A frequent complication of surgical procedures in the cerebellopontine angle (CPA), e.g., for treatment of vestibular schwannoma (VS), is damage to the cranial nerves. The facial and auditory nerves are especially at risk, damage to which can cause facial nerve palsy and deterioration of hearing. Even a comparatively small decline of hearing thresholds may affect speech perception. In such cases, participating in a conversation becomes challenging, especially in situations with several speakers and background noise. Hearing aids alleviate the impact, but do not entirely solve the issue. In fact, patients with normal preoperative hearing describe postoperative

---

**Intraoperative auditory steady-state monitoring during surgery in the cerebellopontine angle for estimation of postoperative hearing classes**

Stefan Rampp, MD, Leonhard Rensch, Sebastian Simmermacher, MD, Torsten Rahne, PhD, Christian Strauss, MD, and Julian Prell, MD

Departments of Neurosurgery and Otorhinolaryngology, Head and Neck Surgery, University Hospital Halle (Saale), Germany

**OBJECTIVE** Brainstem auditory evoked potentials (BAEPs) have been used for intraoperative monitoring of the auditory nerve for many years. However, BAEPs yield limited information about the expected postoperative hearing quality and speech perception. The auditory steady-state response (ASSR) enables objective audiograms to be obtained in patients under anesthesia. These ASSRs could be used for intraoperative estimation of hearing classes to improve the postoperative outcome and quality of life. Studies investigating the clinical use of ASSRs during total intravenous anesthesia are currently lacking. The work presented in this article therefore investigates the application of ASSRs for intraoperative estimation of hearing classes.

**METHODS** In 43 patients undergoing surgery for vestibular schwannoma, ASSR measurements were performed at the beginning and end of the surgical procedure. ASSR stimuli consisted of 80-dB hearing level amplitude-modulated tones with 5-minute duration, 90-Hz modulation, and 3 different carrier frequencies: 500, 1000, and 2000 Hz. Stimulation was performed unilaterally with and without contralateral masking, using single and combined carriers. Evoked responses were recorded and analyzed in the frequency domain. ASSRs were compared with extraoperative hearing classes and BAEPs using ANOVA, correlation, and receiver operating characteristic statistics.

**RESULTS** ASSRs yielded high and consistent area under the curve (AUC) values (mean 0.83) and correlation values (mean 0.63), indicating reliable prediction of hearing classes. Analysis of BAEP amplitude changes showed lower AUC (mean 0.79) and correlation values (0.63, 0.37, and 0.50 for Waves I, III, and V, respectively). Latencies showed low AUC values (mean 0.6) and no significant correlation. Combination of several carriers for simultaneous evaluation reduced ASSR amplitudes and respective AUC values. Contralateral masking did not show a significant effect.

**CONCLUSIONS** ASSRs robustly estimate hearing class in patients under total intravenous anesthesia, even when using short measurement durations. The method provides a diagnostic performance that exceeds conventional BAEP monitoring and enables objective and automated evaluation. On the basis of these findings, continuous intraoperative auditory monitoring could become a promising alternative or adjunct to BAEPs.

https://thejns.org/doi/abs/10.3171/2016.7.JNS16460

**KEY WORDS** intraoperative monitoring; vestibular schwannoma; cerebellopontine angle; CPA; hearing class; diagnostic technique
hearing impairment as the most debilitating complication, even compared with moderate or severe facial palsy, with a severe impact on quality of life (QOL).\(^{19}\)

Intraoperative monitoring of brainstem auditory evoked potentials (BAEPs) has been used since the 1980s to prevent such postoperative functional deficits. BAEPs show increases in latency and decreases of amplitude when the cochlear nerve is damaged. As a consequence, the surgical strategy may be altered or vasoactive treatment may be started\(^2\) to enable high postoperative QOL.

However, whether changes of BAEPs can be used to estimate the quality of speech perception is questionable. Click sounds are typically used for auditory stimulation, which provide high response amplitudes but also limit the frequency specificity. Reliance on subjective interpretation further limits the correlation of specific BAEP parameters to a certain hearing class. Bischoff et al.\(^3\) even describe patients with seemingly stable BAEPs but nevertheless considerable deterioration of hearing.

A different type of auditory evoked response, the auditory steady-state response (ASSR),\(^{10,37}\) may represent aitory steady-state response (ASSR),\(^{10,37}\) with a severe impact on quality of life (QOL).\(^{19}\) the humming characteristic of the stimulus, i.e., the modulation rate. In the present study, modulation rates of approximately 90 Hz were used. This means that the intensity of the sound (in this study, 80 dB hearing level [HL]) changed 90 times per second, from 0% to 100% and back to 0%. The result is a humming sound with a carrier of 1000 Hz and a modulation rate of 90 Hz.

