

NEWBORN HEARING SCREENING AND ASSESSMENT

**Guidance for Auditory Brainstem Response testing in  
babies**

Version 2.1

March 2013

**NHSP Clinical Group**

**Graham Sutton<sup>1</sup>, Guy Lightfoot<sup>2</sup> (Co-editors)**

**Contributors: John Stevens<sup>3</sup>, Rachel Booth<sup>4</sup>, Siobhan Brennan<sup>5</sup>, Rachel Feirn<sup>6</sup>,  
Rhys Meredith<sup>7</sup>**

1. Newborn Hearing Screening Programme Centre, London, UK
2. Dept of Medical Physics and Clinical Engineering, Royal Liverpool University Hospital, Liverpool, UK.
3. University of Sheffield, Sheffield, UK.
4. Audiology Dept, Central Manchester Hospitals, Manchester
5. Regional Department of Neurotology, Sheffield Teaching Hospitals, Sheffield, UK
6. Formerly of Children's Hearing Centre, Bristol, UK
7. Audiology Dept, Abertawe Bro Morgannwg University Health Board, Swansea, UK

**With thanks to others who made contributions to this and earlier versions, including Inga Ferm, Amanda Hall, Rob Low, Steve Mason, David Stapells, Clive Elliott, Dave Parker and Mike Vidler**

Correspondence to Newborn Hearing Screening Programme, c/o Newborn Hearing Screening Programme Centre, 344-354 Grays Inn Rd, London WC1X 8BP, UK.

ABR guidance - History:

Version	Date	Amendment History
Click ABR version 1.0	1999-2001	First version created after extensive correspondence with national and international scientists/audiologists.
Tone pip ABR Version 1.0	2001	
Bone conduction version 1.0	2001	
Click ABR version 2.0	March 2007	Comments from NHSP Clinical Group and others incorporated.
Click ABR version 2.1	April 2008	Reviewed by members of NHSP clinical group and external experts
Combined ABR guidance 1.0	October 2009	Click ABR, tone pip ABR and bone conduction ABR protocols combined to produce this guidance. Placed on NHSP website and circulated to relevant professionals for consultation
1.1	March 2010	Revision of 1.0 following consultation. It also incorporated the changes made in click ABR protocol version 2.0 (2008).
2.1	March 2013	This document – major revision and restructure

Approvals:

This document requires the following approvals:

Name	Signature	Title / Responsibility	Date	Version
NHSP Clinical Group	Graham Sutton	Chair Clinical Group	March 2013	2.1
Adrian Davis	by email	Director NHSP	March 2013	2.1

# CONTENTS

<b>1. INTRODUCTION .....</b>	<b>5</b>
<b>2. SCOPE .....</b>	<b>5</b>
<b>3. PATIENT PREPARATION.....</b>	<b>5</b>
3.1 Test Environment .....	5
3.2 Precautions against cross-infection .....	5
3.3 Choice of electrodes and application .....	5
3.4 Electrode location – AC and BC.....	6
<b>4. STIMULUS .....</b>	<b>6</b>
4.1 Stimulus and stimulus rate – air conduction and single channel bone conduction.....	6
4.2 Earphone .....	8
4.3 Warning – insert earphones .....	8
4.4 Bone vibrator .....	8
4.5 Placement of bone vibrator .....	9
4.6 Pressure to apply to bone vibrator .....	9
4.7 Effect of age on the bone conduction stimulus .....	9
<b>5. DATA COLLECTION AND ANALYSIS.....</b>	<b>9</b>
5.1 Amplifier and artefact rejection level .....	9
5.2 Blocking of stimulus artefact.....	9
5.3 Recording Filters .....	9
5.4 Use of digital filters (e.g. smoothing filter on the averaged waveform) .....	10
5.5 Notch filter .....	10
5.6 Window length and averaging .....	10
5.7 Display.....	10
5.8 Masking .....	10
5.9 Criteria for accepting the presence of a response .....	11
5.10 Resolving inconclusive results .....	13
5.11 Definition of ABR threshold for NHSP .....	14
5.12 Gold Standard thresholds.....	14
5.13 Testing at other frequencies.....	14
5.14 Reporting thresholds (including those which are not gold standard) .....	15
5.15 Baseline drift and the use of blocked-stimulus runs.....	16
5.16 Post-auricular myogenic (PAM) responses.....	16
<b>6. CALIBRATION .....</b>	<b>16</b>
<b>7. ARTEFACTS .....</b>	<b>17</b>
7.1 Recording system checks .....	17
7.2 Control recordings during testing .....	17
<b>8. GLOSSARY.....</b>	<b>18</b>
<b>APPENDIX A: Document revision history .....</b>	<b>19</b>
<b>APPENDIX B: Two-Channel BC Recording.....</b>	<b>20</b>
<b>APPENDIX C: More detailed advice and examples of ABR waveforms meeting the response criteria <i>CR</i>, <i>RA</i> and <i>Inc</i> .....</b>	<b>22</b>

**APPENDIX D: Artefact rejection level and number of sweeps per average ..... 32**

**APPENDIX E: Objective measures for ABR interpretation in babies ..... 34**

**REFERENCES ..... 38**

# 1. INTRODUCTION

This document has been prepared as part of the advice and guidance from the Newborn Hearing Screening Programme (NHSP) in England. It should be read in conjunction with the NHSP document 'Guidelines for early assessment and management of babies referred from the Newborn Hearing Screening Programme' 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 NHSP website (<http://hearing.screening.nhs.uk>) in particular the guidelines for testing for Auditory Neuropathy Spectrum Disorder (ANSD).

A summary of the changes made in each version of this guidance is given in Appendix 1. In this version a major revision and restructuring has been carried out. There are new Appendices on artefact rejection, on objective measures (Fsp and residual noise) and on 2-channel BC recording (previously in main document). The criterion for a minimum response size for a **CR** has been reduced from 50 to 40nV.

## 2. SCOPE

These guidelines are for the use of ABR in assessing hearing in babies in the first few months of life. Frequency-specific information is required<sup>a</sup>. For more details, please refer to the NHSP guidance on early assessment.

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

## 3. PATIENT PREPARATION

### 3.1 Test Environment

Threshold ABR tests should ideally be performed in a sound-proofed room or environment which meets the same standards as used in pure tone audiometry. The minimum standard should be an environment in which the lowest air conduction and bone conduction stimulus levels that are to be used (typically 10dBnHL) can be clearly heard by a normally hearing adult. Fan noise from the equipment can cause masking of stimuli at low stimulus levels: if this is the case the equipment should be sited further away from the test subject. Also, levels of electrical interference (e.g. 50Hz mains) should be sufficiently low such that the signal baseline is not adversely affected. Test rooms should not be sited close to potential sources of interference such as high powered mains equipment, transformers, or plant equipment.

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. Careful selection of the local test area or room may be necessary in order to achieve satisfactory environmental conditions.

### 3.2 Precautions against cross-infection

All local procedures should be adhered to. These should cover hygiene, use of equipment and electrodes

### 3.3 Choice of electrodes and application

The following sterile procedure is recommended. The skin should be gently and carefully abraded using a suitable sterile abrasive electrode paste and a clean gauze. An alternative is the use of a disposable abrasive pad. Disposable electrodes are recommended.

Artefact size from induced electrical interference is proportional to the difference in the electrode impedances. This difference in impedances is most easily minimised by ensuring all electrodes have low impedances. The impedance, as measured between each electrode pair should be under 5000

---

<sup>a</sup> This may include tone pip ABR, narrowband chirp ABR and ASSR.

ohms and similar across electrodes. However in good recording conditions and in a screened room higher electrode impedances can be tolerated. High impedance would also give an unacceptably large stimulus artefact at high stimulus levels; particularly for bone conduction ABR.

### 3.4 Electrode location – AC and BC

A single channel recording is recommended for AC and BC with electrodes located as follows:

- Positive electrode: **high forehead** as near to Cz<sup>b</sup> as possible and midline. The fontanelle should be avoided but the electrode should be placed as close as possible to this otherwise the ABR response 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 the positive and negative electrode connections should be reversed.

Alternative montages.

- There is some evidence that the nape of the neck rather than the mastoid gives a larger wave V response although it is reported (Stevens et al 2013a) that there is little difference in test efficiency. If the nape of the neck is used for the negative electrode, the common electrode can be placed on the forehead (at least 4cm from the positive electrode) or either mastoid, whichever is the more practical in the individual case. The mastoid should be used for the negative electrode if wave I or CM is required for neuro-diagnostic purposes.

Two-channel recording of BC may also be considered. This records the contralateral BC response in an attempt to determine which cochlea is generating the ABR. Note that this technique has limitations that should be understood if it is to be used. Details are provided in Appendix B for those interested. If this option is being considered then the common electrode can be placed on the forehead as the ipsilateral and contralateral electrodes will be used in the 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<sub>10</sub> complex which is crucial to paediatric threshold testing.

### 4.1 Stimulus and stimulus rate – air conduction and single-channel 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). Narrow band (pip-like) chirps (NBchirp) can also be used (see Elberling and Don 2010). The envelope for the rise, plateau and fall phases of the tone pip can be Blackman<sup>c</sup> or linear (Blackman is preferred). Some equipment specifies a Blackman envelope by stating the total number of cycles (this should be 5).

<sup>b</sup> Cz is the standard position used in adults. It is defined in the 10-20 electrode system for electroencephalography. For the purpose here 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).

<sup>c</sup> 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.

**TABLE 1**  
**SUMMARY OF RECOMMENDED ABR PARAMETERS<sup>d</sup>**

	Click, NBchirp & 2kHz / 4kHz tone pip/	0.5kHz / 1kHz tone pip
Electrode location <sup>e</sup>	Positive : High forehead (as close to vertex as possible but avoiding fontanelle) Negative : Ipsilateral mastoid Common : Contralateral mastoid	
Stimulus type	Alternating polarity	
Stimulus timing	Click: 100µs. Tone pip: 2-1-2 cycles (linear rise–plateau–fall) or 5-cycle Blackman	
Stimulus rate <sup>f</sup>	45.1 - 49.1/s 17.1 - 19.1/s for wave I on BC <sup>g</sup>	35.1 - 39.1/s
Calibration values for OdBnHL	Refer to NHSP calibration data	
Amplifier reject levels	±3 to ±10µV <sup>h</sup> peak-to-peak. Start at ≤ ±5µV peak-to-peak	
Amplifier filters	Low frequency: 30Hz High frequency: 1500Hz	
Window length <sup>i</sup>	20ms	25ms
Number of sweeps averaged per replication	Typically: 2000 click & NBchirp, or 3000 for TP Minimum: 1500 click & NBchirp, or 2000 for TP	
Display scales	Within range 25-100nV ≡ 1 ms See equipment specific settings.	
Display	Wave V up	

110

<sup>d</sup> Note that some equipment offer more advanced features or stimuli, not covered in this table. See the NHSP equipment-specific parameter document for details

<sup>e</sup> Alternative electrode montage: Negative – nape; Common – forehead at least 4cm from positive or negative electrodes (low forehead or to one side). Note that some manufacturers label positive and negative as active and reference respectively. Referring to common as Ground or Earth is technically incorrect; indeed it is dangerous to ground a patient. For two-channel BC see Appendix B.

<sup>f</sup> 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.

<sup>g</sup> If wave V asymmetry is being used in place of wave I presence, then stimulus rate is not reduced

<sup>h</sup> See Appendix D for note on using 10 µV rejection setting.

<sup>i</sup> 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.

115 The frequencies used for frequency-specific testing should be 0.5, 1, 2 or 4kHz.

The slower stimulus rates for 0.5kHz and 1kHz tpABR allow for longer window lengths to be used so that the full SN<sub>10</sub> part of the waveform is recorded<sup>l</sup>. A range of recommended values has been given 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 optimal parameters for NHSP work are available; 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<sup>k</sup>. 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.

125 The level of stimulus output should be checked at the start of a session (see the NHSP document on routine ‘stage A’ checks for ABR systems) and monitored by listening to the earphone at critical points during the test, particularly if unexpected results occur.

