Electrode location (two channel)	Positive: High forehead (as close to vertex as possible but avoiding fontanelle) Negative: Ipsilateral and contralateral mastoids Common: Contralateral mastoid forehead, lateral to the positive electrode by at least 4cm.	
Stimulus polarity	Alternating	
Stimulus timing	Click: 100µs Tone pip: 2-1-2 cycles (linear rise–plateau–fall) or 5-cycle Blackman	
Stimulus rate[5]	45.1 - 49.1/s	35.1 - 39.1/s
Calibration values for 0dBnHL	Refer to BSA ABR calibration data	
Amplifier reject levels	±10µV with noise-weighted averaging	
Amplifier filters	Low frequency (high-pass): 30Hz High frequency (low-pass): 1500Hz	
Window length[6]	20ms	25ms
Number of sweeps averaged per waveform	Absolute minimum: 1000 but refer to section 5.8; stopping decisions are to be based on residual noise and meeting CR & RA criteria, not on sweeps	
Display scales	Within range 25-100nV ≡ 1ms See equipment specific settings	
Display	Wave V up	

Table 1: Summary of recommended ABR parameters[7]

Testers must be familiar with the risks associated with the use of transducers in patients with programmable PVP shunts and follow BSA guidance (BSA, 2023).

[5] Most equipment can provide a rate within these ranges for the suggested window length (see equipment-specific parameters). The rate must not be related to 50Hz. If chirp stimuli are used the optimum rate depends on the chirp duration.

[6] These window lengths are nominal values and should be set to the closest value available on the equipment. Chirps should be used with a window length of 20ms regardless of stimulus frequency.

[7] Note that some equipment offer more advanced features or stimuli, not covered in this table. See the NHSP equipment-specific parameter document for details.

4.2 Earphone

This should be able to deliver a stimulus up to 140dBSPL peak (about 107dBnHL for a click stimulus) without distortion. Supra-aural or insert earphones (e.g. type ER-3A) are suitable. The actual stimulus level is more uncertain with insert earphones due to the greater variation in the enclosed volume of a baby's ear canal and this has implications for the precision of ABR results when inserts are used in babies. However, insert earphones reduce the need for masking and attenuate ambient noise more than supra-aural earphones. If insert earphones are used, Care should be taken that wax is not compacted by the probe, so blocking the sound pathway. Supra-aural earphones should be centred over the ear canal to avoid collapsing the ear canal due to excess pressure.

4.3 Warning – insert earphones

Insert phones should not be used above the maximum levels given in the BSA guidelines for early audiological assessment (BSA, 2021). The values are included in appendix D for convenience. This is because a baby has a much smaller ear canal which will lead to a 10-20dB higher stimulus level at certain frequencies compared to the same insert earphone used in an adult. This uplift is thought to diminish over the early months of life as the ear canal grows (see BSA early assessment guidance for more detail). In cases where no ABR is recorded at the maximum recommended stimulus level using insert phones, testers should consider the use of supra-aural earphones up to the maximum recommended levels for those transducers.

4.4 Bone vibrator

This should be able to deliver a stimulus up to 60dBnHL (50dBnHL at 0.5kHz) without obvious waveform distortion. Stimulus levels should not exceed these values unless the bone vibrator has been passed in calibration as being able to deliver higher levels without distortion. The Radioear type B71 or type B81 bone vibrator shall be used as calibration data are not available for other types of bone vibrator. A check should be made that the impedance of the bone vibrator is correct for the equipment being used. A 'Stage A' listening test near threshold, and at 50dBnHL or above should be carried out at the start of each session in which a bone vibrator is used.

4.5 Placement of bone vibrator

The bone vibrator should be placed on the mastoid approximately 1 finger's width above the electrode. If possible, move hair away from where the bone vibrator is to be placed. The bone vibrator lead should be kept away from the electrode and electrode lead.

Placement on the temporal bone slightly posterior to the upper part of the pinna may be a good alternative when the mastoid is difficult to access due to proximity of the electrode (Small, Hatton and Stapells, 2007). See Figure 1.

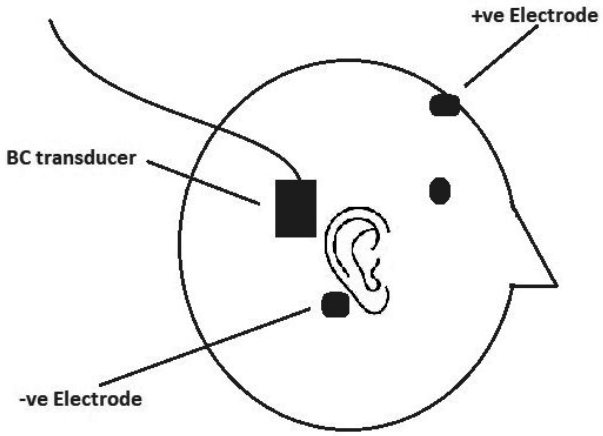

Figure 1 Suggested placement of the bone vibrator for a newborn

4.6 Pressure to apply to bone vibrator

A moderate force ('finger pressure') should be applied to the bone vibrator, but the exact force is not critical - tests on an artificial mastoid have demonstrated an error of no more than 2dB over a wide range of applied forces (Webb, 1993). It is not good practice to ask the parent or carer to hold the bone vibrator as this may lead to changes in position.

4.7 Effect of age on the bone-conduction and insert stimulus

The effective level of the stimulus changes with age. Please refer to the BSA Early assessment guidance (BSA, 2021) and appendix D of this guidance.

5 Data collection and analysis

5.1 Characteristics of the ABR

Interpretation of the ABR, covered later in this section, relies on knowledge of the features and characteristics of the ABR.

The ABR waveform is a series of seven peaks, by convention labelled using roman numerals, and troughs. However, attention focuses on ABR waves I, III & V, whose generators are anatomically located in the inner ear, lower brainstem and upper brainstem respectively. The latency (time delay from the stimulus to the peak) of the peaks depends on several factors but is roughly in the range 1 to 12ms. The amplitude of the ABR is tiny in comparison to other electrophysiological signals: typically less than 0.5µV (500nV), by convention measured from the highest peak (usually wave V) to the lowest trough; it becomes smaller as the stimulus level is reduced towards the ABR threshold.

Figure 2 shows a series of ABR waveforms at different stimulus levels – sometimes called an *intensity series*. Replication has not been performed, as that is a separate issue, discussed later. Wave I is evident only at the higher levels whereas wave V can be traced down to threshold. In this example, wave III is not clear. Note that as the stimulus level is reduced, the size of the response reduces and the latency of the response increases. This is a characteristic of the ABR (and all auditory evoked responses) and testers should expect this. This knowledge can aid response identification but at threshold, where the response can be tiny, there is a danger of wrongly interpreting a bump of noise as a genuine response.

The response morphology also changes with stimulus level; the clarity and crispness of the peaks becomes less distinct near threshold.
Less distinct responses and longer latencies[8] are also associated with lower stimulus frequencies.

[8] Longer latencies are not seen with lower frequency chirp stimuli, which compensate for the cochlear travelling wave delay.

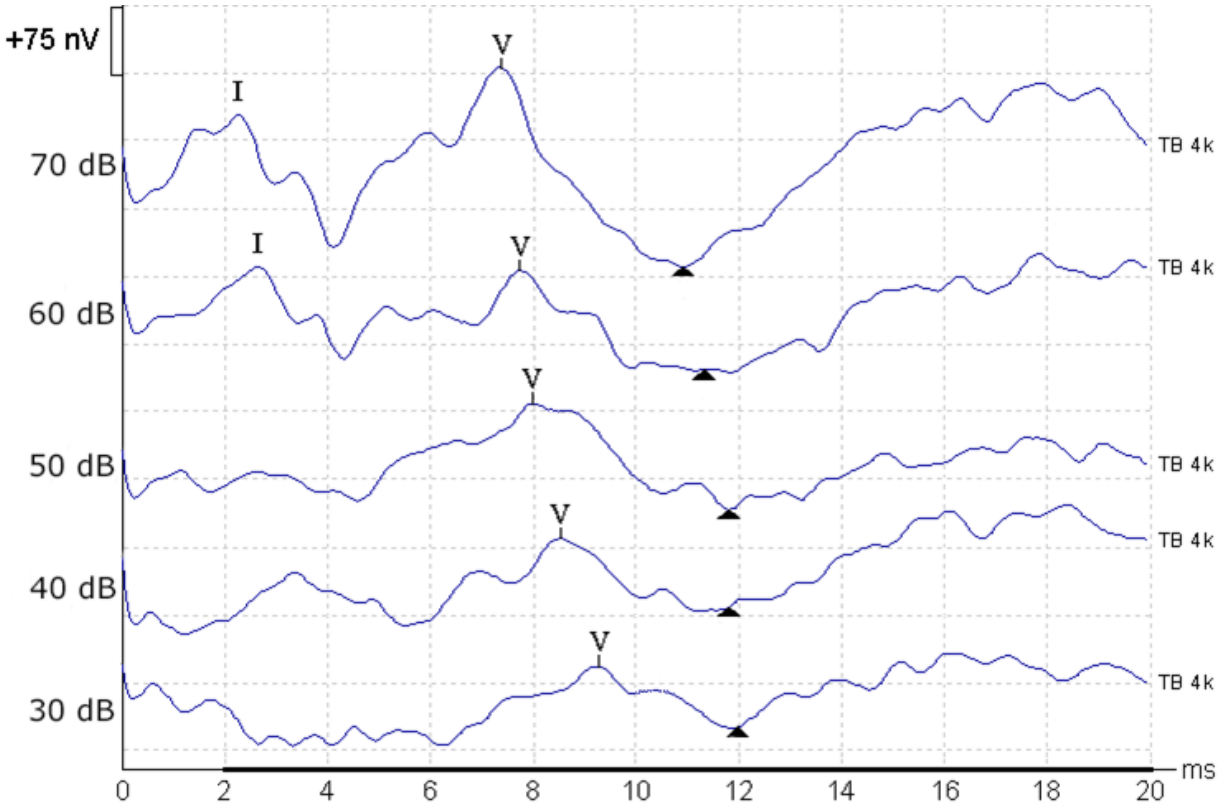

Figure 2. An ABR intensity series in response to 4kHz tone pips in a normal newborn. Stimulus values in dBnHL. Triangles mark the trough used as a reference point for response amplitude measurements.

5.2 Artefact rejection (AR) level

In the following, reference to "signal" refers to the ABR whereas "noise" refers to everything else, which can be a combination of EEG, ECG, muscle activity and other electrical interference. A large ABR can be around 0.5µV (500nV) at suprathreshold levels, whereas near threshold it can be 0.1µV (100nV) or even less (the minimum acceptable amplitude is taken as 0.04µV or 40nV). The noise can be as little as 2µV in a sleeping baby (EEG) to over 20µV (muscle activity). This means that the ABR signal is far smaller than the competing noise and in order to correctly identify the presence of the ABR an averaging technique must be used in order to reduce the noise so that it is a fraction of the ABR.

Averaging reduces the residual noise (and so improves the SNR) by the square root of the number of sweeps. For a 10-fold reduction in noise 100 sweeps is needed; for a 100-fold improvement this increases to 10,000 sweeps. The number of sweeps needed depends on the relative sizes of the response and the noise. Testers should always try to maximise the signal and minimise the noise. This involves optimal electrode positioning and avoiding sources of noise, both electrical and from the baby.

Artefact rejection simply rejects sweeps (stops them from being added to the average) if the peak-to-peak amplitude of the activity anywhere within the sweep is more than a defined level: this is the artefact rejection (AR) level. This ensures that data are collected only when the instantaneous test conditions are favourable.

This guidance assumes that noise-weighted averaging (e.g. Bayesian averaging) is available and used. In noise-weighted averaging, the system allows the final average to be dominated by low-noise periods, thus avoiding the need to manually pause the recording when the baby is restless. The previous guidance advocated a default AR level of ±5µV and changing the level interactively. This is no longer recommended for systems with noise-weighted averaging. By default, testers should use noise-weighted averaging with an AR level of ±10µV. AR levels less than ±10µV can be used at the discretion of the tester (and should not attract criticism by peer reviewers). When ECG activity is evident in the incoming signal (where the EEG waveform "jumps" about twice a second; a baby's typical heart rate), reducing the AR from ±10µV to ±7µV or even ±5µV will exclude this source of noise and be advantageous. Also see appendix F on baseline drift.

Where the system does not have noise-weighted averaging a default AR level of ±5µV (max ±10µV) and changing the level interactively is recommended.

5.3 Recording filters

Low frequency (high pass): a value around 30Hz is recommended. This has been found to give the best signal to noise ratio of wave V near threshold. In normal testing environments, higher values should be avoided; although less electrical and myogenic noise is recorded the response is also attenuated, making interpretation difficult. The FAQ section (appendix G) includes advice on this.

High frequency (low pass): a value around 1500Hz is recommended. There is little response energy above this frequency. A higher value generally adds more electronic noise from the amplifier.

5.4 Use of digital / display filters on the averaged waveform

Digital display filtering in the range 50 Hz to 1000 Hz is not recommended for routine use as it is likely to result in a change in the waveform shape and response amplitude which makes interpretation and audit of the waveforms more difficult. For the same reason smoothing filters should not be used and should not be necessary as the 1500Hz filter used in the recording is sufficient to remove any unwanted high frequency noise.
The exception to this is when a sloping baseline is recorded, making the placement of markers problematic for wave V and SN_{10} for the measurement of response amplitude. If digital (display) filtering is available, it may be helpful to apply a high-pass filter at 50-75Hz to reduce the extent of the baseline slope and so obtain a more accurate estimate of response amplitude. Appendix F provides more advice on dealing with baseline drift. If digital filtering is used, details should be included in the clinical report, peer review document and S4H entry.

5.5 Notch filter

This will not be required under normal recording conditions and with good electrode practice as 50Hz mains artefact should be absent or minimal. If mains artefact levels are high, it is better to identify and remove the source of the problem rather than rely on the use of the notch filter. However, if there is an unusual and exceptional degree of mains interference which cannot be eliminated the temporary use of a notch filter is preferable to raising the high pass filter or abandoning the test. When a notch filter is used this must be noted in the clinical report. The available evidence (Lightfoot *et al.*, 2014) is that notch filtering does not distort the newborn ABR, with the exception of testing at 0.5kHz where waveform distortion has been observed and could compromise waveform interpretation. At 0.5kHz therefore the notch filter must not be used.

5.6 Display

Always adhere to the convention of plotting wave V upwards. The display should be always set at a fixed number of µV or nV per division (1µV=1000nV). The amplitude (y) and time (x) scales should be such to ensure that small waveforms near threshold are visible. The broad range of acceptable values is 25- 100nV (0.025-0.1µV) on the response amplitude (y) axis to match 1ms on the time (x) axis. Please refer to the equipment specific settings for how to achieve these in practice.

In some equipment the display aspect ratio of the on-screen and printed waveforms are not the same. It is important to ensure the printed waveforms' aspect ratio is within the ranges recommended above. This is of particular importance to facilitate peer review or re-interpretation at a later date. Information on specific scales to use for different types of equipment is available (ERA Training and Consultancy, 2023).

An automatic display gain function must not be used as this may set an inappropriate display gain for assessment of the response.

5.7 Criteria for accepting the presence or absence of a response

This relates to the **precision** phase referred to in section 5.8.
The primary method of establishing the presence and absence of a response is visual interpretation. When objective measurements are available these can be valuable in helping us to be confident in our interpretation and in deciding when to stop averaging (see Appendix E).

Replication of waveforms contributing to the reported result (as defined below) is essential if a correct visual interpretation is to be made. Replication is not needed at other stimulus levels. For example, in the initial phase referred to in section 5.8, if the first stimulus level is 40dBnHL and a waveform is obtained with no likely response then the best use of time is not to replicate until a response is observed at higher stimulus levels and it is clear which levels need to be replicated to determine threshold.

Decision criteria for the result at each stimulus level

For each stimulus level the result should be marked in one of three ways.

- ***CR***: Clear Response present
- ***RA***: Response Absent
- ***Inc***: Inconclusive.

The reader should make reference to Appendix C which contains more detailed advice and examples on this process. The rules for marking the results as ***CR***, or ***RA*** require two waveforms which are optimally superimposed (displayed on a common baseline representing the average value of all points in the averaged waveform). For systems that do not show a baseline for each waveform, waveforms should be positioned such that the gap between them, averaged across the entire latency range, is minimised. Decision criteria should be applied to the two superimposed waveforms only. Where there are more than two recordings, these should be combined to form two waveforms by using a weighted add (or merge) function to obtain two waveforms in a 'matched pairs fashion' – i.e. combine waveforms 1 and 3 (to form waveform 5) and waveforms 2 and 4 (to form waveform 6). The 2 new waveforms (5 and 6) shall then be superimposed and compared visually using the same mechanism as below. Where more than 4 waveforms have been recorded, the odd numbered waveforms shall be added together and compared to the addition of the even numbered waveforms. (see also later note about when a further pair of recordings have been made). It should not normally be necessary to display the original waveforms.

For ***CR*** there must be a high degree of correlation between the replications and the waveforms should show the expected characteristics in terms of amplitude, latency and morphology.

The size/amplitude of the response (as judged from the wave III/V to the following SN_{10} trough, averaged across replicates – see Figure 3 below) should be a minimum of 40nV, on average, and at least 3 times the background noise level (the noise level should be estimated from the average gap between the waveforms; do <u>not</u> use the residual noise figure reported by the equipment when calculating the response signal to noise ratio).
If wave III is higher than wave V, use wave III/ SN_{10} as the response amplitude.
The waveform should be judged over the whole time window excluding any stimulus artefact.
For an example, see Figure 3.

Waveforms should be compared with those at other stimulus levels (where available) to confirm that they follow the expected changes with stimulus level. See Appendix C.
See also Appendix E Section 3.1 Modified Criteria for Single (unreplicated) waveform.

The interpretation process should be carried out according to the flow diagram in Figure 5.
Firstly, does the result meet the criterion for a clear response (***CR***)

This criterion ensures a high degree of confidence (around 97%) of the presence of an ABR. Examples showing where this criterion has been met and where it has not been met are given in Appendix C.