Due to stimulus processing of this humming characteristic, the evoked response on electroencephalography (EEG) contains a component that exactly corresponds to the harmonics of power line noise at 100 Hz and an electrical artifact at approximately 80 Hz. Finally, combined carrier tones with masking were used to evaluate the impact of simultaneous testing and masking on the overall detectability of responses. Single carrier tones were generated at 500, 1000, or 2000 Hz. Stimuli were amplitude modulated with a modulation depth of 100% at a frequency of 90 Hz. A combined multiple carrier tone with all of the 3 carriers was constructed, using modulation frequencies of 88, 90, and 92 Hz. Finally, combined carrier tones with masking were created by adding white noise at 40 dB SPL only to the channel presented to the ear contralateral to ASSR stimulation. The modulation frequencies were selected to avoid the harmonics of power line noise at 100 Hz and an electrical artifact at approximately 80 Hz that were occasionally observed in our setup.

The intensity of stimuli was subjectively calibrated to 80 dB HL using a Nicolet Viking IV (CareFusion). This intensity was chosen for loud stimulation but without risk of potential hearing damage and reflects results from a previous study.\(^{36}\) Sounds were presented in randomized

### Methods

#### Patient Demographic Data

Data from 43 consecutive adult patients undergoing surgery in the CPA due to VS (27 women, mean age 50 years, range 24–73 years) were investigated. In all patients, BAEPs were recorded for monitoring of the auditory nerve as part of the clinical routine in CPA surgery. Measurements from deaf patients (hearing Class V) were also included to evaluate the assumed lack of ASSRs in such cases. Individual pre- and postoperative hearing quality was evaluated using pure tone and speech audiometry on the affected side and categorized into hearing classes according to Gardner and Robertson.\(^{13}\)

The local ethics committee approved the study and all patients gave their written informed consent.

#### Anesthesia Procedures

For premedication, all patients received 7.5 mg midazolam orally. All patients underwent routine monitoring by means of electrocardiogram, pulse oximetry, and invasive continuous blood pressure monitoring; they received intravenous anesthesia with a total intravenous anesthesia technique with propofol and remifentanil. Anesthesia was induced with 0.5 mg/kg remifentanil per minute and propofol (usually 2 mg/kg) until loss of consciousness. Orotracheal intubation was facilitated by rocuronium 0.6 mg/kg; thereafter, no muscle relaxants were administered. Anesthesia was maintained by 0.2–0.5 mg/kg remifentanil per minute and 4–5 mg/kg propofol per hour as clinically required. Doses of narcotic drugs were titrated by the reaction of the patient to surgical stimulus and hemodynamic parameters such as blood pressure and heart rate. Overall management complied with standard anesthesia as usually practiced in our institution.

#### Auditory Stimulation

**ASSR Stimulation**

Auditory stimulation was performed using 5 different 5-minute amplitude-modulated sounds: single carrier tones, a combined carrier tone without masking, and a combined carrier tone with masking. The 2 latter were used to evaluate the impact of simultaneous testing and masking on the overall detectability of responses. Single carrier tones were generated at 500, 1000, or 2000 Hz. Stimuli were amplitude modulated with a modulation depth of 100% at a frequency of 90 Hz. A combined multiple carrier tone with all of the 3 carriers was constructed, using modulation frequencies of 88, 90, and 92 Hz. Finally, combined carrier tones with masking were created by adding white noise at 40 dB SPL only to the channel presented to the ear contralateral to ASSR stimulation. The modulation frequencies were selected to avoid the harmonics of power line noise at 100 Hz and an electrical artifact at approximately 80 Hz that were occasionally observed in our setup.

The intensity of stimuli was subjectively calibrated to 80 dB HL using a Nicolet Viking IV (CareFusion). This intensity was chosen for loud stimulation but without risk of potential hearing damage and reflects results from a previous study.\(^{36}\) Sounds were presented in randomized
order as single 5-minute WAV files using a 24-bit external sound card (Edirol UA-25; Roland) and Nicolet TIP-300 (CareFusion) in-ear earphones, which deliver the stimulus via transducers coupled to the ear by plastic tubes, electrically decoupling the transducer from the patient. Stimulation was performed monaurally to the left and right ears separately; masked stimuli applied masking only to the ear contralateral to ASSR stimulation.

