How Good Are We in Evaluating a Bedside Head Impulse Test?

Athanasia Korda,1 John Patrick Carey,2 Ewa Zamaro,1 Marco Domenico Caversaccio,1 and Georgios Mantokoudis1

INTRODUCTION

Clinicians performing a horizontal head impulse test (HIT) are looking for a corrective saccade, a corrective eye movement toward the direction of a deficient slow phase response of the vestibulo-ocular reflex (VOR); however, the detection of such saccades is a challenge. Experts performing HITs improve their accuracy by performing unpredictable, fast, and large head movements to enhance the visibility of saccades (Tjernström et al. 2012). It is known that high acceleration reveals VOR deficits better and elicits larger overt catch-up saccades (Weber et al. 2008). The Halmagyi–Curtthoys HIT is one of the most important clinical tests in patients with an acute vestibular syndrome (Halmagyi & Curtthoys 1988). The HIT in conjunction with nystagmus and test of skew (also known as “HINTS”) is important for the discrimination between central and peripheral causes of dizziness (Newman-Toker et al. 2013a,b). Any improvement of clinical bedside performance is therefore of great importance.

HIT sensitivity was reported to range between 34 and 74%, whereas the specificity was between 67 and 89% (Beynon et al. 1998; Jorns-Häderli et al. 2007; Mahringer and Rambold 2014; Celebisoy 2018). Clinical HIT (cHIT) is relying on the subjective interpretation of the clinician whether he is seeing a corrective saccade or not. cHIT accuracy is related to the examiner’s experience but also to the degree of canal paresis (Beynon et al. 1998) and vestibular compensation mechanisms (Mantokoudis et al. 2016). “Experts” tend to evaluate a cHIT as normal if the quantitative HIT (scleral search coils) is almost normal (Jorns-Haderli et al. 2007). Quantitative video-HIT (vHIT) devices offer, however, a noninvasive and objective way for the assessment of VOR and refixation saccades.

We sought to study the accuracy of detecting saccades in cHIT performed by an expert compared to vHIT in patients with vestibular dysfunction. We hypothesized that latency and amplitude of the corrective saccades play an important role for the clinical detection of an abnormal hHIT.

MATERIAL AND METHODS

Study Population

In this prospective, observational cohort study, we collected 658 cHITs from seven patients with a deficient vestibular function before and after surgery (labyrinthectomy following gentamicin injection or superior semicircular canal dehiscence plugging) at a tertiary referral hospital (see Supplemental Digital Content 1, http://links.lww.com/EANDH/A672). Patients received additional tests of auditory and vestibular function such as vHIT, bithermal caloric irrigations, c- and oVEMPs, pure-tone audiometry, video-nystagmography, or rotational chair examinations.

We did not focus on a specific underlying pathology, degree of vestibular dysfunction or degree of vestibular compensation, because we assessed the ability of a human eye to detect corrective saccades independent of the underlying disease and its compensation. The aim was to collect normal and pathological HITs with various saccade latencies, amplitudes, and velocities. All patients had normal vision based on static visual acuity or a normal corrected vision. Abnormal vision, vertical skew deviation, and spontaneous nystagmus were exclusion criteria.
Ethical Considerations
All patients provided written informed consent as approved by the local institutional review board.

Clinical Head Impulse Testing
One neurootologist with 20 years’ experience manually imposed several/repeated passive, unpredictable, at varying head angular accelerations (>2109°/s) and velocities between 92 and 279°/s and low amplitude (< 20°) lateral head movements in the yaw plane of the horizontal canal by holding the patient’s head in the temporal-parietal region and instructed them to keep their eyes on a stable target at about 1.5 m away. The patients were seated either in bed or on a chair, and the examiner was sitting in front. Head movements were performed at a random cadence and direction to prevent prediction. Target peak head velocity was monitored by the build-in gyrosopes. All cHITs have been recorded by a room camera and microphone. The clinician indicated orally whether the test was abnormal or not just by visual inspection of the eyes. He was aware of patient’s disease, but not of the degree of vestibular deficit or the degree of vestibular compensation.