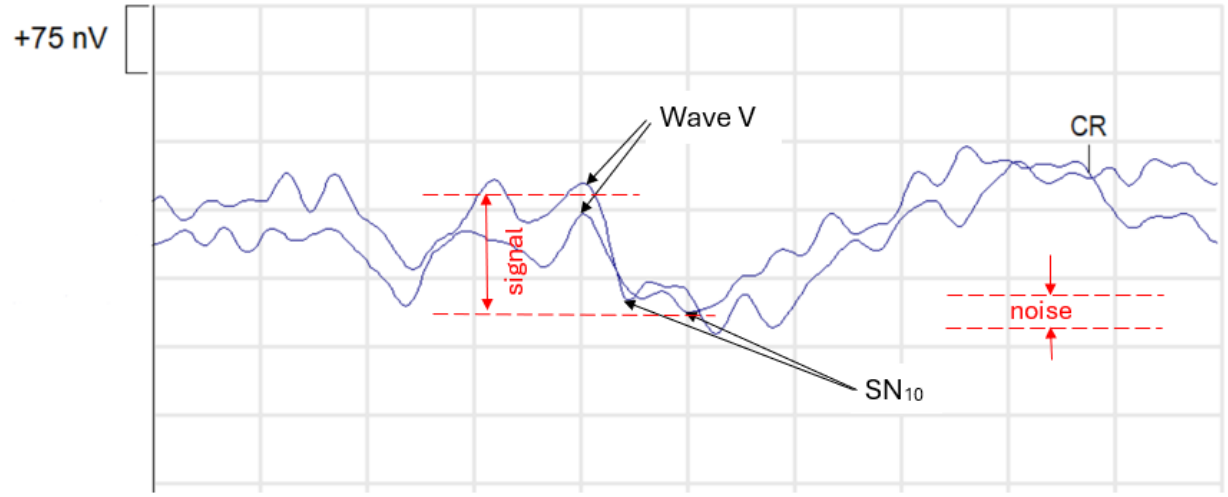

Figure 3: Example of a clear response (CR) using the wave V peak and SN_{10} trough to measure signal amplitude. Time scale: 2ms/division. The signal (response) amplitude is measured from the vertical mid-points of the two wave V peaks and the two SN_{10} troughs. The noise is an estimation of the average gap between the optimally superimposed waveforms across the entire latency range.

In Figure 3 the two waveforms have been superimposed so as to minimise the average gap between them. Some systems have a "superimpose" function which does this optimally. If superimposing the waveforms manually do not simply align the wave V peaks. Rather, position the waveforms so as to minimise the average gap between them (ignore any region containing a stimulus artefact). The "signal" is taken as the average of the two waveforms' wave III/V to SN_{10} trough amplitude: in this example it is just over one and a half 75nV divisions (120nV or 0.12µV). The noise is the gap between the waveforms, averaged across the entire window, excluding any region of stimulus artefact, and in this example it is just over half a 75nV division (40nV). The signal to noise ratio in this example is therefore 120/40 (working in nV) which is about 3:1 – on the limit of acceptability for ***CR***. Note that for ***CR*** there is no need for the noise to be less than a specific value – it just needs to be less than a third of the response amplitude. This, together with a characteristic waveform of more than 40nV amplitude allows us to regard this as a clear response (***CR***).

If the result does not meet the criterion for a clear response, the question should be asked 'Is this a response absent (***RA***)?'

For ***RA*** the waveforms must contain no evidence of a response (even an abnormal or uncharacteristic response), and the average gap (noise) between a pair of optimally superimposed waveforms should be less than or equal to 25nV (using the same method for measuring the background noise for ***CR*** described above).
For an example, see Figure 4.

This criterion ensures a high degree of confidence (around 97%) of the absence of an ABR.
This average difference (noise) criterion is designed to ensure a small response is not being obscured by noise. In practice, although a waveform may contain no evidence of a response, it may not be completely flat - see Appendix F on baseline drift. As with ***CR,*** the principle underpinning ***RA*** is that there must be a high degree of confidence that a response is genuinely absent.

Figure 4 shows a pair of waveforms meeting the criteria for ***RA***. The residual noise is estimated as the average gap between optimally superimposed waveforms (superimposed so as to minimise the average gap). Any difference between the waveforms is simply noise, including the peak and trough in one waveform in the 14–16ms range. If that feature had been common to both waveforms then it could be evidence of an abnormal (late latency) vestigial response and ***RA*** could not be claimed.

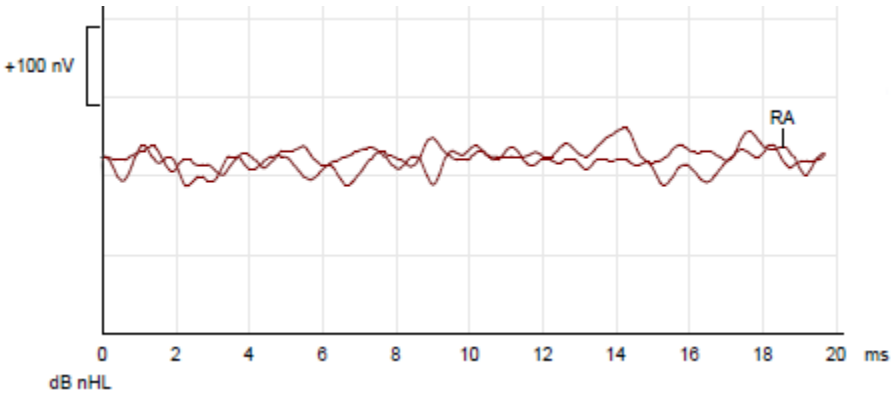

Figure 4. An example of RA

Finally, if the result does not meet the criteria for either a clear response (***CR***) or response absent (***RA***), the result should be marked as inconclusive (***Inc***).

Inconclusive waveforms should not contribute to the definition of threshold. Waveforms that appear to show a likely response or likely absence of response need not be replicated if they do not contribute to the definition of threshold as defined in section 5.12, below, but they should be labelled as ***CR*** or ***RA*** only if replicated and satisfy the criteria for ***CR*** or ***RA***. Waveform peaks and the SN_{10} trough should be marked on each waveform rated ***CR*** (and where appropriate ***Inc***) to facilitate latency & amplitude measurements for peer review purposes. The placement of peak markers should not be construed as meaning ***CR***.
Appendix C gives examples of inconclusive waveforms.

Each threshold measurement should continue until there is a very high degree of confidence, with any inconclusive results being resolved.

The above criteria and decision-making process are summarised in Figure 5. Note that Appendix E Section 3 details an optional variation in the flow chart below when objective measurements are used to help determine ***CR*** and ***RA***.

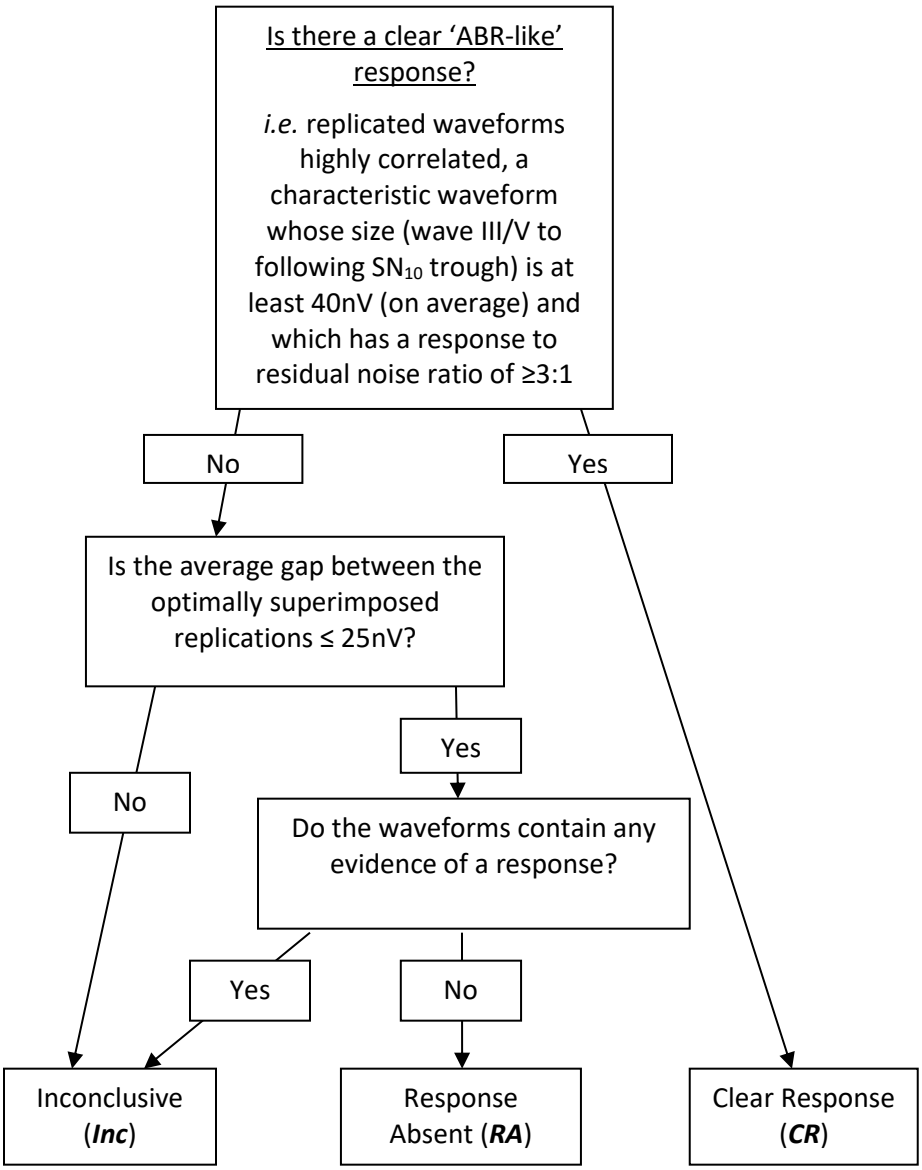

Figure 5: Flow chart showing summary of decision making process (see text for details)

NOTE: If the decision on threshold relies on a very low amplitude ***CR*** (between 40 and 50nV), confirmatory tests should be carried out 5 or 10dB higher, where the ***CR*** response amplitude should be at least 10nV more than that at the lower level. If this result is obtained at the maximum stimulus level then a blocked stimulus run should be carried out to confirm this is a genuine physiological response.

5.8 Advice on when to stop averaging based on residual noise

The guidance on the number of sweeps necessary to obtain waveforms on which clinical decisions can be based has changed. The number of sweeps needed to acquire waveforms of acceptable quality depends on a waveform's residual noise (RN). All ABR systems available in the UK now show the RN in the waveform, during averaging, and the emphasis has switched from sweeps to RN, which is a more direct and reliable index of quality. The decision to terminate the acquisition of an average should therefore be based on the RN of the waveform, and the features the waveform reveals, rather than the number of sweeps (but note that any waveform must contain at least 1000 accepted sweeps).

Testers need to know what number of sweeps to specify in the equipment test protocols. Select a large number such as 4000 or 8000 so the tester, rather than the equipment, makes an informed decision to terminate an average. In the rare event that this number is insufficient, perform additional runs and combine in a pair-wise fashion (see Section 5.10 on resolving inconclusives).

The most efficient test strategy is to perform detailed ABR testing around threshold. It is therefore important to get a clear idea of where that threshold may be quickly. A two-stage approach is suggested. In the initial "quick look" phase, the object is to perform tests at stimulus levels that inform the tester where the approximate ABR threshold is likely to be, but at this stage, formal ***CR*** or ***RA*** criteria need not be met and waveforms should not be labelled ***CR*** or ***RA*** unless the formal criteria are satisfied. For this initial phase, average until the RN falls to the number below. (This is twice the ***RA*** noise criterion for the ABR system being used (it varies across systems but is given below):

Biologic Nav Pro	30nV
Vivosonic Integrity V500	40nV
NeuroAudio	34nV
GSI Audera Pro	120nV
Interacoustics Eclipse (software version 4.4.2 or later):	30nV
Interacoustics Eclipse (software version 4.2):	50nV

This will take typically about a quarter of the time needed to reduce the noise down to the ***RA*** criterion. This should normally be sufficient to be able to identify whether the waveform contains a *likely* response or not. No waveform should contain fewer than 1000 sweeps. Objective measurements of response presence, such as Fmp, may be used at this stage (see Appendix E) and averaging can be stopped early (even if the residual noise does not meet the above values) if the objective measurement suggests a response and a characteristic ABR waveform is obtained. If it is unclear whether a response is present, continue averaging to reduce the RN further. At this stage, it is not usually necessary to replicate. If no likely response is seen at the initial stimulus level, increase in 20dB steps.

In the **precision** phase, time should be invested in obtaining waveforms using sufficient sweeps that allow the ***CR*** and ***RA*** criteria to be satisfied visually and the ABR threshold defined. See section 5.7. If the first run in the initial phase shows a likely response, the tester can immediately move to the precision phase. Example:

- Initial (quick look) phase: 40dB: likely absent; 60dB: likely absent; 80dB: likely response; 70dB: likely response. The threshold is likely to be around 70dB but at this stage we are not certain.
- Precision phase: 70dB x2: =***CR***; 60dB x2: =***RA***; 80dB x2: =***CR.***
- Reported threshold: =70dB (gold standard – see section 5.13).

A more detailed example of the use of initial and precision test phases is also included in Q16 of the Frequently Asked Questions appendix.
In both the **initial** and **precision** phases, if a likely response or a ***CR*** is seen, it is helpful if wave V and SN_{10} markers are used to quantify response latency and amplitude for peer review purposes.

Note that monitoring the reported RN is useful in the initial phase but candidacy for ***CR*** in the precision phase should be based on estimating the noise visually, to estimate the signal to noise ratio. This is because these two estimates of noise differ and the >3:1 signal to noise ratio criterion for ***CR*** is based on the average gap method.

5.9 Ear-specific recording

As with pure-tone audiometry, in certain circumstances the stimulus level is high enough to cross to the other cochlea and produce a response; a crossed (or "shadow") response may mislead as to the true threshold. A masking noise calculator spreadsheet designed by Guy Lightfoot is available to download from https://www.eratraining.co.uk/resources.html, and can be used to alert when there is a risk of cross-hearing. If this is the case, the tester has three options:

(a) Apply masking to the non-test ear;
(b) Use 2-channel recording to inform us if cross-hearing is actually taking place;
(c) Look for the presence of ABR wave I to inform us if cross-hearing is actually taking place.

Details of the background and evidence used for masking is given in an appendix of the BSA guidelines for Early Assessment (BSA, 2021). As a practical guide, the following tables may be helpful. It is based on the Noise Calculator spreadsheet 2024 and assumes the non-test ear is normal. Cross-hearing may occur when using stimuli at or above the following levels (in dBnHL) for babies of 0 to 12 weeks corrected age (consult the spreadsheet for other ages).

Transducer	0.5kHz	1kHz	2kHz	4kHz
Supra-aural, dBnHL	55	55	55	75
Insert, dBnHL	55	55	60	75
BC, dBnHL	-5*	-5*	15	20

Table 2: Stimulus levels for tone pips at or above which masking needs to be considered.

Transducer	0.5kHz	1kHz	2kHz	4kHz
Supra-aural, dBnHL	50	50	50	70
Insert, dBnHL	50	50	55	70
BC, dBnHL	-10*	-10*	10	15

Table 3: Stimulus levels for CE-Chirps at or above which masking needs to be considered.

The classical means of ensuring ear-specific results is to apply a masking technique. In many instances there is a risk that a response could be crossed but is in fact it is not crossed. As an alternative to masking therefore it may be sufficient to demonstrate that a response, that could in theory be crossed, is actually not crossed, and is being generated by the ipsilateral cochlea. This is particularly attractive in cases of "masking dilemma", where cross-masking is a possibility, such as in a bilateral conductive loss.

* Because these levels are so low, BC testing at 500Hz & 1kHz must be accepted as not ear-specific unless one of the following techniques are used to ensure thresholds are ear-specific.

Masking

If the masking noise calculator warns of the risk of cross-hearing, it is better to initially establish the not-masked ABR threshold on both sides first, as masking may not be needed. If masking is used, present the masking noise to the non-test ear at the level suggested by the noise calculator. When masking, always obtain gold standard thresholds and account for the estimated air-bone gap and BC threshold in the non-test ear. Always assume the worst-case value for the non-test ear BC threshold, for example if the non-test BC threshold is ≤10dB, select 0dB.

2-channel recording of wave V

The purpose of two channel recording is to record both the ipsilateral and the contralateral response in an attempt to determine which cochlea is generating the ABR. Here, we look at the wave V asymmetries between the ipsilateral and contralateral recordings; the channel with the larger amplitude and earlier latency wave V is likely to be the side generating the response. If this is ipsilateral to the stimulated side then the response is not crossed and no masking is needed. If the larger & earlier response is in the channel contralateral to the stimulated side (or if the result is unclear) then crossover may be occurring and masking should be used to determine the test-ear threshold. This method is applicable to both AC and BC testing; although the evidence in table B1 is derived from bone conduction testing, the neural pathway is the same for AC and BC and the 2-channel characteristics will be the same.
This method does not identify the correct side in 100% of cases, may falsely label some unilateral sensorineural losses as conductive and although it may be useful, masking should be regarded as the definitive method. It is important to note that the 2-channel method is helpful only in children under two years corrected age (note the ages of patients in the studies shown in Appendix B) and may give erroneous results in older children, where masking should be used.

Figure 6 gives an example where the BC stimulus generates the ABR via the ipsilateral ear. The contralateral waveforms have a lower amplitude and longer latency wave V so in this case it is reasonable to ascribe the responses to the left ear without recourse to masking.