**BAEP Stimulation**

BAEPs were recorded intraoperatively using a Viking II (Nicolet Instrument Corp.) until 2012 and an Endeavor system thereafter (Natus) with Nicolet TIP-300 in-ear earphones. Auditory stimulation corresponded to standard V.20,42 Stored BAEP traces were analyzed and interpreted visually. Changes of amplitudes and latencies of Waves I, III, and V at the end of the surgical procedure compared with the beginning were evaluated. Amplitude changes were categorized into 3 groups: no/minimal decrease, subtle decrease < 50%, and clear decrease ≥ 50% compared with amplitudes at the beginning.

**ASSR and BAEP Recordings**

ASSRs without BAEP stimulation were recorded during the initial phase of the operation (before or during opening of the skull), as well as at the end of the surgical procedure. The recordings were limited to these 2 time points to avoid interference with routine BAEP monitoring procedures, which were considered highly relevant in the patients who were investigated. Although there are techniques available to combine BAEPs and ASSRs,23 these are not yet available as part of a certified medical device. It was thus not possible to record ASSRs continuously due to ethical and patient safety reasons.

Data were recorded with a 2-channel EEG montage routinely used for BAEPs and ASSRs: A1 versus Cz and A2 versus Cz. EEG voltages were amplified using a biosignal amplifier (15LT, Grass-Telefactor; Astro-Med GmbH), high-pass filtered (analog filter, 5 Hz), and sampled at 1024 Hz. Only data ipsilateral to the surgical procedure were analyzed.

**Preprocessing Procedures**

Recorded data were processed offline using MATLAB software (version 2009b; MathWorks). Raw data were subdivided into 2-second epochs. Epochs with amplitudes exceeding 40 μV after bandpass filtering between 30 and 300 Hz were rejected, corresponding to values in the literature, ranging from approximately 30 to 60 μV.20,42 The remaining epochs were pooled for further analysis, which are addressed as raw epochs in the remainder of this article.

**ASSR Measurement Paradigm**

The measurement of ASSR amplitudes consisted of several steps, following published procedures.23 The first step consists of averaging to improve the signal-to-noise ratio of the ASSRs: 8 epochs are concatenated into a “sweep,” which then consists of 8*2 seconds, corresponding to 16,384 samples (at a sample rate of 1024 Hz). In the present study, 2 averaging durations were evaluated: the maximum number of sweeps from all available data of a stimulation, and only 4 sweeps. The latter was performed to evaluate the prognostic value of ASSRs from short averaging durations, which would be needed for real-time analysis and feedback.

In a second step, the averaged sweep data were submitted to frequency spectrum calculation using the fast Fourier transform. The spectral power at the modulation frequency (the ASSR power) was extracted, which will be addressed as ASSR amplitude in the remainder of this article.

**Statistical Evaluation**

**Comparison With Hearing Class**

The relationship between ASSR amplitudes and hearing was evaluated by calculating the Spearman correlation to the respective Gardner-Robertson hearing class. The ASSRs at the beginning of the surgical procedure were compared with the preoperative hearing classes, and ASSRs at the end were compared with the postoperative hearing classes. Carriers were investigated separately, and correlation values were calculated over all stimulations, i.e., without differentiation of the individual carrier frequencies.

For evaluation of sensitivity and specificity of ASSRs for hearing classes, receiver operating characteristic (ROC) analysis was performed. The ability to discern different hearing class ranges was investigated by calculating the area under the curve (AUC). AUC values range from 0 to 1. Values of approximately 0.5 show a performance comparable to chance. Higher values demonstrate diagnostic utility regarding a certain outcome. By convention, AUCs < 0.7 are considered poor, ≥ 0.7 to < 0.8 fair, and ≥ 0.8 to < 0.9 good. Higher values demonstrate excellent diagnostic performance.

**Effects of Masking and Combination of Multiple Carriers**

The influence of masking and the combination of multiple carriers within 1 stimulus was evaluated using ANOVA analyses of ASSR amplitude. Tested factors were carrier frequency (500, 1000, and 2000 Hz), masking (yes/no), number of averaging sweeps, Gardner-Robertson hearing class, and multiple versus single stimulation.

**Evaluation of Short Averaging Times**

Performance of ASSRs using relatively few data is highly interesting due to its potential applications for real-time monitoring. For these applications, short averaging times directly translate to faster feedback to the surgeon. Therefore, a second ANOVA analysis was performed using a fixed number of only 4 sweeps, amounting to 1 minute of data per stimulation. Investigated factors were carrier frequency, hearing class, and multiple versus single stimulation. Masking was not included as a factor, because it did not show a significant influence in the first ANOVA analyzing all available data.