Video-Oculography and Video HIT
The cHIT of all participants was additionally recorded simultaneously with a portable, lightweight video-oculography device (vHIT) (EyeSeeCam, Munich, Germany) (Schneider et al. 2009). We used the build-in inertial accelerometers for recording head movements, a 250-Hz infrared video camera for right eye tracking and a laser mount for calibration. The goggles were secured tightly with adjustable straps to prevent slippage during the HIT. The gold standard for recording accurately and objectively HITs is the magnetic scleral search coil technique (Eibenberger et al. 2016); however, such recording techniques are invasive, technically demanding and not generally available. Quantitative vHIT recordings are therefore more widespread used in dizziness clinics. Direct comparisons between the magnetic scleral search coil technique and the vHIT showed that the latter is a reliable method for both saccade detection and VOR Gain estimation (Imai et al. 2005; MacDougall et al. 2009; Agrawal et al. 2014).

Saccade Analysis
All hHITs collected and accepted by the algorithm of the device software were stored and assessed for saccades by a single, masked trained rater. We off-line processed raw quantitative data exported from the HIT device by using Matlab R2014a (Mathworks, Natick, MA, USA). Saccades were detected automatically and verified by an expert rater masked to the patient’s results and diagnosis. Minimal saccade velocity amplitude cutoff was >40°/s. Saccade latency (time from the onset of the HIT until the onset of the saccade) and amplitude for the first compensatory saccade were determined for each HIT. The onset of the head impulse was defined as the time when head velocity exceeded 20°/s.

Statistical Analysis
All vHITs were segregated into normal and abnormal tests using a VOR gain cutoff >0.68 based on normative data (MacDougall et al. 2009). Mean VOR gain (slow-phase velocity/head velocity) was calculated between 40 and 60 ms after HIT onset. This early window in the course of the head impulse captures the eye movement response that is most likely vestibular in origin and minimizes the influence of any catch-up saccades.

Mixed effects logistic regression was used to test whether the physician identified a saccade more frequently with increasing latency, amplitude, gain, and velocity (univariately and all together), with patient and session within patient as random effects. Additionally, whether the HIT toward the pathological side was included in additional models. Only HITs with a single corrective saccade were included in the analysis except for gain, where we included all HITs.

Sensitivity and specificity of saccade amplitude, gain, and speed to predict the physician identifying a saccade were derived using mixed effects logistic regression. Optimal cutoffs were selected using Jouden’s J to optimize both sensitivity and specificity (selects point closest to top-left corner of the ROC curve).

RESULTS
We collected totally 658 cHITs and excluded 293 because of artifacts and invalid HITs rejected by the VOG algorithm. We present here data from 365 cHIT tests (seven patients with an average of 52.1 tests per individual, range of 16 to 96) collected during 1 to 5 separated sessions (mean 8 postoperative days, range 1 to 54 days). Two hundred four cHIT tests showed a single saccade, 81 cHITs had no saccade, and 80 cHITs had multiple saccades. Forty-nine percent were covert (occurring during head movement, latencies ranged from 42 to 233 ms) and 51% overt (occurring after head movement, latencies range from 114 to 299 ms). Figure 1 shows two vHIT recordings with a covert (A) or overt (B) saccade.

Saccade amplitude was statistically the most significant parameter for saccade detection (p < 0.001). An increment of one degree almost doubled the probability of the physician identifying a saccade (Table 1, odds ratio [OR] = 1.77). This result remained robust with multivariate analysis (OR 1.8, p = 0.004, Table 1).

In the univariate analyses (Table 1), saccade latency did not statistically significantly influence test accuracy; however, if adjusted for HIT direction toward the pathological side, we observed a statistically significant effect on the probability of saccade detection (p = 0.036, OR 8.05, Table 2).

