Test–Retest Reliability of Video Head Impulse Test in Healthy Individuals and Individuals with Dizziness

Niraj Kumar Singh*
Rajeshwari Govindaswamy†
Nirmala Jagadish*

Abstract

Background: Video head impulse test (vHIT) is widely accepted as a test for the assessment of functional integrity of semicircular canals (SCCs). It allows for the evaluation of the functioning of all six SCCs independent of each other. It works on the principle of the vestibulo-ocular reflex (VOR). In individuals with vestibular pathologies, the VOR is impaired, and hence, the use of vHIT may provide vital information about the functional status of SCCs and the VOR pathway originating from them.

Purpose: In the recent past, studies reported excellent test–retest reliability of vHIT in healthy individuals. However, these studies used analysis of variance or the nonparametric counterpart Wilcoxon signed-rank test, which are insufficient statistical methods for conclusions about test–retest reliability. Further, because vHIT assesses VOR function in individuals with vestibular pathologies, it is important to assess test–retest reliability in the pathological group as well. Therefore, the present study aimed to evaluate test–retest reliability of vHIT in healthy individuals and individuals with vestibular pathology.

Research Design: Repeated measures.

Study Sample: Twenty healthy individuals with no history of vestibular pathology and 20 individuals with known vestibular pathology were included.

Data Collection and Analysis: Each participant underwent vHIT testing for all three SCCs of both sides on four different occasions. VOR gain and the presence of pathological saccades were noted and analyzed for each recording.

Results: Intraclass correlation coefficient (ICC) revealed excellent test–retest reliability for VOR gain in both groups (ICC $\geq$ 0.76). Kappa coefficient analysis for the presence of refixation saccades demonstrated moderate to excellent agreement between test sessions ($K \geq$ 0.63) for the lateral canal. For the anterior and posterior SCC, there was large variability between sessions for refixation saccades.

Conclusions: This study provides evidence about test–retest reliability of VOR gain and refixation saccades assessed using vHIT in healthy individuals and individuals with vestibulopathies. These findings suggest that both measures are highly reliable and replicable across test sessions, except refixation saccades in vertical canals which varied between sessions in some individuals.

Key Words: saccades, test-retest reliability, vHIT, VOR gain

Abbreviations: BPPV = benign paroxysmal positional vertigo; ICC = intraclass correlation coefficient; LARP = left anterior right posterior; MD = Meniere's disease; RALP = right anterior left posterior; SCC = semicircular canals; SD = standard deviation; vHIT = video head impulse test; VOR = vestibulo-ocular reflex; VNG = video nystagmography

*Department of Audiology, All India Institute of Speech and Hearing, Mysore, India; †Department of Otorhinolaryngology, All India Institute of Speech and Hearing, Mysore, India

Corresponding author: Niraj Kumar Singh, Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, Mysore-570006, Karnataka, India; Email: niraj6@gmail.com

INTRODUCTION

The visual image is maintained on the fovea of the retina during angular head movement due to the vestibulo-ocular reflex (VOR) initiated in the semicircular canals (SCCs). When the head rotates with a certain speed and direction, the VOR causes the eyes to rotate with the same speed but in the opposite direction. Thus, the combination of head and eye rotation ultimately produces a near zero displacement of the object's image on the retina. In cases of a deficient VOR, steady gaze during head motion is not maintained. Hence, the VOR functions to maintain stable visual images during active head rotation. However, defects caused by vestibular pathology in these reflex pathways would result in various physiological signs and symptoms.

The VOR pathways originating in the lateral SCC can be assessed by caloric irrigation, rotary chair testing, and scleral search coils (Halmagyi and Curthoys, 1988). However, there was paucity of tests for the objective assessment of SCC function in humans without obvious discomfort to the participants (MacDougall et al, 2013) until Halmagyi and Curthoys (1988) developed a bedside screening tool to assess the lateral SCC function. This was later expanded for evaluating the posterior and anterior SCCs (Aw et al, 1996). Now the same head impulse test has been developed into a video-based test called the video head impulse test (vHIT). vHIT allows for assessment over a broader range of VOR frequencies in a comfortable way rather than either caloric stimulation, rotary chair, or scleral search coils.

vHIT is a relatively new clinical tool for assessing the functionality of each of the six SCCs independently. In this test, the eye movements generated by the VOR in response to brief, unpredictable, passive head rotations are recorded and computed along with the head movement to give a measure of the VOR gain. To do so, it uses a high-speed digital video camera to record the eye movement during head rotations and a sensor to measure the head velocity (Manzari et al, 2011; Murnane et al, 2014). The brief head impulses delivered along horizontal and vertical planes are used for the assessment of the lateral SCC and vertical SCC (anterior and posterior), respectively. The vertical canals are oriented at approximately 45° away from the midsagittal plane of the head. The anterior SCC of one side is oriented approximately parallel to the posterior SCC of the other side and, therefore, forms functional pairing (MacDougal et al, 2013). Owing to this, the head is turned 45° to the right for testing the left anterior and right posterior (LARP) SCCs and 45° to the left side for assessing the right anterior and left posterior (RALP) SCCs (MacDougal et al, 2013). Therefore, vHIT allows for the assessment of all six SCCs. This was validated against the scleral search coil test, the gold standard test for VOR measurement, and found to yield the same results (MacDougall et al, 2009; 2013; Agrawal et al, 2014).