**Changes of BAEPs**

Pre- and postoperative hearing classes were compared
with BAEP amplitude and latency differences between the beginning and end of the surgical procedure. Due to the ordinal scaling of the Gardner-Robertson hearing classes, as well as our categorization of BAEP amplitude changes, Spearman correlation coefficient and significance were calculated (MATLAB version 2009b, MathWorks). Partial correlations to the postoperative hearing class were evaluated, controlling for the preoperative hearing class.

Because only changes of BAEP responses are interpreted for intraoperative monitoring, the ability to detect deterioration of a certain hearing quality, i.e., hearing class range, was evaluated. Similar to the statistical analysis of ASSR results, ROC analyses were calculated for amplitude changes and latency differences of Waves I, III, and V for different hearing ranges: normal hearing (Class I), serviceable hearing (Classes I and II), and any hearing (Classes I–IV).

Results

Recordings at the beginning of the surgical procedure were possible in all 43 patients; at the end, recordings in 2 patients were not possible due to logistical reasons. Preoperatively, 22 patients had hearing that corresponded to Gardner-Robertson Class I, 11 Class II, 8 Class III, and 2 Class IV. Postoperatively, 22 patients had hearing that corresponded to Class I, 4 Class II, 3 Class III, 2 Class IV, and 29 Class V.

A total of 75.1% of all stimulation epochs were considered valid with residual noise under 40 nV; the remainder were rejected due to excessive noise or artifacts. On average, 13.6 (± 3.6 SD) sweeps per stimulation were used for analysis (range 1–18 sweeps), corresponding to 218 (± 58 SD) seconds of data (range 16–288 seconds).

Influence of Masking and Combination of Multiple Carriers

Contralateral masking did not reveal a significant influence on ASSR amplitudes when carrier frequency, number of averaging sweeps, hearing class, and single versus multiple carriers were taken into account (Table 1). In contrast, simultaneous stimulation with multiple carriers exhibited lower amplitudes. This was the case both when all available data were used for ASSR analysis and when data from only 4 sweeps were used (Table 2): 35.3 nV versus 26.3 nV (p < 0.01; t-test, all data) and 42.0 nV versus 32.6 nV (p < 0.05).

Relationship Between ASSRs and Hearing Quality

Correlations of ASSR Amplitudes to Hearing Classes

Correlations using all available data reached a value of −0.78 (high amplitudes are correlated with lower, i.e., better hearing classes) for stimulation of 1000 Hz separate from other carrier frequencies (Table 3). Correlations were generally higher using separate versus combined carriers. The latter achieved a best correlation of −0.66, also for 1000-Hz stimulation (as part of a combined 500, 1000, and 2000-Hz stimulus). The worst correlation of −0.47 resulted from analyzing 500 Hz after combined stimulation.

Using only 4 sweeps of data, the best correlation of −0.65 was also achieved with separate 1000-Hz stimulation. The worst correlation of −0.38 was also observed with 500 Hz as part of a combined stimulus.

Differentiation of Hearing Classes

The ability to differentiate between different hearing classes using amplitude thresholds was evaluated with ROC statistics. The AUC values of separate carriers were generally high and ranged from 0.80 to 0.93 when all available data were used and from 0.69 to 0.85 when only 4 sweeps were analyzed (Table 4). In both settings, the best diagnostic performance was achieved by stimulation with 1000 Hz for differentiation of hearing Classes I–III from Classes IV and V.

Combination of carriers for simultaneous stimulation again yielded lower results, which ranged from 0.70 to 0.86 using all data and from 0.65 to 0.77 using only 4 sweeps. Analyzing 1000-Hz responses for differentiation of hearing Classes I–III from Classes IV and V also produced the

### Table 1. ANOVA results of the influence of stimulation factors on ASSR amplitudes using all available data

<table>
<thead>
<tr>
<th>Factor</th>
<th>Sum Square</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier</td>
<td>23,381.1</td>
<td>2</td>
<td>11,690.5</td>
<td>13.6</td>
<td>0.0000</td>
</tr>
<tr>
<td>Masked</td>
<td>512.0</td>
<td>1</td>
<td>512.0</td>
<td>0.6</td>
<td>0.4409</td>
</tr>
<tr>
<td>Sweeps</td>
<td>39,990.0</td>
<td>17</td>
<td>2,352.4</td>
<td>2.7</td>
<td>0.0002</td>
</tr>
<tr>
<td>Hearing class</td>
<td>156,525.2</td>
<td>4</td>
<td>39,131.3</td>
<td>45.4</td>
<td>0.0000</td>
</tr>
<tr>
<td>Combined</td>
<td>6,163.9</td>
<td>1</td>
<td>6,163.9</td>
<td>7.2</td>
<td>0.0076</td>
</tr>
<tr>
<td>Error</td>
<td>607,893.1</td>
<td>706</td>
<td>861.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>941,123.2</td>
<td>731</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DF = degrees of freedom; F = F-statistic.
* p values < 0.05 were considered statistically significant.