130 Stimulus levels should be recorded in dBnHL. The “nHL” can be taken to imply the use of either ISO 389-6 (2007) or NHSP-recommended calibration values (see NHSP website)<sup>l</sup>. For testing carried out under the NHSP in England, the advice in the NHSP early assessment guidelines should be adhered to.

#### 135 **4.2 Earphone**

This should be able to deliver a stimulus up to 140dB SPL peak (about 107dBnHL for a click stimulus) without distortion. TDH39/49 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. However insert earphones reduce the need for masking and attenuate ambient noise more than supra-aural earphones. If insert earphones are used, take care 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.

#### 145 **4.3 Warning – insert earphones**

Insert phones should not be used above the maximum levels given in the NHSP guidelines for early audiological assessment.<sup>m</sup> This is because a baby has a much smaller ear canal which will lead to a 10-20dB higher stimulus level 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 NHSP early assessment guidance for more detail).

#### 150 **4.4 Bone vibrator**

This should be able to deliver a stimulus up to 60dBnHL (50dBnHL at 500Hz) without obvious waveform distortion. Stimulus levels should not exceed these values unless the bone conductor has been passed in calibration as being able to deliver higher levels without distortion. The Radioear type B-71 bone vibrator should be used as calibration data are available for this and not for other types. 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.

160

---

<sup>j</sup> There is however evidence that equal test efficiency can be achieved at both lower and faster rates (Stevens et al 2013a)

<sup>k</sup> 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 & Don (2010).

<sup>l</sup> If using nonstandard stimuli, one may require local nHL values referenced to the average psychoacoustic threshold of a group of normally hearing young adults

<sup>m</sup> The only exception would be if the equipment included a microphone to automatically adjust the stimulus level for ear canal volume - but this is not yet available.



#### 4.5 Placement of bone vibrator

The bone vibrator should be placed on the mastoid, as described under electrode placement above. The bone vibrator lead should be kept away from the electrode and electrode lead. This placement gives a higher stimulus level when compared to a forehead placement in young babies (Webb 1993). A mastoid location also takes full advantage of the inter-aural attenuation which is at least 20dB for clicks in babies under 12 weeks corrected age (Webb 1993).

Placement on the temporal bone slightly posterior to the upper part of the pinna may be a good alternative where the mastoid is difficult due to proximity of the electrode (Small, Hatton & Stapells (2007).

#### 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).

#### 4.7 Effect of age on the bone conduction stimulus

The effective level of the stimulus changes with age. Please refer to the NHSP Early assessment guidance.

## 5. DATA COLLECTION AND ANALYSIS

### 5.1 Amplifier and artefact rejection level

The key to successful testing is a relaxed and sleeping baby. To ensure that unwanted electrical activity does not contaminate the recording we recommend that recording rejection levels are set between  $\pm 3$  and  $\pm 10\mu\text{V}$ . An initial value of no more than  $\pm 5\mu\text{V}$  should be set up in test protocols. If occasionally the background activity is above  $\pm 5\mu\text{V}$  for long periods, it is usually best to wait until the activity reduces. If this does not happen then the rejection level can be raised, but not generally to more than the maximum of  $\pm 10\mu\text{V}$ . Of course doing so allows more noise into the recording, requiring a substantial increase in the number of sweeps. Appendix D gives more detailed advice. Consult the equipment manuals for the value of amplifier gain/sensitivity required to achieve these rejection levels. In some equipment, for each amplifier gain/sensitivity, a range of reject levels can be set.

The use of higher rejection levels when recording conditions are difficult (e.g. high muscle activity) will lead to a poorer signal to noise ratio in the averaged signal.

### 5.2 Blocking of stimulus artefact

At high stimulus levels the stimulus artefact may exceed the artefact rejection level and so prevent recording. Some equipment will allow the artefact rejection to be ignored for a set time 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. Where separate ABR test protocols exist for each frequency then the blocking value should be set to the duration of the stimulus (1.5 ms for clicks). Refer to equipment-specific recommendations for the blocking period when using chirps.

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. See appendix C for advice on the display options for the blocking period. 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.

### 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. Higher values should be avoided; although less electrical and myogenic is recorded the response is also attenuated, making interpretation difficult.

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.

220 **5.4 Use of digital filters (e.g. smoothing filter on the averaged waveform)**  
Digital filtering in the range 50 Hz to 1000 Hz is not recommended as it is likely to result in a change in the waveform shape which makes audit of the traces 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.

225 **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 the notch filter, which may distort or attenuate the slower components of the recorded ABR waveform. 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.

230 **5.6 Window length and averaging**  
Recommended window lengths are given in Table 1 These values ensure collection of the complete waveform including the SN<sub>10</sub> component, taking into account tone pip frequency, age of baby and stimulus level.

240 The number of sweeps per waveform accepted should be varied depending on both the size of the response and the level of background activity. The aim is to achieve a clear response or response absence rating (see later). The number of sweeps required to achieve this will normally vary between 1500 and 3000, although it may be higher when the responses are small or the background noise is high. Typically a figure of 2000 sweeps is recommended (minimum of 1500) for ckABR & narrowband chirp ABR and 3000 (minimum of 2000) for tpABR.

245 Exceptionally there may be such a large ABR response or low background activity that fewer sweeps can be used (subject to a minimum of 1500 for tpABR and 1000 for ckABR & narrowband chirp ABR). Any waveform must still be judged against the full clear response (**CR**) criteria, described later.

250 In order to resolve inconclusive waveforms additional replications may be needed. To judge these, the waveforms should be combined in a pair-wise fashion, with interpretation based on a single pair of optimally superimposed waveforms.

255 More detail on the number of sweeps to use is given in appendices D & E

**5.7 Display**  
Always adhere to the convention of plotting wave V upwards. The display should be always set at a fixed number of  $\mu\text{V}$  or  $\text{nV}$  per division ( $1\mu\text{V}=1000\text{nV}$ ). 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 (0.025-0.1 $\mu\text{V}$ ) 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.

265 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 if the results are likely to be subject to peer review or re-interpretation at a later date. Information on specific scales to use for different types of equipment is available on the NHSP website.

270 Do not use an automatic display gain as this may set an inappropriate display gain for assessment of the response.

**5.8 Masking**  
As with pure tone audiometry, masking of the contralateral ear is required in certain circumstances where the stimulus level is high enough to cross to the other cochlea and produce a response. If masking is not used a crossed shadow response may mislead as to the true threshold. A masking calculator spreadsheet designed by Guy Lightfoot is available to download from the NHSP website, and this should be used to alert when masking may be needed and to calculate the level of masking

280 noise required. Details of the background and evidence used is given in an Appendix of the NHSP  
guidelines for early assessment .

### 5.9 Criteria for accepting the presence of a response

285 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

290 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 if  
the first stimulus level is 40dBnHL and a flat trace is obtained then the best use of time may be 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

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

**CR:** Clear Response present

**RA:** Response Absent \*, or

**Inc:** Inconclusive.

300 \*Note that the term **RA** was chosen for simplicity and clarity in clinical reporting.

305 The reader should make reference to Appendix C which contains more detailed advice and examples  
on this process. The rules for marking the results require two waveforms which are optimally  
superimposed (displayed on a common baseline representing the average value of all points in the  
averaged waveform). Where there are more than two recordings these should be combined to form  
two waveforms (see also later note about when a further pair of recordings have been made).

310 The interpretation process should be carried out according to the flow diagram in Figure 1.

315 Firstly does the result meet the criterion for a clear response (**CR**)?

315 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.

320 The size/amplitude of the response (as judged from the wave III/V to the following SN<sub>10</sub> trough – refer  
to appendix for examples) should be a minimum of 40nV and at least 3 times the background noise  
level (the noise level can be estimated from the average difference between the traces

325 The waveform should be judged over the whole time window excluding any stimulus artefact.

Waveforms should be compared with those at other stimulus levels (where available) to confirm that  
they follow the expected changes with stimulus level.

330 This criterion ensures a high degree of confidence of the presence of an ABR response. Examples  
showing where this criterion has been met and where it has not been met are given in Appendix C,  
The use of a weighted addition of the two waveforms as a more accurate method of measuring the  
size/amplitude of the response, is also covered in Appendix C.

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

335 For **RA** the waveforms must be appropriately flat, with no evidence of a response and the average  
difference (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).

340 This average difference (noise) criterion is designed to ensure a small response is not being obscured by noise. The waveform should ideally be flat but this is not always achieved - see the note below 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.

345 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**).

350 Waveforms may be inconclusive for a number of reasons. The most common is where there is excessive residual noise ruling out a **CR** or **RA** rating. Another is where the residual noise is low yet the criteria for **CR** are not fully met. In this case it is suggested that testing is focused at 5dB higher and 5dB lower stimulus levels where the results are likely to be **CR** and **RA**.

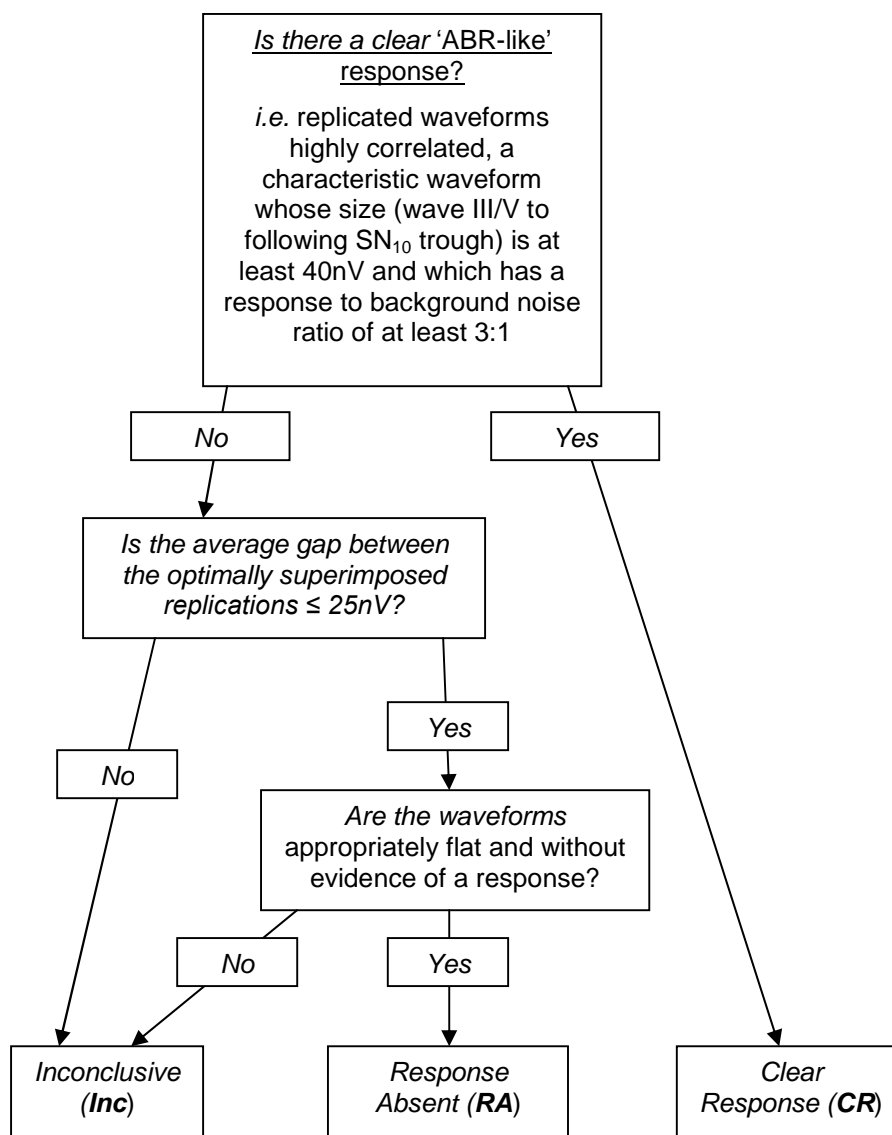
355 Inconclusive waveforms should not contribute to the derivation of threshold, and in this context that includes all unreplicated waveforms. Waveforms that appear to show a likely response or likely absence of response need not be replicated if they do not bracket the threshold, but they should only be labelled as **CR** or **RA** if replicated.

