Published in final edited form as:

Otol Neurotol. 2019 March ; 40(3): e215–e224. doi:10.1097/MAO.0000000000002092.

Cervical Vestibular Evoked Myogenic Potentials in Meniere's Disease: A comparison of response metrics

Kimberley S. Noij, MD¹, Barbara S. Herrmann, PhD^{2,3}, John J. Guinan Jr., PhD^{2,4}, Steven D. Rauch, MD^{1,2}

¹Department of Otolaryngology, Massachusetts Eye and Ear Infirmary

²Department of Otolaryngology, Harvard Medical School

³Department of Audiology, Massachusetts Eye and Ear Infirmary

⁴Eaton Peabody Lab, Massachusetts Eye and Ear Infirmary

Abstract

Objective: The cervical vestibular evoked myogenic potential (cVEMP) has been used to evaluate patients with Meniere's Disease (MD). Studied cVEMP metrics include: amplitude, threshold, frequency tuning and interaural asymmetry ratio (IAR). However, few studies compared these metrics in the same set of MD patients, and methodological differences prevent such a comparison across studies. This study investigates the value of different cVEMP metrics in distinguishing one set of MD patients from age-matched controls.

Study design: Prospective study

Setting: Tertiary care center

Patients: Thirty patients with definite unilateral MD and twenty-three age-matched controls were prospectively included. All underwent cVEMP testing at 500, 750, 1000 and 2000 Hz on each side. Ears were separated into three groups: affected MD, unaffected MD and control.

Main outcome measures: Sound level functions were obtained at each frequency, and normalized peak-to-peak amplitude (VEMPn), VEMP inhibition depth (VEMPid), threshold, frequency-tuning ratio and IAR were calculated. For all metrics, the differentiation between MD and control ears was compared using ROC curves.

Results: 500 Hz cVEMP threshold, VEMPn and VEMPid were similarly good at distinguishing affected MD ears from healthy ears, with ROC area under the curves (AUCs) of >0.828 and optimal sensitivities and specificities of at least 80% and 70%. Combinations of these three metrics yielded slightly larger AUCs (>0.880). Tuning ratios and IAR were less effective in separating healthy from affected ears with AUCs ranging from 0.529–0.720.

Corresponding Author Steven D. Rauch, MD, 243 Charles Street, Boston, MA 02114, Phone number: 617-573-3644, Fax number: 617-573-3914, Steven_Rauch@meei.harvard.edu.

Conflict of interest: None

This study was approved by the Human Studies Committee of the Massachusetts Eye and Ear Infirmary. Protocol number: 13–097H. Principal Investigator: Steven D. Rauch.

Conclusion: The cVEMP metrics most useful in distinguishing MD patients from healthy controls are threshold, VEMPn and VEMPid, using 500 Hz stimuli.

Keywords

cVEMP; Meniere's disease

Introduction

The cervical vestibular evoked myogenic potential (cVEMP) is used to evaluate the function of the saccule and inferior vestibular nerve. During this test, the saccule is acoustically or mechanically stimulated resulting in a primarily ipsilateral inhibition of the sternocleidomastoid (SCM) muscle that is measured by SCM electromyography (EMG) (1,2).

cVEMPs have been studied in patients with superior semicircular canal dehiscence (SCD) syndrome and Meniere's disease (MD) (3-6). In SCD patients, the cVEMP seems a promising screening tool achieving high sensitivities and specificities (7). The value of cVEMP in differentiating MD patients from healthy subjects is less clear. Most studies agree that MD patients have lower cVEMP amplitudes (8-14), higher thresholds (3, 14), altered frequency tuning (3,8,11,13-17) and higher interaural asymmetry ratios (IARs) (8, 14, 18). There is, however, considerable overlap in cVEMP metrics between MD and healthy subjects. The variety of methods used to obtain, analyze and evaluate cVEMPs makes it challenging to compare studies. For example, most studies investigating cVEMP in MD patients did not describe methods to normalize cVEMPs to account for differences in muscle contraction (3, 10, 13, 15, 16, 18-21). Muscle contraction has a large effect on cVEMP amplitude (1, 22-27). Finding either a difference or no difference between groups using non-normalized amplitudes could potentially be due to differences in muscle contraction between groups. Studies that failed to normalize cVEMPs for muscle contraction rarely compare muscle contractions between groups, which makes these studies difficult to interpret and compare. Furthermore, many studies evaluating MD patients lack an age-matched control group (3, 9, 13, 16, 17, 19). cVEMP amplitudes decrease, and thresholds increase with age (22, 28-30). Because similar changes are seen in MD patients, it is essential to include an age-matched control group.

Several studies of cVEMP metrics in MD patients describe sensitivities and specificities and/or provide receiver operating characteristic (ROC) curves and areas under the curves (AUCs) (8, 11, 15, 18, 19, 21). The evaluated metrics differed across studies and included 500/1000Hz or 1000/500 Hz amplitude ratios (8, 11, 18), shifting of the most sensitive frequency (lowest threshold) to 1000 Hz (15), a combination of amplitudes, 500/1000 Hz ratio, IARs and audiogram data (18), and assessing cVEMPs as "normal" versus "abnormal" (19, 21).

The use of many different metrics, inconsistent use of normalization across studies, and the lack of healthy age-matched controls in many studies make it difficult to compare the results from these studies to determine the value of cVEMPs in detecting saccular dysfunction in MD patients. The present study compares the usefulness of different cVEMP metrics in

differentiating one set of MD patients from healthy age-matched controls by providing sensitivities and specificities for each metric. For reliable comparison, all metrics were obtained in the same patients and controls, using the same methods.