### Table 2. ANOVA results of the influence of stimulation factors on ASSR amplitudes using only 4 sweeps of data

<table>
<thead>
<tr>
<th>Factor</th>
<th>Sum Square</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier</td>
<td>26,035.2</td>
<td>2</td>
<td>13,017.6</td>
<td>10.6</td>
<td>0.0000</td>
</tr>
<tr>
<td>Hearing class</td>
<td>167,198.4</td>
<td>4</td>
<td>41,796.5</td>
<td>34.1</td>
<td>0.0000</td>
</tr>
<tr>
<td>Multiple</td>
<td>10,921.6</td>
<td>1</td>
<td>10,921.6</td>
<td>8.9</td>
<td>0.0030</td>
</tr>
<tr>
<td>Error</td>
<td>588,878.8</td>
<td>487</td>
<td>1,226.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>792,473.2</td>
<td>487</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p values < 0.05 were considered statistically significant.

### Table 3. Correlation values of ASSR amplitude with hearing classes

<table>
<thead>
<tr>
<th>Carrier</th>
<th>All Available Data</th>
<th>4 Sweeps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single</td>
<td>Combined</td>
</tr>
<tr>
<td>All</td>
<td>−0.68</td>
<td>−0.57</td>
</tr>
<tr>
<td>500</td>
<td>−0.58</td>
<td>−0.47</td>
</tr>
<tr>
<td>1000</td>
<td>−0.78</td>
<td>−0.66</td>
</tr>
<tr>
<td>2000</td>
<td>−0.73</td>
<td>−0.59</td>
</tr>
</tbody>
</table>

All correlations were found to be significant at the p < 0.0001 level.
best AUC values. Figure 1 shows ROC curves for some clinically interesting scenarios.

Differentiation of Normal Hearing (Class I)

Normal hearing (Class I) was best detected using separate 1000-Hz stimulation. Analyzing all available data yielded an AUC value of 0.89. The “optimal operating point” (i.e., the best compromise between sensitivity and specificity, assuming that both are equally important) was found with an amplitude threshold of 94 nV. This yielded a perfect specificity of 100% and a sensitivity of 42% (circle marker in Fig. 1). The positive predictive value (PPV) in this case amounts to 100%, with a corresponding negative predictive value (NPV) of 79%. This means that an ASSR amplitude after 1000-Hz stimulation of $\geq 94$ nV always indicated normal hearing. Lower amplitudes were associated with worse hearing in 79% of patients.

Lowering the threshold to 29 nV yielded a perfect sensitivity and a specificity of 0.74 (triangle marker in Fig. 1), a PPV of 63%, and an NPV of 100%. In this case, an ASSR amplitude < 29 nV always marked a deterioration of normal hearing, whereas higher amplitudes correctly showed preserved normal hearing in 63% of patients.

Using only 4 sweeps for fast detection and feedback to the surgeon still enabled a maximum AUC value of 0.84 using single 1000-Hz stimulation. The optimal operating point was found at an almost identical threshold of 93 nV. Sensitivity was 95% and specificity was 48%.

Combined stimuli showed decreased AUC values both when all or only 4 sweeps of data were used (Table 4).

Differentiation of Serviceable Hearing (Classes I and II)

Using all available data and single-carrier stimuli, serviceable hearing was differentiated best with 1000 Hz, resulting in an AUC of 0.93. The optimal operating point at 29 nV yielded a sensitivity of 88% and specificity of 88%.

<table>
<thead>
<tr>
<th>ASSR Amplitude</th>
<th>Hearing Class I</th>
<th>Hearing Classes I–II</th>
<th>Hearing Classes I–IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>All*</td>
<td>0.85</td>
<td>0.87</td>
<td>0.86</td>
</tr>
<tr>
<td>500 Hz</td>
<td>0.80</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td>1000 Hz</td>
<td>0.89</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>2000 Hz</td>
<td>0.88</td>
<td>0.89</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Combined

| All*           | 0.78           | 0.81                 | 0.80                 |
| 500 Hz         | 0.70           | 0.78                 | 0.75                 |
| 1000 Hz        | 0.82           | 0.86                 | 0.86                 |
| 2000 Hz        | 0.80           | 0.80                 | 0.78                 |

Class I = normal hearing; Classes I & II = serviceable hearing; Classes I–IV = any preserved hearing.

* The “all” condition takes all ASSR amplitudes of all carriers in a measurement into account, i.e., a single correlation for all amplitudes versus hearing classes was calculated.