360 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 1.

365

**FIGURE 1**  
Flow chart showing summary of decision making process (see text for details)



NOTE: If the decision on threshold relies on only one very low amplitude **CR** (between 40 and 50nV), confirmatory tests should be carried out 5 or 10dB higher, where the response amplitude should be at least 50nV.

### 5.10 Resolving inconclusive results

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 is around 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 (we recommend that waveform 1 be added to waveform 3 and that waveform 2 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**.

On some occasions one may only be able to obtain only one rather than two further waveforms. In this case we recommend that waveform 1 be added to waveform 3 and superimposed with waveform 2 to test against the criteria.

### 5.11 Definition of ABR threshold for NHSP

395 ABR threshold has been defined in the NHSP early assessment guidelines<sup>1</sup> as the lowest level at which a clear response (**CR**) is present, with a response absent (**RA**) recording at a level 5 or 10dB below the threshold<sup>n</sup>, obtained under good recording conditions.

400 This is the definition that should be used for entering results onto eSP. Refer to the NHSP early assessment guidelines<sup>1</sup> for the use of ABR threshold in management.

Independent auditing of the results should not give thresholds more than 10dB different from those originally recorded.

405

### 5.12 Gold Standard thresholds

410 In order to minimise errors in estimating ABR threshold (see note below) for at least one AC ABR threshold and one BC ABR threshold (where carried out) for each ear there should be a minimum of a clear response (**CR**) at threshold

a clear response (**CR**) at 10dB or 5dB above threshold, and

415 a response absent (**RA**) at 10dB or 5dB below threshold.

The term 'gold standard' has been given to this combination of criteria for easy reference.

420 The minimum of one AC and one BC threshold in each ear to the gold standard will normally apply to one of the frequency-specific ABR thresholds, ideally 4kHz. Where it was not possible to obtain a frequency-specific threshold it should be applied to a click ABR threshold.

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

(a) where the minimum discharge criterion of  $\leq 30\text{dBHL}$  for 4kHz AC ABR is met; Gold standard is met by recording a **CR** at  $30\text{dBHL}$  and at 35 or  $40\text{dBHL}$ <sup>o</sup>. An **RA** at 10dB or 5dB below threshold is not required. If for some reason testing is carried out at other frequencies then the 'gold standard' is not required at these frequencies

430 (b) where it is not possible to test at 10dB or 5dB above threshold due to the maximum stimulus level having been reached. In this case at least one further recording should be carried out at threshold but displayed slightly above or below the superimposed pair. This allows estimation of the residual noise from a single gap between the first two waveforms yet still facilitates a comparison of response features.

435 (c) where there is **RA** at the maximum stimulus level. The **RA** should be of good quality but 2 recordings suffice

### 440 5.13 Testing at other frequencies

One should usually achieve a gold standard threshold at 4kHz before moving on to other frequencies.

<sup>n</sup> The reason for including a replicated response at 10dB or 5dB 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 is likely to 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 10dB or 5dB and prevent any serious error.

<sup>o</sup> For the derivation of the  $\text{dBHL}$  value from the  $\text{dBnHL}$  value see section entitled 'prediction of PTA from the ABR threshold' in the NHSP early assessment guidelines.

Once a gold standard has been achieved at one frequency (for at least one AC ABR threshold, and one BC ABR threshold where carried out) for each ear, this can be relaxed for other frequencies.  
445 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 10dB or 5dB above threshold (the confirmatory level).

As with the 'gold standard', waveforms should always be replicated if they are used in the definition of the ABR threshold. Some examples will help to illustrate this. If there is any doubt in the result then revert to the 'gold standard procedure'.

**Example of threshold measured to the 'gold standard'**

See Appendix C, Figure 5.

455

**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

460 The prime strategy is to test at a given stimulus level until a decision of **CR** or **RA** can be made. Sometimes this may not be possible and an alternative strategy might be to use the time to determine if the response is present at a 5dB higher stimulus level (if not already tested at this level): For example if a **CR** is obtained at 75dBnHL, an **RA** at 60dBnHL and an **Inc** result (despite very low residual noise) at 65dBnHL, it may be better to test at 70dBnHL rather than attempt to resolve the  
465 inconclusive result at 65dBnHL.

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. We cannot cherry-pick waveforms simply because they demonstrate a favoured result.

470

**5.14 Reporting thresholds (including those which are not gold standard)**

Results should be clearly marked clearly 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.

475

For example

'=45dBnHL' means **CR** at 45dB (and 5-10dB above for 'gold standard'), and **RA** at 35 or 40 dB

'<=45dBnHL' means **CR** at 45dB but not tested (or inconclusive) below this level.

480

'>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:

485

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.

b) if no **CR** is obtained above an **RA** result, report threshold > highest **RA**;

e.g. **Inc** at 70dB, **inc** at 60dB, **RA** at 50dB, report threshold as >50dBnHL.

490

c) If an **RA** response is obtained but not within 10dB of the lowest **CR**, report threshold as <=lowest **CR** and > highest **RA**;

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

495

1. 15dB or 20dB gap between lowest **CR** and highest **RA** -

e.g. **CR** at 70dB, **Inc** at 60dB, **RA** at 50dB. In eSP: enter =70dBnHL (in preference to <=70dBnHL), with a note in the session summary that the threshold is in the range 55 to 70dBnHL.

500

2. More than 20dB gap between lowest **CR** and highest **RA** -

e.g. **CR** at 70dB, **Inc** at 60dB, **RA** at 40dB. In eSP: enter <=70dBnHL, with a note in the session summary that the threshold lies in the range 45 to 70dBnHL.

505 d) If no **RA** is obtained below a **CR**, report threshold as  $\leq$ lowest **CR**;  
e.g. **CR** at 70dB, **Inc** at 60dB, **inc** at 50dB, report threshold  $\leq$ 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.

510 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

515 This term 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. A moderate amount of baseline drift is acceptable if it does not affect the ability to observe an ABR response and it cannot be taken for a false ABR response (e.g. baseline drift due to stimulus artefact should end before the first key component of any ABR response). If these rules regarding baseline drift are not met then the result should be considered as inconclusive. Where doubt exists in the possibility of a genuine  
520 response a blocked-stimulus run can sometimes help resolve the matter. Appendix C 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

525 Occasionally a post-auricular myogenic response will be recorded when the ABR response does not meet the **CR** criteria. This could occur in a baby who is not asleep or relaxed. The question arises as to whether the PAM response should be used to help estimate of the baby's hearing threshold. If the PAM response is of the expected waveform and latency and it meets the 3:1 part of the criteria used for **CR** (with peak and trough latencies adjusted accordingly) then it can be considered as a **CR**.

530 However the PAM threshold cannot be used in the same way as the ABR threshold to give the estimated hearing level (eHL). The difference between the PAM threshold and the hearing level is much more variable than for ABR. The clinical report should clearly state that the result is based on the PAM response and that it is only possible to infer that the result indicates hearing down to the  
535 lowest level at which the PAM result showed a **CR**. If the result is entered into eSP a note should be made to this effect and the eHL value produced by eSP should not be used as it is only valid for ABR.

## 6. CALIBRATION

540 A subjective 'Stage A' check should be carried out on the ABR equipment prior to use. Details of this are available on the NHSP website.

545 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 (2007). In July 2005 a provisional set of reference levels for AC and BC ABR was agreed for NHSP; the latest version of these figures, including reference threshold values for narrowband chirps, is published on the NHSP website, with the ISO389-6 values used where  
550 appropriate. These values should be used to calibrate equipment used for hearing assessment.

555 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. Bear this in mind or perform the listening check at 20 per second.

560 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 NHSP can advise which instruments require users to apply a correction for this problem. Further information on this is available on the NHSP website (Audiology homepage).



## 7. ARTEFACTS

565

### 7.1 Recording system checks

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

570

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).

575

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. Do this on a baby where clear ABR responses have been obtained at discharge levels. Set the stimulus to 30dBnHL and block the sound from reaching the ear (see next section on how to do this (30dBnHL is chosen as it may not be possible to completely acoustically block a high stimulus level). Obtain replicated traces with the stimulus still on (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.

580

### 7.2 Control recordings during testing

585

Control recordings should be carried out whenever the ABR response 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 ABR response.

590

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.*

595

For AC ABR the acoustic block can be in the form of a cover 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 if that has a much better threshold.

600

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 conduction should be covered to reduce the air born sound.

605

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.

610 **8. GLOSSARY**

	ABR	Auditory brainstem response
	AC	Air conduction
	ANSD	Auditory Neuropathy Spectrum Disorder
615	BC	Bone conduction
	ckABR	Click-evoked ABR
	CM	Cochlear microphonic
	CR	Clear Response
	dBeHL	Estimated PTA from electrophysiological thresholds
620	dBnHL	Decibels Hearing Level (the “n” is a hangover from days before an international calibration reference was available, and the scale used was derived from “nominal” or “normal” studies. nHL has been retained by convention to distinguish it from dBHL used for long duration tonal stimuli)
	ECochG	Electrocochleography
625	eSP	e-Screener Plus (Electronic record system for NHSP)
	Inc	Inconclusive
	NBchirp	Narrowband chirp
	NHSP	Newborn Hearing Screening Programme (England)
	RA	Response Absent
630	RETFL	Reference Equivalent Threshold Force Level
	RETSPL	Reference Equivalent Threshold Sound Pressure Level
	tpABR	Tone pip evoked ABR

## Appendix A

### 635 Document revision history

#### Major changes in this document (version 2.1, 2013)

This version has undergone major revision and restructuring.

640 The criterion for a minimum response size for **CR** has been reduced from 50nV to 40nV. The proviso is that where there is a very small (40—50nV) **CR** at only one level then a level 5 or 10 dB above should be tested and the response there should exceed 50nV.

We have included information on narrowband chirps and encourage their use.

645 More detailed guidance has been added on reporting of thresholds.

A paragraph on masking has been added with a cross-reference to detail in the Early Assessment guidance.

650 There have been changes to the examples in Appendix C, including changes in interpretation for some, based on the revised criteria.

655 New appendices have been added on the use of Fsp/Fmp and of residual noise values. Also on use of blocked-stimulus control runs.

A section has been added to Appendix E on the use of rarefaction and condensation subaverages instead of true replication

#### 660 Major changes in version 1 (2010)

The click, tone pip and bone conduction ABR protocols were brought together in a single document.

665 The criteria for accepting the presence of a response and establishing the ABR threshold were renamed and modified from (a), (b) and (c) to “clear response” (**CR**), “response absent” (**RA**) and “inconclusive” (**Inc**).

The method for assessing residual noise in distinguishing between a **RA** and **Inc** was changed to the same method as that used in establishing a clear response.

670 The criterion for **RA** was changed from requiring a single recording of sufficient ‘flatness’. *i.e.* ‘a peak-to-peak noise level of less than 50nV over any 4ms segment of the waveform’ to ‘replicated traces with less than 50nV residual noise (average gap)’. This was a slight relaxation from the previous version.

675 The appendix was expanded to give more examples of results and their interpretation. A section was included on how to deal with stimulus artefact and baseline drift when interpreting responses.

680 We kept the recommendation to have a ‘gold standard’ threshold for at least one AC ABR threshold and one BC ABR threshold (where carried out) for each ear. However this was relaxed for ‘other thresholds’.

685 The section on display was expanded. The previous tone pip protocol had recommended the use of two fixed scales to cover the range of sizes of ABR responses recorded. This now applies to all tests. A single summary table of recommended parameter values for all ABR test was added

#### Major changes in Click ABR version 2.1 (2008)

690 Additions/modifications to the text were made on the use of insert earphones, accepting the presence of a response, definition of threshold and some of the test parameter values (e.g. amplifier reject levels). The section on calibration referenced the separate NHSP calibration document and ISO 389-6 (2007) which gives data on reference equivalent threshold values for transient stimuli. New advice on how to assess the signal-to-noise ratio of an ABR recording was included to assist in response identification.