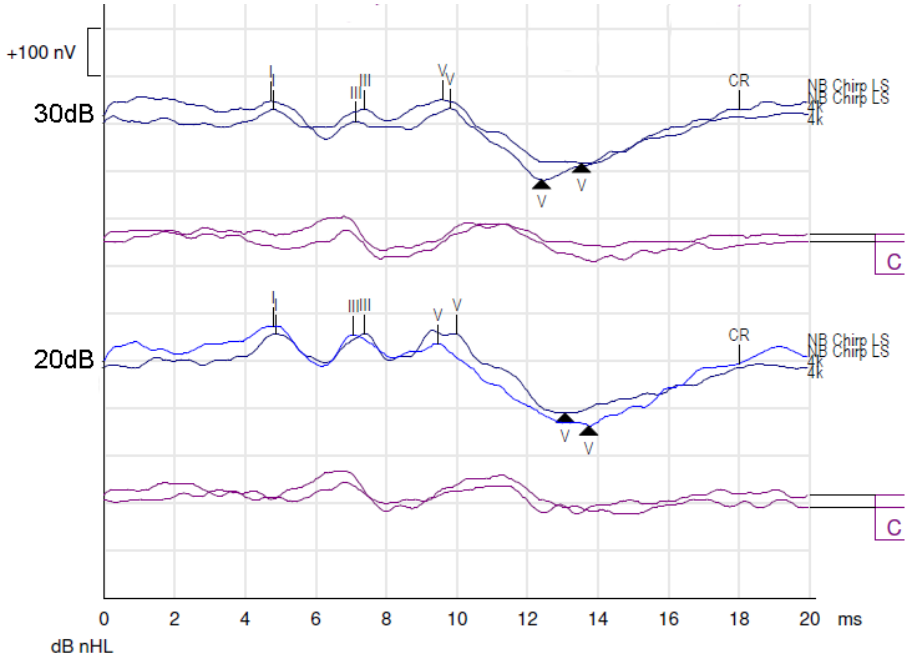

Figure 6: 4kHz BC waveforms (top to bottom):
30dB ipsi;
30dB contra (marked "C");
20dB ipsi;
20dB contra (marked "C")

Figure 7 illustrates a crossed pattern – the BC stimulus generates the ABR via the contralateral ear, whose response is larger than that in the ipsilateral channel. Masking would be needed to establish the true ipsilateral BC threshold. Note that the contralateral ear threshold cannot be inferred from such responses since inter-aural attenuation attenuates the stimulus, the extent of which is not precisely known. However, it will be no greater than 30dB, the level at which the contralateral response is seen.

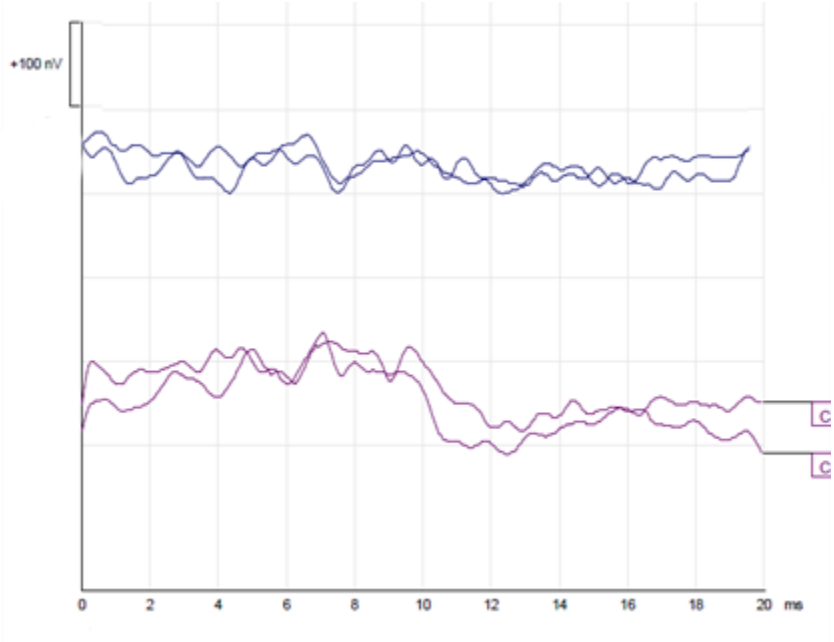

Figure 7: 4kHz BC waveforms (top to bottom):
30dB ipsi;
30dB contra (marked "C")

Identifying wave I

Wave V is generated in the lateral lemniscus in the upper brainstem, anatomically close to the midline and therefore it can be considered as a "far field" response, usually recordable in both ipsilateral and contralateral channels. Conversely, the generator of wave I is anatomically in the cochlea, far closer to the ipsilateral mastoid electrode than the contralateral mastoid electrode so is considered as a "near-field" ipsilateral response and usually cannot be seen in the contralateral channel. Identifying wave I in the ipsilateral channel is therefore confirmation that the response is **not** crossed. The absence of wave I in the ipsilateral channel cannot be taken as evidence of a crossed response since wave I is not always recordable (in either channel), especially close to threshold.

This method was originally described as a 2-channel technique in much the same way as the 2-channel wave V method, but a study (Ferm and Lightfoot, 2016) demonstrated the utility of a single-channel variant (using normal stimulus rates) and identified that the use of NB-Chirps increases the likelihood of identifying wave I close to threshold compared to tone pips at the same stimulus frequency. Thus, when there is a risk of cross-hearing, one should first examine the waveform and if wave I is unequivocally present then it is likely that the response is not crossed. Wave I should be present in both waveforms and large enough to be confident it is not residual noise.

Figure 6 shows wave I in the ipsilateral channel at both 30dB and 20dB but not in the contralateral channel.

Figure 8 is an example of single-channel recording (4kHz BC, left) in which wave I is seen. The right ear was normal so this could have been a crossed response. For the purposes of this example, masking has also been applied, confirming that the response is not crossed.

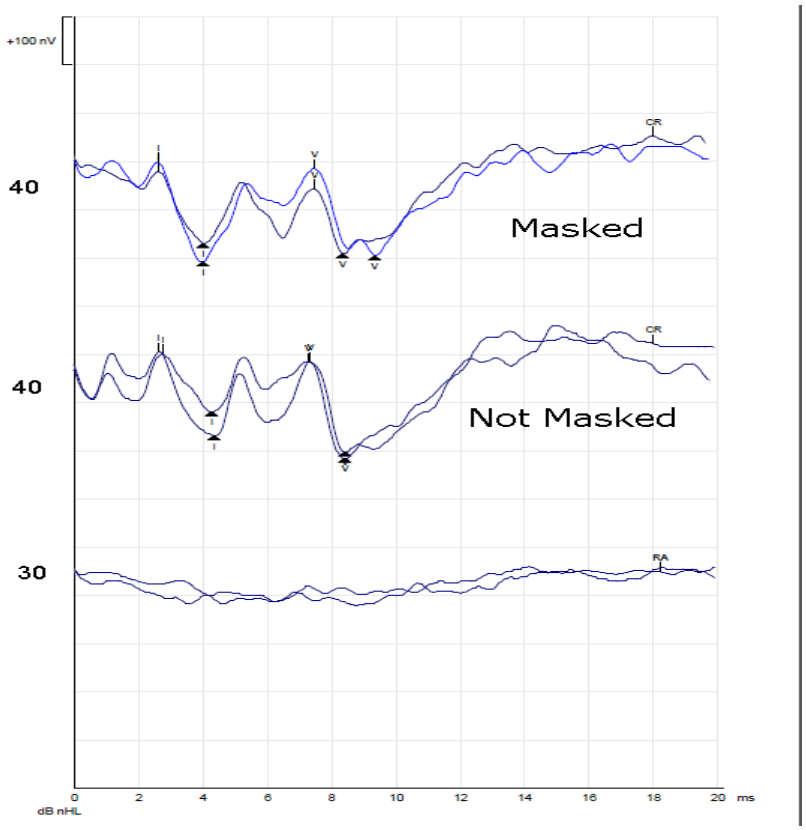

Figure 8: Single channel 4kHz BC waveforms

Reporting ear-specific results

When masking has been used to ensure an ABR threshold is ear-specific we should insert "(M)" after the result (e.g. =45dBnHL(M)) in the clinical report, S4H entry or peer review spreadsheet. Likewise, if 2-channel wave V or wave I tests have confirmed that a response is not crossed "(M)" should follow the result to denote that the result is ear-specific. A clinical note should be made to provide details. Similarly, if a response could be crossed (as warned by the masking noise calculator for example) yet masking, wave I and two-channel methods have not been used to confirm the response is ear-specific, then the result should be reported using the (NM) qualifier, thus warning the reader of the possibility that it could be crossed. Where 2-channel recordings have been used to confirm an ear-specific or a crossed response, these should be printed and used in the clinical notes and made available for peer review purposes.

5.10 Resolving inconclusive results

The primary strategy is to test at a given stimulus level until a decision of ***CR*** or ***RA*** can be made.

Where there is some evidence of a response but criteria for ***CR*** are not met, further replications may help resolve the result by reducing the residual noise (this is needed only where the level defines threshold). For example, two further recordings could be carried out and the waveforms could be added in pairs, using weighted addition. The waveforms should also be added in a fixed order to avoid bias (waveform 1 should be added to waveform 3 and waveform 2 should be added to waveform 4). This will produce two summed waveforms, each the average of a higher number of sweeps. If the waveforms are in equally good recording conditions the effect should be to reduce the noise level, and the new pair of waveforms can be examined to see if they meet the 3:1 condition for ***CR***, or whether they meet the criteria for ***RA***. Additional replications may be needed, and combined as above, if the result remains inconclusive.

On some occasions one may be able to obtain only one rather than two further waveforms. In this case waveform 1 should be added to waveform 3. The new weighted add waveform should be superimposed with waveform 2 to test against the criteria.
When reducing the noise by performing additional runs does not resolve an inconclusive result then a blocked stimulus run (see Section 5.15 and also 6.2) can be helpful. This technique is particularly valuable when a likely ***RA*** or ***CR*** is recorded at the maximum available stimulus level.

In making a decision at a given stimulus level all recorded waveforms at that stimulus level should be considered. A waveform may be rejected only where there is a good technical reason or where the noise levels (e.g. electro-myogenic activity) were untypically high. Waveforms should not be selected simply because they demonstrate a favoured result.

5.11 Definition of ABR threshold for NHSP

ABR threshold has been defined in the BSA early assessment guidelines (BSA, 2021) as the lowest level at which a clear response (***CR***) is present, with a response absent (***RA***) recording at a level 5dB or 10dB below the threshold[9], obtained under good recording conditions.

[9] The reason for including a replicated response at 5dB or 10dB above threshold as well as one at threshold is that it provides a safeguard against the possibility that the response at threshold is spurious. The response at 10dB or 5 dB above threshold should be larger and clearer than that at threshold. If, on audit, there is disagreement about the clear response at threshold, then the response at the higher level will normally limit the disagreement to 5dB or 10dB and prevent any serious error. If the response morphology across levels is abnormal a blocked stimulus run may be necessary to confirm the validity of the responses.

This is the definition that should be used for entering results onto S4H. The standard discharge level for NHSP is a clear response bilaterally at <= 30dBeHL at 4kHz. Refer to the BSA early assessment guidelines for the use of ABR threshold in management and other specialist cases.
Independent auditing / peer review of the results should not give thresholds more than 10dB different from those originally recorded.

5.12 Gold standard thresholds

The term 'gold standard' has been given to the combination:

- an ear-specific clear response (***CR***) at threshold
- a clear response (***CR***) at 5dB or 10dB above threshold, and
- a response absent (***RA***) at 5dB or 10dB below threshold.

To minimise errors in estimating ABR threshold, the gold standard should be obtained for at least one ear specific AC ABR threshold and one ear specific BC ABR threshold for each ear. See Q4 in appendix G. Testers or departments have the option to conduct testing to the gold standard for all stimuli if they wish.

Figure 9 shows a correctly displayed and interpreted ABR series using 4kHz air conduction tone pips. The vertical scale is 100nV/div. Stimulus levels are in dBnHL. The waveforms at 70dB and 60dB are ***CR***. At 50dB there is no candidate response and the average gap between the waveforms is about an eighth of a division i.e. 12nV so it meets the maximum noise criterion, with no evidence of a response, so 50dB can be considered as ***RA***.
The ABR threshold is therefore =60dB and 'gold standard' since there are ***CR***s at 60dB and at 5 to 10dB above (in this case at 70dB) and there is an ***RA*** at 5 to 10dB below (in this case at 50dB). Responses are ear-specific.

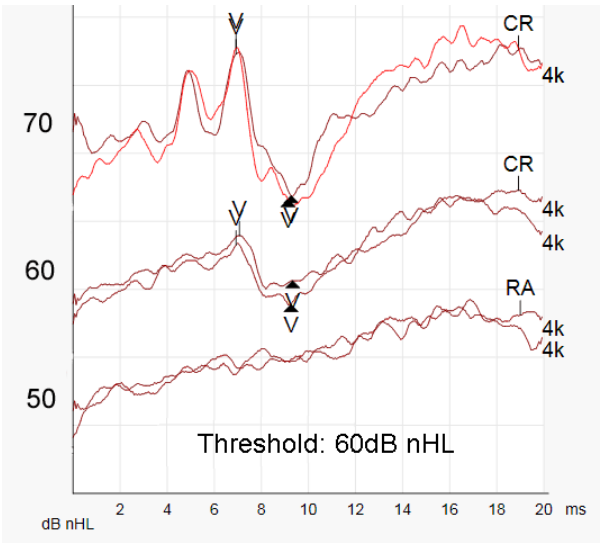

Figure 9: Waveforms meeting the 'gold Standard' for a 4kHz AC Tone Pip.

Note that in Fig 9 there is a slightly rising baseline but this does not compromise or invalidate interpretation.

The definition of the gold standard needs to be modified slightly in the following situations:

(a) Where the minimum discharge criterion of <=30dBeHL for 4kHz AC ABR is met; gold standard is met by recording a ***CR*** at 30dBeHL and at 35 or 40dBeHL[10]. An ***RA*** at 5dB or 10dB below threshold is not required in these circumstances.

(b) Where the minimum criterion of ≤20dBeHL for 4kHz BC ABR is met and there is evidence that the BC response is ear-specific, gold standard is met by recording a ***CR*** at 20dBeHL and at 25 or 30dBeHL. An ***RA*** at 5dB or 10dB below threshold is not required. Whilst <=20dBeHL (ear-specific) would satisfy the minimum criteria, in practice it may be pragmatic to test to <=15dBeHL to remove the possibility of a crossed response without recourse to masking or other ear-specific techniques in babies <12 weeks old.

(c) Where it is not possible to test at 5dB or 10dB above threshold due to the maximum stimulus level having been reached. In this case perform a blocked stimulus run (see section 6.2 and display as suggested above); if the response is genuine then it will disappear when the stimulus is blocked.

(d) Where there is ***RA*** at the maximum stimulus level. The ***RA*** should be of good quality, but 2 recordings suffice. The gold standard status for this scenario remains, but the implication this has for tests at other frequencies has been changed.
(e) In situations where results on a single ear yield a RA at one frequency and a response at another frequency, gold standard criteria should apply to both frequencies. This is to avoid the situation where the gold standard associated with the ***RA*** at the maximum stimulus level removes the requirement for a gold standard where a response is seen.

5.13 Testing at other frequencies

A gold standard threshold at 4kHz should be obtained before moving on to other frequencies.
Once a gold standard has been achieved at one frequency (for at least one AC and one BC (if tested) ABR threshold for each ear), this can be relaxed for other frequencies, if desired, provided that tests at these other frequencies are obtained within a maximum of 14 days of the gold standard test and that there is no evidence that there has been a change in hearing status. If these conditions are not satisfied, then the gold standard should be established at the next frequency to be tested. There should still be a high degree of certainty for each threshold measurement and any inconclusive results resolved, but there is no requirement to also test at 5dB or 10dB above threshold (the confirmatory level).

[10] For the derivation of the dBeHL value from the dBnHL value see Appendix D and the section entitled 'Prediction of the estimated hearing level threshold' in the BSA early assessment guidelines.

Where the discharge criterion has been met at the gold standard at 4kHz (where no ***RA*** is required), the gold standard at other frequencies (for example for 1 kHz AC ABR) can be claimed (if the tester wishes) by recording a ***CR*** at 30dBeHL or below and at 5dB or 10dB above. An ***RA*** at 5dB or 10dB below threshold is not required in these circumstances.

An example of threshold measured to the 'gold standard' is shown in figure 9 above.

An example of threshold not measured to the 'gold standard':
If the AC 4kHz tpABR threshold is measured using the 'gold standard', then if testing at AC 1kHz tpABR, a ***CR*** at 70dBnHL and ***RA*** at 60dBnHL is sufficient to report the threshold as =70dBnHL.

5.14 Reporting thresholds (incl those which are not gold standard)

Results should be clearly marked using the symbols '=', '≤' or '<=', and '>', and notes should always be made of any limitations or caveats about interpretation so this information is available to those who may carry out further tests. A threshold cannot be reported as meeting the gold standard if it is not ear-specific.

For example
'=45dBnHL' means ***CR*** at 45dB (and 5-10dB above for 'gold standard'), with ***RA*** at 35 or 40dB. Note that a "=" threshold can be reported only if there is an ***RA*** at 5dB or 10dB below threshold.
'<=45dBnHL' means ***CR*** at 45dB but not tested (or inconclusive) below this level.
'>80dBnHL' means ***RA*** at 80dB, but not tested (or inconclusive) above this level.

Where the gold standard defined above has not been achieved, threshold should be reported as follows:

a) If no 'confirmatory' ***CR*** Is obtained at 5 or 10dB above threshold, report threshold = lowest ***CR*** obtained *e.g.* ***CR*** at 70dB, ***RA*** at 60dB, threshold =70dBnHL (not gold standard).

b) If no ***CR*** is obtained above an ***RA*** result, report threshold > highest ***RA***;
e.g. ***Inc*** at 70dBL, ***inc*** at 60dB, ***RA*** at 50dB, report threshold as >50dBnHL (not gold standard).

c) If an ***RA*** response is obtained but not within 10dB of the lowest ***CR,*** report threshold as <=lowest ***CR*** and > highest ***RA***; for example <=60dB and >40dB (not gold standard).

For reporting in S4H (because only a single value can be entered) apply the following rules

1. 15dB or 20dB gap between lowest ***CR*** and highest ***RA*** -
e.g. ***CR*** at 70dB, ***Inc*** at 60dB, ***RA*** at 50dB. In S4H: enter =70dBnHL (in preference to <=70dBnHL), with a note in the session summary that the threshold is in the range 55 to 70dBnHL.

2. More than 20dB gap between lowest ***CR*** and highest ***RA*** -
e.g. ***CR*** at 70dB, ***Inc*** at 60dB, ***RA*** at 40dB. In S4H: enter <=70dBnHL, with a note in the session summary that the threshold lies in the range 45 to 70dBnHL.

d) If no ***RA*** is obtained below a ***CR***, report threshold as <=lowest ***CR***;
e.g. ***CR*** at 70dB, ***Inc*** at 60dB, ***inc*** at 50dB, report threshold <=70dBnHL

Note that the situations in c) and d) are best avoided, and efforts should always be made to reduce the gap between ***CR*** and ***RA*** to 10 dB or less when testing.

In all these cases, all clinical factors should also be taken into account, particularly where thresholds are being used for fitting hearing aids, and it is important to avoid over-amplification.