Regarding VOR Gain, an increment of 0.1 reduced the probability of a correct saccade detection (Table 1, OR = 0.67, p = 0.01). When we tested all parameters together in the multivariate analysis, we found that the effect of VOR gain was no longer statistically significant (Table 1, p = 0.3). Similar results to the unadjusted analysis in terms of side and pathology were also observed in the univariate analyses that adjusted only for the HIT direction (right/left) or direction to the pathological side (Table 2).

Adjusted analyses for all variables and HIT direction are shown in Table 2. HIT direction appeared to be unimportant and was not statistically significant. Saccade amplitude remained a statistically significant parameter for an accurate HIT detection (p = 0.007).

The sensitivity and specificity for the optimal cutoff of the indicated variable is reported in Table 3. Overall sensitivity for the physician to identify saccades was 20.7%, whereas sensitivity and specificity of HITs with a single saccade was 13.7% (95% confidence interval [CI] 9.3 to 19.2%) and 100% (95% CI 95.5 to 100%), respectively, as at no time did the physician identify a saccade when the video-oculography goggles did not.
The sensitivity of covert saccade detection was 15 and 12.5% for overt saccades. Sensitivity was better (22.2%) for HITs performed to the pathological side. The physician’s sensitivity for detecting a saccade amplitude of >1°, however, was 93%, and specificity was 79%. The optimal cutoff for saccade detection was >1° saccade amplitude, >43°/s velocity, and a latency of >181 ms. Figure 2 shows the receiving operator characteristics for latency, amplitude, velocity, and VOR gain.

**DISCUSSION**

We investigated the accuracy of detecting saccades in cHIT. We found that VOR gain of a HIT, and especially the amplitude of the following corrective saccades, was highly associated with a positive clinical test, while the latency of saccades did not seem to have any important influence. The probability of saccade detection was eight times higher for HIT toward the pathological side; however, the sensitivity of saccade detection by an expert neurootologist was low whereas specificity was perfect: the observer never reported a saccade when actually there was none.

One study supported that these catch-up saccades were only visible after the head movement as they could not be distinguished from the VOR, while the head was still moving (Blödow et al. 2013). Corrective saccades during head movements are often invisible and considered as “covert” saccades. Therefore, bedside HIT depends on the timing and size of the consequent catch-up saccades rather than on the measured VOR gain. Contrary to that knowledge, in our study, latency was not as important as a saccade amplitude. This was unexpected because covert saccades occur early during a head impulse, whereas overt saccades occur later. It is still unclear how far the examiner’s ability to detect saccades is compromised by a simultaneous head movement. Our study showed that it was also possible for an expert to recognize covert catch-up saccades. This unexpected phenomenon should be further investigated in future studies.

The saccade amplitude and consequently saccade velocity seems to play the most important role in saccade detection. Saccade amplitude depends on the amplitude of the excursion and the acceleration of the head, and the degree of vestibular loss showing a linear correlation (Weber et al. 2008). In addition, we found that the probability of saccade detection was eight times higher for HIT toward the pathological side.

The accuracy of cHIT has been measured by Jorns-Häderli et al. (2007) based on VOR gain in healthy subjects and patients with unilateral or bilateral vestibular loss. Both, experts and nonexperts were able to assess accurately impulses presented on video clips; however, 3D visual and tactile cues were missing and it remained unclear whether nonexperts might have performed cHIT examinations correctly. Our study results are in line with previous studies showing that the occurrence of saccades was a more reliable predictor than the gain value in the clinical evaluation of vestibular function (Korsager et al. 2016). Small saccades may also occur in healthy subjects with normal VOR gains (Jorns-Häderli et al. 2007).

Although manually applied HITs would imply some bias in head acceleration and velocity regarding hand dominance, we could not find any statistical difference between right and left impulses.

