Assessment of Inter- and Intra-Audiologist Agreements on Auditory Brainstem Response (ABR) Waves Interpretation and ABR Variability

Ahmad Aidil Arafat DZULKARNAIN, PhD; Nordiyana CHE AZID, BAud(Hons)

Department of Audiology and Speech-Language Pathology, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Kuantan, Pahang, Malaysia

SUMMARY
Aim of study: This study investigated the consistency in Auditory Brainstem Response (ABR) waveform evaluations between two audiologists (inter-audiologist agreement) and within each of the audiologist (intra-audiologist agreement).

Methods: Two audiologists from one of the audiology clinics in Kuantan, Pahang, Malaysia were involved in this study. Both audiologists were required to identify and mark the presence of Waves I, III and V in 66 ABR waveforms. Over a one-month interval, each audiologist was required to carry out the same procedure on the same ABR waveforms. This process was continued until we had three separate reviews from each audiologist.

Results: There was a high inter-audiologist ABR waveform identification agreement (over the range 81.71–89.77%), but a lower intra-audiologist ABR waveform identification agreement (over the range 50%–78%) for both audiologists. Our results also showed a high intra-audiologist ABR latency agreement within 0.2 ms (>90%), but a slightly lower inter-audiologist latency agreement (75–84%) within 0.2 ms.

Conclusion: Our results support the need for the clinic to implement further strategies for improving the respective lower agreements and consistencies. These include conducting a continuous education program and using an objective algorithm to support their interpretations.

KEY WORDS:
Auditory brainstem response, ABR variability, inter- and intra-audiologist agreement, subjective judgment, waveform analysis

List of Abbreviations
ABR: Auditory Brainstem Response
IIUM: International Islamic University Malaysia
dBnHL: decibel normalized-hearing level

INTRODUCTION
Auditory Brainstem Response (ABR) is a tool used to estimate hearing sensitivity and to determine site of disorders within the auditory pathways. The ABR is an objective test, because the patient is not required to respond behaviourally to indicate being able to hear to the acoustic stimulus, such as a click, tone burst and chirp. However, from the audiologist's perspective, it is considered as being a subjective test, because the interpretation of the ABR waveforms requires certain clinical skills.

Audiologist needs to perform several steps for successfully interpreting the ABR results. The first step is to determine whether the ABR is present or absent, which needs a basic knowledge in separating the true ABR signal and the noise. This follows with the second step where the audiologist or the observer is required to identify the clinically-important ABR Waves I, III and V. The third step is for the observer to select the location of the peaks for each ABR wave. The marker of the ABR peak in the third step will subsequently obtain the absolute latency values for the respective waves. For these three steps mentioned above, the audiologists would need to use their knowledge and various strategies to ensure accurate interpretations.

According to Elberling and Don, there are four techniques used for detecting the presence or absence of ABR waveforms. Those techniques are response replication, response judgement by a single observer, response judgement by multiple observers and response tracking (for an extensive review, refer to Elberling and Don, 2007)(2). Various techniques are also used in identifying ABR waves, such as using normative data of absolute latency according to the sequence of the wave, examining the morphology of the wave, using multiple-channel recordings and estimating the location of each respective wave by the approximation of 1 ms intervals between the preceding and following waves. In addition, three techniques that can be used to select the ABR peaks that include selecting the point with the maximum amplitude as the peak of the wave component, choosing the last point followed by a negative slope as the peak of wave component or a combination of the first two approaches. The use of different techniques in determining the presence or absence of the ABR, identifying the ABR waves and selecting the ABR peaks could lead to slight differences in judgement and thus, slight differences in the final ABR findings.

With the variation in the techniques used to interpret the ABR, few researches since the mid-1980s have been conducted to investigate how consistently 2–3 independent audiologists can analyse the same ABR waveforms (inter-
Consequently, the exposure in conducting an ABR test, extensive studies to understand the inter- and intra-audiologist agreements. Rossmann and Cashman investigated the inter-audiologist agreement between two audiologists in the identification of the presence or absence of the ABR waveforms in 100 otoneurologic patients. Their results revealed that both audiologists were in agreement for 98% of the time when identifying a patient with normal ABRs. The agreement decreased to 94%, when identifying patient with abnormal ABRs.