FIG. 1. ROC curves for amplitudes of single-carrier ASSRs without masking for normal and serviceable hearing, as well as any remaining hearing. All available data were used for the figure. For the purpose of brevity, the respective results for combined carriers, masking, and 4 sweep–only stimulation are not displayed, but they are qualitatively similar. Table 3 provides AUC values for a quantitative comparison. GR = Gardner-Robertson.
With 4 sweeps and a single carrier, the AUC decreased to 0.85 at 1000 Hz. Optimal operating point sensitivity at 43 nV was 93%, and specificity was 71%. The remaining results are summarized in Table 4.

Differentiation of Any Hearing (Classes I–IV)
Detection of hearing of any degree in contrast to loss of hearing was possible using all available data with an optimal AUC value of 0.91, again using 1000 Hz as the carrier. The optimal cutoff was 15 nV, yielding a sensitivity of 79% and a specificity of 86%. Using only 4 sweeps of data still allowed differentiation, with an AUC of 0.84 at 1000 Hz and a comparable threshold at 14 nV. Here, sensitivity was 43% and specificity was 96%. A combination of carriers also decreased AUC values by approximately 0.05–0.1 (Table 4).

**Relationship Between BAEPs and Hearing Classes**

Partial correlations of amplitude changes and postoperative hearing were 0.63 for Wave I (p < 0.001), 0.37 for Wave III (p < 0.05), and 0.50 for Wave V (p < 0.005). Correlations of latency changes did not reach statistical significance. Figure 2 shows an overview of BAEP amplitude changes versus postoperative hearing and hearing deterioration.

Deterioration of Normal Hearing (Class I)
Deterioration of normal hearing was detected best by amplitude decreases of Wave V, yielding an AUC of 0.87. Using already slight amplitude decreases as an indicator resulted in a sensitivity of 100%, a specificity of 25%, a PPV of 85%, and an NPV of 100%. The performance of Wave I was only slightly lower, whereas the differences to Wave III were large.

Latency changes of Wave III or V were most useful for recognizing deterioration of normal hearing (AUC 0.71 and 0.69), whereas Wave I yielded an AUC value close to chance (0.55; Table 5).

Deterioration of Serviceable Hearing (Classes I and II)
The highest AUC value for detection of deteriorated serviceable hearing was 0.85, again observed with changes of Wave V, followed by changes of Wave I with an almost similar performance. Wave III yielded the worst performance, with an AUC of 0.74. Latency changes barely exceeded the level of random chance, with Wave I yielding the highest AUC value of 0.64 (Table 5).
Auditory steady-state response hearing class estimation

Loss of Hearing (deterioration to Class V)

Loss of hearing was best detected using amplitude decreases of Wave I (AUC 0.87), whereas both Waves III and V were less useful. The best PPV values for loss of hearing were found with abolished responses, ranging from 75% to 100%. However, NPV values of approximately 40% were rather low. Loss of hearing is thus clearly indicated by abolished responses; however, persisting Waves I, III, or V do not exclude hearing loss.

Latency differences of Wave I reached an AUC of 0.72, whereas Waves III and V again did not exceed the level of random chance to predict a loss of hearing (Table 5).

Discussion

ASSRs were recorded at the beginning and end of surgery in the CPA for treatment of VS. The response amplitudes were related to pre- and postoperative hearing quality and compared with BAEPs at the beginning and end of the procedure. Our findings demonstrate a high diagnostic performance of ASSRs that exceeded the utility of BAEPs to predict the hearing class.

ASSR Amplitude Predicts the Hearing Class

The diagnostic performance of ASSRs was remarkably robust over the whole range of hearing classes. The ROC analysis yielded high AUC values between 0.8 and 0.93. Using 1000-Hz stimulation, it was possible to achieve values ≥ 0.89 for differentiation of normal and serviceable hearing, as well as a complete hearing loss. This high level of diagnostic value may be enabled by the frequency-specific evaluation of hearing. The resulting ASSR audiograms are strongly correlated to pure tone hearing thresholds1,7,38,43 and are therefore a highly sensitive measure for hearing quality. Although we did not evaluate complete audiograms using ASSRs, this property may still contribute to the good results when few carrier frequencies are tested to estimate hearing classes. Our results may also benefit from the fact that ASSRs seem to be advantageous in patients with high-frequency sensorineural hearing loss,18 which is typical for patients with VS.15

The frequency specificity also enables automated analysis, which measures the power at the user-defined modulation frequency. In contrast, the characteristic BAEP waves have varying amplitudes and latencies, which pose a challenge for automated evaluation. Currently, no automated procedure is used in clinical routines. Although such procedures have been available for decades,2 noise, artifacts, and interindividual differences have so far prevented widespread adoption.