5.15 Baseline drift and the use of blocked-stimulus runs

This term (also known as a sloping baseline) is used to describe non-flat recording baselines such as those due to large stimulus artefacts. Other types of artefact may also give rise to baselines that drift such as contamination of the recording with cardiac activity. A moderate amount of baseline drift is acceptable if it does not affect the ability to observe an ABR and it cannot be taken for a false ABR (e.g. baseline drift due to stimulus artefact should end before the first key component of any ABR). If these rules regarding baseline drift are not met then the result should be considered as inconclusive. Where doubt exists in the possibility of genuine response presence or absence a blocked-stimulus run can sometimes help resolve the matter. Appendix F contains a series of examples to help the reader distinguish between what can be called ***RA*** and what is ***Inc*** when 'baseline drift' is present and of the role of blocked-stimulus runs.

5.16 Post-auricular myogenic (PAM) responses

Occasionally a post-auricular myogenic response which occurs around 12-15ms will be recorded when the ABR does not meet the ***CR*** criteria. This could occur in a baby who is not asleep or relaxed. If the state of the baby results in high levels of myogenic noise, and increased numbers of sweeps fail to resolve an ABR waveform to ***CR*** or ***RA*** status, the tester should consider changing the position of the baby and/or electrodes. A PAM (and myogenic interference in general) is more likely when the mastoid muscle is under tension and it is therefore worth checking the baby's head is not turned to one side. Any PAM will be attenuated if the mastoid electrode is replaced by an earlobe electrode. If this does not resolve the situation then a further test session should be considered. A PAM should not be considered in isolation, as a ***CR*** since ***CR*** requires a waveform with a characteristic ABR waveform for the purpose of interpretation.

Figure 10 shows an example of the post-auricular myogenic response, the large positive peak following wave V. The size of the response is very variable but can be very much larger than the ABR. In Figure 10 the dotted line is the common baseline for the two superimposed waveforms.

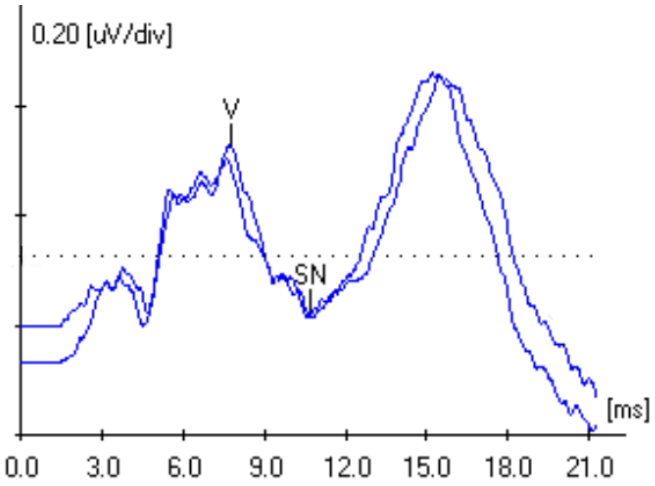

Figure 10. An example of a post-auricular myogenic response to a 35dB nHL click stimulus

6 Artefacts

6.1 Blocking of stimulus artefact

At high stimulus levels the stimulus artefact may exceed the AR level and so prevent recording. Some equipment will allow the artefact rejection to be ignored for a set latency range after the start of the stimulus to prevent this happening. This is referred to as 'blocking' by some manufacturers. It is suggested that if this facility is available it is set to a default value of 1.5ms for clicks and for the duration of the stimulus for tone pips. Please refer to equipment-specific recommendations for the blocking period when using chirps (ERA Training and Consultancy, 2023).

If this 'blocking' facility is not available, it may be possible to delay the start of the recording to the end of the stimulus artefact. If such a delay is introduced, then the time of the delay should be noted in the results and any latency measurements adjusted if necessary. Some ABR equipment offer the facility whereby the recorded waveform has a flat line displayed in the in the blocking period. These options do not affect the blocking function, merely the appearance of the displayed waveform. Systems should display the waveform by default. The 'flat line' option is appropriate only in cases where a stimulus artefact is so large as to cause waveforms to appear over several pages or be truncated when printed. Its inappropriate use can hamper correct interpretation in CM testing and when using stimuli below 2kHz.

6.2 Control (blocked stimulus) recordings during testing

Control recordings should be carried out whenever the ABR is marginal and/or is of the form that could be an artefactual response - e.g. mains artefact that is time locked to the stimulus could result in replicated waveforms that mimic some types of response.

When carrying out the control recording the stimulus should remain at the test level but prevented from stimulating the ear. If the response is artefactual, and is not an electro-physiological response to the sound stimulus, it will still be present. Turning the stimulus level right down is not appropriate.

Note on how to achieve stimulus blocking.
For AC ABR the acoustic block can be in the form of a disc of acoustic putty for earphones, or a tubing clamp for insert earphones, and should give a substantial reduction (>30dB) in the sound level. Note that as control recordings are carried out where responses are marginal and therefore close to the ABR threshold, a 30dB reduction is normally sufficient although one may see a response from the contralateral ear (when using 2-channel testing) if that has a much better threshold.

For BC ABR the bone conductor can be lifted a few millimetres from the scalp to prevent transmission of the sound. Note that a response may still occur by air-conduction of the stimulus and if necessary the bone conductor should be covered to reduce the airborne sound.

Touching the baby's skin may change the extent of mains-related activity recorded, but it may be very difficult to avoid this when undertaking a control recording.

Action required
If artefactual responses are observed, then it is essential to determine their source and remove them from the recording process. Advice should be sought and if necessary, manufacturers contacted, so that the source of the artefacts can be eliminated.

Figure 11 gives an "intensity series" for BC stimuli (note the stimulus artefact that one often sees when using a bone vibrator). Three waveforms should not have been used at 45dB. The use of 5dB steps is rarely justifiable clinically but is helpful here, for the proposes of demonstration. Clear responses are evident at 45dB & 40dB but how do we interpret the waveforms at 35dB & 30dB? Both have a residual noise of less than 25nV but are these ***CR*** or ***RA***? At 35dB the replications are highly correlated, the response is a fraction under half a division: 60nV, and, given the waveforms at higher levels, an acceptable response morphology. The noise is 20nV so the signal to noise ratio is 3:1. This level can therefore be accepted as a ***CR***. The same cannot be said of the 30dB waveforms since the wave V- SN_{10} response is 40nV, the signal to noise ratio is below 3:1 and the response morphology is more questionable. The noise is about 25nV but given the possible response, it cannot be ***RA***. The activity before wave V is higher (perhaps the vestiges of wave III) and if we allow this to determine the response amplitude the signal to noise ratio is >3:1.

This is a genuinely difficult case. The blocked stimulus run, labelled "control" (with the vibrator slightly lifted from the baby's mastoid) was performed in an attempt to resolve the 30dB waveforms. These are flatter than the result at 30dB and this supports the possible presence of physiological activity in the 30dB result. As a consequence, 30dB can be taken as ***CR***. Had the control run not been performed (or had not produced flat waveforms) 30dB would have to be interpreted as ***Inc***. The current definition of ***RA*** (requiring no evidence of a response) and the use of wave III together with the blocked stimulus run allows us to accept this as ***CR*** and report the threshold as ≤30dB.

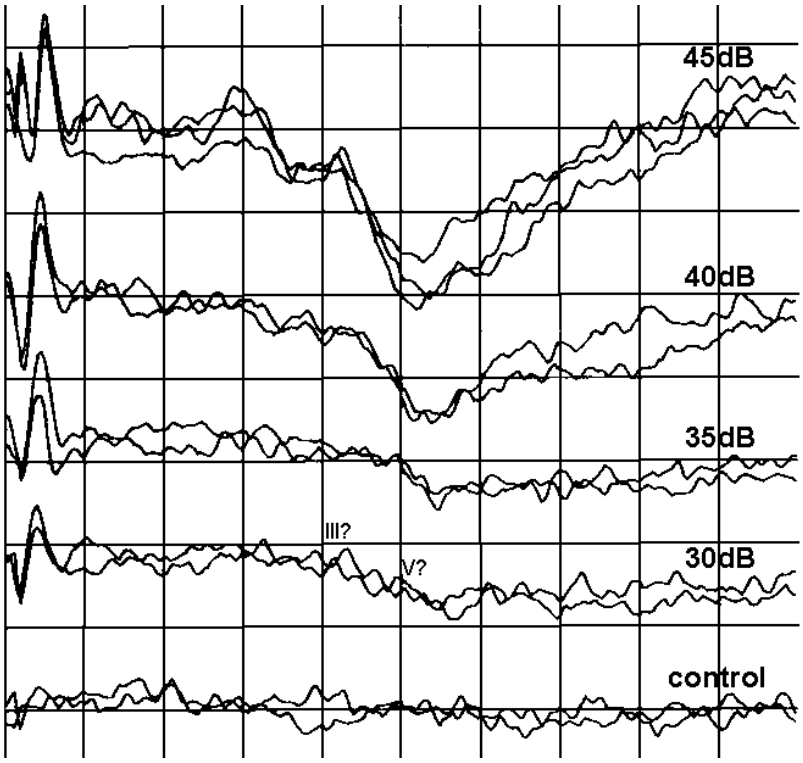

Figure 11: A BC "intensity series" showing a CR at 45dB, 40dB, 35dB (and also 30dB confirmed by the control run). Scales: 120nV / div. 2 ms / div.

A blocked stimulus run may not need to be replicated, if it resolves the uncertainty, but if any doubt remains, replication might assist interpretation. In Figure 11 above, the "control" was replicated because the first blocked run seemed to show a positive peak where wave III was seen at 30dB. Replication was performed, which revealed that the peak was not repeatable and therefore probably just residual noise.

7 Records Management

ABR reports (including waveforms) should be saved in an appropriate Electronic Patient Record or Patient Database. Furthermore, back-ups of patient records should be carried out regularly. Safekeeping and disposal should adhere to Records Management Code of Practice 2021 (NHSX, 2021) which states that adult health records should be retained for a minimum of 8 years and children's records up to 26th birthday (exceptions apply).

Appendix A: Major changes in this recommended procedure

- A new section (2.4) has been added on staff training, accreditation and peer review.
- The requirement to calibrate equipment on an annual basis has been added.
- "Stage A" checks has been included in the section on the preparation for testing.
- Equal emphasis is now given to single and two-channel methods.
- Reference to forehead BC placement has been removed.
- Artefact rejection levels are now given for systems with weighted averaging, which if available, is to be used by default.
- A new section outlines the characteristics of the ABR.
- New advice is given on when to stop averaging during the first run at a new stimulus level.
- The figure and text explaining the ***CR*** 3:1 SNR has been moved to the main text section.
- The term "appropriately flat" has been deleted from the definition of ***RA***.
- A new example of ***RA*** has been added in the main text.
- The gold standard now requires the threshold to be ear-specific.
- An example of the gold standard being met has been moved from Appendix C to section 5.12 of the main text.
- NM (not masked) has been introduced to denote the possibility that a reported threshold could be crossed.
- Much of the material in the former Appendices B, C & D have been transferred to the main text, to improve clarity and readability.
- A new example of 2-channel with a crossed response has been added to section 5.9.
- Appendix C includes new advice on latency variability.
- Appendix D gives the dBnHL to dBeHL corrections and maximum stimulus levels for babies up to 12 weeks.
- The interpretation of figures C4, F1 and F6 have been changed.
- The previous appendix on the use of rarefaction and condensation sub-averages in lieu of true replication has been deleted.
- Equipment-specific advice is given for Fsp/Fmp.
- Appendix E (Objective measurements) now requires replication of waveforms for ***RA***.
- Advice on interpretation and dealing with baseline drift (appendix F) has been revised.
- A previous appendix section on the use of A+B and A-B waveforms has been deleted.
- A new "Frequently Asked Questions" appendix G has been added.

Appendix B: 2-channel recording of wave V

Study	N	Age of infants	Hearing status	ABR or ASSR	Stimulus Freq (Hz)	Amplitude	Latency	Difference in threshold	% with absent contra responses
Foxe and Stapells 1993	21	2 weeks to 13 months	Normal	ABR	0.5 & 2k	Mean contralateral response significantly smaller (size of the difference not reported)	0.5kHz: Mean approx 2ms later in contra 2kHz: Mean approx 0.5-1 ms later in contra*		
Stapells and Ruben 1989	48	2 weeks to 2 years	Normal and conductive hearing loss	ABR	0.5 & 2k	0.5k Hz: Mean contra approx 0.1uV less in contra*. 94% had smaller contra response 2kHz: 93% had smaller contra response	0.5kHz Mean approx 1ms later in contra *		0.5kHz At 40 dBnHL 0% At 30 dBnHL 5% At 20 dBnHL ~ 25% At 10 dBnHL ~50% At 0 dBnHL ~75%
Small and Stapells 2008	14	8 to 44 weeks	Normal (passed DP screen)	ASSR	0.5 to 4k	Mean contra response was 57-73% of mean ipsi (all frequencies)		4kHz: Mean 14.2 dB worse in contra ear SD 5.2 Range 10-20 N=12	34% (unclear which frequency)

Table B1 Literature summary: comparison of ipsilateral and contralateral Bone-Conduction ABR waveforms in infants
(Foxe and Stapells, 1993); (Stapells and Ruben, 1989); (Small and Stapells, 2008)
*No range or SD presented. References for table B1: (details in References section).

Appendix C: More detailed advice and examples of ABR waveforms meeting the response criteria CR, RA and Inc

C1. Establishing a clear response (CR), satisfying the 3 to 1 signal to noise criterion

It is important to remember, as detailed in the main text, that as well as meeting the 3:1 signal to noise criteria, the waveforms should be inspected to see if they show the expected characteristics in terms of amplitude, latency and morphology. The residual noise in averaged waveforms can lead to minor variations in the measured latency and amplitudes of the response. In a recent study, (Lightfoot, 2024) reported that the typical wave V latency difference between replicates near threshold can be up to around 0.6 ms. The study also analysed the latency shift in wave V as the stimulus was changed from threshold plus 10 dB to threshold. A mean shift of around 0.5 ms was typical (threshold being later than 10 dB above threshold) but the 95% range included negative values, i.e. threshold was occasionally seen at an earlier latency than 10 dB above. Such is the effect residual noise or transducer movement can have on latency measurements. For this reason, responses should not be automatically rejected as spurious if latency repeatability or shift across stimulus levels is not entirely as expected. Rather, in such cases, further averaging (which will reduce the residual noise) or a blocked stimulus run should be performed to validate or reject a response.

The points taken to measure the response amplitude are usually the wave V peak and the SN_{10} trough but there are a few issues that should be highlighted. Firstly, when measuring the vertical position of a peak (or trough) take the average vertical position of each replicate's peak, not the highest (or lowest). Secondly, wave III is occasionally more positive than wave V and under such circumstances it is appropriate to use wave III instead of wave V as the positive point from which amplitude is measured since it is a valid component of the ABR. Figure C1 illustrates such a case. Finally, it is important to note that SN_{10} is not necessarily the first dip or trough following wave V. In Figure C1 there is such a dip, marked X (at higher test levels this would probably be the trough between waves V & VI) but, for the purposes of amplitude measurement, we should mark SN_{10} as the lowest point in the waveform following wave V (providing the waveform is not sloping – see appendix F). Although this definition of SN_{10} may not be what is seen in textbooks, we use the term as a convenient label for the reference point when measuring response amplitude.

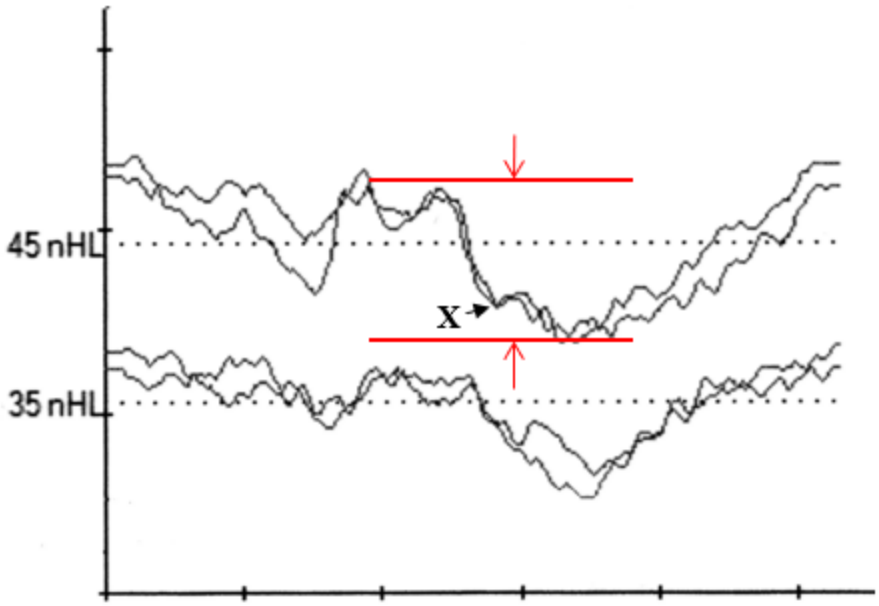

Figure C1: Example of a clear response (CR) using the wave III peak and SN_{10} trough to measure signal amplitude Scales: 300nV / div. 4 ms / div

C2. Estimating residual noise: deciding between response absent (RA) and inconclusive (Inc)

As with the assessment of clear responses, the assessment of waveforms for ***RA*** & ***Inc*** status should be across the entire window but excluding any region of stimulus artefact. With the waveforms optimally superimposed estimate the residual noise from the average gap between the pair of waveforms. For ***RA*** status at a given test level the average gap must be no more than 25nV.

Figure C2 contains a possible response and the waveforms have been superimposed in order to evaluate the signal to noise ratio and residual noise level. The signal to noise ratio as illustrated in Fig C2 is <3:1 so it cannot be classified as ***CR***. Note that one "response" is twice the size of the other - this difference is quite possibly just the effects of noise. The average gap between the waveforms is about 40nV so it fails the 25nV ***RA*** noise criterion too, in addition to a possible response being present. We must therefore categorise this as inconclusive (***Inc***).

If it were important to resolve this level into ***CR*** or ***RA*** then a further pair of averaging runs would be needed, to reduce the noise level. In order to interpret the resulting four waveforms they should be combined pair wise using a "weighted add" function and the resulting two waveforms' signal and noise should be reassessed. When combining waveforms it is important not to choose which pairs to combine

on the basis of their shape. For consistency it is recommended that four waveforms are always combined thus: 1 & 3; 2 & 4. See section 5.10.