## Appendix B

### Two-Channel BC Recording

695

The purpose of two channel recording is to record the contralateral response in an attempt to determine which cochlea is generating the ABR. There are two approaches: either to record wave I which will be present in the recording channel using the mastoid electrode of the ear responding to the stimulus, or to look at the wave V asymmetries between the ipsilateral and contralateral recordings (reported earlier latency and larger amplitude in the ipsilateral channel - Stapells (2000)).

700

705

Note that the latter technique (use of wave V) is not 100% reliable and has limitations that should be understood if it is to be used. It may be less reliable in individuals than group data suggests. It does not identify the correct side in 100% of cases, may falsely label some unilateral conductive losses as sensorineural and although it may be useful, masking should be regarded as the definitive method.

710

Two-channel recording can be helpful if there is a masking dilemma; but note that if two channel recording reveals that the cochlea generating the ABR is not the ear of interest then masking (or ECochG) will be necessary.

715

- 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 response will be reduced in size
- Negative electrodes: ipsilateral mastoid and contralateral mastoid. Use a low mastoid as for one channel recording.
- Common electrode: forehead (at least 4cm from the positive electrode).

720

The above electrode configurations should result in wave V being plotted upwards on the display. If this is not the case then the positive and negative electrode connections should be reversed.

725

The mastoid location has been recommended rather than the nape for the negative electrode for the two channel recording as this allows for the recording of wave I, if present.

730

#### **Stimulus and stimulus rate – two-channel bone conduction**

For two-channel recording, where an attempt is being made to record wave I, the stimulus rate should be lower. A provisional rate of 19.1/s is recommended. The lower rate is not required where the latency of wave V rather than the presence of wave I is being used to help determine which cochlea is being stimulated. This lower rate should also be used if an attempt is being made to record wave I for a single-channel recording.

735

#### **Literature review**

A tabular summary review of the literature on 2-channel BC ABR prepared by Amanda Hall (Bristol) is presented on the next page.

740

**Literature summary: comparison of ipsilateral and contralateral Bone Conduction ABR waveforms in infants (thanks to Amanda Hall)**

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)	<u>500Hz</u> Mean approx 2ms later in contra <u>2kHz</u> 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	<u>500 Hz</u> Mean contra approx 0.1uV less in contra*. 94% had smaller contra response <u>2kHz</u> 93% had smaller contra response	<u>500Hz</u> Mean approx 1ms later in contra *		<u>500 Hz</u> 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 )		<u>4kHz</u> Mean 14.2 dB worse in contra ear SD 5.2 Range 10-20 N=12	34% (unclear which frequency)

\*No range or SD presented

Foxe JJ, Stapells DR. *Normal infant and adult auditory brainstem responses to bone-conducted tones.* Audiology. 1993;32(2):95-109.

745 Small SA, Stapells DR. *Normal ipsilateral /contralateral asymmetries in infant multiple auditory steady-state responses to air- and bone-conduction stimuli.* Ear Hear. 2008 Apr;29(2):185-98.

Stapells DR, Ruben RJ. *Auditory brain stem responses to bone-conducted tones in infants.* Ann OtolRhinolLaryngol. 1989 Dec;98(12 Pt 1):941-9.

## Appendix C

### 750 More detailed advice and examples of ABR waveforms meeting the response criteria CR, RA and Inc

#### 755 Example of a clear response (CR), satisfying the 3 to 1 signal to noise criterion

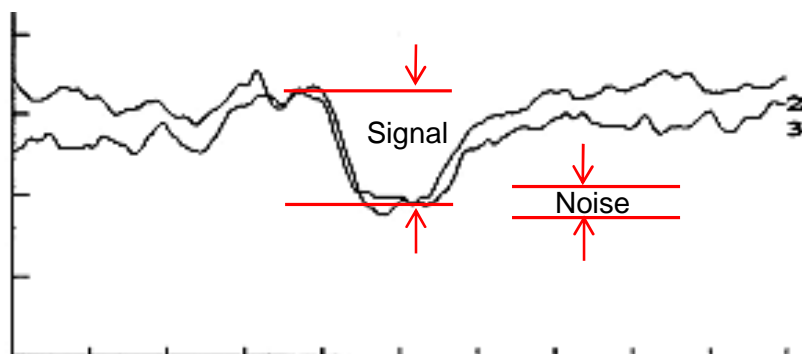


Figure 1

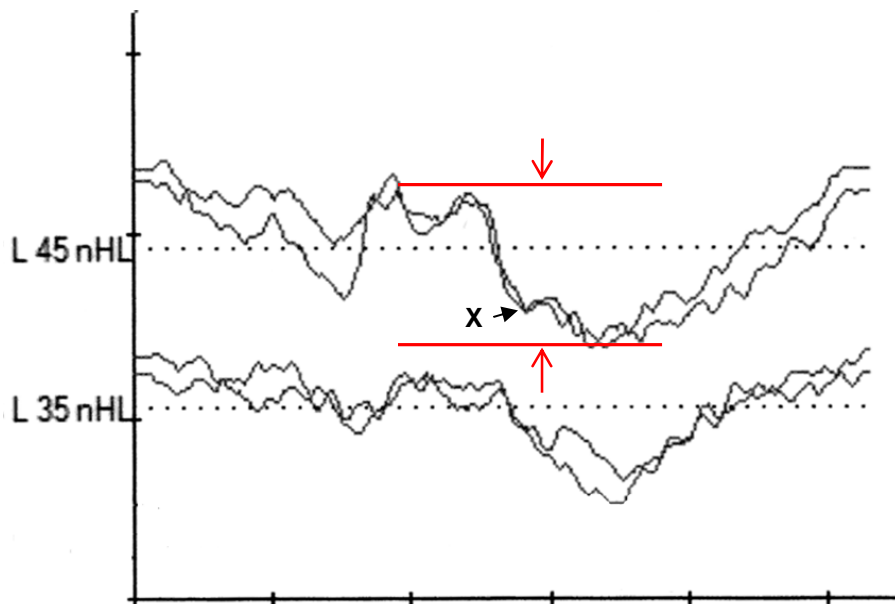
Scales:  
150nV / div  
2 ms / div

760 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 must show the expected characteristics in terms of amplitude, latency and morphology.

765 In Figure 1 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 total area between them (ignore any region containing a stimulus artefact). The “signal”<sup>p</sup> is taken as the average of the two waveforms’ wave III/V to SN<sub>10</sub> trough amplitude: in this example it is one and a third divisions (200nV or 0.2μV). 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 about a third of a division (50nV). The signal to noise ratio in this example is therefore 200/50 (working in nV) which is about four to one. 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**).

775 The points taken to measure the response amplitude are usually the wave V peak and the SN<sub>10</sub> 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 2 illustrates such a case. Finally, it is important to note that SN<sub>10</sub> is not necessarily the first dip or trough following wave V. In Figure 2 there is such a dip, marked X (at higher test levels this would probably be the trough between waves V & VI) but it is the broader trough we should take as SN<sub>10</sub> when measuring response amplitude.

<sup>p</sup> The term “signal” is used to denote the ABR response and has nothing to do with the sound stimulus. “Signal to noise ratio” can be interpreted as the “response to noise ratio”.



**Figure 2**

Scales:  
300nV / div  
4 ms / div

790

**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. This is thought to be approximately equivalent to a peak-to-peak amplitude of 70nV of a single waveform measured over a 4ms interval.

795

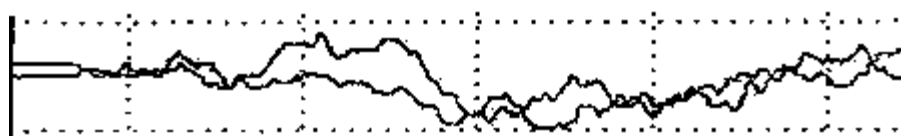
800

Figure 3 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 is <3:1 so it cannot be classified as *CR*. Note that one "response" is twice the size of the other - this is quite possibly just noise. The average gap between the waveforms is about 40nV so it fails the 25nV *RA* noise criterion too. We must therefore categorise this as inconclusive (*Inc*).

805

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 waveforms are always combined thus: 1 & 3; 2 & 4.

810



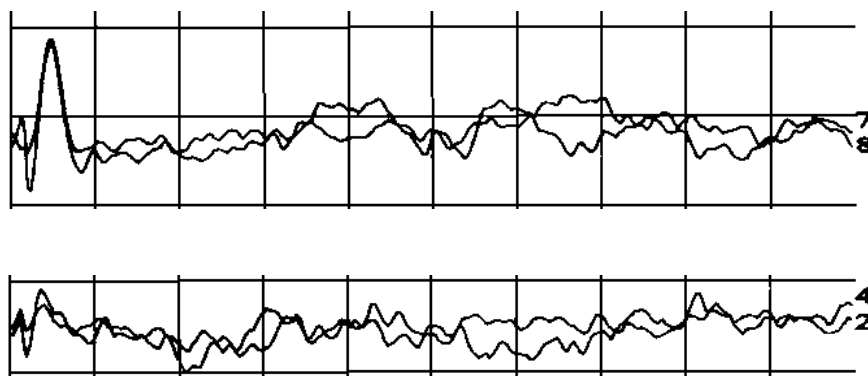
**Figure 3**

Scales:  
150nV / div  
4ms / div

815

820 Figure 4 below gives two examples of replicated waveforms that just meet the requirements for **RA**  
*i.e.* the conditions for a **CR** have not been met and the average gap between optimally superimposed  
 waveforms is less than 25nV. Note that it is the *average* gap we must estimate, excluding any region  
 of stimulus artefact. There are several points at which the gap is zero (where the waveforms cross)  
 and other points where the gap is relatively large. In both these examples the average gap is around  
 one fifth of a vertical division: (25nV).

825



**Figure 4**

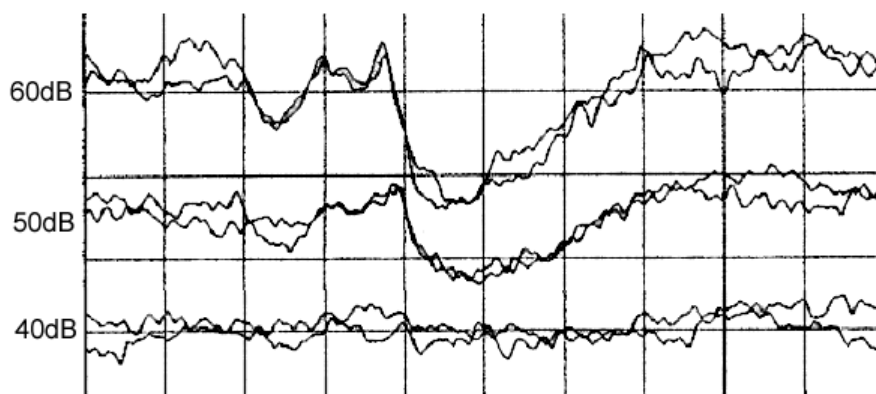
Scales:  
 120nV / div  
 2 ms / div

830 **An example of ABR threshold being found ('Gold standard')**

Figure 5 shows a correctly displayed and interpreted ABR series using 1kHz tone pips (note the  
 2.5ms / div scale giving a 25ms window). Stimulus levels are in dBnHL. The waveforms at 60dB and  
 50dB are **CR**. At 40dB there is no candidate response and the average gap between the waveforms  
 is about a fifth of a division or 24nV so it meets the maximum noise criterion and 40dB can be graded  
 as **RA**.

835

The ABR threshold is therefore 50dB and 'gold standard' since there are **CRs** at 50dB and at 5 to  
 10dB above (in this case at 60dB) and there is an **RA** at 5 to 10dB below (in this case at 40dB).