BAEP Evaluation

Evaluation of BAEP amplitude changes also yielded high AUC values, but comparably low correlations. This corresponds well to the clinical experience that BAEPs are valuable to detecting a potential deterioration of hearing (high AUC), but have only limited value to differentiating between hearing classes (low correlations).

Practical issues cause further limitations. BAEP monitoring uses amplitude and latency changes compared with a baseline recording at the beginning of surgery. In contrast to the ASSR, a specific BAEP amplitude is not directly interpreted. Similar to the ASSR, it is also questionable if a given amplitude difference has the same meaning for different ranges of baseline BAEPs. For example, a small difference may be negligible when normal, high-amplitude BAEPs have been measured. When already reduced BAEPs occur, the same small difference may indicate a critical deterioration of hearing. These properties result in a considerable amplitude variability, which poses difficulties for interpretation. Noise, artifacts, and the need for subjective visual evaluation further increase these difficulties.

Our results show that already quite subtle amplitude changes can be warning signs for hearing deterioration. Guidelines for BAEP interpretation9 suggest using a reduction of 50% as a threshold. In light of our results, this clinical practice may overlook and protract considerable damage. However, its merit certainly lies in the easy application. In contrast, the potentially important subtler changes are difficult to detect robustly by visual interpretation alone.

Changes of BAEP latencies demonstrated only very limited use for predicting hearing deterioration. Correlations were not significant, and a majority of AUC values did not exceed the level of random chance. A threshold of 1 msec, as suggested by guidelines9 and literature,32 was not exceeded, even with deterioration to hearing Class III and worse. Polo et al.32 observed latency differences of approximately 1 msec in patients in whom there was a hearing deterioration of 20 dB. However, they investigated BAEPs during microvascular decompression for hemifacial spasm. In these cases, the auditory nerve was therefore unaffected, in contrast to patients with VS. A potential explanation for the low prognostic utility of BAEP latencies may thus be chronic alterations of the auditory nerve due to the impact of the tumor. Such changes may affect both baseline latencies and their intraoperative behavior in response to nerve damage.

Aihara et al.2 investigated hearing in cases of VS. They compared patients with hearing Class II with patients with hearing Class III. The rationale was that QOL suffers severely when normal or serviceable hearing cannot be preserved. The study provided AUC values for different

### TABLE 5. AUC values for BAEP amplitude decrease and latency increase predicting different Gardner-Robertson hearing class ranges

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hearing Class I</th>
<th>Hearing Classes I &amp; II</th>
<th>Hearing Classes I–IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitude decrease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave I</td>
<td>0.85</td>
<td>0.83</td>
<td>0.87</td>
</tr>
<tr>
<td>Wave III</td>
<td>0.69</td>
<td>0.74</td>
<td>0.68</td>
</tr>
<tr>
<td>Wave V</td>
<td>0.87</td>
<td>0.85</td>
<td>0.72</td>
</tr>
<tr>
<td>Latency difference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave I</td>
<td>0.55</td>
<td>0.64</td>
<td>0.72</td>
</tr>
<tr>
<td>Wave III</td>
<td>0.71</td>
<td>0.53</td>
<td>0.56</td>
</tr>
<tr>
<td>Wave V</td>
<td>0.69</td>
<td>0.51</td>
<td>0.48</td>
</tr>
</tbody>
</table>

* Also see Table 4.
Combining multiple carrier frequencies shortens ASSR recording times. However, if low-intensity stimuli are used, e.g., to robustly estimate hearing thresholds, the remaining measurement duration is still in the range of 10–20 minutes.28 For real-time auditory nerve monitoring, such durations are too long and certainly exceed the time needed for BAEP estimation. Intraoperative ASSR threshold estimation could thus only be applied intermittently or in parallel to BAEPs with updates every 10–20 minutes. However, our results provide evidence that real-time ASSR monitoring may be viable when the specific requirements of such an application are considered.

First, intraoperative monitoring does not necessarily require exact threshold estimations. Instead, current clinical routine demonstrates that it is mostly sufficient to evaluate the likely hearing classes, which represent a coarser classification of hearing quality. Our data show that in most cases, this can be achieved even with only 4 sweeps. Four sweeps correspond to approximately 1 minute of artifact-free recording. Taking the percentage of 75% valid and thus artifact-free data epochs into account, this means that approximately 75 seconds can produce enough data to estimate hearing classes. Of course, more data would attain a more robust estimation. Providing feedback every 4 sweeps could therefore be added to by calculating a longer trend. The implementation of ASSR monitoring could thus provide, for example, 3 parameters: ASSR amplitude after 4 sweeps, after 10–20 minutes of data collection, and possibly an intermediate estimation, e.g., using 4–5 minutes.