Methods

Subjects

Thirty patients with definite unilateral Meniere's disease, per Lopez et al. 2015, and twenty-three age-matched controls were prospectively included (31). The Lopez et al. criteria for definite Meniere's Disease include: A) The presence of two or more spontaneous episodes of vertigo, each lasting 20 minutes to 12 hours, B) Audiometrically documented low- to medium frequency sensorineural hearing loss in one ear, defined as a bone-conducted threshold of at least 30 dB HL at each of two contiguous frequencies below 2000 Hz, on at least one occasion, C) Fluctuating hearing loss, tinnitus or aural fullness in the affected ear and D) Not better accounted for by another diagnosis (31). Patients who previously received invasive treatment, including intratympanic gentamicin or corticosteroid injections and surgery were excluded. Exclusion criteria for the healthy control group were a self-reported history of hearing loss, vertigo and/or balance problems. This study was approved by the Human Studies Committee of the Massachusetts Eye and Ear Infirmary (#13-097H, PI: S.D. Rauch).

Audiometry

All subjects underwent pure tone audiometry with air- and bone-conduction tonal thresholds measured at octave frequencies from 250 to 4000 Hz. If the difference between air- and unmasked bone-conduction thresholds was larger than 10 dB HL, bone-conduction thresholds were masked. The air-bone gap (ABG) was calculated at each tested frequency by subtracting the bone-conduction threshold from the air-conduction threshold. Patients and healthy subjects with an ABG >10 dB at any of the frequencies used for cVEMP testing (500, 750, 1000 and 2000 Hz) were excluded from this study, because this could indicate the presence of middle ear pathology, which can influence cVEMP outcomes (decrease amplitude and increase threshold).

cVEMP

A custom-programmed evoked potential system was used to generate tone bursts and record cVEMPs. The time between the start of symptoms and the obtained cVEMP varied from 12 to 333 months (average: 83 months). During cVEMP testing, subjects sat up straight and turned their head away from the stimulated ear to contract the ipsilateral SCM (i.e. the left ear alone was acoustically stimulated while contracting the left SCM). EMG activity was recorded from four surface electrodes: a non-inverting electrode on the middle belly of each SCM, an inverting electrode at the midpoint between SCM attachments to the sternum, and a ground electrode on the midline forehead. Ipsilateral SCM EMG was monitored while subjects contracted their SCM to produce >45 μ V root mean square (rms) EMG. This minimum muscle contraction of 45 μ V rms was chosen based on a previous study, which concluded that for contractions that produced 45–300 μ V rms EMG, the contraction strength had little effect on the metrics used in the current study (normalized peak-to-peak amplitude,

VEMP inhibition depth and cVEMP detection).(22) EMG activity was amplified and bandpass filtered between 10 and 750 Hz using the bioamplifier of the Eclipse EP15 (Interacoustics). The output of the bio-amplifier was sampled at 50 kHz with a 16-bit analog-to-digital converter (National Instruments).

cVEMPs were elicited using 500, 750, 1000 and 2000 Hz tone bursts generated by custom-programmed evoked-potential software (National Instruments 16-bit digital I/O board) using a Blackman gating function with two cycle rise and fall times (4.0 ms at 500 Hz, 2.5 ms at 750 Hz, 2 ms at 1000 Hz, 1 ms at 2000 Hz) and no plateau. Tone bursts were presented monaurally via circumaural headphones (Telephonics TDH-49) at a repetition rate of 13 bursts/s. At least 200 cVEMP responses were obtained and averaged for each recording. In the MD group, tone bursts were presented at 103, 113, 123 and 128 dB peak sound pressure level (peSPL), while in the healthy control group, tone bursts were presented at 93, 103, 113 and 123 dB peSPL (123 dB peSPL is equivalent to 90 dB nHL). This protocol difference between the MD and healthy control group was chosen because MD patients generally have higher cVEMP thresholds. Numerical values for sound level, frequency and side were placed in a table which was randomized (separately for each subject) across all three variables and which set the presentation order for these variables.

cVEMP metrics

The collection of sound level functions at all frequencies and the saving of all individual responses allowed for the calculation of multiple metrics, including VEMP_n (normalized peak-to-peak amplitude), VEMP_{id} (VEMP inhibition depth), threshold, frequency tuning and interaural asymmetry ratio (IAR).

VEMP_n

We have recently completed a study comparing different methods of normalizing cVEMPs and found that several methods were equivalently good while others were inferior (van Tilburg et al. 2018, *Ear and Hearing*, in Press). Here we make use of these results and use just one of the good methods which serves a proxy for all of the good normalization methods. Van Tilburg et al., (in Press) found that the best cVEMP normalizations used EMG quantification from individual-trace EMGs, either by averaging the EMG and applying this average EMG to normalize the cVEMP waveform, or by applying the EMG trace by trace. Rectified and rms EMG were usually equivalent. In addition, EMG measurement windows close to the time of the cVEMP response were best, and inclusion of the time when cVEMP occurred had a negligible effect because the cVEMP modulation is a small fraction of the EMG amplitude and the EMG amplitude near the cVEMP time is the most relevant. In contrast, normalizing the average cVEMP amplitude by an EMG metric derived from the averaged cVEMP waveform (e.g. from the 20 or 50 ms period before the VEMP response) was distinctly inferior. Since the good normalization methods were equivalent, we use only one of them.

Trace-by-trace normalization was performed, meaning that each raw EMG trace was divided by the overall, 77 ms long rms value of the same trace (32). VEMP_n was the amplitude

difference between the first positive (P1) and the first negative (N1) peak of the average normalized cVEMP waveform (32).