vHIT provides a mechanism to quantify the VOR gain and record the presence of refixation saccades. The presence of refixation saccades has long been identified as an indicator of deficient VOR and, thus, vestibular impairment (Halmagyi and Curthoys, 1988; Aw et al, 1996; MacDougall et al, 2009; Manzari et al, 2011; Murnane et al, 2014). Refixation saccades are of two types— overt and covert saccades. When the eye movement fails to maintain the target during brief head movements/ impulse, the eyes miss the target only to quickly refixate after completion of the head movement. Such saccades are called overt saccades. These can be visualized by an experienced clinician. In certain individuals, such saccadic corrections are generated while the head impulse is still in process and these are called covert saccades (Judge et al, 2017). The presence of either of these two types of refixation saccades is considered a positive indicator for vestibulopathy (MacDougall et al, 2009; Manzari et al, 2011; Murnane et al, 2014). However, to establish a diagnosis based on vHIT, it is imperative to establish that results on the test do not change when repeating the test on the same individual (test–retest reliability). Studies on healthy participants with no vestibular pathologies have shown excellent test–retest reliability by showing no significant difference in VOR gain between test sessions using repeated measures analysis of variance (repeated measures ANOVA) or Wilcoxon signed-rank test for within-participant comparisons (Murnane et al, 2014; Bansal and Sinha, 2016; Ross and Helminski, 2016). Nonetheless, these measures compare the central tendencies of measurements obtained between two sessions on the same participant and, therefore, may not accurately provide a measure of test–retest reliability of vHIT, especially when used in isolation (i.e., without using other methods such as intraclass correlation coefficient [ICC] or Cronbach's alpha) (Bruton et al, 2000). Unfortunately both the aforementioned studies (Murnane et al, 2014; Bansal and Sinha, 2016) used comparison of means or mean ranks in isolation, which are reported to be inappropriate measures of test– retest reliability when used in isolation (Zaki et al, 2013). An alternate method of assessing test–retest reliability is ICC (Brown, 1997; Zaki et al, 2013). The modern version of the ICC calculates the coefficient values using variance estimates through partitioning of the total variance between and within subject variance (Bruton et al, 2000; Weir, 2005; Zaki et al, 2013) which makes this a more suitable measure of test–retest reliability, more so when compared against methods using comparisons of measures of central tendencies. This was clearly demonstrated by the example shown in Table 2 of the study by Zaki et al (2013). In that example, the authors clearly illustrated that despite poor test–retest

reliability of the data, procedures using comparison of means showed no significant difference between sessions, causing a misleading perception of high degree of test–retest reliability. For a more detailed explanation, readers may refer to Zaki et al (2013). Even though Korsager et al (2016) did use ICC as a measure of the test–retest reliability, they examined it only for the lateral SCC. The vertical canals were not assessed in their study. Furthermore, the extent to which the results can be reliably replicated in individuals with vestibular pathologies is yet to be investigated. Hence, the aim of the present study was to establish the test–retest reliability of VOR gain and presence of refixation saccades obtained through vHIT in healthy individuals and individuals with known vestibular pathologies.

METHOD

Participants for the Study

The study was approved by the institutional review board as part of an intramural funded project. The study included 20 healthy individuals in the age range of 18–30 years (mean = 22.2 years), which included 14 female and 6 male participants. These participants had no history or complaint of vestibular, otological, or neurological problems. The study also included 20 consecutive participants in the age range of 21–80 years (mean = 45 years) with vertigo as the primary complaint. Of these, nine participants were later diagnosed with definite Meniere's disease (MD) (five bilateral, two right, and two left ears), nine with benign paroxysmal positional vertigo (BPPV) in the posterior canal (two bilateral, two right, and five left ears), one had labyrinthitis, and one had vertebrobasilar insufficiency. None of them had received any treatment for these problems, except the participant with labyrinthitis, before or during the course of the study. The participant with labyrinthitis was treated with antibiotics and other necessary medications during the hospital stay; however, he continued to experience residual vertigo during the chronic phase.

These participants were diagnosed by a team of experienced professionals consisting of an audiologist, an otorhinolaryngologist, and a neurologist. The diagnosis of MD was made using the criteria put forth by the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS, 1995), whereas the diagnosis of BPPV was made following the criteria established by Bhattacharyya et al (2008). Diagnosis of labyrinthitis was accomplished through the presence of unilateral sudden hearing loss and history of sudden onset of persistent vertigo, followed by vomiting for three days until hospitalized and treated for infection. He also reported history of upper respiratory tract infection in days immediately preceding the vertigo attack. The neurological

and radiological investigations showed negative results for space-occupying lesions, stroke, or ischemia. The diagnosis of vertebrobasilar insufficiency was made after an agreement between the otorhinolaryngologist and neurologist on reviewing the radiological and other laboratory test profiles.

INSTRUMENTATION AND CALIBRATION

All participants underwent vHIT by using the ICS Impulse equipment that has video goggles with a monocular camera capable of recording at a speed of 250 frames/second. The goggles were secured tightly using an elastic band around the head, ensuring adequate visibility of the pupils. Each participant was seated at distance of 1 m from the wall, and a fixed target point (a large dot) was placed on the wall at the eye level at a distance. The device was calibrated for 10° eye movement using laser pointers that projected visual targets alternately at an angle of 10° on either side of the center target point on the wall.

Procedure

The clinician's both hands were placed on the head of the participant for lateral canal testing, whereas one hand was placed on the head and the other hand was supporting the chin of the participant for RALP and LARP. Then, small rapid, passive, unpredictable head rotations were made along the plane of the SCC being tested while maintaining head displacements between 10° and 20° and peak head velocity between 100° and 250° per second for the lateral canal and 50°–250° for vertical canals. The participant's task was to maintain gaze on the center target point on the wall. Each participant underwent vHIT testing on four separate occasions to establish test–retest reliability. The first vHIT testing was carried out when the patient came to the outpatient department with the complaint of vertigo. The second vHIT test was performed when audiological evaluations were carried out. The third evaluation was performed on the day the participant came to make an appointment for the video nystagmography (VNG) test battery in the vertigo clinic, and the fourth evaluation was performed on the day of VNG test, but before undergoing VNG testing. On average, four to six days were taken to complete all four sessions, and there was a gap of minimum a day and maximum two days difference between sessions. Therefore, participants in the clinical group were not deprived of treatment for the sake of the study, and none of them were on medications between the first and fourth vHIT evaluation. Participants in the control group (healthy individuals) were tested with similar time gaps between sessions. There was a gap of a day between sessions for the control group, and in four days, all four sessions were completed. All

participants in the healthy and pathological groups were tested by the same clinician in all four sessions.