**Figure 5**

Scale:  
 120nV / div  
 2.5 ms / div

840

**The use of blocked-stimulus control runs to assist interpretation**

845 Figure 6 gives an "intensity series" for BC clicks (note the stimulus artefact that one often sees when  
 using a bone vibrator). 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**?  
 For **CR** there needs to be a high degree of correlation between the replications, a characteristic  
 waveform, a minimum response amplitude of 40nV and a signal to noise ratio of at least 3:1. At 35dB  
 850 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<sub>10</sub> 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



855 response it is not sufficiently flat to 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 control run (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 at **Inc**. [Note that in the 2010 version of this document the 30dB was accepted as **RA**]. The current definition of **RA** (requiring no evidence of a response) and the use of wave III together with the control run allows us to accept this as **CR**.

865

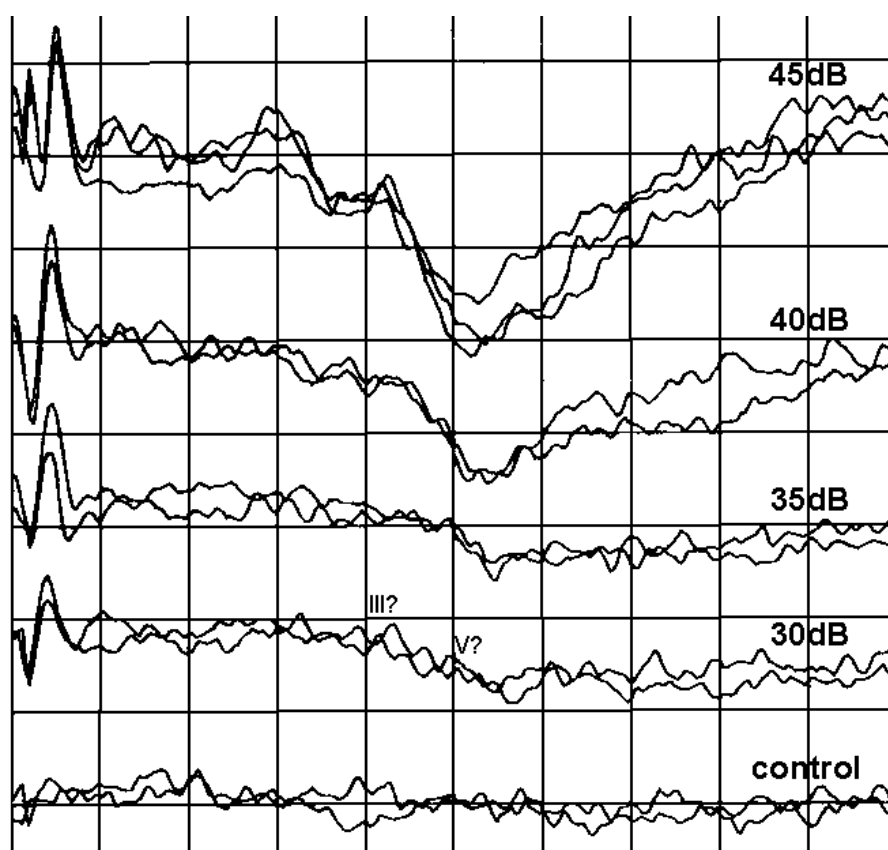


Figure 6

Scales:  
120nV / div  
2 ms / div

870 Figure 7 (4kHz AC) shows a **CR** in the upper waveforms but interpretation of the lower waveforms (at 10dB lower stimulus level) is more challenging. There is a high degree of correlation between the replications but the expected characteristics are not met in terms of latency as required under **CR** criteria. 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.

875 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 control run may have resolved this case.

880

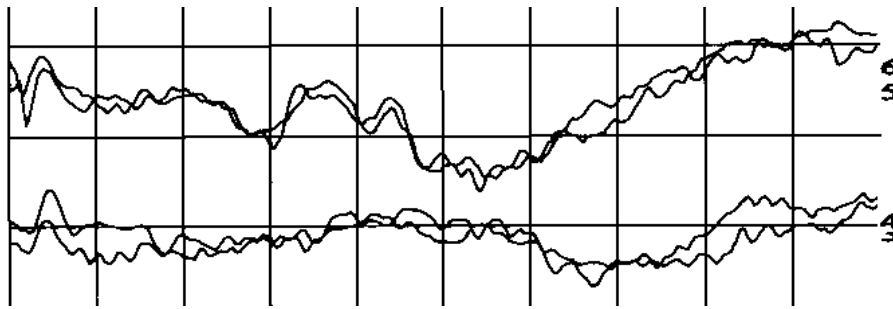


Figure 7

Scales:  
120nV / div  
2 ms / div

885

**An example where the response is 'flat' enough to accept the conclusion of RA**

Figure 8 shows **CRs** 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?  
 890 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 instrument or mains-related. All suggestion of wave V and SN<sub>10</sub> have disappeared. Unlike the case in figure 7, this example is sufficiently flat in this region and we can accept it as **RA**. A blocked-stimulus run could have been used to increase confidence if required. The threshold is therefore =45dBnHL.

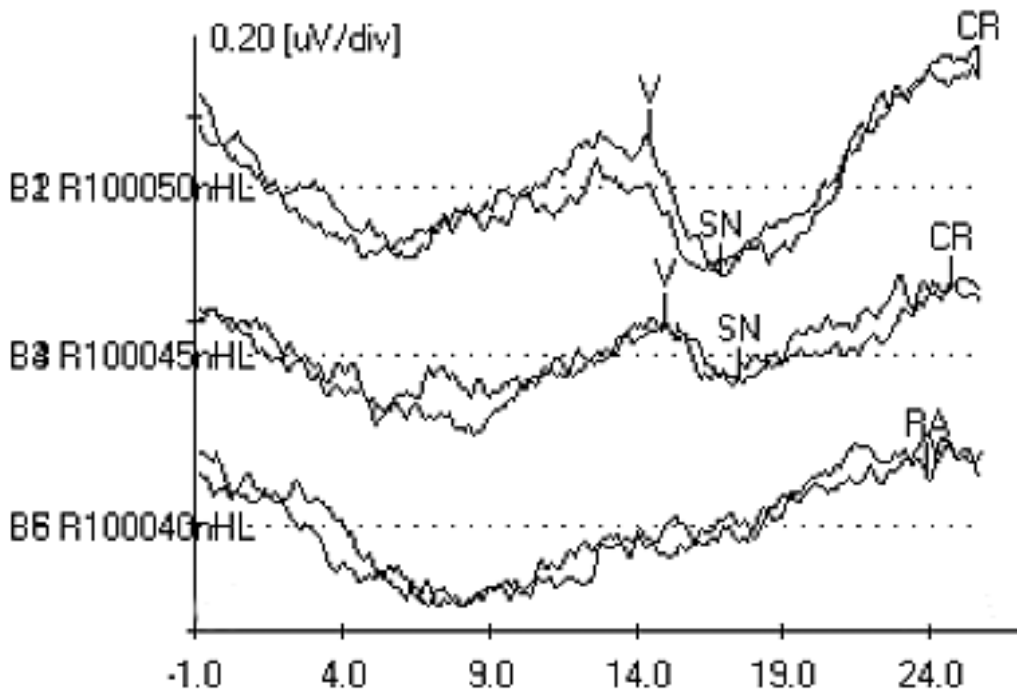


Figure 8

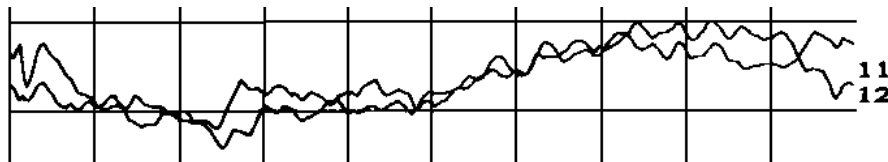
Scales:  
200nV / div  
5 ms / div

895

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

'Non-flat' Example 1

900 Figure 9 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 control recording to see if the baseline drift remains. See also example 6.

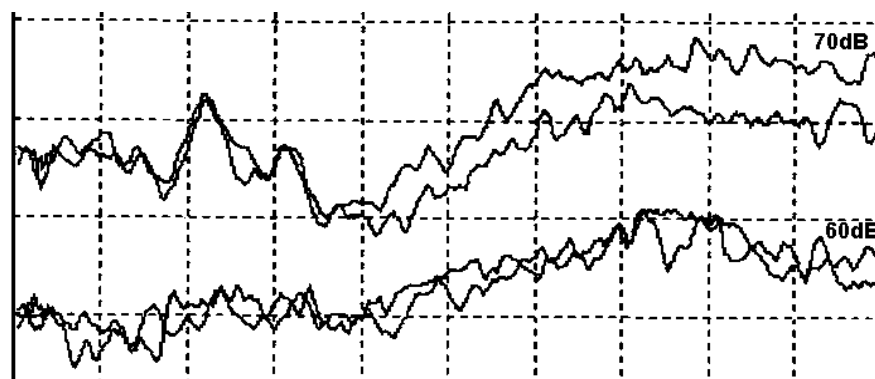


**Figure 9**

Scales:  
120nV / div  
2 ms / div

'Non-flat' Example 2

910 In figure 10 the AC 4kHz waveforms marked at 70dB (stimulus levels are in dBnHL) 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<sub>10</sub>. That being the case it is obvious that at 70dB there is a **CR**. 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. It is difficult to know what the cause of this baseline drift is. It could be due to the equipment, residual mains artefact time locked to the stimulus or even possibly a physiological response (but very un-ABR like). 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.

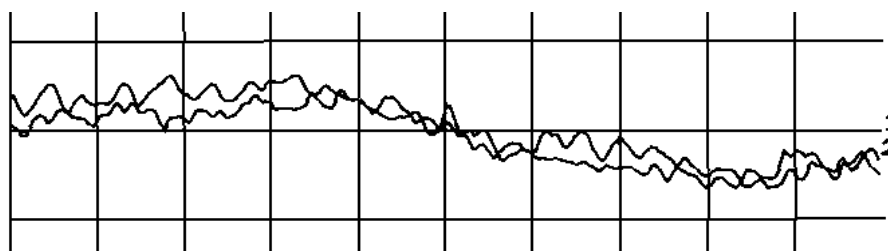


**Figure 10**

Scales:  
120nV / div  
2.5 ms / div

'Non-flat' Example 3

925 A further example of **RA** with 'baseline drift' is illustrated in Figure 11. 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.



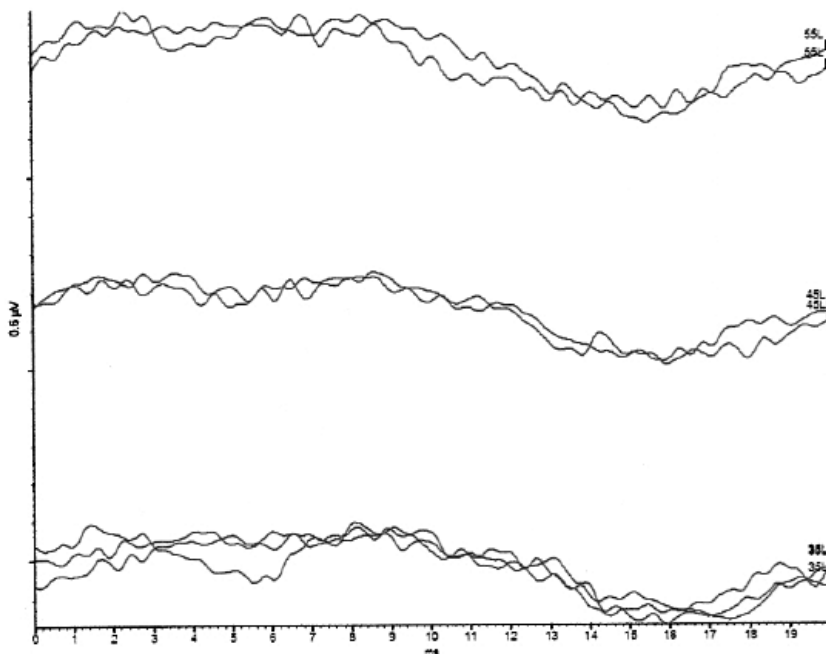
**Figure 11**

Scales:  
120nV / div  
2 ms / div

935 'Non-flat' Example 4

The waveforms of Figure 12 might be accepted as clear responses because they have an 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. Tests at a higher level and a blocked-stimulus control run (producing flat traces) would have helped resolve this case.