Figure C2: An inconclusive response due to the 3:1 signal to noise rule for a *CR* not being met and the <25nV noise between waveforms not being met for a *RA* and a possible response being present. Scales: 150nV / div. 4ms / div

Figure C3 illustrates when the ***RA*** requirement to "*contain no evidence of a response (even an abnormal or uncharacteristic response)*" has not been satisfied, even though the residual noise criterion is met. The tester sought to ignore the activity >16ms because wave V and SN_{10} are expected at shorter latencies, yet this could be an abnormal response. Note how a blocked stimulus run or filtering would not have helped resolve this case. Inconclusive is the only safe interpretation.

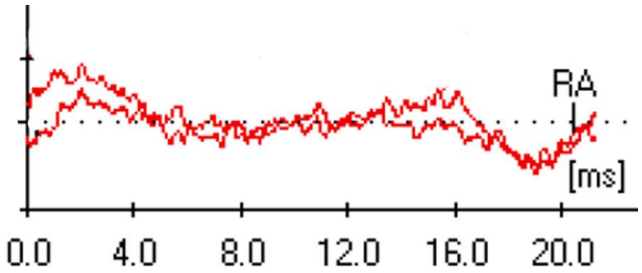

Figure C3. Failure to observe the "no evidence of a response" requirement.
Scales: 200nV / div. 4 ms / div

Figure C4 is a further example of the *RA* requirement to "*contain no evidence of a response (even an abnormal or uncharacteristic response)*" has not been satisfied, this time when short-latency activity is present beyond the stimulus artefact range. This might be wave I or perhaps the summating potential.

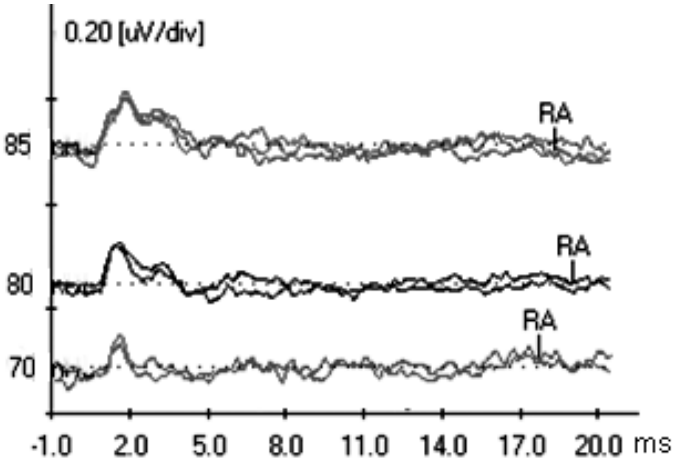

Figure C4. Failure to observe the "no evidence of a response" requirement.
Scales: 200nV / div. 4 ms / div

Neither figure C3 nor C4 provide us with the required high degree of certainty that a physiological response, even an abnormal one, is truly absent. Figure C4 should be regarded as an abnormal response and testing for ANSD should be considered.

C3. The use of blocked-stimulus control runs to assist interpretation

Figure C5 (4kHz AC) shows a ***CR*** at 50dBbut interpretation of the 40dB waveforms is more challenging. There is a high degree of correlation between the replications, but the latency shift is larger than expected. The next stage in the process is to ask if the criteria for ***RA*** have been met. The noise content (based on the average gap between replicates) is about a sixth of a division (20nV) so the noise criterion for ***RA*** is satisfied. However, we cannot be as certain that a response is truly absent. The only safe interpretation here is inconclusive (***Inc)*** even though the criteria for ***CR*** have been partially met (signal to noise ratio, response size (if considered a response) and the residual noise) and the criterion for residual noise has been met for ***RA***. In other words, this is a very difficult example but the tester should be guided by the question: 'Do you have a very high degree of confidence that your interpretation is correct?' In this case the answer has to be 'no' so ***Inc*** is the only safe conclusion. A blocked stimulus (control) run may have resolved this case.

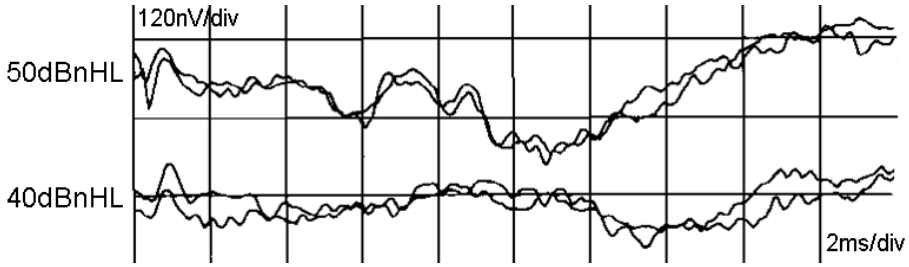

Figure C5: 4kHz AC waveform showing *CR* in the upper waveforms and an *Inc* in the lower waveforms which may have been resolved if a 'control run' had been performed. Scales: 120nV / div. 2 ms / div.

Appendix D: ABR dBnHL to dBeHL corrections for babies

In the table below, combined corrections should be added to the thresholds in dBnHL to give the estimated threshold in dBeHL. Taken from BSA Early Assessment guidance (BSA, 2021), which includes corrections values for older age groups.

AC - INSERTS	Tone pip/click ABR					Chirp			
Frequency (Hz)	500	1000	2000	4000	Click	500	1000	2000	4000
Correction (dB)	-15	-10	-5	0	5	-10	-5	0	5
AC - SUPRA-AURAL EARPHONES	Tone pip/click ABR					Chirp			
Frequency (Hz)	500	1000	2000	4000	Click	500	1000	2000	4000
Correction (dB)	-20	-15	-10	-10	-5	-15	-10	-5	-5
BC	Tone pip/click ABR					Chirp			
Frequency (Hz)	500	1000	2000	4000	Click	500	1000	2000	4000
Correction (dB)	5	5	-5	0	#	10	10	0	5

Table D1. ABR dBnHL to dBeHL corrections for babies ≤12 weeks

Refer to BSA Early Assessment guidance

.

Frequency (Hz)	500	1000	2000	4000	Click
Recommended maximum for tone pips & clicks	100	100	95	85	85
Recommended maximum for narrow band CE-Chirps	95	100	95	85	

Table D2. Normal maximum stimulus levels for inserts in dBnHL (≤12 weeks)

These values are designed to ensure the peak-peak SPL in the ear canal of newborns never exceeds 135dB

Frequency (Hz)	500	1000	2000	4000	Click
Recommended maximum for tone pips & clicks	110	115	110	105	100
Recommended maximum for narrow band CE-Chirps	105	110	105	100	

Table D3. Normal maximum stimulus levels for phones in dBnHL (all ages)

Appendix E: Objective measures for ABR interpretation in babies

This appendix contains some provisional advice on the use of response confidence measures (Fsp or Fmp) and residual noise measures. These assist the user to decide whether and when sufficient sweeps have been acquired, rather than using a simple fixed number of sweeps approach.

E1. Use of response confidence measures (Fsp/Fmp/SNR) in ABR testing in babies

E1.1 Introduction

Fsp is one of a number of measures available to determine the degree of confidence in the presence of an ABR (Elberling and Don, 1984). It compares the variance of the averaged waveform to the variance of the background noise level. The variance of the averaged waveform is a measure of the size of the ABR (if present) plus any residual noise. The higher the Fsp value the greater the response compared to the background noise and the greater the confidence of a clear response. Fmp (Don and Elberling, 1994) is a slightly more sophisticated (multiple point rather than single point) version of Fsp and for simplicity, will be referred to interchangeably in this document.

In this BSA Guidance for ABR testing in babies, the standard measure used to estimate this confidence is a visual estimate of the signal-to-noise ratio[11] (SNR) which has to reach a value of 3:1 as one of the conditions for a clear response. Note that there are other criteria, such as 'ABR-like waveform and good correlation between replications' that also have to be met before the result can be considered a clear response. The Fsp and Fmp statistics offer an alternative to measuring the SNR by inspection of a pair of replicated waveforms. An obvious use of Fsp and Fmp is to guide the tester when to stop averaging, providing the other criteria for ***CR*** have been satisfied and the minimum number of sweeps have been acquired. This has the potential to save time, especially when the response is large.

Note that Fsp & Fmp can help us decide whether a response is likely to be present; they cannot be used to tell us about response absence. For that, a measure of residual noise is needed, together with other ***RA*** criteria.

E1.2 Validation of Fmp and advice

The following refers to Fmp available on the Eclipse system with software version 4.6.1 or later. For earlier Eclipse software and for the Biologic NavPro system, it is now recommended that Fmp/Fsp is not used because of misleadingly high values that are associated with sloping baselines.

[11] Because of the way in which this is estimated, this signal-to-noise ratio is not a true SNR; the visually estimated 3:1 criterion probably corresponds to a true SNR of between 1.0 and 1.5

The Interacoustics Eclipse Fmp calculation was changed for software version 4.6.1, reducing the previous vulnerability to erroneous Fmp values associated with sloping baselines, caused by the 30Hz high-pass filter advocated in this guidance. A study to validate the updated Fmp was conducted (Lightfoot et al, 2023) and suggested that for a response certainty of 97.5%, the Fmp criterion should be >2.2. A certainty of 99% was provided by a criterion of 2.8. Note that other ***CR*** criteria also must be met, namely that the waveform morphology, amplitude and latency must be convincingly ABR-like and that the response amplitude must be at least 100nV (not the >40nV used for the primary ***CR*** criterion). At stimulus levels above the ABR threshold, a single (unreplicated) run many suffice when the Eclipse (software version 4.6.1 or later) Fmp exceeds the criterion and amplitude is at least 100nV, (see section E3 below) but at the stimulus level being taken as the ABR threshold, replication is always required, showing repeatable responses, with either:

- Fmp at or above the criterion and amplitude at least 100nV in both waveforms, or
- the pair of waveforms meet the BSA >3:1 criterion and the amplitude, averaged across the replicates, is at least 40nV.

The Fmp study also evaluated the sensitivity of Fmp to identify responses deemed to be present using the reference BSA (3:1) criterion and concluded that Fmp offers reasonable sensitivity at 10dB above the ABR threshold but at threshold, Fmp has limited sensitivity (51% at 4kHz and 32% at 1kHz). This means that Fmp values below 2.2 might be seen when a response meets the BSA 3:1 criterion. Expressed in a different way, a near-threshold response may require more sweeps for Fmp to reach 2.2 than for the BSA 3:1 criterion to be satisfied. If the criteria for ***CR*** have been met, low Fmp values do not disqualify the ***CR***.

Important: A low value of Fmp can never be taken to imply response absence.
Note also that if the stimulus artefact is close to the Fmp window or there is any other artefact in the waveform the Fmp result should not be used.

E1.3 Recommended settings for the Fmp range

Stimulus	Fmp range
Click	5 to 15ms
4kHz tone pip	5 to 15ms
2kHz tone pip	7 to 17ms
1kHz tone pip	10 to 20ms
0.5kHz tone pip	10 to 20ms
CE Chirps LS	5 to 15ms

E1.4 Recommended Fsp/Fmp criteria

The following criteria should correspond to a confidence of response presence of 97% or more.

Biologic Nav Pro	Do not use
Vivosonic Integrity V500 (manufacturer's data, version 8.12.0 or later))	>1.25
Neurosoft NeuroAudio (report by ERA Training)	>2.5
Interacoustics Eclipse (software version 4.6.1 or later): (Lightfoot et al, 2023)	>2.2
Interacoustics Eclipse (software version 4.4 or earlier):	Do not use

E2. Residual noise

E2.1 Background

Residual noise in ABR testing is a measure of the background electrical activity weighted by the averaging process. A baby when asleep will typically have a level of electrical activity (mainly EEG) of about 2 to 4µV peak to peak or about 0.75 to 1.5µV root-mean-square (RMS). Averaging will reduce this by the square root of the number of sweeps -e.g. after 2500 sweeps the value will be reduced by a factor of 50. So, for a typical baby the value would be expected to be between 15 and 30nV. This assumes that the artefact rejection level was adjusted such that little or no myogenic or electrical artefacts at all were allowed into the average. Considerably greater noise levels will result if any such activity is allowed into the average.

Both the amplitude of the background electrical activity in babies and the ABR vary considerably. In babies with a large ABR the level of residual noise required to achieve a clear response will not be as low as for babies with a small ABR.

The way in which ABR equipment calculates residual noise is not the same as the BSA "average gap between superimposed replicates" method. When using the residual noise figure offered by ABR systems it is therefore necessary to estimate the value equivalent to the BSA 25nV average gap.

E2.2 Comparison of equipment residual noise to average gap measures

Three studies were undertaken, comparing the residual noise values reported by the Nav Pro and Eclipse in order to establish equipment-specific noise values that correspond to the BSA "average gap" ***RA*** criterion of 25nV. When combined the studies concluded that these are 18nV for the Nav Pro and for the Eclipse (software version 4.4.2 or later). However there was some variability in the relationships and the above values represent the 50% point (the most likely). If we were to use these values as a surrogate for the BSA 25nV gap method then noise would be underestimated 50% of the time. A more cautious approach is to use slightly lower values, as given below under 'Advice'. These values should only be used as target values for when to stop averaging: they are not meant to be true estimates of residual noise. A further study was undertaken to determine the equivalent of the BSA 25nV ***RA*** noise criterion for the Vivosonic Integrity V500, NeuroAudio and GSI Audera Pro by feeding the same raw EEG data into those systems and an Eclipse system with any noise-weighted averaging disabled. The value for the Integrity is also included below under "Advice".

E2.3 Advice: Use of residual noise measures in determining RA

Residual noise values (if provided by the equipment) may be used as a guide of when to stop averaging if the outcome of the test appears to be a candidate for ***RA*** status. The recommended target values for five types of equipment are as follows.

Biologic Nav Pro	15nV
Vivosonic Integrity V500	20nV
NeuroAudio	17nV
GSI Audera Pro	60nV
Interacoustics Eclipse (software version 4.4.2 or later):	15nV
Interacoustics Eclipse (software version 4.2):	25nV

Even lower values of residual noise will make achievement of the BSA ***RA*** noise criterion more secure.

Note that both of the waveforms defining ***RA*** should have a residual noise value no more than the values shown above. Waveforms still have to meet the 25nV average gap criterion and the other ***RA*** criteria in section 5.9. However, there may be some instances where the <25nV average gap criterion has been satisfied but the residual noise reported by the ABR system is greater than the criterion shown above. In these cases, be guided by the 25nV average gap criterion, providing it has been judged accurately.

E3. Using objective confidence measurements as an alternative to waveform replication

The primary purpose of waveform replication is to allow the tester to judge whether a response is reliably present or absent, using the criteria developed for ***CR*** and ***RA***. However, the availability of objective response confidence and residual noise values could, theoretically, make replication unnecessary, providing other important criteria are satisfied. What follows is ***optional*** advice for those with access to Fsp or Fmp (for ***CR***). Fsp or Fmp and residual noise values, if relied on for interpretation, should be available on the printout for peer review.

E3.1 Clear Response

The definition of a Clear Response in section 5.9, relying on waveform replication, remains the reference and continues to be a requirement for the ***CR*** at the threshold level or for testing at discharge level (or lower levels being used in lieu of discharge level).

Single (unreplicated) waveforms qualify as ***CR*** at other stimulus levels (including those levels needed to define the gold standard) if the following modified criteria are satisfied:

- The waveforms should show the expected characteristics in terms of amplitude, latency and morphology.

- The size/amplitude of the response (as judged from the wave III/V to the following SN_{10} trough) should be a minimum of 100nV and Fsp/Fmp should exceed the criterion given in section 1.4 of this appendix.
- Waveforms should be compared with those at other stimulus levels (where available) to confirm that they follow the expected changes with stimulus level. Note however, that the latencies of observed peaks are influenced by residual noise and occasionally, peaks can have very similar latency for stimuli 10dB apart.

The principle underpinning ***CR*** is that there must be a high degree of confidence that a response is genuinely present. If the only recordable ***CR*** is at the maximum stimulus level, replication is required. Replication is essential if the waveform appears sloping or has a low-frequency component since such waveforms, with certain software, can have a large of Fsp even without a genuine response.

An example of the above modified criteria applied to a raised threshold is shown in Figure E1, using an Interacoustics Eclipse system (software version 4.6.1). At 55dB, the Fmp was 10.6 (***CR*** criterion >2.2) and the response amplitude was 300nV.

E3.2 Response Absent

The definition of Response Absent in section 5.9, relying on waveform replication, remains the reference. The previous version of this guidance (2019) allowed ***RA*** to be established from a single (unreplicated) waveform. This advice has been reversed; experience has revealed that without replication, it is often very difficult to ensure that a vestigial response is not present. Replication is therefore mandatory for ***RA***. Clinicians may continue to monitor the reported residual noise during acquisition, as a guide in deciding when to terminate an average.
Figure E1 also shows the use of residual noise for ***RA*** waveforms. The waveforms at 35dB are ***RA***; they are free of evidence of any response, have an average gap less than 25nV, supported by reported residual noise values of 12nV and 13nV (an Eclipse system was used, for which the ***RA*** criterion is <15nV in each waveform).