A low bedside saccade detection sensitivity of 13%, but a high specificity of 100% means that physicians remain unsure if the examination looks normal. However, they feel confident, if they observe a corrective saccade. The sensitivity of bedside cHIT was lower than reported in the literature (Beynon et al. 1998; Jorns-Häderli et al. 2007; Mahringer and Rambold 2014; Celebisoy 2018). This is mainly due to the high sensitivity of VOg devices in detecting very small saccades, which are not visible by human eyes. However, small saccades might not be clinically relevant the reason why the sensitivity of saccade detection is not equal to the sensitivity of a true pathologic HIT. In addition, we assessed the sensitivity for the detection of one single saccade in one single head impulse. Clinicians can increase their HIT sensitivity by applying multiple head impulses. We did not assess the accuracy of the HIT in relation to the underlying disease or degree of vestibular dysfunction but rather focused on the ability of the human eye to detect saccades at various latencies and amplitudes.

To our knowledge, only a few studies have tried to compare cHIT simultaneously with more objective measures like the evaluations of vHIT (Yip et al. 2016). Our study had several

![Fig. 1. vHIT examples. It shows two recordings of vHIT velocity profiles for head- and eye movements. The eye trace is mirrored. The neurootology expert recognized the large covert saccade (CS) correctly (A) but missed the smaller overt saccade (OS) in (B). vHIT indicates video head impulse test.](image-url)
limitations. First, the expert was aware of the patient’s history. Experts use the knowledge of the patients’ history to increase the sensitivity of pathological bedside HIT and—at least partially—to correctly revise a pathological pre-cHIT to a truly normal post-cHIT (Helmchen et al. 2017). The examiner in our study, however, was masked regarding the horizontal semicircular canal (SCC) function, which could have been normal or abnormal in patients before or after surgical intervention. Vestibular function of the horizontal SCC could have ranged between low/absent, moderate/residual, or normal/borderline after gentamicin treatment or after superior SCC plugging. While the plugged vertical canal remains deficient after surgery, there is a reported 80% recovery rate of the horizontal canal function over time (Mantokoudis et al. 2016). It might be conceivable that the sensitivity for covert saccade detection could have been overestimated, although it was already very low at 15%. Furthermore, the examiner was masked regarding the degree of vestibular compensation.

In addition, beats of nystagmus toward the direction of the deficient slow phase VOR are difficult to distinguish from corrective saccades; however, our patients had no evidence of spontaneous nystagmus under visual fixation. Second, we analyzed the evaluations of only one expert to avoid inter-rater variability, however, our results may not be generally applicable to all professionals dealing with such patients. It remains unknown whether different experts or even novices make statistically important different evaluations. In addition, fixation at the nose of the examiner leads to eye vergence, increasing the expected saccade amplitude in a patient with a deficient VOR. In our study, however, the viewing distance was >1.5 m, which would not produce a requirement for an increased VOR needed for stabilizing the images of close targets.

Our findings have practical implications for clinical care and future research. Because the saccade amplitude was the most important HIT parameter affecting the correct interpretation of the HIT, we recommend performing high amplitude and high acceleration head movements, to elicit larger corrective saccade amplitudes toward the opposite direction of the slow-phase VOR. The larger the amplitude, the more likely a physician would be able to detect the saccade or to distinguish a saccade from a potential small nystagmus beat. The direction, however, should be from head eccentric position back to the center, primary position. Such movements are more predictable,