VEMPid

VEMPid is a metric that estimates the percentage reduction in spike-rate of SCM motoneurons elicited by acoustic stimulation of the saccule. This metric uses a template correlation method (34). The VEMPid is larger when the cVEMP response is strong and smaller when the cVEMP response is weaker. In this study, VEMPid was calculated using a generic template created from cVEMP responses of healthy subjects (33). The first step in computing VEMPid, is to calculate template correlation values (TCVs) using the point-by-point correlation of each individual cVEMP trace with the generic template. VEMPid was then calculated by dividing the mean of all (at least 200) TCVs by the standard deviation of the TCVs and multiplying by 0.2. While the original VEMPid calculation used a subject-specific template, it was later determined that VEMPid can be calculated using a generic template if the generic template's latency is set to the latency of the patient's response (33, 34). In patients, the use of a generic is preferred over a subject-specific template because patients may not have a robust cVEMP suitable for a template.

Threshold

After cVEMPs at the four preset sound levels were obtained, additional 5dB steps were obtained and if necessary repeated, to find the threshold (the lowest sound level at which a cVEMP was present). If no cVEMP was present at our equipment limit (133 dB peSPL), threshold was defined as 10 dB higher than this limit.

Frequency tuning

cVEMP tuning was evaluated using VEMPn and threshold. 500/1000 Hz ratios were calculated using these metrics. VEMPid was not used to calculate 500/1000 Hz, because this metric is not suitable for such a computation (VEMPid can be negative and/or very small).

IAR

The IAR was calculated using VEMPn. In previous studies a variety of formulae have been used to compute the IAR (8, 12, 16, 18-21, 35, 36) and we decided to use all options in the current study:

Healthy control group:

$$\begin{aligned} \text{Control L-R: } IAR &= 100 * \frac{VEMP_{nleft} - VEMP_{nright}}{VEMP_{nleft} + VEMP_{nright}} \\ \text{Control R-L: } IAR &= 100 * \frac{VEMP_{nright} - VEMP_{nleft}}{VEMP_{nright} + VEMP_{nleft}} \\ \text{Control Abs: } IAR &= 100 * \frac{|VEMP_{nleft} - VEMP_{nright}|}{VEMP_{nleft} + VEMP_{nright}} \end{aligned}$$

MD group:

$$\begin{aligned} \text{MD U-A: } IAR &= 100 * \frac{VEMPn_{unaffected} - VEMPn_{affected}}{VEMPn_{unaffected} + VEMPn_{affected}} \\ \text{MD Abs: } IAR &= 100 * \frac{|VEMPn_{unaffected} - VEMPn_{affected}|}{VEMPn_{unaffected} + VEMPn_{affected}} \end{aligned}$$

Controls Abs and MD Abs represent the absolute values used in the corresponding formulae.

Data analysis

MD and control-group ages were compared using an independent-samples t-test. Audiometric data and cVEMP metrics (VEMPn, VEMPid, threshold, 500/1000 Hz ratios and IAR) among affected, unaffected and healthy control ears were compared using full-factorial analyses of variance (ANOVA) with group and frequency considered fixed factors and subject considered a random factor. *Post-hoc* pairwise group comparisons used Bonferroni adjustment for multiple comparisons. The ability of each cVEMP metric to distinguish between affected and healthy control ears was assessed with ROC curves. Statistical analyses were performed using SPSS (version 24.0; Chicago, IL). A *p* value of <0.05 was considered statistically significant. When a Bonferroni correction was used, the significance criterion was 0.05 divided by the number of comparisons.

Results

Patient characteristics

Thirty patients with unilateral definite Meniere's disease (31) were included (12 female, mean age: 55.6 years; range: 28–75). Ears were separated into two groups: affected and unaffected. Additionally, twenty-three age-matched healthy controls were included (12 female, mean age: 54.8 years; range: 33–73). No significant difference in age was found between control and MD groups (*p*=0.783).

Audiograms

Bone conduction thresholds were recorded for all groups (Fig. 1A). Average low-frequency bone-conduction thresholds were calculated using 250, 500 and 1000 Hz data (Fig. 1B). Affected ears had significantly higher average low-frequency bone-conduction thresholds compared to the unaffected and healthy control ears, as did unaffected vs. control ears (*p*<0.001 for each comparison).

cVEMP

On average, VEMPn and VEMPid were largest in control ears, next largest in unaffected ears and smallest in affected ears (Fig. 2). To evaluate differences in VEMPn and VEMPid among the three groups, we used the highest sound level available for all groups (123 dB peSPL).

For VEMPn, there was a significant interaction between group and frequency (*F*=2.979, *p*=0.007; Fig. 2). At 500, 750 and 1000 Hz, VEMPn was significantly higher in control vs. affected ears (*p*<0.001 for each frequency; Bonferroni-adjusted significance criterion is 0.0167). For 500 Hz, VEMPn was significantly higher in control vs. unaffected ears

($p < 0.001$), and the difference observed between affected and unaffected ears (Fig. 2) did not reach significance ($p = 0.064$). At 750 and 1000 Hz, differences that did not reach Bonferroni-adjusted significance were also observed for control vs. unaffected ears ($p = 0.044$ and $p = 0.052$) and for affected vs. unaffected ears ($p = 0.031$ and $p = 0.127$). At 2000 Hz, no significant differences in VEMPn were found for any group combinations.

For VEMPid, there was also a significant interaction between group and frequency ($F = 2.170$, $p = 0.045$; Fig. 2). At 500, 750 and 1000 Hz, VEMPid was significantly larger in control vs. affected ears ($p < 0.001$ for each frequency; Bonferroni-adjusted significance criterion is 0.0167). For 500 and 1000 Hz, VEMPid was significantly higher in control vs. unaffected ears ($p < 0.001$ and $p = 0.011$), while the difference between affected and unaffected ears at these frequencies did not reach significance ($p = 0.078$ and 0.129). For 750 Hz, the difference between control and unaffected ears ($p = 0.041$) and between affected and unaffected ears ($p = 0.038$) did not meet the Bonferroni-adjusted significance criterion. At 2000 Hz, no significant differences in VEMPid were found for any group combinations.