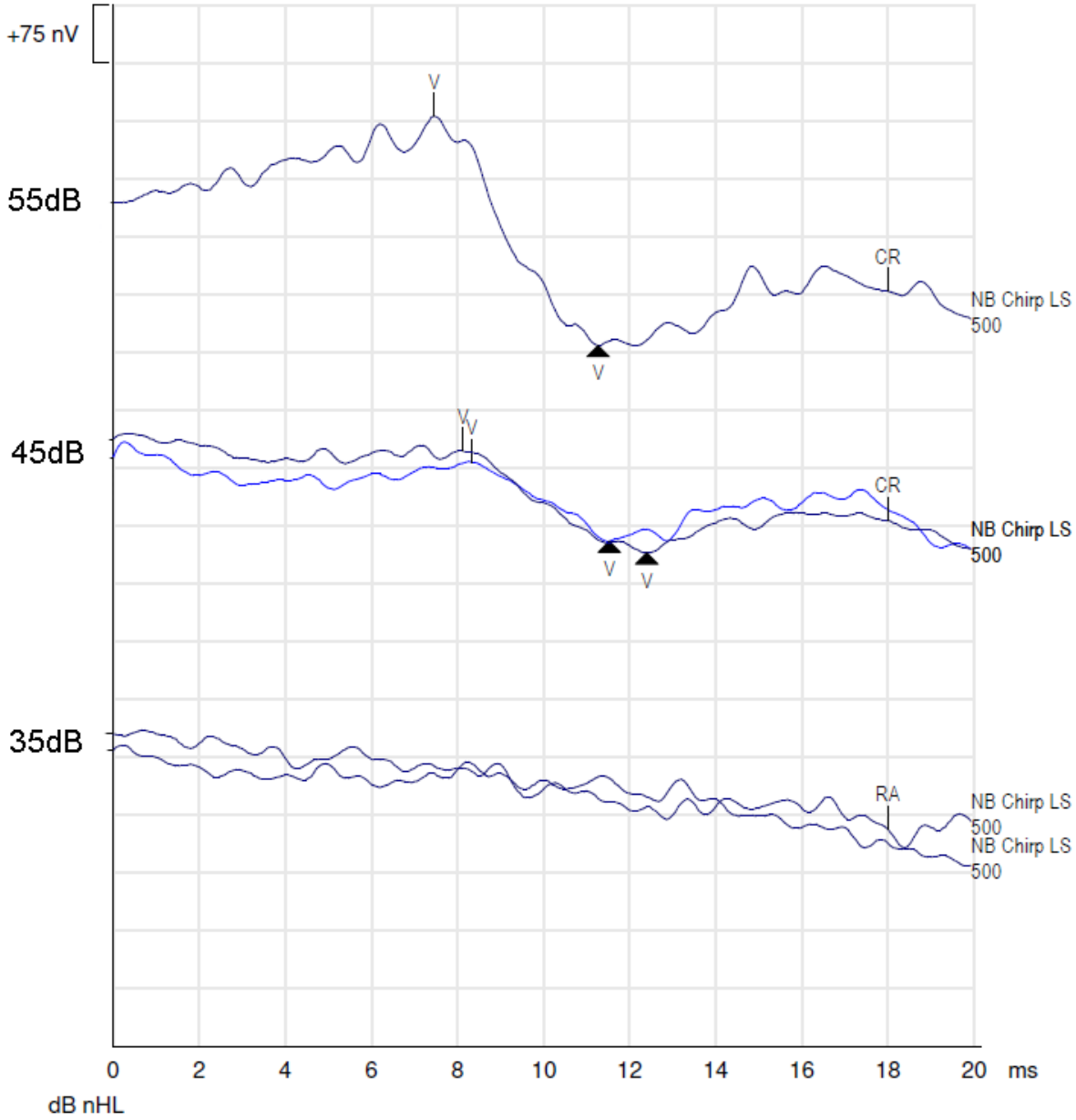

Figure E1. An example where Fmp was used to establish CR above threshold.

Appendix F: Baseline drift

This appendix expands on section 5.16

F1. Possible mechanisms

There are a number of possible mechanisms that could result in a sloping ABR baseline, and the stimulus artefact has often been cited, principally because a more likely cause had not been identified. However, a study (Lightfoot, 2017) has suggested that a probable cause in many cases is the cardiac activity of the patient, which is unintentionally recorded by the ABR electrodes. The degree of the resulting slope is influenced by the artefact rejection level, which the tester could change to investigate whether this makes the slope steeper or flatter (it could be either).

F2. Strategies available to the tester to deal with baseline drift

If changing the artefact rejection level leads to a change in the recorded slope or if performing a blocked stimulus run does not eliminate the slope, then this would support the conclusion that the slope is not ABR-related. Simply knowing this can be helpful in waveform interpretation.

The ability of ABR systems to reject "common mode" signals like mains is degraded if the recording electrode impedances are dissimilar, so an obvious first step is to check (and if necessary improve) electrode contact.

Every attempt to eliminate at source the noise causing the baseline drift should be undertaken. This may include turning off all other non-essential electrical equipment in the room, balancing the impedance of the electrodes to as low as possible, good electrode management (plaited and away from the transducer or other cables) & use of a grounding lead if available. However, if baseline drift continues and makes interpretation problematic, then raising the high-pass filter above 30Hz (for example to 75Hz or 100Hz) may help. Some equipment offers digital "display" filters which can be applied to existing waveforms, avoiding the need to obtain further averaging runs. Note: Whilst the application of filters is unlikely to change the clinical threshold significantly, this will have the effect of reducing the amplitude of any wave V - SN_{10} response so may increase the uncertainty of measurement. Filtered waveforms should therefore be identified as such in the clinical report and be interpreted with caution.

F3. Examples of 'baseline drift' and other waveforms with non-flat baselines that represent interpretation difficulties

Figure F1 shows an example where the baseline is not flat. There are ***CR***s for AC 1kHz tone pips at 50dB & 45dB. At 40dB the average gap is less than one-eighth of a division (25nV) but does this qualify as ***RA***? Are there response features present? The first half of the waveform appears highly correlated to the 50dB and 45dB waveforms, but it is doubtful that this is part of the response; it is probably cardiac related. All

suggestion of wave V and SN_{10} have disappeared at 40dB. If a blocked-stimulus run had been performed at 40dB and showed the same pattern, then ***RA*** could be claimed but the tester should appreciate that this approach may increase the uncertainty of measurement of threshold for this recording. However, without a blocked stimulus run at 40dB, 40dB should be regarded as ***Inc*** and the result reported as ≤45dB.

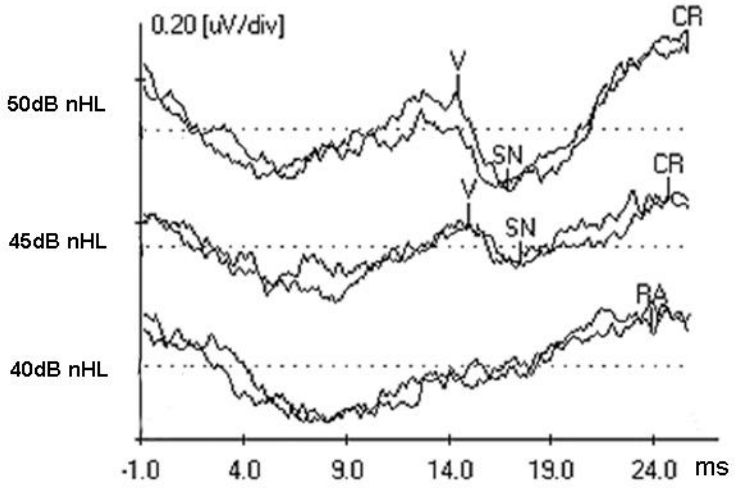

Figure F1: AC 1kHz waveforms at 50dB, 45dB and 40dB.
Scales: 200nV / div. 5 ms / div.

Note that in the previous guidance, 40dB had been deemed ***RA*** and the result given as =45dB.

Figure E1 (see previous section) did not require a blocked stimulus run at 35dB since the slope of the baseline was slight and free of any suggestion of a response.

Figure F2 illustrates downward-sloping baselines in which the two waveforms at 25dB have residual noise comfortably below the residual noise criterion for ***RA***, as measured by the equipment (in this an Eclipse: 12nV and 13nV), yet the average gap (left panel) is >25nV. Application of a 75Hz high-pass display filter (right panel) removes the low frequency drift, allowing ***RA*** to be concluded at 25dB. Note how, at 35dB and 45dB, the display filter sightly reduces the response amplitude and changes where SN_{10} markers would be placed – these should have been repositioned. The use of the display filter should be noted in the clinical report.

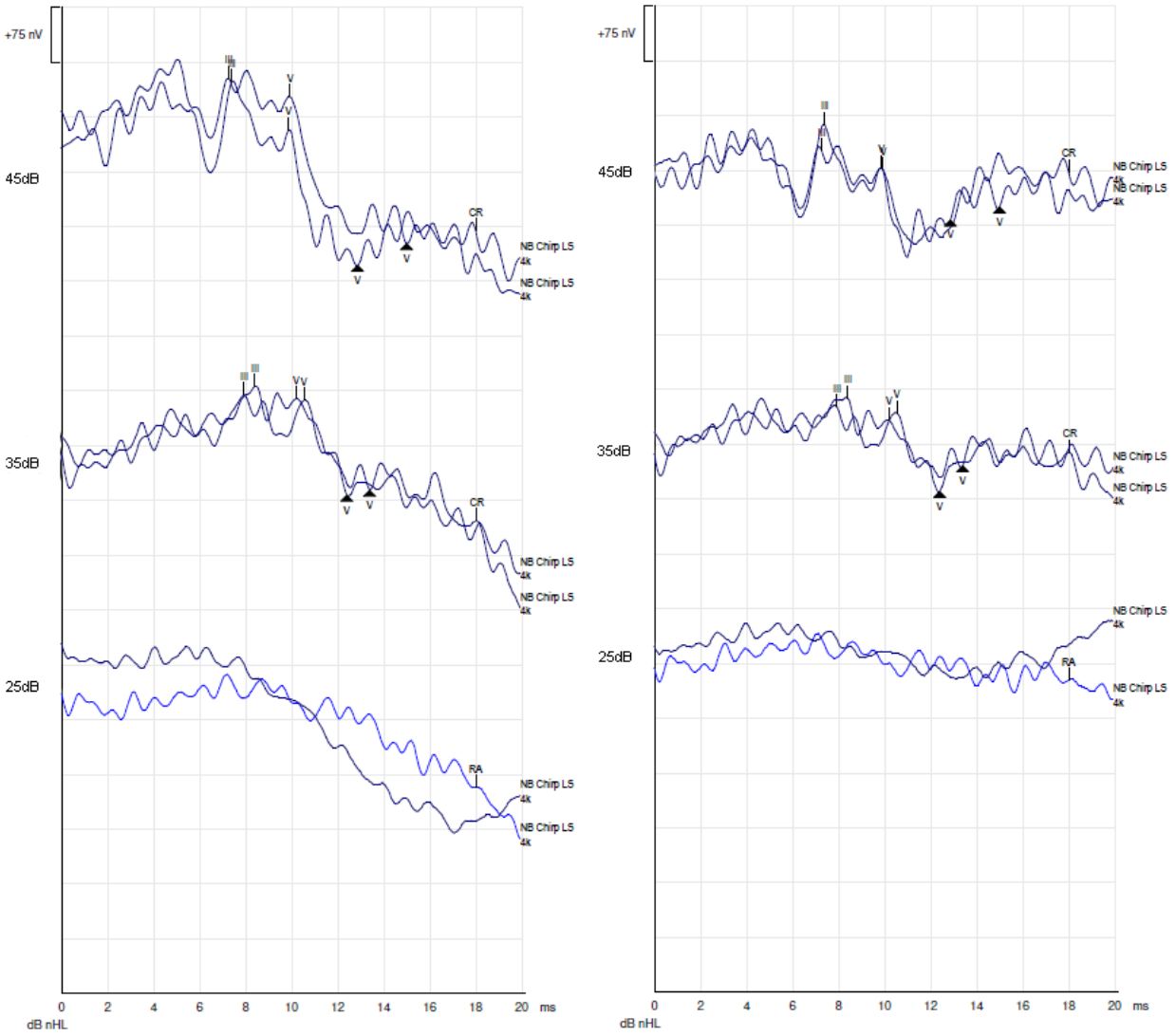

Figure F2. An example of the appropriate use of a digital high-pass display filter. Left panel: unfiltered waveforms; right panel: filtered waveforms.

Figure F3 meets the requirements for ***RA*** status but has a non-flat baseline. It has no features characteristic of an ABR and because of this, ***RA*** can be accepted. Nevertheless, if such a result is critical (for example if obtained at the maximum available stimulus level, suggesting a very elevated threshold) then confidence in this interpretation should be increased by attempting to remove the cause of the baseline drift and obtaining a further pair of replications. It may be helpful to run a blocked stimulus run to see if the baseline drift remains.

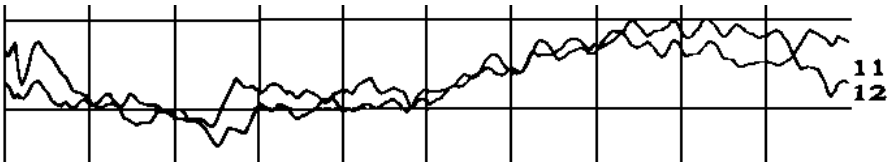

Figure F3: Non-flat baseline that meets the criteria for RA. Scales: 120nV / div. 2 ms / div.

In Figure F4 the AC 4kHz waveforms marked at 70dB have been positioned so as to superimpose the ABR peaks III & V which is tempting but not ideal in terms of assessing the residual noise. Another interesting feature is that in this example, the size of the response should be judged from wave III to SN_{10}. That being the case it is obvious that at 70dB there is a CR, even though superimposition is not ideal. However, our main concern in this example is how to interpret the 60dB waveforms. The noise is sufficiently low for ***RA*** yet there is a feature that some observers may interpret as a tiny wave V (if so, the signal to noise ratio is clearly below the required 3:1). Note that the baseline drift is apparent at both test levels. A previous version of this guidance stated: "The waveform at 60dB in this example may be taken as ***RA*** unless the outcome was critical or was at variance with results obtained for other stimuli." We wish to change this advice as the 60dB waveforms may contain a vestigial wave V so must be taken as ***Inc***. A blocked stimulus run could be obtained and ***RA*** concluded only if the feature remained.

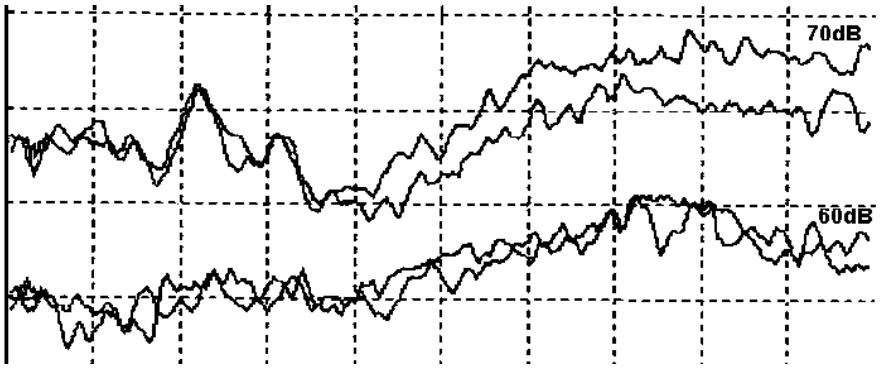

Figure F4: AC 4kHz waveforms at 70dBnHL and 60dBnHL. Scales: 120nV / div. 2.5 ms / div.

A further example of a possible ***RA*** with 'baseline drift' is illustrated in Figure F5. This again illustrates that it is sometimes difficult to distinguish non-flat baselines from near-threshold responses particularly for low frequency stimuli. The tester must rely on their judgement of what constitutes a "characteristic waveform" and how the features of this waveform (particularly latency) relate to those obtained at higher test levels. If a similar pattern was recorded with a blocked stimulus, ***RA*** would be acceptable but the tester should appreciate that this approach may increase the uncertainty of measurement of threshold for this recording.

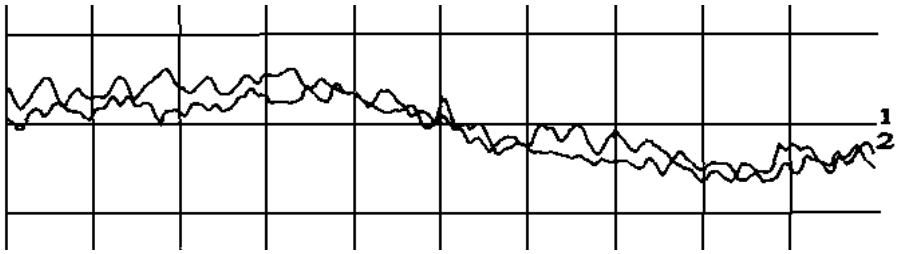

Figure F5: Non-flat baseline. Scales: 120nV / div. 2 ms / div.

The waveforms of Figure F6 might be incorrectly accepted as clear responses because they have a signal-to-noise ratio well in excess of 3:1 and some features of an ABR waveform of the form often seen with lower frequency stimuli close to threshold. As such the case could have been discharged. However closer inspection reveals no apparent increase in amplitude above threshold. This may be 50Hz noise that just happens to be in a phase that looks very response-like, or it may be the ECG artefact. Tests at a higher level and a blocked-stimulus control run (producing flat waveforms) would have helped resolve this case. Without that additional information these waveforms must be regarded as inconclusive.

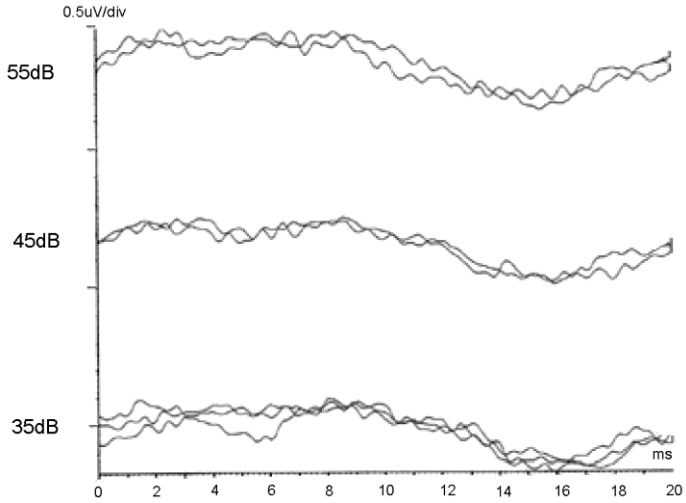

Figure F6: An inconclusive case that needs a blocked-stimulus control run to be performed to distinguish between a true response and possible artefact. Stimulus levels in dB nHL. Scale: 500nV / div; 1 ms / div.

BC waveforms are illustrated in Figure F7, with a ***CR*** evident at 50dB (stimulus levels in dBnHL). How do we interpret the waveforms at 40dB & 30dB? Both have residual noise below 25nV which on this scale is an eighth of a division. The "characteristic waveform" requirement is arguably just satisfied at 40dB and is probably not satisfied at 30dB but the degree of confidence in either conclusion is not high. We must try to ensure that any disagreement between independent observers is no more than 10dB. With that in mind it would be reasonable to interpret 40dB as ***CR*** only if a blocked stimulus run abolished the "response". If the pattern persisted, then an ***RA*** could be considered. This example is very challenging but fortunately examples like this are not very common in clinical practice. The attraction of a blocked stimulus run is obvious in this case. For BC, this is performed by holding the vibrator just clear of the normal application site yet with the stimulus being delivered to the vibrator (so that any stimulus artefact is maintained).