940



**Figure 12**

Scale:  
100nV / minor div  
1 ms / div

945

'Non-flat' Example 5

BC waveforms are illustrated in figure 13, 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 probably 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 reasonable to interpret 40dB as **CR** and 30dB as **Inc**, reporting the result as  $\leq 40$ dB. This example is very challenging but fortunately examples like this are not very common in clinical practice!

950

955

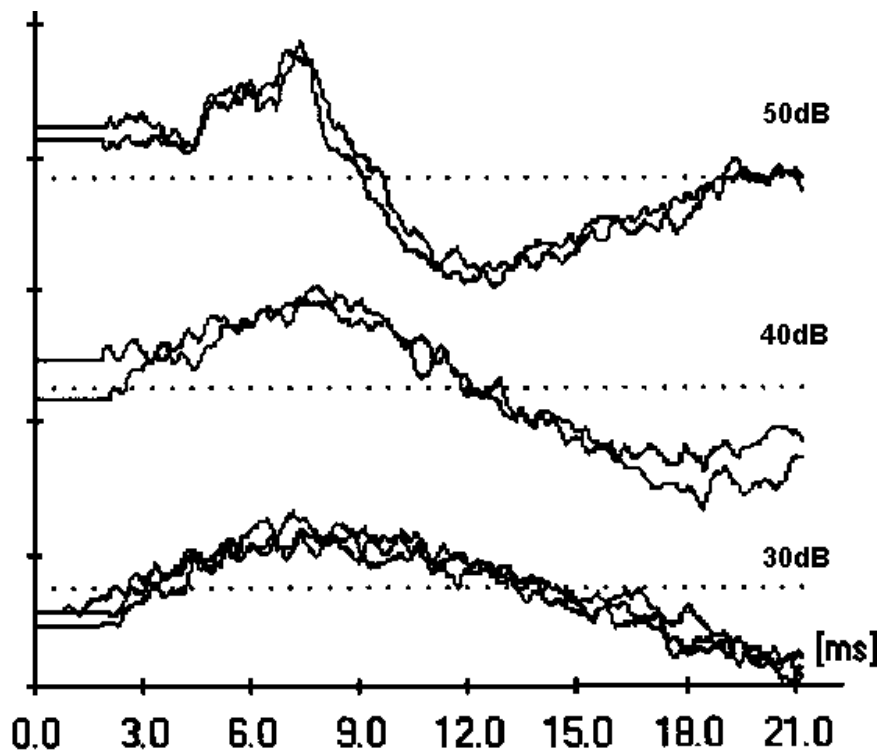


Figure 13

Scale:  
200nV / div  
3 ms / div

#### The use of A+B & A-B facilities

960 Most ABR system offer the facility to *add*, *merge* or *combine* waveforms (often referred to as A+B) and *subtract* waveforms (referred to as A-B or *difference*). Both can be “weighted” or “unweighted” calculations, which lead to different results. A weighted addition/subtraction scales the contribution of each averaged waveform to the sum by the number of sweeps in each average. An unweighted addition/subtraction does not use this scaling. **It is very important that the correct type (i.e. weighted or unweighted) is used as specified in the following advice .**

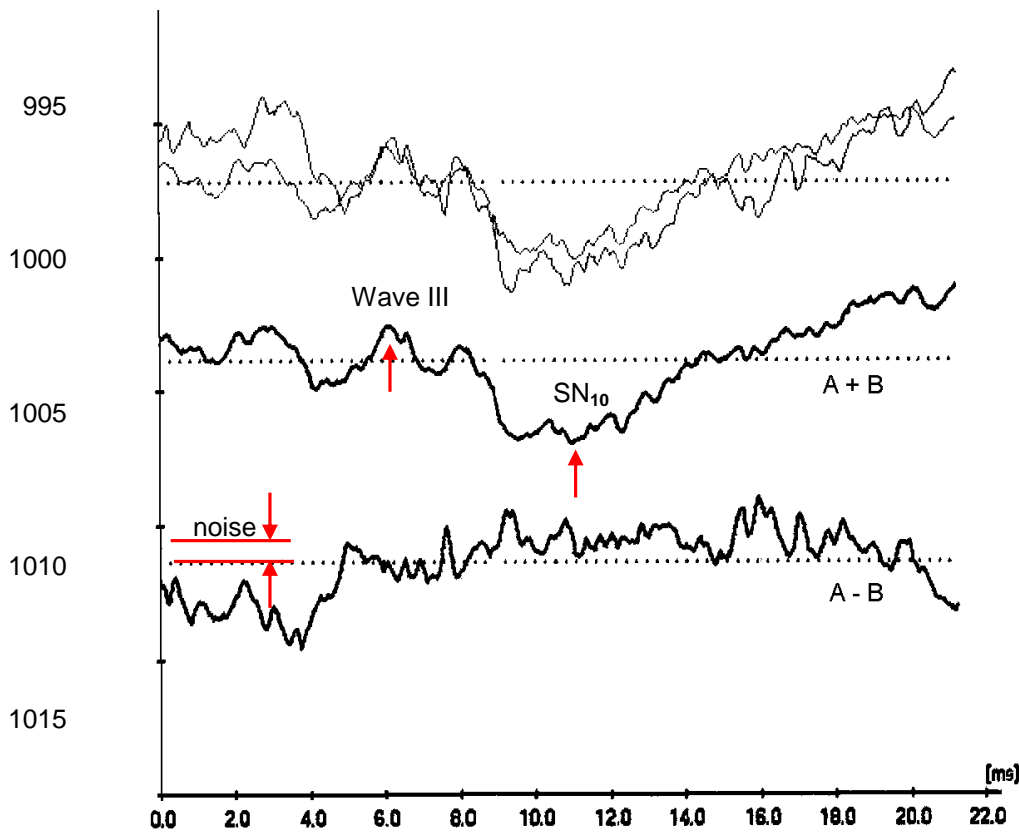
The addition of waveforms using a **weighted** A+B function can be useful when assessing the size of the response.

970 The use of **unweighted** A-B offers an alternative method for assessing the magnitude of residual noise as used in **CR** and **RA**. The two measurements are equivalent and the same criterion of  $0.025\mu\text{V}$  (25nV) is used for **RA**. This is not considered to be any easier than the method described in the main text on estimating the average gap between waveforms as it still requires the average of a waveform to be visually estimated and ,as decisions need to be made quickly whilst testing, may use up valuable test time. However the process is now described for those who wish to use this method.

#### 'A+B/A-B' Example 1

980 Figure 14 (top) illustrates a pair of waveforms recorded to 55dBnHL clicks that are to be assessed. The requirements for **CR** are met in terms of a characteristic waveform but we need to be sure the signal to noise ratio is at least 3:1 and that the response is  $>0.05\mu\text{V}$  (50nV). The middle waveform is the weighted addition (A+B) of the two top waveforms. In this case wave III is more positive (higher) than wave V so we can use the wave III –  $\text{SN}_{10}$  amplitude to represent the size of the response, which is 7/8 of a division or 175nV. The lower waveform is the unweighted difference (A-B). This represents the residual noise in the two upper waveforms and to assess the extent of the noise we must estimate the average height of the waveform above and below the baseline, averaged across the entire width of the trace. This is shown in the figure and is about 1/6<sup>th</sup> division, or 30nV. This is the same as the average gap between the original waveforms. The signal to noise ratio is therefore 175/30 or just under 6:1, easily exceeding the minimum 3:1 requirement. We are therefore able to classify this as a clear response (**CR**).

990

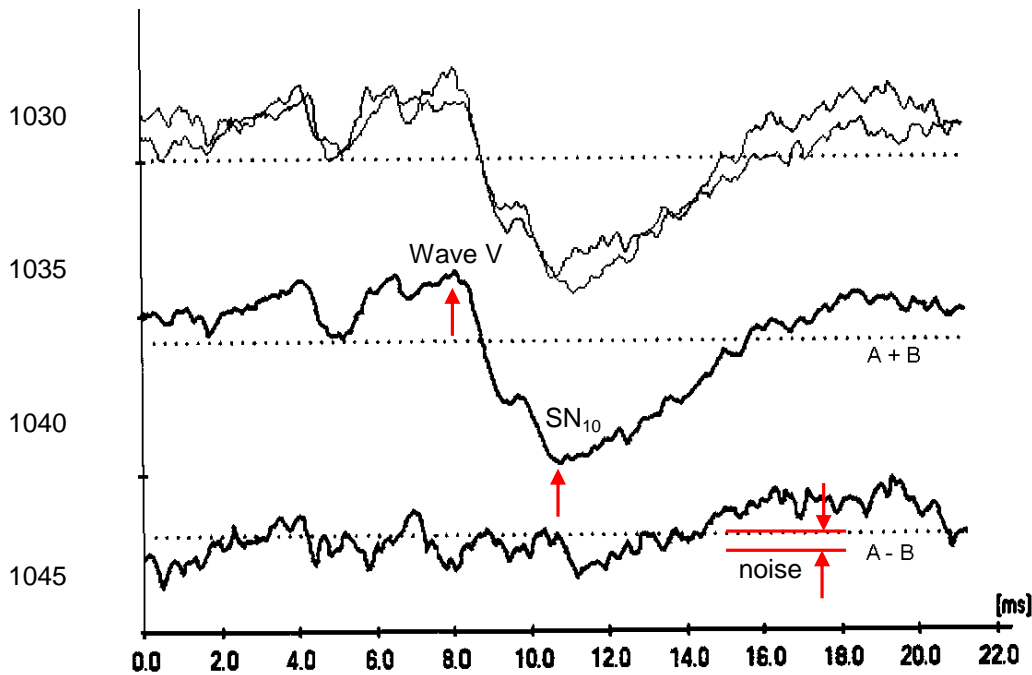


1020

'A+B/A-B' Example 2

The waveforms in Figure 15 (top) were evoked by 4kHz tone pips at 40dBnHL and there is no question they can be classified as a **CR**. The response (middle waveform, weighted A+B) is 1.3 divisions so is 260nV. The noise (lower waveform, unweighted A-B) is estimated as shown, at 1/8<sup>th</sup> division or 25nV. The signal to noise ratio is therefore over 10:1.

1025



1050

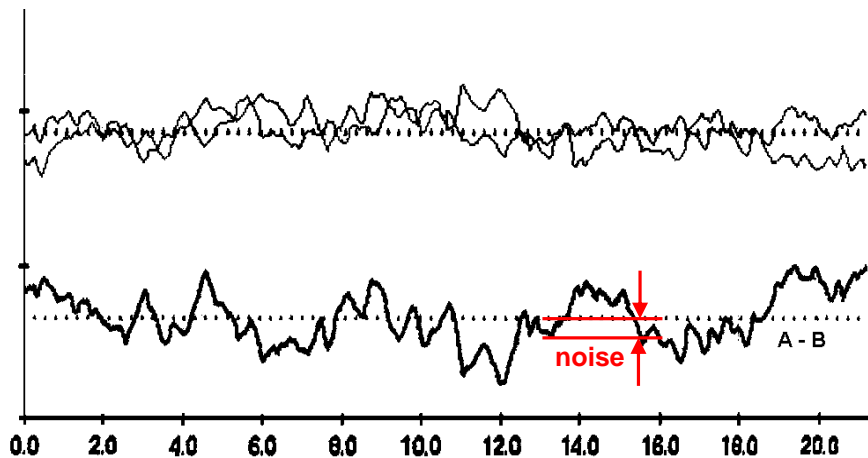
**Figure 15**

Scales:  
200nV / div  
2 ms / div

### 'A+B/A-B' Example 3

1055

The residual noise in the upper waveforms of Figure 16 needs to be less than 25nV to qualify for **RA** status otherwise we cannot be certain that the noise is not obscuring a small response. The lower waveform is the unweighted A-B and its average value (and the average gap between the upper waveforms) is again about 1/8<sup>th</sup> division or 25nV. It therefore just qualifies for **RA** status.