| TABLE 2. Analysis Adjusting for HIT Direction |
|---------------------------------------------|
| **Univariate Analysis**                      |
|                                           | OR (95% CI) | p   |
| First saccade latency (per 10 ms)           | 1.00 (0.91–1.10) | 0.99 |
| Pathological side                           | Healthy     | Ref. |
| Pathological                               | Pathological |      |
| First saccade amp [°]                       | 8.05 (1.15–56.44) | 0.036 |
| Pathological side                           | Healthy     | Ref. |
| Pathological                               | Pathological |      |
| VOR gain (per 0.1 increase)                 | 0.76 (0.53–1.09) | 0.13 |
| Pathological side                           | Healthy     | Ref. |
| Pathological                               | Pathological |      |
| Head velocity (per 10 increase)             | 0.91 (0.77–1.08) | 0.29 |
| Pathological side                           | Healthy     | Ref. |
| Pathological                               | Pathological |      |
| **Multivariate analysis**                   | OR (95% CI) | p   |
| First saccade latency (per 10 ms)           | 0.99 (0.90–1.09) | 0.88 |
| First saccade amp [°]                       | 1.73 (1.16–2.58) | 0.007 |
| VOR gain (per 0.1 increase)                 | 0.93 (0.63–1.38) | 0.72 |
| Head velocity (per 10 increase)             | 0.84 (0.70–1.01) | 0.07 |
| Pathological side                           | Healthy     | Ref. |
| Pathological                               | Pathological |      |
| CI indicates confidence interval; HIT, head impulse test; OR, odds ratio; VOR, vestibulo-ocular reflex. |

| TABLE 3. Sensitivity/Specificity of Each Variable on the Physician Detecting a Saccade (Univariate Only) Using Only Those With One Saccade |
|-------------------------------------------------|-----------------|-----------------|
| Cutoff | Sensitivity | Specificity |
|-----------------|-----------------|-----------------|
| First saccade amp [°]  | 1 | 92.9 (77.4–98.0) | 79.0 (73.6–83.5) |
| Gain 40–60 | 1 | 100.0 (87.9–100.0) | 67.3 (61.4–72.8) |
| 0.80 | 8.5 (3.6–18.4) | 99.0 (97.2–99.7) |
| 0.68 | 8.5 (3.7–18.4) | 99.3 (97.6–99.8) |
| First sac [°/s] | 43 | 100.0 (87.9–100.0) | 72.0 (66.2–77.1) |
| First saccade latency [ms] | 181 | 100.0 (87.9–100.0) | 64.6 (58.6–70.2) |

Fig. 2. ROC analysis. It shows a ROC curve for saccade latencies, saccade amplitude, saccade velocity, and VOR gain. ROC indicates receiver operating characteristics; VOR, vestibulo-ocular reflex.
but can be masked by slow side-to-side movements of the head (Tjernström et al. 2012). Quantitative, noninvasive vHIT measurements would offer an alternative solution to physicians to perform an accurate, fast and cost-efficient HIT examination. Video-oculography is less prone to wrong saccade interpretations because it evaluates the slow-phase VOR directly and usually provides an unambiguous documentation of a saccade correction. Finally, future diagnostic accuracy studies using vHIT should include a comprehensive saccade analysis.

CONCLUSION

The overall sensitivity for the physician to identify saccades in pathologic head impulses was low; however, the saccade amplitude was the most important factor for accurate saccade detection in clinically performed head impulse tests. The probability of saccade recognition doubles with an increment of 1° saccade amplitude. Contrary to current knowledge, saccade latency and VOR gain play a minor role in saccade detection. This was unexpected because covert saccades are more difficult to detect compared to overt saccades with longer latencies. The role of saccade latency in relation to the head movement and saccade patterns should be further investigated in future studies.

ACKNOWLEDGMENTS

This study was supported by the Swiss National Science Foundation #320030_173081. The Authors would like to thank Aaron Wong for the MATLAB saccade detection script.

The data that support the findings of this study are available from the corresponding author (G.M.) upon reasonable request.

All authors contributed to this work, discussed the results and implications, and commented on the article at all stages. A.K. assisted in data analysis, drafted the manuscript, reviewed and critically edited the manuscript, and approved the final version. J.C. assisted in study design/conception, data collection, reviewed and critically edited the manuscript, and approved the final version. E.Z. assisted in data analysis, reviewed and critically edited the manuscript, and approved the final version. M.C. assisted in data analysis, reviewed and critically edited the manuscript, and approved the final version. J.C. assisted in study design/conception, data collection, reviewed and critically edited the manuscript, and approved the final version.