For the lateral canal, horizontal head impulses were made while the clinician held the participant's head firmly, ensuring no contact with the headband of the goggles, and turned the head briskly to the left/right randomly while the participant maintained their gaze on the target.

For vertical canals, the head was turned to left or right by 45° and RALP and LARP of the other modules, respectively, were used. RALP was used to test the functionality of the RALP SCC, whereas LARP was used to investigate the function for integrity of the LARP SCC. For testing these vertical SCCs, the participant was instructed to maintain gaze at the pointer on the wall. The vertical head rotation impulses were delivered in such a way that the movement was along the anterior canal of one side and the posterior canal of the other side. Feedback about the SCC being stimulated was provided by the software, and hence, the appropriate angle of head rotation was obtained to ensure the noninvolvement of other SCCs. Responses for 40 impulses, 20 in each direction, were recorded for each SCC pair (lateral, LARP, and RALP). Rest breaks of approximately five minutes were provided between tests for lateral, RALP, and LARP planes.

Measures

Outcome measures were VOR gain and the presence of refixation saccades. The VOR gain was measured as the ratio of the eye velocity to the head velocity. The VOR gain was calculated automatically by the equipment, which it does by using the area under the curve for each head impulse (with desaccading). Finally, an average value of VOR gain for all impulses in a particular direction of head impulse is shown on the screen on completion of the test. These were considered the final VOR gain. Furthermore, the presence of refixation saccades was automatically identified by the software as an eye movement that lagged the head movement and was color coded as red. Failure of the eyes to remain focused on the target because of their displacement caused by head rotation away from the target triggers saccadic eye movement to refocus on the target. This is called "refixation saccade." The presence of overt and covert saccades during or soon after the head impulses in the VOR graph was detected as refixation saccades. To eliminate the artifactual responses being processed as a true refixation saccade, the clinician monitored the video recording of the eye movement during the impulses. Any such artifactual responses were deleted from the final responses that were considered for reporting the outcomes. The presence of such a refixation saccade on >50% of the head impulses was operationally defined "as an indicator" of pathology.

Statistical Analyses

Descriptive statistical analyses were used to obtain mean and standard deviation (SD) of VOR gain for each SCC. Test–retest reliability was evaluated separately in healthy individuals and individuals with vestibular pathologies. Two-way mixed-effects model of the ICC was used with an absolute agreement definition. This model was used because the two-way mixed-effects model has been reported to be most suitable for investigating test–retest reliability when a single examiner has provided multiple scores, as it would be unreasonable to generalize one examiner's scores to a larger population of examiners (Shrout and Fleiss, 1979; Koo and Li, 2016). This model was also used because repeated measurements by the same examiner on the same sample cannot be considered as randomized samples (Portney and Watkins, 2000). Finally, the absolute agreement definition was chosen because measurements would be pointless if there is lack of conformity between the repeated measurements. More detailed explanations for the need to choose the two-way mixed-effects model for evaluating test–retest reliability when multiple scores are given by a single examiner are available in a publication by Koo and Li (2016). ICC values were noted for each of the three planes vHIT (lateral, RALP, and LARP). The presence of refixation saccades and their test–retest reliability were analyzed using kappa coefficient analyses.

RESULTS

The study participants consisted of 20 healthy individuals (40 ears). The data for both ears were combined for analysis because of finding no significant difference in VOR gain between the ears in any session when evaluated using the Wilcoxon signed-rank test. A table containing the Z-values and p-values as the outcome of the Wilcoxon signed-rank test for within group between ears comparisons of VOR gain is given as Supplemental Table S1 (available with the online version of this article). The study also consisted of 20 individuals with unilateral/bilateral vestibular pathologies (n = 28 ears). There were 12 individuals with unilateral vestibulopathies, and hence, there were 12 unaffected ears of individuals with vestibular pathologies. These individuals of both the groups underwent vHIT using lateral, LARP, and RALP plane modules. Figure 1 shows head and eye velocity recordings across four sessions from one representative healthy individual, whereas Figure 2 shows a similar pattern of recordings from one individual with bilateral MD. It can be observed from Figure 1 that there were no refixation saccades in any SCC plane and the traces for the head (represented in blue for left and orange for right side impulses) and eye (represented in green, irrespective of the eye movement direction)
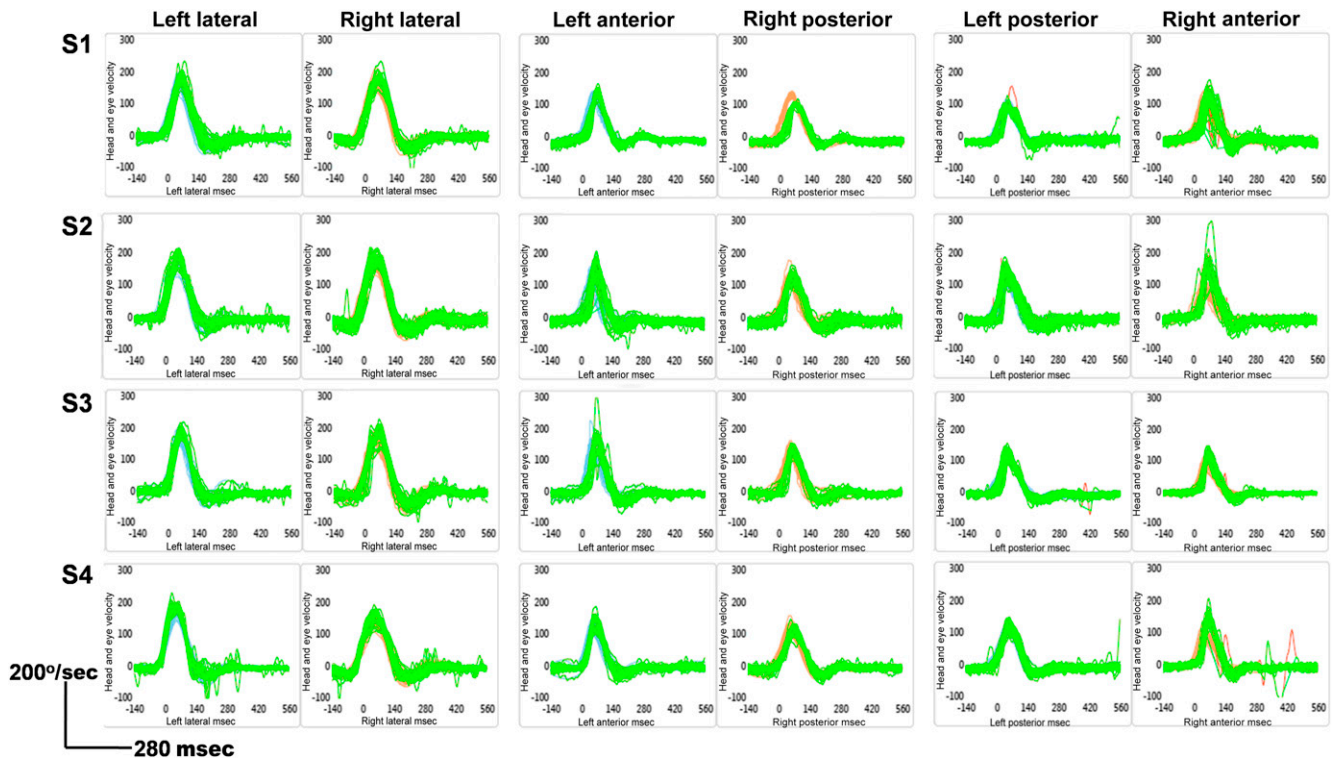
Figure 1. Representative vHIT recordings of four different sessions for lateral, LARP, and RALP SCC in a healthy individual. Note that blue and orange colors represent traces from left- and right-sided head impulses, respectively, whereas green color traces represent eye movement irrespective of the direction of movement. (This figure appears in color in the online version of this article.)