A further strategy would be to use higher stimulus intensities. These would result in higher response amplitudes, allowing for faster detection. The relationship between amplitude and hearing at a certain carrier frequency is known to be nonlinear,33 i.e., the same amplitude difference does not have the same meaning for low and high ASSR amplitude ranges. However, this well-known property did not noticeably affect our results; at least, it did not prevent clinical utility. Again, the reason may be that we did not aim to estimate hearing thresholds, but only hearing classes. Because ASSRs use narrowband stimuli even with several carriers, the total amount of energy delivered to the inner ear is much lower than when broadband BAEP clicks are used. The risk of damaging inner-ear hair cells with high-intensity stimuli should thus be rather low.

Analysis of several carrier frequencies surprisingly did not result in a better estimation of the hearing class. In fact, the highest AUC values were achieved using 1000-Hz stimulation. This means that it may not be necessary to test different frequencies, either separately or combined into a single stimulus, which also saves measurement time. A potential explanation may be that neither pre-existing hearing deterioration nor deterioration due to nerve damage is limited to narrow-frequency bands. Instead, deterioration at the different carrier frequencies may be highly correlated.

Furthermore, patients with VS typically display a sensorineural hearing loss of higher frequencies, i.e., above 1000 Hz.13 The remaining hearing of frequencies in a band around 1000 Hz could then represent the main sensory input with information about high-frequency content. Also,
as a result of losing this information, patients trying to understand language-related sounds may lack the last crucial components, with a severe impact on hearing quality. Remaining hearing may therefore be highly sensitive to further impairment, especially at approximately 1000 Hz. This hypothesis, however, also suggests that the diagnostic value of the ASSR amplitude with a 1000-Hz carrier may be specific for this patient population and not easily transferable to those with other reasons for hearing impairment.

Furthermore, due to the lack of data regarding the intraoperative time course, it remains an open question whether acute damage to the nerve is detected first at 1000 Hz, whether there is a uniform decrease of ASSR amplitudes over a broader frequency range, or whether there are different patterns entirely. Due to the principle of ASSR generation and measurement, changes will only occur after the hearing nerve has been damaged, comparable to BAEPs. Acute severe damage can thus not be predicted. In cases of less damage cumulating over the course of manipulation, the sensitivity of ASSRs may enable early detection. A change of surgical strategy or, for example, vasoactive treatment, may then contribute to avoidance of a permanent clinically significant deficit.

The learning curve of ASSRs should not be very steep, because the analysis is objective and fully automated. The interpretation of intraoperative changes should benefit from the experience with BAEPs. However, studies using continuous intraoperative recordings are still required. The present results now serve as a basis and motivation to implement such investigations.

Limitations

The main limitation of our study is the restriction to BAEP and ASSR evaluation at the beginning and end of the surgical procedure. We aimed to avoid interference with the clinical routine, which prevented ASSR recordings during surgery. This approach certainly overlooks the time course of both ASSRs and BAEPs and their relationship to the postoperative hearing quality. Furthermore, although we investigated a series of consecutive patients, data analysis and comparison with the clinical outcome was retrospective. The interpretation of BAEPs, especially, may have thus been suboptimal, although patients’ hearing quality was not known to the investigator. For the interpretation of ASSRs, the retrospective nature is not expected to have an impact, because ASSRs are analyzed using objective and automated procedures. Future prospective studies, which should include at least intermittent intraoperative ASSR measurements, are required to further evaluate the clinical value of ASSRs for intraoperative monitoring of the auditory nerve. Similarly, more sophisticated methods for ASSR analysis should be evaluated, such as weighted averaging and taking spectral harmonics as well as response phases into account.21

Conclusions

Our results demonstrate that ASSRs robustly estimate hearing quality in total intravenous anesthesia, even when using short measurement durations. The method showed a diagnostic performance that exceeds conventional BAEP monitoring and enables objective and automated evaluation. On the basis of these findings, continuous intraoperative auditory monitoring could become a promising alternative or adjunct to BAEP monitoring.

Acknowledgments

We thank Mrs. Anke Dietz for performing the intraoperative recordings and numerous pre- and postprocessing steps.

References


34. Raftopoulos C, Abu Seriib B, Duprez T, Docquier MA, Guérit JM: Microsurgical results with large vestibular schwannomas with preservation of facial and cochlear nerve function as the primary aim. *Acta Neurochir (Wien)* 147:697–706, 2005