**Figure 16**

Scales:  
200nV / div  
2 ms / div

1060

### The use of rarefaction and condensation sub-averages in lieu of true replication

1065

When alternating polarity stimuli are used, some ABR systems allow the separate display of the rarefaction and condensation waveforms. The question arises: can these separate polarity sub-averages be used as an alternative to conventional replication? This would be valid only if there is no observable difference in response morphology to these separate polarity stimuli. Provisional data from Rhys Meredith using near-threshold very low noise waveforms suggests this is indeed the case and response morphology does not appear to be influenced by stimulus polarity at near-threshold levels. The use of separate polarity sub-averages is therefore considered acceptable if they are superimposed and analysed as suggested for true replications and providing that each sub-average contains the number of sweeps suggested in section 5.6. If differences in response morphology or large stimulus artefacts are encountered then the acquisition and display of replicated alternating polarity waveforms is necessary.

1070

1075

### Advice on the use of waveform display options for the “blocking” period

1080

Some ABR equipment offer the facility whereby the latency range over which the artefact rejection system is disabled during stimulus presentation (commonly referred to as the “blocking” period) can be set to either display the recorded waveform or display a flat line. 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 when printed. The inappropriate use of the ‘flat line’ option can hamper correct interpretation in CM testing and when using stimuli below 2kHz.

## 1085 **Appendix D**

### **Artefact rejection level and number of sweeps per average**

1090 This appendix aims to provide testers with an insight of the relationship between the artefact rejection (AR) level and the number of sweeps needed in the waveforms to reach a satisfactory signal-to-noise ratio (SNR). This knowledge should guide testers' choice of AR level and sweeps - something that may need to be adjusted during the course of testing in response to changes in test conditions. This advice applies only to babies tested before 12 weeks corrected age.

1095 In the following, reference to "signal" refers to the ABR we want to record whereas "noise" refers to everything else, which can be a combination of EEG, ECG, muscle activity and other electrical interference<sup>q</sup>. A large ABR can be around 0.5µV whereas near threshold it can be 0.1µV or even less (the minimum acceptable amplitude is taken as 0.04µV). The noise can be as little as 2µV in a sleeping baby (EEG) to over 20µV (muscle activity). This means that the signal we want to record is far smaller than the competing noise and in order to correctly identify the presence of the ABR we have to use an averaging technique in order to reduce the noise so that it is a fraction of the ABR.

1100 Averaging improves the SNR by the square root of the number of sweeps. For a 10-fold improvement in SNR we need 100 sweeps; for a 100-fold improvement we need 10,000 sweeps. The number of sweeps needed depends on the relative sizes of the response and the noise. We should 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.

1110 Artefact rejection simply rejects sweeps (stops them from being added to the average) if the peak-to-peak amplitude of the activity in the sweep is more than a defined level: the artefact rejection level. This ensures that data are collected only when the instantaneous test conditions are favourable. The NHSP recommended range for AR is ±3 to ±10µV with a default starting level of no more than ±5µV (±3µV for CM tests) so these values should be used when setting up test protocols. Some equipment allows the AR level to be changed in an interactive fashion, whilst averaging; in others it is more cumbersome and time consuming. Bayesian averaging (see below) should be used if available.

1115 In non-ideal recording conditions the background activity may be above ±5µV most of the time, leading to a high proportion of sweeps being rejected. This is usually caused by the baby's muscle activity. Since bursts of muscle activity are often short lived it is usually best to wait for the baby to settle. However if this does not happen then the rejection level can be raised but since this allows higher amplitude noise into the recording, we must increase the number of sweeps if we are to achieve the necessary SNR in the final recording.

1120 An example will help. We will assume a satisfactory SNR and a **CR** is achieved using 3000 sweeps and a rejection level of ±5µV with a modest number of rejections. If the baby becomes slightly restless and we choose to raise the rejection level to ±7µV, how many sweeps are needed to achieve the same SNR? The answer can be as much as  $3000 \times (7\mu\text{V} / 5\mu\text{V})^2$  or 5880, *i.e.* roughly double the original 3000. This is the worst case scenario but Lightfoot & Stevens (2013) report that 4500 to 6000 will be needed to achieve the same SNR. Similarly, changing from ±5µV to ±10µV means using 6000 to 12000 sweeps. The cost of increasing the rejection level therefore carries a heavy penalty. Another study (Stevens *et al*, 2013a) also provides evidence that use of ±5µV is more test efficient than ±10µV reject level.

1125 Testers need to know what to do in clinic. In non-ideal (but not impossible) test conditions which of the following is most efficient?

1135

- Stick to ±5µV (but many sweeps will be rejected, extending test time) or
- Relax the rejection level (allowing more noise into the average and use many more sweeps to deal with this).

1140 The recent study by Lightfoot & Stevens (2013) tried to identify the best clinical strategy to use. The conclusion was that ±5µV should be used so long as no more than around 30%<sup>r</sup> of sweeps were

<sup>q</sup> Separate advice is available to help testers identify the sources and minimize the effects of electrical noise.

<sup>r</sup> A 33% rejection rate is where the number of rejected sweeps is half the number of accepted sweeps.



rejected. If 30% is exceeded, pause to see if the baby will settle. If the baby does not settle but test conditions are not impossible increase to  $\pm 7\mu\text{V}$  but expect to use 4500 - 6000 sweeps. Doing this will reduce the rejection rate. If the rejection rate falls below, say, 10% because the baby has settled then return to  $\pm 5\mu\text{V}$ . If the baby is very restless and substantially more than 30% rejections occur at  $\pm 7\mu\text{V}$  stop the test and wait for the baby to settle. Exceptionally, testing may not be possible without raising the AR level up to  $\pm 10\mu\text{V}$  but it should be realised that it will be very difficult to record an acceptable ABR even with very high numbers of sweeps.

1145

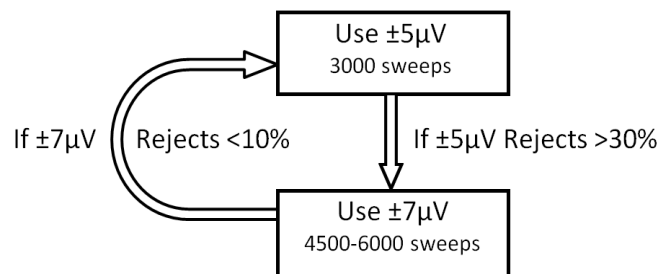
The speed or frequency at which these AR changes are made depends to a large extent on the facility to make AR level changes during averaging and whether the instantaneous reject rate can be easily estimated. Users of equipment that is unable to do this must balance any benefit from changing the AR level with time lost in making an AR change.

1150

Appendix E deals with objective measures that can guide us in our decision of when to stop averaging but if these are not available or not used then the numbers of sweeps given above should be our guide, as well as observation of the averaged waveform.

1155

This strategy is summarised in Fig 1.



1160

Fig 1 A proposed artefact rejection / sweeps strategy

### Bayesian Weighted Averaging

This technique is available on some ABR systems. Rather than a fixed artefact rejection level, the idea is to use a more lenient level but allow the software to place more importance on sweeps with a low noise content and less importance on noisier sweeps. This waters down the destructive effects of the noise and can be thought of as “shades of grey” rather than the “black and white” technique of simple artefact rejection. The use of Bayesian averaging combined with an artefact rejection level of  $\pm 10\mu\text{V}$  performs well and is certainly easier than implementing the above interactive strategy.

1165

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  $\pm 10\mu\text{V}$  to  $\pm 7\mu\text{V}$  or even  $\pm 5\mu\text{V}$  will exclude this source of noise and be advantageous.

1170

However we need to know how many sweeps are needed to achieve an acceptable SNR. The use of an objective response presence confidence measure such as Fmp is appropriate as a guide when a response is present whilst an objective residual noise figure is appropriate as a guide when a response is absent (see appendix E). Both require accurate tester judgement to confirm that all requirements of **CR** and **RA** are satisfied.

1175

### Advice

**Use  $\pm 5\mu\text{V}$  artefact rejection level and typically 3000 sweeps as the starting point and use a strategy based on Fig 1, including an increase in the number of sweeps as suggested.**

1180

**If Bayesian weighted averaging is available together with objective measures of response confidence and residual noise then use Bayesian averaging, with an artefact rejection level of no more than  $\pm 10\mu\text{V}$ ; less if cardiac activity is recorded.**

1185

## 1190 **Appendix E**

### **Objective measures for ABR interpretation in babies**

The NHSP clinical group has considered the availability and reliability of objective measures for ABR response detection and the measurement of residual noise on current commercial ABR equipment. 1195 This appendix contains some provisional advice on the use of response confidence measures (Fsp in the Natus Biologic Nav Pro EP system; Fmp in the Interacoustics Eclipse system) and residual noise measures. These assist the user decide whether and when sufficient sweeps have been acquired., rather than using a simple fixed number of sweeps approach.

#### 1200 **1. Use of response confidence measures (Fsp / Fmp) in ABR testing in babies**

##### **1.1 Introduction**

Fsp is one of a number of measures available to determine the degree of confidence in the presence of an ABR response (Elberling & 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 response (if present) plus any residual noise. The higher the Fsp value the greater the ABR response compared to the background noise and the greater the confidence of a clear response. Fmp (Don & Elberling, 1994) is a slightly more sophisticated (multiple point) version of Fsp. 1205

In the NHSP Guidance for ABR testing in babies, the standard measure used to estimate this confidence is a visual estimate of the signal-to-noise ratio<sup>s</sup> (SNR) which has to reach a value of 3:1 as one of the conditions for a clear response. Note that there are other conditions, 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. 1210

However at this stage it is still recommended that the SNR is estimated by visual inspection of a pair of replicated waveforms, using the Fsp or Fmp value as supporting information and as a guide to when to stop averaging. The waveforms should not be interpreted using Fsp or Fmp alone. 1215

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. 1220

##### **1.2 Validation of Fsp and Fmp**

In order to validate the Fsp & Fmp technique and confirm that the measurement agrees with the theoretical value we need to obtain Fsp and Fmp values from a number of babies when no stimulus is present. Two studies were undertaken. For Fsp John Stevens & Siobhan Brennan (Sheffield) collected 42 no-stimulus tests from a group of 19 babies using the Biologic Nav Pro ER system. For Fmp, Inga Ferm (Croydon) collected 47 no-stimulus tests in a group of 40 babies using the Interacoustics Eclipse system. Figure 1 shows the cumulative distributions for Fsp and Fmp in these studies. The theory based on an F distribution of F(5,250) suggests that the value of these statistics should be 0.87 at the 50% point. The results of the studies gave a lower than expected Fmp value for the Eclipse system whilst those for the Biologic Nav Pro ER system were higher than expected. The origin of these differences is being investigated. The two measurements agree very well and have the closest match to F(5,250) if Fsp is multiplied by 0.5 and Fmp is multiplied by 1.4 (figure 2). 1230

The two distributions are very similar and a reasonable match to the F(5,250) distribution when Fsp and Fmp are scaled as suggested above. 1235

---

<sup>s</sup> Because of the way in which this is estimated, the NHSP signal-to-noise ratio is not a true SNR; the NHSP 3:1 probably corresponds to a true SNR of around 1.5 1240