Pratt et al examined intra- and inter-audiologist agreements in three audiologists over 252 ABR waveform s from 50 non-tumours ears and 13 ears with a confirmed VIII nerve tumour. Their results from the ipsilateral montage revealed that the inter-audiologist agreement ranged from 82–94% for Waves 1, III and V. In addition, the intra-audiologist agreements were reported to range from 87–94% for all ABR waves.

Olsen et al conducted another study to determine the consistency in the latency measurements and the interpretation of the ABR waveforms. In this study, three audiologists were required to interpret ABR waveform s from 30 otoneurologic patients twice within a four-month interval. From the study, Olsen et al. found that the latency agreement within 0.2 ms exceeded 90% for both the inter- and intra-audiologists’ agreements for Waves I, III and V. Only 10% or less of the latency differences were more than 0.3 ms. The inter-audiologist agreement indicating the ABR as normal or abnormal reached 90%, whereas the intra-audiologist agreements ranged from 97 to 100%.

Despite the high inter- and intra-audiologist agreements in the ABR interpretation reported in the literature, there are no reports (to our knowledge) of the intra- and inter-audiologist agreements in a developing country such as ours (Malaysia). Those findings in the literature may not be similar to our context, because the educational levels, scope of practice and minimum requirements to practice as an audiologist (for example, minimum clinical hours) are different between countries, in particular, between developed and developing countries. This is further complicated by the non-existence of a national newborn hearing screening in a developing country that could be a main source for an audiologist to gain more experience in analysing ABR waveform s. Consequently, the exposure in conducting an ABR test, performing ABR waveform analysis and the number and type of cases involving ABRs could be different between audiologists working in different countries and in different clinical settings.

The limited knowledge of our own ABR inter- and intra-audiologist agreements warrants further investigation. In particular, we need to determine the consistency in our ABR interpretations to ensure an accurate interpretation is being made. A lower consistency might urge that further remedial action is required by the respective audiology clinic to improve the situations. This present study is aimed at describing the consistency in the ABR waveform evaluations (inter- and intra-audiologists) in one of the audiology clinics in Kuantan, Pahang, Malaysia. We hypothesize that this study’s inter/intra-agreements will be higher than 90%, similar to those reported in the literature.

MATERIALS AND METHODS

Participants

The subject of this study was sampled from two audiologists at one of the audiology clinics in Kuantan, Pahang, Malaysia. Both audiologists graduated from an audiology program and institution that is recognized by the Ministry of Higher Education of Malaysia and had more than three years’ experience working with cases involving ABR tests.

Procedure

The study protocol is approved by the university ethics committee. This study was divided into two phases. Phase 1 involved the process of collecting ABR waveform s to be used in Phase 2. Phase 2 is the actual study that involved the interpretation of the ABR waveform s by both audiologists to compare the analyses between them (inter-audiologist agreement) and their consistency in waveform identification and latency measurements over time (intra-audiologist agreement).

Phase 1 of this study was the retrieval of the ABR waveform s from the auditory-evoked potential (AEP) system at the International Islamic University, Malaysia (IIUM) Hearing and Speech Clinic. In this phase, we edited the ABR waveform s by removing the latency markings for the purposes of this study. To initiate this study, ABR waveform s from 50 patients at various intensity levels had been taken from the Grason-Stadler Inc. (GSI) Audera system at the IIUM Hearing and Speech Clinic. Sixty-six ABR waveform s from these patients were selected (including the absence waves). The ABR was recorded for either a site-of-lesion purpose or a threshold-seeking purposes using our clinic’s ABR protocol. The recording was made using single-channel recording (ipsilateral) with a 15 ms time-window using a band-pass of 100–3000 Hz or 30–3000 Hz (depending on whether the subject was adult or infant) with a 12-decibel (dB)/octave slope. The number of sweeps for the testing was 2500, if the purpose of the test was for sites-of-lesions, and 2000 up to 4000 sweeps if the purpose was for threshold-seeking. The ABR waveform s distributions according to the intensity levels were as follows, (i) 17 waveform s for the 80–95 decibel normalized-hearing level (dBnHL), (ii) 24 waveform s for the 50–75 dBnHL and (iii) 25 waveform s for the 20–45 dBnHL.