movement nearly completely overlapped, indicating normal VOR gains. Figure 2 clearly shows the presence of covert refixation saccades (red color traces) on both side lateral SCC in all sessions, but only on the right posterior SCC in session 1.

VOR gains were calculated and subjected to descriptive statistical analyses for obtaining mean and SD. The number of ears with the presence of refixation saccades was also counted for each test session. Refixation saccades were analyzed with respect to the plane of vHIT excitation and laterality lesion. Table 1 shows mean and SD of VOR gain obtained for all six SCCs in all four test sessions. The table also shows the number of ears with the presence of significant refixation saccades in all four sessions in each canal's plane.

In healthy individuals (nonpathological group), the ICC values were found to be 0.775 [$F_{(39, 120)} = 4.443$, $p < 0.001$], 0.863 [$F_{(39, 120)} = 7.288$, $p < 0.001$], and 0.768 [$F_{(39, 120)} = 4.308$, $p < 0.001$] for lateral, anterior, and posterior SCCs, respectively. In individuals with vestibular pathologies, data were separately analyzed as affected ears and unaffected (asymptomatic) ears. The ICC values in 28 affected ears were 0.959 [$F_{(27, 84)} = 24.349$, $p < 0.001$], 0.929 [$F_{(27, 84)} = 14.163$, $p < 0.001$], and 0.939 [$F_{(27, 84)} = 16.465$, $p < 0.001$] for lateral, anterior, and posterior SCCs, respectively. Similarly, the ICC values in the unaffected ears for lateral, anterior,

and posterior SCCs were 0.816 [$F_{(11, 36)} = 5.448$, $p < 0.001$], 0.899 [$F_{(11, 36)} = 9.906$, $p < 0.001$], and 0.736 [$F_{(11, 36)} = 3.786$, $p < 0.01$], respectively.