On average, cVEMP thresholds were highest in the affected ears, followed by unaffected ears and were lowest in control ears (Fig. 3). A significant interaction was found between group and frequency ($F = 0.268$, $p = 0.014$). For 500, 750 and 1000 Hz, threshold was significantly lower in control vs. affected ears ($p < 0.001$ for each frequency) and for control vs. unaffected ears ($p < 0.001$, $p = 0.003$ and $p = 0.002$ respectively). No significant difference in threshold was found between affected and unaffected ears at any frequency. At 2000 Hz no significant threshold differences were found for any group combinations.

Tuning

Half of the affected ears had a higher threshold at 500 Hz than at 1000 Hz (Fig. 4A). In the unaffected ears thresholds at 500 and 1000 Hz were most often equal (Fig. 4A). In contrast, most control subjects had a higher threshold at 1000 Hz than at 500 Hz (Fig. 4A). For the 500/1000 Hz threshold ratio (Fig. 4B), a significant effect of group was found ($F = 4.300$, $p = 0.016$), and pairwise comparisons revealed significantly higher threshold ratios in affected vs. control ears ($p = 0.005$; Bonferroni adjusted criterion is 0.0167). The differences in threshold ratio between affected and unaffected ($p = 0.173$) or unaffected and control ears ($p = 0.157$) were not significant. For VEMPn tuning, 123 dB peSPL recordings were used to calculate the 500/1000 Hz ratio and no significant differences were found between groups ($F = 1.348$, $p = 0.264$) (Fig. 4C).

IAR

VEMPn IAR was studied using the 123 dB peSPL recordings. ANOVA found no significant effect of frequency ($F = 0.213$, $p = 0.887$), so data were pooled across frequency. To include all of the calculations used in previous reports, five IAR formulae were used (see Methods). Formulae using ear-differences always yielded smaller IARs than formulae using the absolute value of the difference (Fig. 5). For both types of formulae, the MD group had higher IARs than the control group at 500, 750 and 1000 Hz. In controls, the right ears had slightly higher VEMPn's, and IAR reversed sign for formulae Control L-R vs. Control R-L. Comparing control and MD groups, Control L-R vs. MD U-A groups were significantly

different ($p < 0.001$, Bonferroni-adjusted criterion is 0.0167), but Control R-L vs. MD U-A groups ($p = 0.077$), Control Abs vs. MD U-A groups ($p = 0.018$) and Control Abs vs. MD Abs were not ($p = 0.321$).

ROC curves

To evaluate the ability of the cVEMP metrics to distinguish affected MD ears from control ears we created ROC curves. The 2000 Hz data were not included because this frequency did not separate the subject groups (Figs. 2, 3). Based on the ROC area under the curve (AUC), 500 Hz was the best frequency to distinguish groups for threshold (AUC=0.828), VEMPn (AUC=0.846) and VEMPid (AUC=0.854) (Fig. 6). The sensitivity and specificities that could be achieved with these metrics were also similar, with sensitivities of 80.0, 82.6 and 80.4% respectively, and corresponding specificities of 76.1, 70.0 and 73.3% (Fig. 6). Both threshold tuning ratio and VEMPn tuning ratio were less valuable metrics, with AUCs of 0.697 and 0.578 (Fig. 6). The IAR was also inferior as evaluated by ROC curves for 1) Control L-R vs. MD U-A, 2) Control R-L vs. MD U-A, 3) Control Abs vs. MD U-A, and 4) Control Abs vs. MD Abs. The 500, 750 and 1000 Hz AUCs for each combination were 1) 0.659, 0.720, 0.646, 2) 0.533, 0.622, 0.641, 3) 0.639, 0.567, 0.557 and 4) 0.529, 0.525 and 0.600 (ROC curves not shown).

The tuning-ratio and IAR analyses used only 123 dB peSPL recordings. However, the MD group also had 128 dB peSPL recordings available. We calculated ROC curves from tuning ratios and IARs comparing MD 128 dB peSPL recordings vs. control 123 dB peSPL recordings. These analyses yielded results similar to those from all-123-peSPL comparisons (data not shown).

Because the 500 Hz data for threshold, VEMPn and VEMPid gave the best results, we explored if a combination of these three metrics would improve the ability to distinguish MD-affected from control ears. The metrics were combined using logistic regression models. The following combinations were used: 1) Threshold and VEMPn, 2) Threshold and VEMPid, 3) Threshold, VEMPn and VEMPid, providing the following logistic regression models:

1. $\beta_1 = -7.428 + (\text{Threshold} \times 0.078) + (\text{VEMPn} \times -7.291)$
2. $\beta_2 = -7.480 + (\text{Threshold} \times 0.072) + (\text{VEMPid} \times -0.514)$
3. $\beta_3 = -7.258 + (\text{Threshold} \times 0.072) + (\text{VEMPn} \times -1.888) + (\text{VEMPid} \times -0.396)$

These equations were used to calculate the beta (β) for each subject. The β values were then used to create ROC curves for each combination of metrics, yielding the following AUCs: 1) 0.880, 2) 0.889, and 3) 0.891 (Fig. 6). Corresponding sensitivities and specificities were: 1) 80% and 73.9%, 2) 83.3% and 80.4% 3) 83.3% and 80.4%.

Discussion

The purpose of this study was to investigate how well different cVEMP metrics separate MD from normal ears. To reliably compare these metrics, all were obtained in the same MD

patients and age-matched controls, with consistent methods used to obtain and analyze the cVEMP.