Phase 2 was the actual study where the ABR waveform s were shown to the audiologist for interpretation. The audiologists were involved in this study during their routine clinic times and were not allowed to discuss their answers. Each audiologist was required to interpret the 66 total ABR waveform s, each needed to identify the absence or presence of an ABR waveform and mark the ABR peaks (Waves I, III and V). The peak selection for each ABR waveform would be used as the marker for the absolute latency values. The researcher then noted the latency value and the identification of the ABR peaks. The audiologist was asked to leave the waveform s, if they felt the ABR peaks were absent or there were no ABR in the waveform s.
After approximately one month, the ABR waveforms were shown again to each audiologist for a second review. Then, one month later, the same procedure was repeated for the third review. After three consecutive reviews by the audiologists, the consistency in the waveform identification and latency values were assessed between and within the audiologists by the researcher.

**Data analysis**

The variables obtained from this study were, (i) the number of ABRs identified as present and absent by both audiologists at least twice in the three reviews, (ii) the number of ABRs waves I, III and V identified by each audiologist at least twice in the three reviews, (iii) the percentage of intra- and inter-audiologist agreements in identifying Waves I, III and V and (iv) the percentage of inter- and intra-audiologist agreements on the absolute latencies for Waves I, III and V within certain intervals in milliseconds (ms).

The inter-audiologist agreements for the waveform identifications for the three primary waves (Waves I, III and V) were obtained by determining whether both audiologists had made similar judgements when identifying the Waves I, III and V. Because there were three separate reviews taking place, the agreements in the waveform identification were counted for each ABR peak (I, III and V) in each of the three separate reviews. The agreements between both audiologists were shown as percentages and were averaged for these three reviews. In addition, the inter-audiologist latency agreements were also obtained. This was achieved by taking the maximum differences in the latencies between the three separate reviews (Review 1 vs 2, Review 2 vs 3 and Review 1 vs 3) for each of the ABR peaks. Next, we would determine the percentage of inter-audiologist latency agreement in the 0.1 ms intervals.

On the other hand, the intra-audiologist waveform identification agreements were obtained when waves individual audiologist made a similar judgement in identifying the presence or absence of Waves I, III and V in all three reviews. Similar to the inter-audiologist agreements, the total agreements for each of the peaks I, III and V by the same audiologist in the different reviews were computed and presented in the form of a percentage. The latency agreement for the same audiologist over three separate reviews was calculated by examining the maximum difference between the latency values in the three sessions. This was accomplished by subtracting the lower value from the higher value of the latency, between the three review pairs (Review 1 vs 2, Review 2 vs 3 and Review 1 vs 3) for both audiologist 1 and audiologist 2. Next, we determined the percentage of the intra-latency agreement in the 0.1 ms intervals.

**RESULTS**

Table I shows the number of ABR peaks (I, III and V) identified by both audiologists from the ABRs being identified as present (at least twice across the three reviews). There was a slight difference in the waveform identifications between both audiologists, with four peaks not being identified by the second audiologist for both Waves III and V and one for Wave I.

Table II shows the mean ± 1 SD percentages of inter-waveform identification agreements for Waves I, III and V between both audiologists for the three separate reviews. The overall findings showed that the inter-audiologist waveform identification agreements between audiologist 1 and audiologist 2 for Waves I, III and V were respectively 89.77, 81.71 and 86.99%.

Table III shows the percentages of the intra-audiologist waveform identification agreements for both audiologists across the three separate reviews. For audiologist 1, there was a relatively high agreement in waveform identification; the agreements were 73, 76 and 78% for Waves I, III, and V, respectively. In contrast, a slightly lower percentage of agreements were observed in the three repeated reviews, especially for Wave III, with 50% of the time being for audiologist 2. The rest of the waves that were being consistently identified by audiologist 2 were 67% and 72% of the time for Waves I and V, respectively. For at least 70% of the time, Wave V was the most consistent wave identified by each audiologist in the repeated reviews of the tracings.

Figure 1 shows the percentages of the inter-audiologist ABR latency agreements for Waves I, III and V across the three sessions for the different latency intervals. Overall, the latency agreements within 0.1 ms were about 50% or more for Waves I, III and V. There was a higher agreement in the latency differences within 0.3 ms with 79, 88 and 92% for Waves I, III and V, respectively. In addition, the latency agreement for both Waves I and V never exceeded 0.4 ms. For Wave III, the latency difference exceeded 0.4 ms 9% of the time. Generally, for almost 80% or more of the time, both audiologists agreed with one another within 0.3 ms for all component waves and only a small percentage (9%) of the latencies differed by up to 0.6 ms.