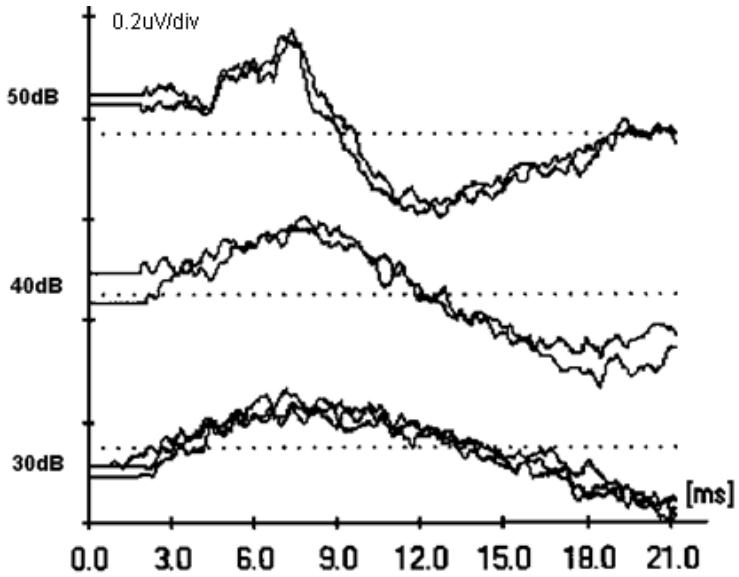

Figure F7: BC waveforms at 50dBnHL, 40dBnHL and 30dBnHL.
Scale: 200nV / div. 3 ms / div.

Figure F8 is an example of the good use of a blocked stimulus run, in which the downward pattern persists when blocked at 65dB, confirming it is not physiological and allowing ***RA*** to be concluded.

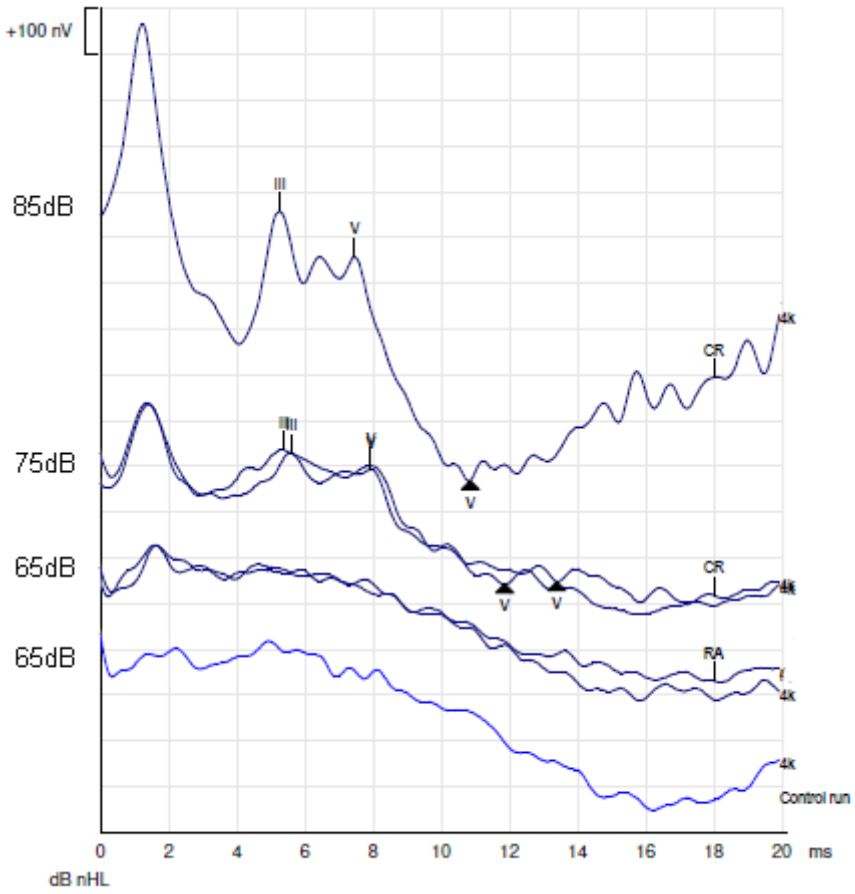

Figure F8. The blocked stimulus (control) run allows an RA to be concluded at 65dB

Appendix G: Frequently asked questions

These questions and answers attempt to address typical practical issues faced by testers. They are written in a less formal style than the main body of the recommended procedure, with the intention of being more accessible. Some of the answers in this appendix illustrate the intended degree of flexibility in the guidance when decisions can be scientifically justified.

Common mistakes

Q1: What are the most common mistakes made when using ABR in newborns and what should we do to avoid them?

A1: In no particular order,

- Placing the "active / positive" recording electrode too low on the forehead. The ABR is roughly vertically orientated, and our active and mastoid electrodes should consequently be vertically orientated (when the head is upright). Cz (the vertex) gives the biggest response but that seems to have fallen out of fashion because it is more difficult to securely attach an electrode where there is hair. The more horizontal the recording electrode orientation, the smaller the response, so ensure the active electrode is placed at the hairline, as high as possible but avoiding the fontanelle.
- Artefact rejection (AR) is too lax. It is not correct to say that it doesn't matter for systems that offer noise-weighted (e.g. Bayesian) averaging. Always keep the AR level at or below ±10µV (but see Q2 below). One possible source of confusion is how particular systems refer to how the pre-amp is set up. In addition to the AR level, the pre-amp's gain must be specified, in order to make the most of the amplifier's performance. The old Biologic NavPro called this "gain", for which a value of 240,000 was recommended for newborn tests. There are a number of other names the gain can be called. The Interacoustics Eclipse specifies this parameter in µV (the recommended value is ±40µV for AC and ±80µV for BC tests). The Eclipse printout shows both the AR and pre-amp sensitivity in µV and it's important not to confuse the two.
- Excess residual noise in the waveform (caused by using insufficient sweeps) to resolve the level into a ***CR*** or RA. In the olden days, a fixed number of sweeps was recommended (typically 2000) but more recently we believe it is more efficient and accurate to use as many sweeps as we need to answer the question, using objective measurements to help to make the decision. When the baby is restless and we use an AR of ±10µV, reducing the residual noise down to the ***RA*** criterion can occasionally require over 8000 sweeps. The problem is that testers also become restless and intolerant of the time it is taking, and they stop too soon, leaving them with an "***Inc***".
- Poor interpretation. This can mean wrongly assigning "***RA***" to a waveform with too much noise or ignoring a possible response. Conversely, mistakes are sometimes made in the assessment of the

signal-to-noise ratio. Both ***CR*** and ***RA*** should carry a certainty of at least 97%. In particular, ***RA*** does not simply mean the ***CR*** criteria haven't been met, it means we must be 97% sure we are genuinely below the ABR threshold.

- Abandoning tests at one frequency and trying another, usually because responses are "poor" (that usually means the recording conditions are poor). Flitting from one test to another without nailing the threshold is likely to leave you with a handful of inconclusives, with nothing you can report with certainty. Don't sacrifice quality for quantity. It is far better to be able to report one threshold with confidence than many that are questionable.
- Underuse of 2-channel recording. Masking is not easy and is not bullet-proof. The latest information on the minimum values of inter-aural attenuation at 1kHz and 500Hz now suggests cross-hearing is almost always possible and masking is needed. In recent years, some testers have made effective use of the contralateral channel to inform them whether a response is crossed or not. Remember that this method may not be reliable in children over 2 years (typically sedation or theatre cases).
- Not accounting for the non-test ear air-bone gap when using the noise (masking) calculator. Let's say the 4kHz tone pip better ear AC (insert) threshold is =40dBnHL and the better ear threshold BC is ≤15dBnHL. Some might enter 25dB as the air-bone gap (40-15) but the BC threshold could be as good as 0dBnHL so 40dB (40-0) should be entered for the air-bone gap.

Waveform interpretation, test settings and testing in difficult conditions

Q2: If I have Bayesian averaging, should I attempt to use the AR setting as close to ±5µV as I can, adjusting it whilst testing, or should I just select ±10µV and leave it there.

A2: set the AR to around ±10µV and leave it there. Use objective measurements to help you decide when to terminate the average. The only exception to that is if there is evidence of ECG (there is a slope to the baseline), when it is worth adjusting the rejection level to around 7µV, to see if the slope of the baseline reduces.

Q3: I could do with some advice on which test parameters to change when the baby is restless. In those circumstances, is it okay to use an artefact rejection of ±20µV, as I find it much easier to get a result?

A3: I assume you've tried waiting for baby to settle, but that isn't working. Be prepared to abandon the session and re-book. However, it might help to obtain something in this session as you cannot guarantee the baby will be settled at the next session. You may be able to record some useful information and the order you should change parameters are:

- Artefact rejection – increase to ±10µV if it was initially less than this but avoid using a higher value. In the initial phase adopt a "cheap & cheerful" approach to get in the ballpark of the threshold and only then use sufficient sweeps (which could be a lot) until you obtain ***CR*** and ***RA*** levels. However, if rejections are still near-total then:

- Increase the amplifier high-pass filter from 30Hz to 100Hz. This reduces the myogenic noise from a restless baby but also reduces the size of responses so the accuracy of your threshold may be somewhat poorer than normal. If you do change the filter, be sure to note this in your report.
- A restless baby will take longer to test than normal because more sweeps are needed to reduce the residual noise to acceptable levels. Time can be recouped by using 20dB rather than 10dB stimulus level steps. This will compromise threshold accuracy, but this is far better than compromising the ***CR*** and ***RA*** criteria. Sacrifice threshold precision if necessary, not certainty.
- In situations where it is deemed appropriate: accept that threshold cannot be determined at this stage but aim for suprathreshold testing. If clear responses can be seen, then at least this rules out auditory neuropathy, before next steps can be considered accordingly.

Be sure to detail in the report the reason for any departures from standard guidance.

Q4: Should I aim for gold standard at all costs?

A4. There will be some cases where despite recording 4 waveforms and adding them in pairs at a particular level (and maybe doing a no sound / blocked stimulus trial too), the result is still Inconclusive. That's just life and you need to move on at this point, ideally by getting definite ***CR***s at 10 and 20 dB above and, if applicable, an ***RA*** at 10 dB below. Note that the gold standard now additionally requires the threshold to be ear-specific.

Q5: I've seen a presentation showing how SNR can be measured whilst averaging on the Eclipse. Is it right that we should use an snr criterion of 4 and if so, why isn't it 3?

A5: I think you're referring to the powerpoint presentation by Guy Lightfoot. Lightfoot has chosen to withdraw that presentation and the method is not recommended now that Fmp on the Eclipse system has been improved. In fact, the Eclipse SNR criterion was ≥5.5 for an unreplicated waveform but at threshold, where replication is required, it was ≥4 in each of the two waveforms. It's not ≥3 because of the way the Eclipse measures residual noise.

Q6. The residual noise criterion for RA is 25nV "average gap" or 15nV objectively reported by the Eclipse. I see from appendix E of the ABR guidance that for RA, replication is needed, which is a change from the previous guidance. However, if I do two runs at the RA level, do both need to be ≤15nV? Surely not, because if I do a weighted addition, the residual noise would be much less than 15nV.

A6. Both waveforms should have residual noise at or below the target value given in E2.3. The two runs, if combined, would have lower residual noise. The equation to calculate the residual noise of two waveforms with residual noises x & y is:

$$\text{Noise} = 1/\sqrt{\frac{1}{x^2}+\frac{1}{y^2}}$$

That means two waveforms, each with 15nV noise, would combine to a single waveform with 10.6nV noise but there is no suggestion that you should combine the waveforms, so all that is rather academic. Don't forget that any ***RA*** also needs to contain no evidence of a response (even an abnormal or uncharacteristic response).

Q7. *There are two ways to measure residual noise for a potential RA – the "average gap" visual method and the objective version reported by the ABR equipment. My peer reviewer recently disagreed with my claimed RA (I used the eyeballing method) because the objective noise was more than the criterion for my equipment. Who's right?*
A7: Yes, that can happen. Providing you're confident your estimation of the average gap is <25nV and the waveforms contain no evidence of a response then you have a valid RA; the objective noise measurement can be ignored. However, your judgement of the average gap needs to be good!

Q8. I sometimes obtain ABRs with a sloping baseline. If it is downwards (from left to right) then the slope will make the measured response bigger than it really is and if the slope is upwards, then it will make the measured response smaller. That affects the estimation of SNR (3:1). How can I get round that, is there anything I can do to avoid the slope and what causes the slope in the first place?

A8.
A8. Firstly, check that the electrode impedances are satisfactory. We now think that those slopes are often a side-effect of the baby's ECG being picked up by the ABR electrodes. See Lightfoot (2017). If it is the ECG, changing the artefact rejection level often alters the slope (for better or worse!) so a quick change of AR level might be worth a try, to see if that reduces the slope. If that doesn't help, then raising the high-pass amplifier filter above 30Hz (for example to 75Hz or 100Hz) may help. Another alternative to remove or reduce the slope is to select a high-pass ***display*** filter of 50-75Hz. However, raising the high-pass filter or applying display filters may also reduce the size of the response slightly, especially for low-frequency tests, so use these only when you need to, and note them in your clinical report.

Q9. When I get a sloping waveform, the Fmp value seems to be very high compared to testing when the waveform doesn't have a slope. Can I accept this value when thinking about testing without replicating?

A9. That happened with the Biologic NavPro (Fsp) and the Interacoustics Eclipse (Fmp) with software earlier than 4.6. It's a technical issue concerning our use of the 30Hz high-pass filter. Interacoustics fixed this problem with the introduction of software version 4.6.1. It doesn't change the appearance of the waveforms, but the Fmp value is no longer affected by a sloping baseline. The NavPro Fsp facility is no longer recommended. Don't overlook the other criteria for ***CR*** – in particular, it needs to have convincingly ABR-like morphology. If you're unsure with single waveform, the solution is obvious: replicate!

Q10. A case of mine recently came back from peer review with a suggestion I think is wrong. My ABR system doesn't quote % rejections, it gives the number of accepted and rejected sweeps. The ABR guidance suggests relaxing the artefact rejection (AR) level if rejections exceed 30% and increasing the number of

sweeps to compensate. I used ±5µV throughout and the baby's state was fairly consistent – settled but sucking a dummy. Typical sweeps were 3000 accepted and 1000 rejected which the reviewer said is over 30% and that I should have changed the AR level whereas I think it is under 30%. Who's right?

A10. You're right. The rejection % is the number of rejections divided by the number of stimuli presented (in this case 3000+1000). However, the guidance relating to artefact rejection has changed and it is no longer recommended to change it, so this issue has gone away. Instead of trying to predict how many sweeps to use, monitor the residual noise and use as many sweeps as needed to reduce it to allow the ***CR*** & ***RA*** criteria to be met.

Q11. *I use a Vivosonic system and the manufacturer suggests using an artefact rejection level of 20, not 10uV. Which should I use? Also, there are two values of sweeps reported – number of stimuli and noise-adjusted sweeps. Which of these need do I go by?*

A11: We have discussed this with Vivosonic and have agreed that in the UK, 10uV should be used for all ABR systems. Use the "noise-adjusted sweeps".

Q12. One of the CR criteria is a minimum response size of 40nV. In a recent case, at threshold, the response was 45nV on one run and 38nV on the other. The 3:1 signal-to-noise criterion was met and there was a good CR at 10dB above. Does this combination allow me to rate this as CR?

A12. Yes. Residual noise can do that, adding to the measured response on one occasion and subtracting it on another – it depends where the bumps of noise happen to be on the waveform. In this case the average response is 41.5nV. Simple averaging of the two response amplitudes is scientifically valid. In other words, both responses need not be at least 40nV, just their average. To facilitate peer review, it is helpful to mark wave V and SN10 on both waveforms, allowing the response amplitude to appear in the report.

Q13. I have a profound PCHI case (CMs and OAE absent, tymps peaked, so probably not ANSD) with >50dBnHL 4kHz BC so I did 1kHz BC. Surprisingly, I got a large and odd-shaped response at 40dB which on the face of it would suggest a large air-bone gap at 1kHz and moderate cochlear sensitivity (not profound). A blocked run suggested it is not an electrical artefact (the response disappeared). I've heard on the grapevine that this pattern has been seen before but I don't know the explanation or how to manage the case. Help!

A13. Yes, this pattern has been reported elsewhere but its origin remains a mystery. It is probably best to manage the case as a PCHI (non-conductive). It rather limits the clinical value of 1kHz BC, when there is a PCHI, and there are some who advocate using only 4kHz AC & BC, together with tymps, to assess the conductive status. Just be aware that there is a diversity of opinion on this.

Q14: Is it okay to add or merge two waveforms to form a single waveform to help with interpretation using objective measures?

A14: That's not the primary intended use of objective measurements, which is to guide you when it is appropriate to stop an average. The idea is that you monitor the objective measurement (residual noise in the case of a potential ***RA*** and Fsp or Fmp in the case of a potential ***CR***) and to stop averaging when the criterion is met, together with the other ***RA*** or ***CR*** criteria.
What you are suggesting is to combine a pair of waveforms that have already been collected and then see what the objective measurements say. In that scenario, you should make a judgement based on the methods detailed in section 5.9, using a pair of superimposed waveforms. Replication is needed for an ***RA*** and combing the two ***RA*** waveforms is unhelpful.

Test strategy

Q15. The guidance suggests doing 4 runs and averaging if there is any doubt, and also suggests if you have only been able to obtain three runs, then to average 1 and 3 and compare to 2. Say I complete two runs that I suspect are almost CR but just a little too noisy, and I run a third. By averaging 1 with 3 and comparing with 2 I now get a CR. Can I stop there even though the baby is sleeping and I could in theory run a fourth? Or is it reasonable to move on with testing at other levels?

A15. If after run 3 and the pairwise addition of runs 1 & 3, you have a comfortable ***CR*** then doing a 4th would probably not be the best use of time if there is more testing to do. However, do a 4th run if the alternative is simply putting the kettle on! As your peer reviewer or mentor, I'd want to satisfy myself that you didn't get into the habit of routinely doing only 3 runs. Good clinical practice involves the judgement of where to most effectively devote test time – the delicate balance between technical precision and maximising clinical information.