For distinguishing affected-MD from control ears, cVEMP threshold, VEMPn and VEMPid metrics at 500 Hz did a similarly good job based on their ROC curves, with AUCs of >0.828 and sensitivities and specificities of at least 80% and 70%. Combinations of the best three metrics (500 Hz: cVEMP threshold, VEMPn and VEMPid) yielded slightly larger AUCs (>0.880) sensitivities and specificities (Fig. 6). Results equivalent to our VEMPn results are expected to be produced by any cVEMP normalization done by rms or rectified EMG from individual traces, whereas normalizations by an EMG metric derived from the averaged cVEMP waveform are expected to be less effective (van Tilburg et al. 2018, In Press). Tuning ratios and IAR were less effective in separating healthy from affected ears with AUCs 0.529–0.720 (Fig. 6).

Although VEMPid has produced results that may seem similar to VEMPn, these are completely different metrics. VEMPn measures the size of a cVEMP while correcting for the effect of muscle contraction and VEMPid estimates the inhibition depth of the saccule based on a computational model applied to the VEMP measurements (34). Although the metrics may produce similar results, the VEMPid is easier to interpret, because it goes down to zero when no response is present (22, 33). Although yet to be proven, VEMPid has the potential of reducing inter-tester variability in determining VEMP threshold. VEMPid measurements may also aid in determining when to stop recording, which would make testing time shorter and decrease patient discomfort (33).

In the current study, well-defined MD patients were compared with healthy controls to investigate how well cVEMP detects saccular dysfunction. Clinically, the ability to detect saccular dysfunction with cVEMPs could potentially be used to identify patients whose symptoms are suggestive, but not diagnostic for MD, as well as to track saccular function over time (38). It appears that different stages of MD may be accompanied by different cVEMP outcomes, with an increase in cVEMP thresholds for patients with more advanced disease (39, 40). Therefore, early stage MD patients may require separate consideration from those included here. Our comparison of well-defined MD patients with healthy controls is a step in this direction and suggests that cVEMP recordings may yield valuable information. A good next step would be to follow unilateral-MD patients over time to determine the value of cVEMP metrics obtained from the non-symptomatic ear in predicting which patients will develop MD in this ear. To investigate how well the cVEMP metrics used in this study are in detecting Meniere's ears from those with other clinical problems, more work is needed in which a similar analysis is performed in patient groups with non-MD clinical presentations. For all of these potential uses of cVEMP measurements, the present study provides a first step in identifying which cVEMP metrics to focus on in future studies.

Comparison with previous studies

Previous studies evaluating AUCs and/or sensitivities and specificities to separate MD from healthy ears used a variety of metrics (8, 11, 15, 18, 19, 21). It is difficult to compare results of these studies with our results because different patient groups and methods were used to

obtain and analyze the cVEMP. Nonetheless, an overview of the analyzed metrics and results can be presented:

1. 500/1000Hz or 1000/500 Hz amplitude ratio (Maxwell et al. 2017: AUC <0.7; Salviz et al. 2015 and 2016: AUC=0.731, sensitivity=76% and specificity=80%.) (8, 11, 18)
2. Shifting of most sensitive frequency (i.e. lowest threshold) to 1000 Hz (Zhu et al. 2014: sensitivity=47%, specificity=64%) (15)
3. A combination of amplitude 500/1000 Hz ratio and asymmetry ratio (Maxwell et al. 2017: AUC=0.814) and audiogram data (Maxwell et al. 2017: AUC=0.906, sensitivity=64%, specificity=93%) (18)
4. Qualifying cVEMP as normal versus abnormal, “abnormal”: Decreased or absent response (Egami et al. 2013: sensitivity=50%, specificity=48.9%) (19); absent response and/or interaural asymmetry >34% (Lamounier et al. 2017: sensitivity=63.5–63.6%, specificity=84.6–93.7%) (21)

The Maxwell et al. study results are the most similar to ours. The AUCs found by Maxwell et al., using a combination of cVEMP and oVEMP 500/1000 Hz amplitude ratio as well as 500 Hz and 1000 Hz cVEMP asymmetry ratio, are similar to the AUCs we found for 500 Hz threshold, VEMPn and VEMPid as separate metrics (and which required cVEMP testing only) (18). Maxwell et al. performed an additional analysis adding audiometric data to the previously mentioned combination, yielding an AUC of 0.906. We did not include audiometric data in our ROC analyses, because patients were pre-selected based on their audiometric data (see Methods, Lopez et al. criteria for definite MD) (31). Of note, the Maxwell et al. study included patients with definite, probable and possible MD and did not mention the use of a normalization technique to correct for differences in muscle contraction (18).

Comparison of our results to Egami et al. 2013 and Lamounier et al. 2017 is difficult. Their rules of what is normal or abnormal are arbitrary, which makes their outcomes challenging to interpret (19, 21). Our ROC curves created with raw data of each metric and directly comparing MD with normal groups provides clearly interpretable results. Other institutions can repeat our analysis and compare their results to ours.

Consistent with previous reports, we found a shift in frequency tuning in MD ears. Most affected ears had higher thresholds at 500 Hz than at 1000 Hz, while the majority of controls had higher thresholds at 1000 Hz than at 500 Hz (3, 15). However, for 500/1000 Hz ratios, a significant difference was only found for threshold ratios between affected and control ears. Control ears had higher VEMPn 500/1000 Hz ratios compared to MD-affected and MD-unaffected ears, but this did not reach the significance level. In contrast to our findings, previous reports did find significant differences between affected, unaffected and/or control ears regarding tuning (8, 12, 18). Differences between our study and that by Maxwell et al. have been previously described and the methodological differences could explain the difference in findings. Similarly, Taylor et al. did not mention the use of a normalization technique to control for differences in muscle contraction, which might explain the

disconcordance with our results (14). In contrast, Salviz et al. 2016 only included patients with definite MD, used an age-matched control group and normalized for muscle contraction (8). It is unclear why our results differ from theirs.