Figure 2 (a, b and c) shows the percentages of the intra-audiologist ABR latency agreements (at various intervals at

---

<table>
<thead>
<tr>
<th>Wave</th>
<th>Audiologist 1 (%)</th>
<th>Audiologist 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>73</td>
<td>67</td>
</tr>
<tr>
<td>III</td>
<td>76</td>
<td>50</td>
</tr>
<tr>
<td>V</td>
<td>78</td>
<td>72</td>
</tr>
</tbody>
</table>

**Table I: Waves I, III and V identified by both audiologists**

<table>
<thead>
<tr>
<th>Wave</th>
<th>Audiologist 1</th>
<th>Audiologist 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>III</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>V</td>
<td>41</td>
<td>37</td>
</tr>
</tbody>
</table>

**Table II: Mean ± 1SD (%) inter-audiologist waveform identification agreements in three reviews**

<table>
<thead>
<tr>
<th>Wave</th>
<th>Agreement (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>89.77 ± 4.44</td>
</tr>
<tr>
<td>III</td>
<td>81.71 ± 9.86</td>
</tr>
<tr>
<td>V</td>
<td>87 ± 1.40</td>
</tr>
</tbody>
</table>

**Table III: Percentage of intra-audiologist waveform identification agreements**
0.1 ms increments) for both audiologists across the three sessions for Waves I, III and V. For Wave I, the maximum difference in the latency measurement between the reviews for audiologist 2 was only 0.1 ms; however, the maximum difference in the latency measurements between the reviews for audiologist 1 was 0.4 ms. For 58% of the time, both audiologists were able to measure Wave I latency identically across the three sessions (with a difference of 0.0 ms). Generally, the consistency in marking Wave I latencies within 0.1 ms was 92% and 100% for audiologists 1 and 2, respectively.

For Wave III, the maximum difference in the latency measurement between the reviews was 0.3 ms and 0.6 ms for audiologist 2 and audiologist 1, respectively. At least 78% of the time, the latency differences between the reviews were within 0.1 ms, and at least 93% were within 0.2 ms, for both audiologists. Both audiologists demonstrated a greater consistency in Wave V latency measurements, where the maximum differences in the latency measurements between the reviews were within 0.3 ms. The consistency in the Wave V latencies within 0.1 ms was observed to be 80% and 90% for audiologists 1 and 2, respectively. On the other hand, for 97% of the time, both audiologists agreed within 0.2 ms, when interpreting the same ABR waves. Overall, for Waves I, III and V, the maximum differences within 0.2 ms in the latency measurements between the sessions for both audiologists were 90% or more.

DISCUSSION

In this study, the inter-audiologist agreements for waveform identification were 89.77, 81.71 and 86.99 % for Waves I, III and V, respectively. These values were close to those reported by Pratt et al. Pratt et al reported that their inter-audiologist agreements were 82, 84 and 87% for Waves I, III and V, respectively. The 10–18% disagreements between both audiologists in the ABR waves identification occurred mostly with the ABR results recorded close to the threshold level. When testing at a lower intensity level, the recorded ABR signal was relatively small compared to the background noise, causing the interpretation to be more difficult. This increased the uncertainty of the audiologist whether or not to accept the respective ABR waves. In general, the greater number of agreements between both audiologists indicated that the ABR results as interpreted were consistent and therefore the reliability was higher.

In addition, our results showed that the intra-audiologist agreements for waveform identification were in the range 50–78% of the time for all ABR peaks. Pratt et al reported that their intra-audiologist agreements ranged from 87–94%.
lower intra-audiologist agreements compared with those reported by Pratt and colleagues could be caused by two factors. First, three repeated reviews were taken in this study compared to the two reviews reported by Pratt and colleagues. The extra reviews in the present study increased the opportunities for both audiologists to amend their judgments in identifying Waves I, III and V (particularly in the extra review session). While testing for intra-rating agreements, having more sessions introduced to the participants would increase the possibility of their responses to change over time (either negatively or positively), because of the many confounding factors, such as changes in their health status, motivations, emotions, fluctuations in memory and attention that could indirectly vary their responses. Second, the lower intra-audiologist agreement could be because both audiologists were using two-channel protocols in their practices (simultaneous ipsilateral and contralateral) compared to the four-channels reported in the literature. The use of multiple channels and multiple electrode montages could assist in waveform identification; this, therefore, could be one of the reasons for the consistent intra-audiologist waveform identification reported by Pratt and colleagues.