Although overall results revealed excellent test–retest reliability over four sessions in clinical and control groups, the individual participant data were closely observed for agreement across the four sessions and some discrepancies could be observed among sessions. In the control group, there was no discrepancy in VOR gain interpretations for any of the SCCs. In the unaffected ears of individuals in clinical groups, the individual participant data again showed no discrepancy among sessions for lateral SCCs; however, there was discrepancy in VOR gain for the anterior SCC in two ears (one participant had normal VOR gain in three sessions, whereas reduced in one session, and another participant had reduced VOR gain in three sessions, but normal in one session). Also, there was discrepancy in VOR gain for the posterior SCC among sessions in two ears, one each of two participants (one participant had normal VOR gain in two sessions and reduced in two sessions, whereas another participant had reduced VOR gain in three sessions but normal in one session). The affected ears showed discrepancy in VOR gain interpretation in two ears for the lateral SCC, two ears for the anterior SCC, and four ears for the posterior SCC. The session-wise findings of VOR gain in individuals
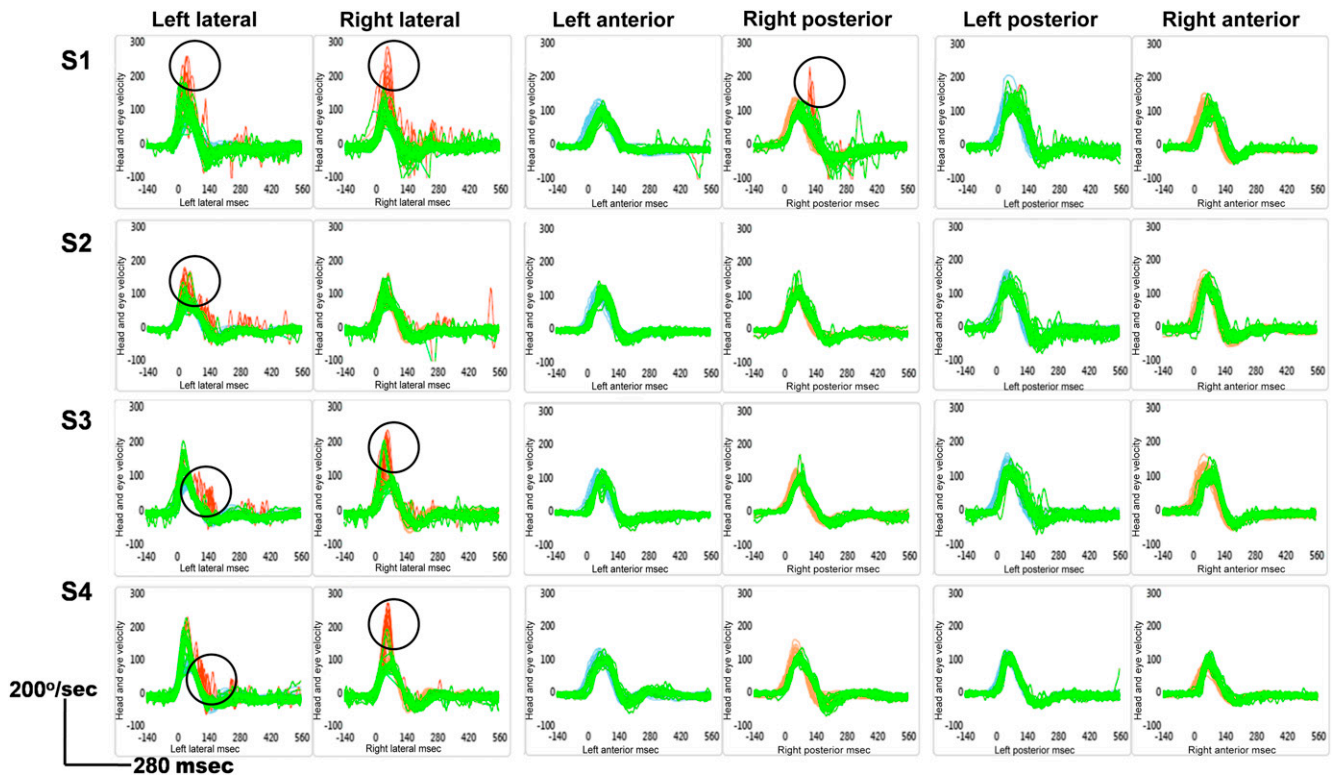
Figure 2. Representative vHIT recordings of four different sessions for lateral, LARP, and RALP SCC in an individual with bilateral MD. Note that blue and orange colors represent traces from left- and right-sided head impulses, respectively, whereas green color traces represent eye movement irrespective of the direction of movement and red color is used to indicate refixation saccades, which is also represented by circles around the traces. (This figure appears in color in the online version of this article.)

with discrepant results among sessions are shown in Table 2. Figure 3 shows the vHIT traces of an individual with a diagnosis of MD in the right ear, who showed discrepancy in VOR gains between sessions (first participant in Table 2).

The refixation saccades were mainly evidenced during the examination of lateral SCC and rarely witnessed during the evaluation of the other SCCs. Kappa coeffi-

cient was carried out to examine test–retest reliability of the presence of significant refixation saccades (presence in >50% of head thrusts for a canal). Results revealed no significant difference for the presence of refixation saccades between the test sessions with a minimum K-value of 0.63 when assessing the lateral SCC. However, on evaluations of anterior and posterior SCCs for the presence significant saccades, a significant

Table 1. Mean and SD of VOR Gain and Number of Ears with the Presence of Refixation Saccades for Lateral, Posterior, and Anterior SCCs Across the Four Test Sessions

Parameter	Session	Healthy Individuals (n = 40 Ears)			Affected Ears of Individuals with Vestibulopathy (n = 28 Ears)			Unaffected Ears of Individuals with Vestibulopathy (n = 12 Ears)		
		Lateral	Posterior	Anterior	Lateral	Posterior	Anterior	Lateral	Posterior	Anterior
VOR gain	1	0.99 (0.12)	0.82 (0.07)	0.86 (0.13)	0.90 (0.20)	0.76 (0.16)	0.81 (0.18)	1.00 (0.12)	0.90 (0.09)	0.87 (0.10)
	2	0.99 (0.14)	0.83 (0.12)	0.84 (0.12)	0.92 (0.23)	0.72 (0.17)	0.76 (0.22)	1.03 (0.13)	0.88 (0.07)	0.94 (0.16)
	3	1.03 (0.15)	0.82 (0.10)	0.85 (0.13)	0.92 (0.21)	0.72 (0.14)	0.71 (0.18)	0.99 (0.12)	0.87 (0.08)	0.99 (0.13)
	4	0.99 (0.13)	0.83 (0.10)	0.85 (0.12)	0.92 (0.20)	0.72 (0.16)	0.73 (0.21)	1.05 (0.12)	0.89 (0.10)	0.90 (0.10)
Refixation	1	0	0	0	13	2	1	9	0	0
saccades	2	0	0	0	16	2	0	9	3	0
	3	0	0	0	15	2	2	9	3	0
	4	0	0	0	13	2	1	9	2	0
Head	1	133.30 (22.24)	125.10 (20.43)	123.20 (18.21)	133.23 (12.46)	124.10 (13.85)	123.17 (13.68)	133.23 (14.49)	123.55 (16.35)	123.44 (18.26)
velocity	2	136.17 (22.19)	123.16 (21.72)	125.43 (18.72)	133.12 (16.32)	122.39 (17.24)	122.26 (14.46)	134.46 (16.86)	122.79 (17.44)	122.21 (16.31)
	3	134.21 (19.69)	128.92 (19.01)	127.13 (19.22)	136.34 (15.93)	123.85 (15.34)	124.78 (13.12)	135.79 (15.27)	124.37 (16.21)	127.57 (18.29)
	4	134.22 (20.17)	123.32 (20.11)	127.10 (17.99)	130.32 (17.12)	125.78 (14.23)	125.67 (14.13)	138.45 (17.27)	129.59 (12.92)	124.19 (18.92)

Note: n = number of ears.