A report from Zhu et al. 2014 [15] included patients with non-MD vestibular pathologies (including migraine related dizziness, BPPD, chronic subjective dizziness syndrome, vestibular neuritis and labyrinthitis) to study the effect of cVEMP threshold tuning. This study found that a “tuning curve shift” defined as the most sensitive frequency being 1000 Hz provided an AUC of 0.560. This study did not look at 500/1000 Hz amplitude ratios which makes it difficult to compare to the other studies. The low AUC value found does not make the use of frequency tuning more promising, and indicates that frequency tuning changes similar to those in MD patients can occur in other vestibular pathologies.

The IAR is a widely used metric to evaluate MD patients, so we included it in this study even though we do not recommend its use in MD patients. Both the current study and previous studies have found that cVEMPs from the asymptomatic ears of unilateral MD patients are different from cVEMPs from normal ears, and the difference is in the direction of the asymptomatic ears being more like symptomatic MD ears, suggesting possible occult changes in the asymptomatic ear of some unilateral MD patients (3). The abnormal cVEMPs in asymptomatic ears of unilateral MD patients would reduce the IAR and make MD-IARs more like IARs from normal ears and thus less useful in detecting MD ears from normal ears. This hypothesis was confirmed by the low AUCs calculated from IARs which ranged from 0.525 to 0.720. The exact AUC value depends on which IAR formula was used, but all of these values are inferior to those obtained by other cVEMP metrics. For this reason, and because approximately one quarter of MD patients suffer from bilateral disease, the IAR should be interpreted with caution and is not used in our clinic (3, 41-47).

Conclusion

The 500 Hz cVEMP threshold, normalized peak-to-peak amplitude (VEMP_n) and VEMP inhibition depth (VEMP_{id}) are most valuable in separating Meniere’s disease (MD) from healthy control ears. The diagnostic accuracy of the 500/1000 Hz threshold and VEMP_n ratio, as well as the IAR, are inferior and therefore these metrics are not recommended for evaluation of patients with clinically definite MD.

Acknowledgements

This work was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Advancing Translational Sciences, National Institutes of Health Award UL1TR002541) and financial contributions from Harvard University and its affiliated academic healthcare centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic healthcare centers, or the National Institutes of Health.

The authors would like to thank Raymond van de Berg from the Department of Otolaryngology and Head & Neck Surgery of Maastricht University Medical Center for reviewing a pre-published version of this manuscript and suggesting edits that improved its precision and comprehensibility.

External funding source: None

References

1. Colebatch JG, Halmagyi GM, Skuse NF. Myogenic potentials generated by a click-evoked vestibulocollic reflex. *J Neurol Neurosurg Psychiatry* 1994;57:190–7. [PubMed: 8126503]
2. Curthoys IS. A critical review of the neurophysiological evidence underlying clinical vestibular testing using sound, vibration and galvanic stimuli. *Clin Neurophysiol* 2010;121:132–44. [PubMed: 19897412]
3. Rauch SD, Zhou G, Kujawa SG, Guinan JJ, Herrmann BS. Vestibular evoked myogenic potentials show altered tuning in patients with Meniere's disease. *Otol Neurotol* 2004;25:333–8. [PubMed: 15129114]
4. Rauch SD, Silveira MB, Zhou G, et al. Vestibular evoked myogenic potentials versus vestibular test battery in patients with Meniere's disease. *Otol Neurotol* 2004;25:981–6. [PubMed: 15547430]
5. Hunter JB, Patel NS, O'Connell BP, et al. Cervical and Ocular VEMP Testing in Diagnosing Superior Semicircular Canal Dehiscence. *Otolaryngol Head Neck Surg* 2017;156:917–23. [PubMed: 28168887]
6. Fife TD, Colebatch JG, Kerber KA, et al. Practice guideline: Cervical and ocular vestibular evoked myogenic potential testing: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* 2017;89:2288–96. [PubMed: 29093067]
7. Noij KS, Duarte MJ, Wong K, et al. Toward Optimizing Cervical Vestibular Evoked Myogenic Potentials (cVEMP): Combining Air-Bone Gap and cVEMP Thresholds to Improve Diagnosis of Superior Canal Dehiscence. *Otol Neurotol* 2018;39:212–220. [PubMed: 29210947]
8. Salviz M, Yuce T, Acar H, Taylan I, Yuceant GA, Karatas A. Diagnostic value of vestibular-evoked myogenic potentials in Meniere's disease and vestibular migraine. *J Vestib Res* 2016;25:261–6. [PubMed: 26890427]
9. Johnson SA, O'Beirne GA, Lin E, Gourley J, Hornibrook J. oVEMPs and cVEMPs in patients with 'clinically certain' Meniere's disease. *Acta Otolaryngol* 2016;136:1029–34. [PubMed: 27121497]
10. Zuniga MG, Janky KL, Schubert MC, Carey JP. Can vestibular-evoked myogenic potentials help differentiate Meniere disease from vestibular migraine? *Otolaryngol Head Neck Surg* 2012;146:788–96. [PubMed: 22267492]
11. Salviz M, Yuce T, Karatas A, Balikci HH, Ozkul MH. Diagnostic value of frequency-associated vestibular-evoked myogenic potential responses in Meniere's disease. *Audiol Neurotol* 2015;20:229–36. [PubMed: 25966621]
12. Taylor RL, Wijewardene AA, Gibson WP, Black DA, Halmagyi GM, Welgampola MS. The vestibular evoked-potential profile of Meniere's disease. *Clin Neurophysiol* 2011;122:1256–63. [PubMed: 21168362]
13. Sandhu JS, Low R, Rea PA, Saunders NC. Altered frequency dynamics of cervical and ocular vestibular evoked myogenic potentials in patients with Meniere's disease. *Otol Neurotol* 2012;33:444–9. [PubMed: 22334161]
14. Taylor RL, Zagami AS, Gibson WP, et al. Vestibular evoked myogenic potentials to sound and vibration: characteristics in vestibular migraine that enable separation from Meniere's disease. *Cephalalgia* 2012;32:213–25. [PubMed: 22259049]
15. Zhu Y, McPherson J, Beatty C, et al. Cervical VEMP threshold response curve in the identification of Meniere's disease. *J Am Acad Audiol* 2014;25:278–88. [PubMed: 25032972]
16. Kim-Lee Y, Ahn JH, Kim YK, Yoon TH. Tone burst vestibular evoked myogenic potentials: diagnostic criteria in patients with Meniere's disease. *Acta Otolaryngol* 2009;129:924–8. [PubMed: 18979269]
17. Node M, Seo T, Miyamoto A, Adachi A, Hashimoto M, Sakagami M. Frequency dynamics shift of vestibular evoked myogenic potentials in patients with endolymphatic hydrops. *Otol Neurotol* 2005;26:1208–13. [PubMed: 16272944]
18. Maxwell R, Jerin C, Gurkov R. Utilisation of multi-frequency VEMPs improves diagnostic accuracy for Meniere's disease. *Eur Arch Otorhinolaryngol* 2017;274:85–93. [PubMed: 27422626]