Our intra-audiologist latency agreement within 0.2 ms was more than 90% for Waves I, III and V; this is consistent with the Olsen et al. findings reporting their agreement exceeding 90%5. On the other hand, our inter-audiologist latency agreement (within 0.2 ms) was in the range of 75–84%; this was slightly lower compared to the 90% or more agreement reported by Olsen et al.7 Lower agreements in determining the latency between the audiologists could be because each audiologist involved in this study used a different method for selecting the ABR peak. Several strategies have been discussed in the literature on how to select the ABR peaks. This includes selecting the point with the maximum amplitude as the peak of the wave component or choosing the last point followed by a negative slope as the peak or combining the two approaches. Using either the first or second approach could result in the ABR latency either becoming shorter or longer, and therefore, if both audiologists were using different approaches, the agreement between them could be slightly different.

The different approaches in selecting ABR peaks are further supported by both audiologists being graduated from different local institutions. Different institutions may emphasize different peak selection criteria in their clinical protocols and have different approaches for training their students in waveform analysis, such as using a traditional method, computer-based simulations training or a combination of these two methods. Currently, there are no international or local standard training modules for ABR waveform analysis that may cause a possible variability in the interpretation of Wave latency between audiologists, as observed in the present study11,17. In addition, the lower inter-audiologist latency agreement could also be influenced by the difference in the years of experience in performing ABR waveform analyses between audiologists 1 and 2. Audiologist 1 had one year more clinical exposure in performing ABR waveform analysis than did audiologist 2 at the time the research was conducted (both audiologists have more than three years clinical experience performing ABR waveform analyses). The differences in the duration of the clinical exposure in interpreting ABR waveforms may influence the inter-audiologist latency agreement as highlighted by Elberling and Don,2 although to date, there are no published results to support this argument.

Among the three components of the ABR waves, Wave V was the most consistent in being identified and the latency value was determined more consistently either between or within the audiologists than for the other waves. This could be because Wave V has the most robust and reproducible ABR peak compared to the other peaks at both the higher and lower intensity levels. Another reason could be because both audiologists were working in an audiology clinic that had more pediatric than adult patients; thus, the purpose of conducting the ABR test was either for threshold-seeking or hearing screening where the focus of the testing would be on the determining the presence of Wave V. This would increase their familiarity with Wave V over the other waves.

The lower agreement in our intra-audiologist ABR waveform identifications and inter-audiologist latency difference indicated that several strategies must be taken by the audiology clinic to improve the consistency in ABR waveform interpretations. The first strategy is to conduct a continuous education among the audiologists in this clinic to standardize their method of interpreting the ABR results. The on-going education sessions can be carried out bi-weekly or monthly to share information between the audiologists on interesting cases, including cases with difficult-to-interpret ABR waveforms, and to discuss strategies for standardizing their interpretations. The second strategy is to advise audiologists to supplement their subjective analyses with an objective analysis, such as F-statistics, at single or multiple points. Using an objective analysis will avoid the audiologist accepting an ABR result when the noise level is relatively high. The third strategy is for the clinic to implement a peer and expert review system. By using this system, the ABR waveform interpretation would be stored in the server and accessible to expert personnel in the ABR waveform analysis. Feedback can be provided by the expert through the system to the audiologist. This will help the audiologists to improve their abilities in the ABR waveform analyses by having this mentoring process. These continuous efforts would lead to an increase in the reliability of the ABR interpretations.

**CONCLUSION**

In general, there were higher inter-audiologist ABR waveform identification agreements (in the range of 82–89%) but lower intra-audiologist ABR waveform identification agreements (in the range of 50–78%). Our results also showed a high intra-audiologist ABR latency agreement (>90%) and a slightly lower inter-audiologist latency agreement within 0.2 ms (75–84%). Our results support the requirement for the clinic to implement further strategies for improving the respective lower agreements. Our conclusions were limited by the study participant’s (two audiologists from one clinic) and the variability of the ABR waveforms (66 ABRs) used in the present study.
ACKNOWLEDGMENT
The authors wish to acknowledge the International Islamic University Malaysia Endowment fund and Ministry of Education, Research Acculturation Grant Scheme (RAGS) for their financial support in conducting this study (Grant numbers: EDW B12304-0782 and RAGS 13-014-0077). We also wish to acknowledge the contributions from two participating audiologists, and Madam Narina Norordin for her opinions concerning this study.

REFERENCES