Table 2. Session-Wise Details of VOR Gain Findings in Individuals with Discrepancy in Findings Among the Sessions

	Right												Left											
	Lateral SCC				Anterior SCC				Posterior SCC				Lateral SCC				Anterior SCC				Posterior SCC			
Diagnosis	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Right MD	0.84	0.76	0.78	0.80	0.73	0.69	0.65	0.66	0.68	0.65	0.64	0.72	0.99	1.06	1.06	0.90	0.78	0.71	0.71	0.68	0.74	0.68	0.68	0.66
Right MD	0.77	0.78	0.91	0.93	0.92	0.87	0.74	0.88	0.91	0.95	0.96	0.85	0.95	0.95	0.96	0.98	0.75	0.59	0.68	0.61	0.74	0.72	0.66	0.67
Bilateral MD	0.89	0.82	0.95	0.90	0.68	0.61	0.67	0.73	0.71	0.73	0.74	0.72	0.96	0.92	0.95	0.88	0.84	0.89	0.86	0.89	0.87	0.85	0.88	0.89
Bilateral MD	0.56	0.58	0.58	0.62	0.68	0.67	0.60	0.64	0.55	0.59	0.74	0.72	0.83	0.87	0.81	0.84	0.79	0.82	0.85	0.78	0.76	0.69	0.69	0.72
Right BPPV	0.87	0.86	0.96	0.95	0.90	0.88	0.77	0.93	0.81	0.77	0.69	0.65	0.87	0.87	0.86	0.88	0.81	0.84	0.81	0.82	0.86	0.88	0.85	0.91
Bilateral BPPV	0.85	0.85	0.83	0.88	0.68	0.64	0.62	0.62	0.82	0.78	0.69	0.67	0.99	0.91	0.91	0.93	0.88	0.89	0.88	0.88	0.87	0.89	0.83	0.86

Note: Digits 1–4 indicate session number.

difference was found between sessions with large variability between session pairs. Hence, the presence of refixation saccades in the lateral SCC was a more consistent finding than in the anterior and the posterior SCCs in the pathological group.

DISCUSSION

In the present study, test–retest reliability of VOR gain measured with vHIT across four sessions for all SCC pairs produced ICC ≥ 0.76. Versino et al (2001) classified the test–retest reliability to be excellent, moderate, and poor when the coefficient values were >0.7, 0.4–0.7, and <0.4, respectively. Therefore, the VOR gain was found to have excellent test–retest reliability for all SCC pairs in both groups of the present

study. Previous attempts at studying the test–retest reliability of the VOR gain assessed using vHIT also reported high test–retest reliability by observing no significant difference in VOR gain between test sessions (Murnane et al, 2014; Bansal and Sinha, 2016). However, their conclusions were based on the findings of repeated measures ANOVA (Murnane et al, 2014) or the Wilcoxon signed-rank test (Bansal and Sinha, 2016). Unfortunately, these are measures of the central tendencies and, therefore, compare means, medians, or ranks and not the performance of the same individual on two or more occasions. Repeated measures ANOVA or the Wilcoxon signed-rank test would yield "no significant difference," which could give a false representation of test–retest reliability. Thus, the lack of group mean, median or rank difference shown by the
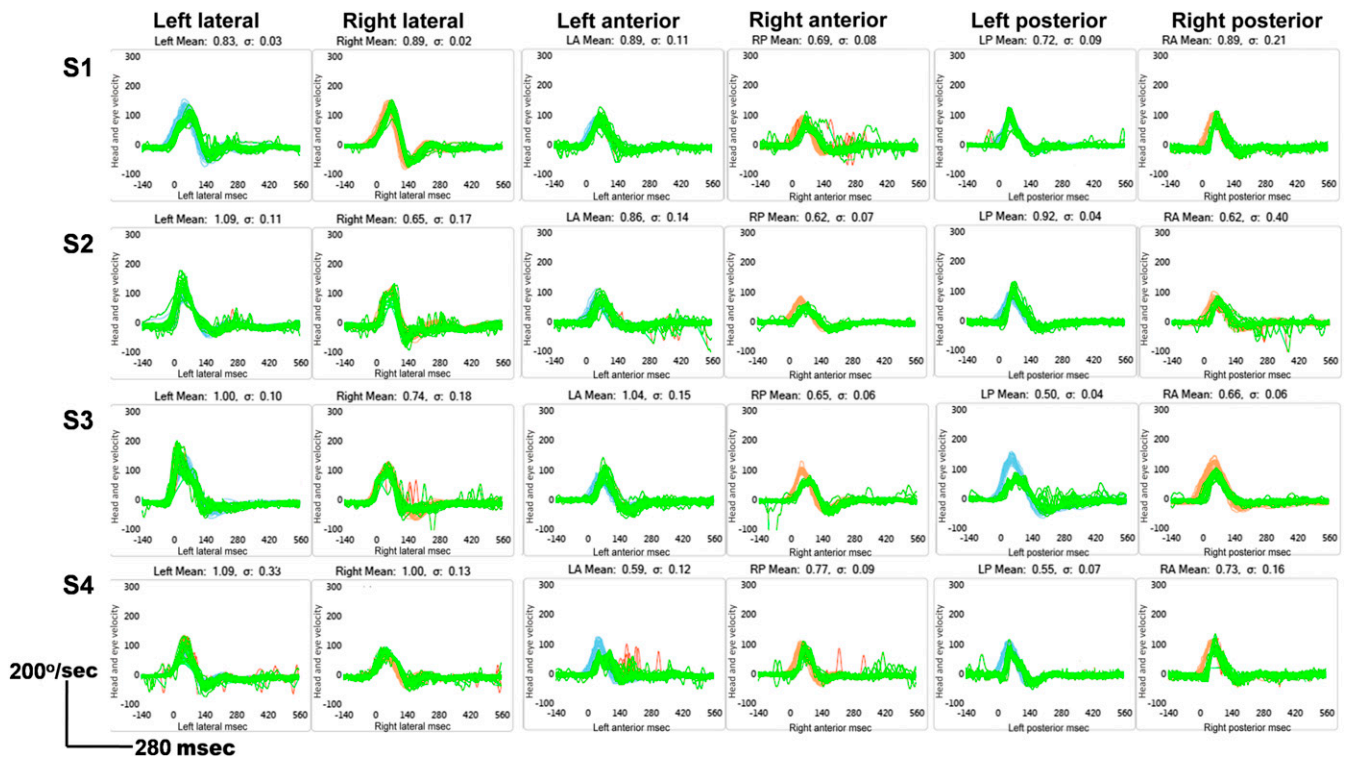