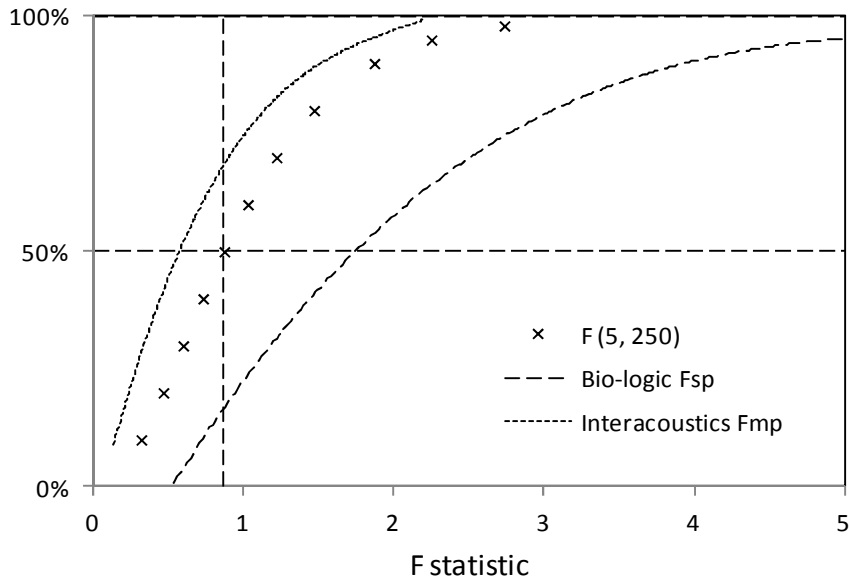


Figure 1 Cumulative distributions for Fsp and Fmp together with the cumulative distribution of F(5, 250). Axes: vertical: proportion of cases; horizontal: F statistic

1245

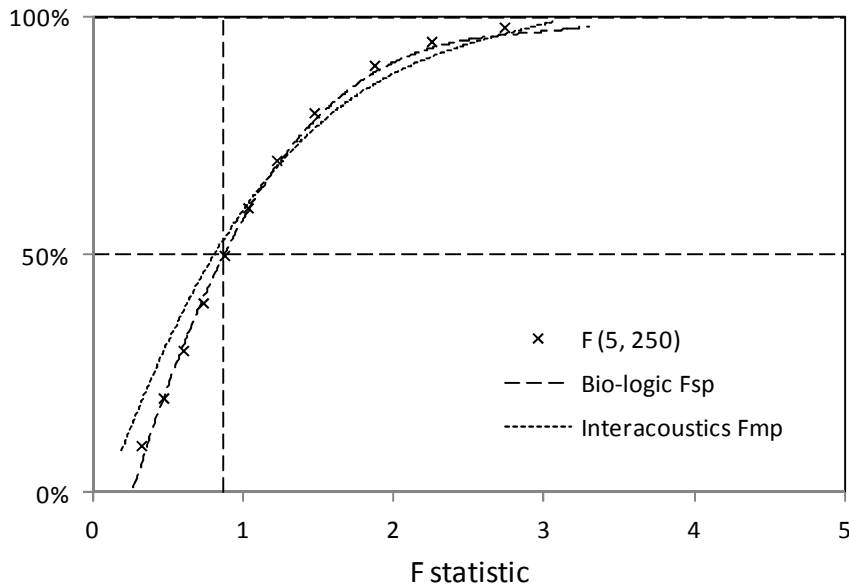


Figure 2 Cumulative distributions after Fsp values are multiplied by 0.5 and Fmp values are multiplied by 1.4. Axes: vertical: proportion of cases; horizontal: F statistic

1250

Based on an F(5,250) distribution, 2.25 corresponds to a 95% confidence in the presence of a response and values of 2.65 and 3.1 correspond to 97.5% & 99% respectively. Although the rescaled plots of the measured Fsp and Fmp shown in Figure 2 closely match this distribution there are limitations in the measured data that limit the direct use of these values. Two of these are (i) the uncertainty due to the relatively small sample size in the measured values and (ii) the underestimation of the low frequency residual noise in the numerator of the Fsp value caused by the relationship between the low frequency filter and the window length. This is difficult to resolve until more studies are undertaken and these are in hand. As an interim measure it is important to adopt a cautious approach and so the 95% confidence limits for the estimate of the mean of the measured Fsp and Fmp values have been used to estimate the confidence in the scaling factors derived from the two studies (1.4 & 0.5).

1255

1260

The results lead to scaled values for the F(5,250) statistic at the 97.5% confidence level:

1265

Interacoustics Eclipse:  $F_{mp}^{97.5\%} = 2.37$

Biologic Nav Pro:  $F_{sp}^{97.5\%} = 6.8$

1270

The scaled values related to the 97.5% confidence level have been chosen, with rounding up to the nearest 0.5.

1275

Occasionally a sloping baseline or low-frequency component is recorded; when this occurred in the no-stimulus trials a high (>10) value of Fsp was recorded. This problem is rare for Fmp since the Eclipse applies a 100Hz high-pass filter to the waveforms from which the calculations are made (not to the waveforms seen on screen) as advised by Elberling & Don, 1984. The Nav Pro does not do this and the user must remain vigilant for suspiciously high values of Fsp in waveforms that do not appear to contain waveform features suggestive of a genuine response. Figures 12 & 13 in appendix C are good examples of waveforms that could have misleadingly high values of Fsp.

1280

### 1.3 Advice on use of confidence measures

1285

**Fmp >2.5 for the Eclipse, and Fsp >7.0 for the Nav Pro  
in both replicated waveforms using the NHSP recommended settings  
support the conclusion that the response probably exceeds the 3:1 condition component of  
the NHSP CR criteria.**

1290

**A CR decision still requires that the conditions detailed in section 5.9 have been met. The  
principal use of Fmp and Fsp is in guiding the user when sufficient sweeps have been  
acquired.**

Note that it is possible to obtain clear responses (**CR**, SNR >3:1) with Fsp and Fmp less than these values in one or both waveforms.

1295

Important Notes:

- Low values of Fsp or Fmp cannot be used to imply response absence.
- High values of Fsp or Fmp should be accepted only if the waveforms appear to contain classic features of an ABR and not just a low-frequency or sloping baseline.
- If the stimulus artefact is close to the Fsp / Fmp window or there is any other artefact in the waveform the Fsp / Fmp result should not be used.

1300

Check that the NHSP protocols for the Biologic Nav Pro and Interacoustics Eclipse are used and that they have the settings shown below for Fsp / Fmp.

1305

*Current settings for the Fsp & Fmp range:*

NB. These values may be already entered in the test protocols available from Biosense Medical and Interacoustics UK for the diagnostic ABR testing of babies referred under the English NHSP programme.

Stimulus	Nav Pro Fsp range	Eclipse Fmp range
Click	6 to 14ms	5 to 15ms
4kHz tone pip	6 to 14ms	5 to 15ms
2kHz tone pip	8 to 16ms	7 to 17ms
1kHz tone pip	11 to 19ms	10 to 20ms
500 Hz tone pip	13 to 21ms	10 to 20ms
Chirps (all types)	N/A	5 to 15ms

1310

Note:

Enabling the Fsp function on the Biologic Nav Pro results in changes in the way the system acquires and displays data. The display is refreshed every 256 sweeps and upon completion of an average, the number of sweeps used is rounded down to an integer multiple of 256. When testing it is

1315 therefore recommended that tests are stopped just after a multiple of 256 sweeps *i.e.* at 1536, 1792, 2048, 2304, 2560, 2816, 3072, 3328, 3584, 3840 or 4096 sweeps.

The figures for the number of sweeps in the current NHSP guidance for ABR testing in babies become typically 2048 (minimum of 1536) for click ABR and typically 3072 (minimum of 2048) for tone pip ABR. It is also recommended that the maximum numbers of sweeps, set in the protocols, are 4096 for tone pips and 3072 for clicks to allow for extended averaging in non-ideal conditions.

1320

## 2. Residual noise

### 2.1 Background

1325 Residual noise, when carrying out an ABR test in a baby, 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 4 $\mu$ V peak to peak or about 0.75 to 1.5 $\mu$ V 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  
1330 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 response vary considerably. In babies with a large ABR response the level of residual noise required to achieve a clear response will not be as low as for babies with a small ABR response.  
1335

The way in which ABR equipment calculates residual noise is not the same as the NHSP “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 NHSP 25nV average gap.  
1340

### 2.2 Comparison of equipment residual noise to NHSP average gap measures

1345 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 NHSP “average gap” **RA** criterion of 25nV. When combined the studies concluded that these are 18nV for the Nav Pro and 30nV for the Eclipse. 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 NHSP 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.  
1350

### 2.3 Advice

#### Use of residual noise measures in determining **RA**

1355 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 two types of equipment as follows.

1360	<b>Biologic Nav Pro</b>	<b>15nV</b>
	<b>Interacoustics Eclipse</b>	<b>25nV</b>

1365 Even lower values of residual noise will make achievement of the NHSP **RA** noise criterion more secure.

NB. Replicated waveforms will still have to meet the 25nV average gap criterion and the other **RA** criteria in section 5.9 and notes on baseline drift.

1370 **REFERENCES**

- Don M.& Elberling C (1994). *Evaluating residual noise in human auditory brainstem responses*. J Acoust Soc Am. 96: 2746 – 2757.
- 1375 Elberling C.& Don M. (1984). *Quality estimation of averaged auditory brainstem responses*. Scand Audiol. 13: 187-197.
- Elberling C, and Don M (2010). *A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses*. J Acoust Soc Am 128, 2955–2964.
- 1380 IEC 60645-3 (2007). *Electroacoustics – Audiometric equipment – Part 3: Test signals of short duration*. Geneva, Switzerland: International Electrotechnical Commission.
- 1385 Newborn Hearing Screening Programme (England). *Routine (stage A) checks for ABR systems*. <http://hearing.screening.nhs.uk/calibration>
- ISO 389-6 (2007). *Acoustics - Reference zero for the calibration of audiometric equipment - Part 6: Reference threshold of hearing for test signals of short duration*. International Organisation for Standardisation, Geneva.
- 1390 Lightfoot G, Sininger Y, Burkhard R, Lodwig A (2007). *Stimulus repetition rate and reference levels for clicks and short tone bursts: a warning to audiologists, researchers, calibration laboratories and equipment manufacturers*. Am J Audiol, 16:94-5.
- 1395 Lightfoot G, Stevens J. (2013). *The effects of artefact rejection and Bayesian weighted averaging on the efficiency of recording the newborn ABR*. Ear & Hearing (submitted).
- Newborn Hearing Screening Programme (England). *Guidelines for early assessment and management of babies referred from the Newborn Hearing Screening Programme*. Newborn Hearing Screening Programme (England) <http://hearing.screening.nhs.uk/audiologicalassessment>
- 1400 Newborn Hearing Screening Programme (England). *ABR calibration specification and ABR calibration data*. <http://hearing.screening.nhs.uk/calibration>.
- 1405 Small SA, Hatton JL, Stapells DR (2007). *Effect of bone oscillator coupling method, placement location, and occlusion on bone–conduction auditory steady state responses in infants*. Ear & Hearing 28(1): 83 -98
- 1410 Small SA, Stapells DR (2008). *Maturation of bone conduction multiple auditory steady-state responses*. Int J Audiol. 47(8): 476-488.
- Stapells DR (2000). *Frequency-specific Evoked Potential Audiometry in infants*. In 'A Sound Foundation Through Early Amplification' .Ed Seewald RC ISBN 3-9522009-0-5
- 1415 Stevens J, Brennan S, Gratton D, Campbell M.(2013a). *ABR in newborns: Effects of electrode configuration, stimulus rate and EEG reject levels on test efficiency*. Int J Audiol (submitted)
- 1420 Stevens J, Boul A, Lear S, Parker G, Ashall-Kelly K, Gratton D. (2013b). *Predictive value of hearing assessment by the auditory brainstem response following universal newborn hearing screening*. Int J Audiol (submitted).
- 1425 Webb HD.(1993). *Auditory screening in high risk neonates: An evaluation of the bone conduction ABR*. M Med Sci thesis University of Sheffield, UK.

**For up-to-date information please visit**

Audiologists' homepage, Newborn Hearing Screening Programme (England).

<http://hearing.screening.nhs.uk/audiology>