Q16: I'd appreciate some advice on the best strategy to use to most efficiently define the ABR threshold. For example, when I'm replicating a level because I reckon it could be threshold, I am able to compare the current waveform to the other one, and I stop when I think the 3:1 SNR criterion has been satisfied. However, when I start a new test or a new stimulus level, there's nothing there for comparison. Are there any suggestions or rules of thumb I can use to help me decide what the most efficient sequence of stimulus levels should be and when to hit the STOP button?

A16: Great question. Section 5.7 offers new advice on this; it refers to "initial" and "precision" phases of the test, but an example should help. Assume the use of 4kHz NB CE-Chirps using an Interacoustics Eclipse system (software version 4.7, for which the Fmp criterion is >2.2), inserts, testing a 3 week corrected age baby, unilateral NHSP referral. In these conditions, the NHSP discharge criterion of 30dBeHL corresponds to 25dBnHL. The baby is sleeping.
Better ear

- Start at 35dBnHL (discharge level + 10dB) in initial mode. After 1000 sweeps we look at the waveform (it appears to show a classic ABR pattern), the RN is (35nV) and Fmp is (2.7). We stop because a likely response is present, as evidenced by Fmp even though the RN is still high. If Fmp hadn't reached the response criterion we would have continued averaging until RN was <30nV.
- Go down by 10dB and test at 25dBnHL. After 1000 sweeps we look at the waveform (a slightly smaller and later ABR than at 35dB), the RN is 37nV and Fmp is 1.9 so we continue averaging until RN falls below 30nV (at 1900 sweeps).
- We replicate at 25dBnHL. We are hopeful that 25dB will satisfy the criteria for ***CR*** so we switch into "precision" mode. Whilst averaging we superimpose the first 25dB run with the current run. After 1000 sweeps we look at the waveform (similar to the first run) and continue averaging until the SNR, judged visually, is over 3:1. We stop at 2200 sweeps. We have a ***CR*** at 25dBnHL.
- Because the 35dB run was >100nV in size and Fmp was over 2.2, we can accept the unreplicated 35dB as ***CR*** for the purposes of the gold standard, and we report ≤25dBnHL. The discharge criterion has been met for this ear.

Referred ear

- Start at 35dBnHL in initial mode. After 1000 sweeps we look at the waveform (no obvious ABR is seen) so we stop when RN falls below 30nV. Officially this waveform is inconclusive but we suspect it is below threshold.
- Go up by 20dB and test at 55dBnHL. After 1000 sweeps we look at the waveform (still no obvious ABR) so again we stop when RN falls below 30nV. Inconclusive.
- Go up by another 20dB and test at 75dBnHL. After 1000 sweeps we look at the waveform (this time we see a classic ABR), the RN is 34nV and Fmp is 1.8 so we continue, and at 2400 sweeps the RN is 29nV, the Fmp is 3.6 and response size is 310nV so we stop. We are confident the threshold is at or below this level.
- Go down by 10dB and test at 65dBnHL. After 1000 sweeps we look at the waveform and see a small but likely ABR. We move into precision mode and continue averaging. We prepare to stop when Fmp rises above 2.2 but it doesn't. By 2600 sweeps, RN is at 28nV so we stop.
- Replicate at 65dBnHL, superimposing the two waveforms. As with the first 65dB run, we stop on the basis of RN and the SNR, which we judge to be about 4:1. This is a ***CR***.
- We return to 55dBnHL, but this time are prepared to continue averaging until RN falls below the ***RA*** criterion of 15nV. We see a fairly flat waveform with no suggestion of a response. We can't utilise the original 55dB for ***RA*** because it was too noisy, though it could be merged with another waveform at that level to achieve <15nV noise level.
- Replicate at 55dBnHL, superimposing with the previous run, again stopping when RN is <15nV. There is no evidence of a response. 55dB therefore meets the criteria for ***RA***.

- We report =65dBnHL, gold standard. We have a replicated ***RA*** at 55dB, a replicated ***CR*** at 65dB and an unreplicated ***CR*** at 75dB on the basis of Fmp.
- Could the 65dB response be crossed? Reference to the 2024 masking noise calculator suggests it won't be crossed in this example.
- We would need to determine whether the referred ear has a conductive or sensorineural loss so next, we would test by BC at 4kHz on the referred ear.
- If the ear-specific BC results on the referred ear showed an elevated threshold, suggesting a sensorineural component, we would need to return to the better ear to demonstrate AC responses at 20dBeHL. For that reason, some testers opt to test the better ear AC at 30 & 20dBeHL from the start.

Q17: My system doesn't offer Fmp so I can't benefit from the options not to replicate above threshold, and I therefore replicate at all levels, not just at the threshold level. A colleague says I'm wasting time by doing that, but I don't want to compromise the accuracy of my results. Who's right?

A17: Your attitude to accuracy is admirable but your colleague could be correct. You're absolutely right to insist on satisfying the ***CR*** and ***RA*** criteria at threshold, at 5-10dB above and at 5-10dB below. However, getting to that neck of the woods, when the threshold turns out to be elevated, shouldn't need replication at the lower stimulus levels which turn out to be well below threshold. Example: let's say you begin at 40dBeHL (because that's discharge level +10dB) and in your first run you see no obvious response. You have a decision to make: do you replicate and spend time to record ***RA*** at 40dBeHL, or do you go up. If you go up without replicating, then that 40dBeHL isn't an ***RA*** but we can live with that for the moment. Go up to 60dBeHL and do the same again. Only begin replicating when you're using stimulus levels that are likely define the result. Section 5.7 refers to **initial** and **precision** phases of testing.

Q18 At my first session for 4 kHz AC, I managed to get 60 dBnHL CR, 50 dBnHL CR and 40 dBnHL Inc (though looked like a possible RA) and I therefore reported ≤50 dBnHL. At the next session should I aim to resolve the 40 dB level or move on to other tests?

A18. Although your current results limit the extent of a loss, so far, your ≤50 dBnHL doesn't convincingly demonstrate there is <u>any</u> hearing loss (the threshold could be normal), so start the next session with the priority of getting an ***RA*** (or not) at 40 before moving to other tests. Also check that the ***CR*** at 50 remains valid.

Masking and 2-channel recording

Q19. I have a bilateral referral, with flat tymps and moderately raised AC 4kHz thresholds so I need to do BC. Using 4kHz chirps, I know that if I get a BC response at 10dBnHL it will be ear-specific. I obtain CRs at 20dB and 10dB bilaterally, and I intend to report ≤10dB but the issue is that without masking, the responses

at 20dB aren't ear-specific. Do I need to demonstrate that the 20dB is ear-specific (using 2-channel or masking) or does the ear-specific response at 10dB let me off the hook?

A19: It's difficult to envisage a scenario in which the 20dB response is crossed whilst the 10dB is not – it would take some pretty unusual non-linearities in the two ears. Pragmatically therefore, the response at 10dB gets you off the hook and there is no need to mask or use 2-channel methods at 20dB. If this is a bilateral conductive (and you don't know that in advance) then masking could be problematic because of the risk of cross-masking, so if you didn't get a ***CR*** at 10dB then 2-channel would be the method of choice, providing the patient is under 2 years.

Q20: The masking noise calculator says I should mask at the stimulus level I'm thinking of using. Should I use masking straight away or only after obtaining the not-masked threshold?

A20: Firstly, think about using 2-channel testing – see section 5.9, and if you routinely attach 4 electrodes, the additional information is effectively free and available at the click of a mouse. However, to answer the question, it may be best NOT to mask initially, because despite what the calculator says, you may not need to, or the level of noise it suggests may be incorrect (not because the calculator is faulty, but you may not have the correct information at this stage).
An example might help. Take the case of a bilateral screening referral; today, OAEs are absent and tymps are non-peaked. You're thinking this is probably a bilateral conductive and statistically, that's most likely. AC 4kHz ABR thresholds are =55dBeHL (Rt) and =60dBeHL (Lt). Time for BC. Because of the likely conductive you start 4kHz BC (Rt) at 25dBeHL but get no obvious response. The 4kHz not-masked BC thresholds turn out to be =35dBeHL (Rt) and =40dBeHL (Lt). A bilateral mixed loss. Those levels are greater than those at which we know BC is ear-specific, so you need to mask, correct? Nope.
Masking is not needed, because the Rt & Lt BC thresholds are 10dB or less apart (that's less than the minimum inter-aural attenuation in a baby at 4kHz) so must, by definition, be not crossed. If you had masked from the start then you wouldn't have known the value of the non-test ear air-bone gap (20dB) or the non-test ear BC threshold (35-40dB), so the calculator would give the wrong advice if you used the default values. Moral: find the thresholds without masking initially, then have a think!

Q21: How do I know if the level of masking noise suggested by the masking calculator is dangerously high?

A21: The masking calculator will give the message "Warning! High noise level" if this is the case. However, that warning is triggered only when the noise level exceeds 100dB SPL so keep an eye on the noise level and use your own judgement. Unlike the stimulus, which is transient and whose peak SPL is a lot higher than the nHL level, which explains the warning we see then using inserts in babies, masking noise is continual and usually calibrated in SPL.

Q22: Have you got any advice for doing BC ABR testing on children with bilateral conductive losses, where there is a risk of cross masking?

A22: In children under 2 yrs, 2-channel recording can be a great help. Identifying wave I is another tip (observing wave I in the ipsilateral channel confirms that the response is being generated by the cochlea on that side). However, if wave I is not seen, we can draw no conclusion, since wave I is often not seen, especially at threshold. In older children, not-masked BC ABR thresholds may be the best you can get (in which case they need to be reported as originating from whichever is the better ear, and that may be unknown). Flag them with the (NM) qualifier.

Q23: Can 2-channel recording be used for air conduction ABR testing as well as BC testing?

A23: Yes, this can be useful in cases of unilateral or asymmetrical hearing loss in children under 2 years to check whether masking is required.

Q24. In a unilateral case (one ear is fine), where the AC threshold at 4kHz of the poorer ear is, say, 60dB, I want to test the poorer ear BC at 4kHz, to find out if it is conductive or sensorineural. I start the BC test at 25dBeHL (on the basis that if I get a CR at 25 and at 15 then I don't need to mask and I've got my answer – it is a conductive loss). If I don't see a BC response at 25 and intend going up to 35, is it worth masking the better ear from that point onwards? After all, if I do get a response then I'm going to have to mask anyway.

A24. You could do that. However, an alternative approach is not to mask and instead look at the contralateral channel to tell you whether any response is crossed. Mask only if the BC response is possibly crossed. When you do apply masking, assume the better cochlea (BC) has a 0dB threshold. If you've got tymp results then if the tymp of the better ear is not peaked (or if you haven't got tymp results) it will be worth selecting a non-test ear air-bone gap value of 20-30dB in the masking noise calculator. In fact, that may be worth doing if you didn't test below 30dBeHL by AC since a minor conductive may be present in the non-test ear. You can get a rough idea of any non-test ear conductive element from the size of the AC response. You say the better ear is fine so I presume a ***CR*** was obtained at the discharge level but if it was small then there may be a 10-20dB conductive on that side, which should be taken into account when masking.

Q25. I have a 3 week old baby and I record flat tymps bilaterally. My expectation is that this is going to be a bilateral conductive. How will that affect my test strategy and what value should I put into the masking calculator for the air-bone gaps?

A25. Follow your normal routine, without masking, until you have obtained the AC and BC 4kHz thresholds in both ears. Then sit back and think it through. Look at the BC thresholds – if they are no more than 10dB apart then neither will be crossed and masking will not be needed for the BC tests. If they are >10dB apart the better one will be valid but there is a risk the poorer one (with the higher threshold) may be crossed, with the real threshold being higher. In cases like this, 2-channel ABR is very helpful and may be sufficient to confirm that the responses are not crossed. If you do need to mask the better ear when testing the worse ear, then go by the air-bone gap you have recorded by ABR but be careful – "≤" ABR thresholds may be lower than you think (see the final part of A1 above). An alternative is to enter 30dB into the noise

calculator for the air-bone gap but note that this is only playing the percentages game and may be incorrect. Just as in conventional pure tone audiometry, there are some audiometric configurations that present a "masking dilemma" that has no satisfactory solution.

Miscellaneous

Q26. I use the "show flat line" option for the blocking period. Why shouldn't I have that on all the time?

A26. It doesn't really matter for 4kHz tests because the blocking period is short (1.25ms) but at 1kHz and 500Hz the flat line obscures quite a lot of the waveform and that can compromise interpretation. In CM tests we do want to see the stimulus artefact in order to distinguish it from any genuine CM so never use the flat line option in a CM test. The most appropriate circumstance for using the flat line option is in high level BC tests, where the stimulus artefact can be so large as to cause the waveform to span several pages or be truncated on the printout.

Q27. The Eclipse ABR system has two ways to combine a pair of waveforms – "add" and "merge". Add retains the original pair whereas merge removes them. Which is recommended?

A27. There is no strict rule on this, but it makes sense to be consistent and perhaps this could be agreed within your local peer review group. It is worth considering the implications. If this is to combine 4 runs into 2, there is usually no need to see the originals because it is equivalent to doing more sweeps, so using the merge function avoids avoidable display clutter and restricting the information available in the printout. However, if seeing the originals was important in order to visualise differences that could have consequences on interpretation, use the add function, but be aware that when the superimpose function is used, all the waveforms at that level are superimposed, which makes interpretation difficult unless they are manually separated.

Q28. If I have used inserts for the AC testing, is there any reason to swap to headphones?

A28. Yes, if you have obtained only an ***RA*** at the maximum recommended stimulus level using inserts, then consider swapping to supra-aural earphones, for which a high level is available.

References

BSA (2019a) *Assessment and Management of Auditory Neuropathy Spectrum Disorder (ANSD) in Young Infants*. Available at: https://www.thebsa.org.uk/wp-content/uploads/2019/01/OD104-85-Recommended-Procedure-Assessment-and-Management-of-ANSD-in-Young-Infants.pdf (Accessed: 5 November 2016).

BSA (2019b) *Cochlear Microphonic Testing*. Available at: https://www.thebsa.org.uk/wp-content/uploads/2019/01/OD104-87-Recommended-Procedure-for-Cochlear-Microphonic-Testing.pdf.

BSA (2019c) *Practice Guidance Principles of external peer review of auditory brainstem response (ABR) testing in babies*. Available at: www.thebsa.org.uk (Accessed: 13 September 2024).

BSA (2019d) *Recommended stimulus reference levels for ABR systems*. Available at: www.thebsa.org (Accessed: 15 March 2023).

BSA (2021) *Guidelines for the Early Audiological Assessment and Management of Babies Referred from the Newborn Hearing Screening Programme*. Available at: https://www.thebsa.org.uk/resources/.

BSA (2023) *Position Statement and Practice Guidance Audiological assessment and hearing aid provision for patients with a programmable ventriculo-peritoneal (PVP) shunt*. Available at: www.thebsa.org (Accessed: 9 October 2024).

Don, M. and Elberling, C. (1994) 'Evaluating residual background noise in human auditory brain-stem responses.', *The Journal of the Acoustical Society of America*, 96(5 Pt 1), pp. 2746–57.

Elberling, C. and Don, M. (1984) 'Quality estimation of averaged auditory brainstem responses', *Scandinavian Audiology*, 13, pp. 187–197.

ERA Training and Consultancy (2023) *Equipment-specific ABR test parameters for newborn diagnostic ABR testing*. Available at: https://www.eratraining.co.uk/resources.html (Accessed: 15 March 2023).

Ferm, I. and Lightfoot, G. (2016) 'ABR wave I presence as an alternative to masking: do NB CE-Chirps offer an advantage over tone pips?', in *Conference presentation at: Hearing Across The Lifespan. Lake Como, Italy.*

Foxe, J. J. and Stapells, D. R. (1993) 'Normal infant and adult auditory brainstem responses to bone-conducted tones', *Audiology : official organ of the International Society of Audiology*, 32(2), pp. 95–109. doi: 10.3109/00206099309071860.

IEC (2020) *IEC 60645-3 Ed. - Electroacoustics - Audiometric equipment - Part 3: Test signals of short duration*. Available at: https://webstore.ansi.org/Standards/IEC/IEC60645Ed2020?gclid=EAIaIQobChMI9Nv91vSO7wIVyu3tCh0qdAkdEAAYAiAAEgJ1WPD_BwE (Accessed: 1 March 2021).

Lightfoot, G. *et al.* (2007) 'Stimulus Repetition Rate and the Reference Levels for Clicks and Short Tone Bursts', *Am J Audiol*, 16, pp. 94–95.

Lightfoot, G. *et al.* (2014) 'The effect of notch filtering on the waveform of the newborn auditory brainstem response.', *International Journal of Audiology*, 53(9), pp. 629–32.

Lightfoot, G. (2017) 'Sloping ABR baselines and the ECG myogenic artefact.', *International journal of audiology*, 56(8), pp. 612–616. doi: 10.1080/14992027.2017.1313463.

Lightfoot, G. *et al.* (2023) 'The Clinical Evaluation of a Modified Fmp in Infant Auditory Brainstem Response Tests', *Ear and Hearing*, 44(4), pp. 924–929. doi: 10.1097/AUD.0000000000001331.

Lightfoot, G. (2024) 'Latency in threshold ABR measurements. British Society of Audiology (BSA) Audacity magazine Issue 23.', *BSA Audacity*. Available at: https://www.thebsa.org.uk/news-and-publications/audacity-magazine/editions/.

'NHSP Recommended stimulus reference levels for ABR systems, 2012' (2012).

NHSX (2021) *Records Management Code of Practice 2021*. Available at: https://transform.england.nhs.uk/media/documents/NHSX_Records_Management_CoP_V7.pdf (Accessed: 17 March 2023).

Small, S. a. and Stapells, D. R. (2008) 'Normal Ipsilateral/Contralateral Asymmetries in Infant Multiple Auditory Steady-State Responses to Air- and Bone-Conduction Stimuli', *Ear and Hearing*, 29(2), pp. 185–198. doi: 10.1097/01.aud.0000305157.83012.f4.

Small, S. a, Hatton, J. L. and Stapells, D. R. (2007) 'Effects of bone oscillator coupling method, placement location, and occlusion on bone-conduction auditory steady-state responses in infants.', *Ear and hearing*, 28(1), pp. 83–98. doi: 10.1097/01.aud.0000249787.97957.5b.

Stapells, D. and Ruben, R. . (1989) 'Auditory brain stem responses to bone-conducted tones in infants.', *The Annals of otology, rhinology, and …*, 98(12), pp. 941–949.

Webb, H. (1993) *Auditory screening in high risk neonates: An evaluation of the bone conduction ABR*. University of Sheffield, UK.