Figure 3. Discrepancy in the VOR gain in vHIT recordings of an individual with MD in the right ear. (This figure appears in color in the online version of this article.)

abovementioned studies does not truly represent test–retest reliability. The computation of ICC is an effective tool for measuring the test–retest reliability of any measure (Hatcher, 1994), and hence, using this statistical analysis is an appropriate method. To the best of our knowledge, the present study is the first attempt at evaluating the test–retest reliability of vHIT in all SCC planes. Findings indicate that VOR gain can be repeatedly obtained with high degree of accuracy and precision. Even though previous studies used different statistics to that used in the present study, the findings of the present study are in agreement with those reported previously.

An interesting finding in the present study was observed when analyzing the individual participant data. Whereas there was no observable discrepancy in "normal" or "abnormal" results between sessions in any SCC in the healthy individuals group, small discrepancies were observed in unaffected and affected ears of individuals with vestibular pathologies. A close examination showed that the discrepancies were restricted to only individuals with MD or BPPV. Although the exact pathophysiology of variations in vHIT results in MD is yet to be clearly understood, the duration since last attack might have played a part in producing the discrepancy. While all participants with MD in the present study were tested during the asymptomatic period, the discrepancies appeared evident in only those in whom the last attack was three to four days ago. In a study, Mahringer and Rambold (2014) reported that the chance of obtaining an abnormal result on vHIT increases in case it is administered within five days of the attack of MD and subsequently the chances of abnormal results on vHIT reduces with increasing duration since the last attack. Because, the four sessions in the present study were separated by up to six days, the findings could have been possibly discrepant between the sessions that fell within five days of the MD attack and those that fell beyond this interval. Although the reasons for such changes in vHIT within or beyond the five days interval are not quite clear, they might be a plausible explanation. Furthermore, such findings in the unaffected ears could be possible because of the "latent" or "occult" MD in the unaffected ears, which has been reported using a number of tests including cervical vestibular evoked myogenic potentials (cVEMP) (Ribeiro et al, 2005; Lin et al, 2006; Fouly et al, 2012), electrocohleography (Visu and Singh, 2012), and oto-acoustic emissions (Magliulo et al, 2004) previously. The discrepancy among sessions was also observed in two cases with BPPV, one unilateral, and one bilateral. These discrepancies were observed in the VOR gain of the posterior SCC, which could be related to probably the pathology state at the time of testing. However, more research is needed on vHIT in BPPV to understand these findings clearly.

The presence of refixation saccades and their test–retest reliability in vHIT was found to be moderate to excellent for lateral SCC evaluation; however, they was found to be poor to moderate for vertical SCC. None of the investigators looked at the test–retest reliability of this parameter. Present study suggests the presence of refixation saccades can be reliably and repeatedly recorded when evaluating lateral SCC but not vertical SCC. A study by Korsager et al (2016) reported better reliability for the presence of refixation saccades than the VOR gain between the clinicians; however, this was shown only for the lateral SCC. Korsager et al (2016) did not examine the test–retest reliability of LARP and RALP, the measures that assess functioning of the vertical canals on either side. Therefore, the present study may be treated as an extension to their work with more details on vertical canals in addition to the lateral canal. Furthermore, Korsager et al (2016) reported that the clinician's experience would affect the VOR gain, but not the presence of refixation saccades in the evaluation of vestibular function. In the present study, all the four sessions were tested by the same clinician, and hence, the variability in clinician could not have been the cause of variations in the VOR gain or the refixation saccades in the present study. In addition, the clinician received supervised training for more than six months before the data collection for the present study and, therefore, the clinician's experience would also not have caused changes between sessions. Nonetheless, it is also not clear as to why the vertical canals vHIT testing was associated with lesser degree of test–retest reliability. Minor variations in the angle of head rotation could be a possible reason for the same. For example, although the head should be rotated to 45° to the left or right for RALP or LARP, respectively, the software takes it as acceptable and flashes green signal if the head is turned by any angle between 35° and 45°. Therefore, changes in the head rotation angle between these two values across sessions could have caused the variation in results, thereby causing poor repeatability between sessions. However, more experimental evidences are needed to substantiate the aforementioned possible reason for poorer test–retest reliability of the vertical canals' vHIT outcomes.

CONCLUSIONS

Results of the present study show excellent test–retest reliability of VOR gain for SCC pairs. In terms of the presence of refixation saccades, the reliability was good only for the lateral SCC. Hence, the VOR gain appears to be more reliable across sessions than the presence of refixation saccades, and the presence of refixation saccades is reliable only for the lateral SCC. Therefore clinicians need to exercise caution when

interpreting results of the vertical SCC based on the presence of refixation saccades.