19. Egami N, Ushio M, Yamasoba T, Yamaguchi T, Murofushi T, Iwasaki S. The diagnostic value of vestibular evoked myogenic potentials in patients with Meniere's disease. *J Vestib Res* 2013;23:249–57. [PubMed: 24284605]
20. Inoue A, Egami N, Fujimoto C, Kinoshita M, Yamasoba T, Iwasaki S. Vestibular Evoked Myogenic Potentials in Vestibular Migraine: Do They Help Differentiating From Meniere's Disease? *Ann Otol Rhinol Laryngol* 2016;125:931–937. [PubMed: 27557910]
21. Lamounier P, de Souza TSA, Gobbo DA, Bahmad F Jr. Evaluation of vestibular evoked myogenic potentials (VEMP) and electrocochleography for the diagnosis of Meniere's disease. *Braz J Otorhinolaryngol* 2017;83:394–403. [PubMed: 27397722]
22. Noij KS, Herrmann BS, Rauch SD, Guinan JJ Jr. Toward Optimizing Vestibular Evoked Myogenic Potentials: Normalization Reduces the Need for Strong Neck Muscle Contraction. *Audiol Neurotol* 2017;22:282–291. [PubMed: 29324449]
23. Akin FW and Murnane OD. Vestibular evoked myogenic potentials: preliminary report. *J Am Acad Audiol* 2001;12:445–52. [PubMed: 11699815]
24. Akin FW, Murnane OD, Panus PC, Caruthers SK, Wilkinson AE, Proffitt TM. The influence of voluntary tonic EMG level on the vestibular-evoked myogenic potential. *J Rehabil Res Dev* 2004;41:473–80. [PubMed: 15543465]
25. Isaacson B, Murphy E, Cohen H. Does the method of sternocleidomastoid muscle activation affect the vestibular evoked myogenic potential response? *J Vestib Res* 2006;16:187–91. [PubMed: 17538207]
26. Lim CL, Clouston P, Sheean G, Yiannikas C. The influence of voluntary EMG activity and click intensity on the vestibular click evoked myogenic potential. *Muscle Nerve* 1995;18:1210–3. [PubMed: 7659119]
27. Ochi K, Ohashi T, Nishino H. Variance of vestibular-evoked myogenic potentials. *Laryngoscope* 2001;111:522–7. [PubMed: 11224786]
28. Ochi K, Ohashi T. Age-related changes in the vestibular-evoked myogenic potentials. *Otolaryngol Head Neck Surg* 2003;129:655–9. [PubMed: 14663431]
29. Basta D, Todt I, Ernst A. Characterization of age-related changes in vestibular evoked myogenic potentials. *J Vestib Res* 2007;17:93–8. [PubMed: 18413902]
30. Singh NK, Kashyap RS, Supreetha L, Sahana V. Characterization of age-related changes in sacculocolic response parameters assessed by cervical vestibular evoked myogenic potentials. *Eur Arch Otorhinolaryngol* 2014;271:1869–77. [PubMed: 23982670]
31. Lopez-Escamez JA, Carey J, Chung WH, et al. Diagnostic criteria for Meniere's disease. *J Vestib Res* 2015;25:1–7. [PubMed: 25882471]
32. van Tilburg MJ, Herrmann BS, Guinan JJ Jr, Rauch SD. Normalization reduces intersubject variability in cervical vestibular evoked myogenic potentials. *Otol Neurotol* 2014;35:e222–7. [PubMed: 24836595]
33. Noij KS, van Tilburg MJ, Herrmann BS, Marciniak P, Rauch SD, Guinan JJ Jr. Toward Optimizing VEMP: Calculating VEMP Inhibition Depth With a Generic Template. *Ear Hear* 2018: doi: 10.1097/AUD.0000000000000579.
34. Prakash SR, Herrmann BS, Milojcic R, Rauch SD, Guinan JJ Jr. Evaluating Inhibition of Motoneuron Firing From Electromyogram Data to Assess Vestibular Output Using Vestibular Evoked Myogenic Potentials. *Ear Hear* 2015;36:591–604. [PubMed: 25811931]
35. Murofushi T, Nakahara H, Yoshimura E, Tsuda Y. Association of air-conducted sound oVEMP findings with cVEMP and caloric test findings in patients with unilateral peripheral vestibular disorders. *Acta Otolaryngol* 2011;131:945–50. [PubMed: 21563874]
36. Taylor RL, Bradshaw AP, Halmagyi GM, Welgampola MS. Tuning characteristics of ocular and cervical vestibular evoked myogenic potentials in intact and dehiscent ears. *Audiol Neurotol* 2012;17:207–18. [PubMed: 22472299]
37. Noij KS, Herrmann BS, Guinan JJ Jr, Rauch SD. Toward Optimizing cVEMP: 2000 Hz Tone Bursts Improve the Detection of Superior Canal Dehiscence (SCD) Poster #268, Association of Research in Otolaryngology MidWinter meeting 2018, San Diego CA.
38. van Tilburg MJ, Herrmann BS, Guinan JJ Jr, Rauch SD. Serial cVEMP Testing is Sensitive to Disease Progression in Meniere Patients. *Otol Neurotol* 2016;37:1614–1619. [PubMed: 27668791]

39. Timmer FC, Zhou G, Guinan JJ, Kujawa SG, Herrmann BS, Rauch SD. Vestibular evoked myogenic potential (VEMP) in patients with Ménière's disease with drop attacks. *Laryngoscope* 2006 5;116:776–9. [PubMed: 16652086]
40. Young YH, Huang TW, Cheng PW. Assessing the stage of Meniere's disease using vestibular evoked myogenic potentials. *Arch Otolaryngol Head Neck Surg* 2003;129:815–8. [PubMed: 12925337]
41. Paparella MM and Griebie MS. Bilaterality of Meniere's disease. *Acta Otolaryngol* 1984;97:233–7. [PubMed: 6720298]
42. Stahle J, Friberg U, Svedberg A. Long-term progression of Meniere's disease. *Am J Otol* 1989;10:170–3. [PubMed: 2750865]
43. Yazawa Y, Kitahara M. Bilateral endolymphatic hydrops in Meniere's disease: review of temporal bone autopsies. *Ann Otol Rhinol Laryngol* 1990;99:524–8. [PubMed: 2195960]
44. Tsuji K, Velázquez-Villaseñor L, Rauch SD, Glynn RJ, Wall C 3rd, Merchant SN. Temporal bone studies of the human peripheral vestibular system. Meniere's disease. *Ann Otol Rhinol Laryngol Suppl* 2000;181:26–31. [PubMed: 10821232]
45. House JW, Doherty JK, Fisher LM, Derebery MJ, Berliner KI. Meniere's disease: prevalence of contralateral ear involvement. *Otol Neurotol* 2006;27:355–61. [PubMed: 16639274]
46. Chaves AG, Boari L, Lei Munhoz MS. The outcome of patients with Menieres disease. *Braz J Otorhinolaryngol* 2007;73:346–50. [PubMed: 17684655]
47. Shojaku H, Watanabe Y, Yagi T, et al. Changes in the characteristics of definite Meniere's disease over time in Japan: a long-term survey by the Peripheral Vestibular Disorder Research Committee of Japan, formerly the Meniere's Disease Research Committee of Japan. *Acta Otolaryngol* 2009;129:155–60. [PubMed: 18607900]
48. Tilburg Van et al. "Normalizing cVEMPs: Which method is the most effective?" *Ear and Hearing* 2018, in press.

Figure 1. Average bone conduction threshold for each group and frequency (A). Low-frequency bone conduction thresholds (averaged across 250, 500 and 1000 Hz) for each group (B). Error bars indicate 95% confidence intervals.

Figure 2.

Average normalized peak-to-peak amplitude (VEMPn; top row) and VEMP inhibition depth (VEMPid; bottom row) for 500, 750, 1000 and 2000 Hz. Comparisons across groups are best done using the thick dashed lines, which show the highest sound level that was used on all groups (123 dB peSPL). The 123 dB peSPL data are also displayed in the panels on the right to facilitate comparison across groups. On average, VEMPn and VEMPid were largest in the control group, followed by unaffected Meniere's disease (MD) ears and were lowest for affected MD ears. Error bars indicate the 95% confidence intervals.

Figure 3.

Average cVEMP threshold for the control (open circles), unaffected (grey circles) and affected (black circles) groups at 500, 750, 1000 and 2000 Hz. Error bars indicate the 95% confidence intervals.

Figure 4.

cVEMP frequency tuning varied across groups. Group threshold tuning was compared by the fraction of ears with 500/1000 Hz ratios below, at, and above unity (A), and by average 500/1000 Hz threshold (B) and VEMPn (C) ratios. Error bars indicate the 95% confidence intervals.

Figure 5. Interaural asymmetry ratios (IAR) calculated in five ways (see Methods). Symbols indicate average IARs using the 123 dB peSPL data for all frequencies; error bars indicate 95% confidence intervals. See Methods for formulas: *Abs* indicates the absolute values of the differences were used.

Figure 6.

Receiver operating characteristic (ROC) curves displaying the ability of various cVEMP statistics to separate Meniere's from normal ears. Each panel shows the sensitivity of detecting Meniere's disease (the true-positive rate) versus the false-positive rate (1 - specificity) for 500, 750 and 1000 Hz data at 123 dB peSPL. Top-row ROC statistics are threshold, normalized peak-to-peak amplitude (VEMPn) and VEMP inhibition depth (VEMPid). Bottom-row ROC statistics are 500/1000 Hz threshold ratio, VEMPn 500/1000 Hz ratio and ROCs created by combinations of the top-row metrics. Because the 500 Hz top-row data yielded the largest AUCs, this frequency was used to evaluate combinations of these metrics (bottom right panel). In each panel, the area under the curve (AUC) and 95% confidence interval (CI) of each line are displayed.