REFERENCES

AAO-HNS. (1995) American Academy of Otolaryngology-Head and Neck Surgery Committee on Hearing and Equilibrium Committee on hearing and equilibrium guidelines for the diagnosis and evaluation of therapy in Meniere's disease. *Otolaryngol Head Neck Surg* 113:181–185.

Agrawal Y, Schubert MC, Migliaccio AA, Zee DS, Schneider E, Lehnen N, Carey JP. (2014) Evaluation of quantitative head impulse testing using search coils versus video-oculography in older individuals. *Otol Neurotol* 35(2):283–288.

Aw ST, Haslwanter T, Halmagyi GM, Curthoys IS, Yavor RA, Todd MJ. (1996) Three-dimensional vector analysis of the human vestibuloocular reflex in response to high-acceleration head rotations. I. Responses in normal subjects. *J Neurophysiol* 76(6): 4009–4020.

Bansal S, Sinha SK. (2016) Assessment of VOR gain function and its test-retest reliability in normal hearing individuals. *Eur Arch Otorhinolaryngol* 273:3167–3173.

Bhattacharyya N, Baugh RF, Orvidas L, Barrs D, Bronston LJ, Cass S, Chalian AA, Desmond AL, Earll JM, Fife TD, Fuller DC, Judge JO, Mann NR, Rosenfeld RM, Schuring LT, Steiner RW, Whitney SL, Haidari J. (2008); American Academy of Otolaryngology-Head and Neck Surgery Foundation. (2008) Clinical practice guideline: benign paroxysmal positional vertigo. *Otolaryngol Head Neck Surg* 139(5):S47–S81.

Brown JD. (1997) Statistics corner: questions and answers about language testing statistics: reliability of surveys. *Shiken: JALT Testing & Evaluation SIG Newsletter* 1(2):17–19.

Bruton A, Conway JH, Holgate ST. (2000) Reliability: what is it, and how is it measured? *Physiotherapy* 86:94–99.

Curthoys IS, MacDougall HG, Manzari L, Burgess AM, Bradshaw AP, McGarvie L, Weber KP. (2011) *Clinical Application of a New Objective Test of Semicircular Canal Dynamic Function—The Video Head Impulse Test (vHIT)*.

Fouly H, Minawi ME, Dessouki TE. (2012) Value of VEMP in detecting saccular affection in the asymptomatic ear in patients with Meniere's disease. *Med J Cario Univ* 80(1):397–403.

Halmagyi GM, Curthoys IS. (1988) A clinical sign of canal paresis. *Arch Neurol* 45(7):737–739.

Hatcher L. (1994) *A Step-by-Step Approach to Using the SAS(R) System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute.

Judge PD, Janky KL, Barin K. (2017) Can the video head impulse test define severity of bilateral vestibular hypofunction? *Otol Neurotol* 38:730–736.

Koo TK, Li MY. (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15:155–163.

Korsager LEH, Schmidt JH, Faber C, Wanscher JH. (2016) Reliability and comparison of gain values with occurrence of saccades in the EyeSeeCam video head impulse test (vHIT). *Eur Arch Otorhinolaryngol* 273(12):4273–4279.

Lin MY, Timmer FC, Oriel BS, Zhou G, Guinan J, Kujawa SG, Herrmann BS, Merchant SN, Rauch SD. (2006) Vestibular evoked myogenic potentials (VEMP) can detect asymptomatic saccular hydrops. *Laryngoscope* 116(6):987–992.

MacDougall HG, McGarvie LA, Halmagyi GM, Curthoys IS, Weber KP. (2013) The video head impulse test (vHIT) detects vertical semicircular canal dysfunction. *PLoS One* 8(4):e61488.

MacDougall HG, Weber KP, McGarvie LA, Halmagyi GM, Curthoys IS. (2009) The video head impulse test: diagnostic accuracy in peripheral vestibulopathy. *Neurology* 73(14):1134–1141.

Magliulo G, Cianfrone G, Gagliardi M, Cuiuli G, D'Amico R. (2004) Vestibular evoked myogenic potentials and distortion-product otoacoustic emissions combined with glycerol testing in endolymphatic hydrops: their value in early diagnosis. *Ann Otol Rhinol Laryngol* 113(12):1000–1005.

Mahringer A, Rambold HA. (2014) Caloric test and video-head-impulse: a study of vertigo/dizziness patients in a community hospital. *Eur Arch Otorhinolaryngol* 271(3):463–472.

Murnane O, Mabrey H, Pearson A, Byrd S, Akin F. (2014) Normative data and test-retest reliability of the SYNAPSYS video head impulse test. *J Am Acad Audiol* 25(3):244–252.

Portney LG, Watkins MP. (2000) *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River, NJ: Prentice Hall.

Ribeiro S, Almeida RD, Cavoilla H, Gananca M. (2005) Vestibular evoked myogenic potentials in affected and asymptomatic ears in unilateral Meniere's disease. *Braz J Otorhinolaryngol* 71(1): 60–66.

Ross LM, Helminski JO. (2016) Test-retest and interrater reliability of the video head impulse test in the pediatric population. *Otol Neurotol* 37(5):558–563.

Shrout PE, Fleiss JL. (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420–428.

Versino M, Colnaghi S, Callieco R. (2001) Vestibular evoked myogenic potentials: test-retest reliability. *Funct Neurol* 16: 299–309.

Visu D, Singh NK. (2012) Action potential latency in individuals with endolymphatic hydrops. *Student research at A.I.I.S.H. Mysore (articles based on dissertations done at AIISH)*. 10:89–99.

Weir JP. (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 19: 231–240.

Zaki R, Bulgiba A, Nordin N, Ismail NA. (2013) A systematic review of statistical methods used to test for reliability of medical instruments measuring continuous variables. *Iran J Basic Med Sci* 16(6):803–807.