The authors have no conflicts of interest to disclose.

Address for correspondence: Georgios Mantokoudis, University Department of Otorhinolaryngology, Head and Neck Surgery, Inselspital Bern, 3010 Bern, Switzerland. E-mail: georgios.mantokoudis@insel.ch

REFERENCES

Agrawal, Y., Schubert, M. C., Migliaccio, A. A., Zee, D. S., Schneider, E., Lehnen, N., & Carey, J. P. (2014). Evaluation of quantitative head impulse testing using search coils versus video-oculography in older individuals. Otol Neurotol, 35, 283–288.

Beynon, G. J., Jani, P., & Baguley, D. M. (1998). A clinical evaluation of head impulse testing. Clin Otolaryngol Allied Sci, 23, 117–122.

Blidow, A., Pannasch, S., & Walther, L. E. (2013). Detection of isolated covert saccades with the video head impulse test in peripheral vestibular disorders. Auris Nasus Larynx, 40, 348–351.

Celebisy, N. (2018). Acute vestibular syndrome: clinical head impulse test versus video head impulse test. J Neurol, 263(Suppl 1), 44–47.

Eibenberger, K., Eibenberger, B., & Rucci, M. (2016). Design, simulation and evaluation of uniform magnetic field systems for head-free eye movement recordings with scleral search coils. Conf Proc IEEE Eng Med Biol Soc, 2016, 247–250.

Halmagyi, G. M., & Curthoys, I. S. (1988). A clinical sign of canal paresis. Arch Neurol, 45, 737–739.

Helmchen, C., Knauß, J., Trillenberg, P., Frendl, A., & Sprenger, A. (2017). Role of the patient’s history of vestibular symptoms in the clinical evaluation of the bedside head-impulse test. Front Neurol, 8, 51.

Imai, T., Sekine, K., Hattori, K., Takeda, N., Koizuka, I., Nakamae, K., Miura, K., Fujikoh, H., & Kubo, T. (2005). Comparing the accuracy of video-oculography and the scleral search coil system in human eye movement analysis. Auris Nasus Larynx, 32, 3–9.

Jorns-Häderli, M., Straumann, D., & Palla, A. (2007). Accuracy of the bedside head impulse test in detecting vestibular hypofunction. J Neurol Neurosurg Psychiatry, 78, 1113–1118.

Korsager, L. E. H., Schmidt, J. H., Faber, C., & Wanscher, J. H. (2016). Reliability and comparison of gain values with occurrence of saccades in the EyeSeeCam video head impulse test (vHIT). Eur Arch Otorhinolaryngol, 273, 4273–4279.

MacDougall, H. G., Weber, K. P., Mcgarvie, L. A., Halmagyi, G. M., & Curthoys, I. S. (2009). The video head impulse test: Diagnostic accuracy in peripheral vestibulopathy. Neurology, 73, 1134–1141.

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

Mantokoudis, G., Saber Tehrani, A. S., Wong, A. L., Agrawal, Y., Wenzel, A., & Carey, J. P. (2016). Adaptation and compensation of vestibular responses following superior canal dehiscence surgery. Otol Neurotol, 37, 1399–1405.

Newman-Toker, D. E., Kerber, K. A., Hsieh, Y. H., Pula, J. H., Omron, R., Saber Tehrani, A. S., Mantokoudis, G., Hanley, D. F., Zee, D. S., & Kattah, J. C. (2013a). HINTS outperforms ABCD2 to screen for stroke in acute vertigo and dizziness. Acad Emerg Med, 20, 986–996.

Newman-Toker, D. E., Saber Tehrani, A. S., Mantokoudis, G., Pula, J. H., Guedc, C. I., Kerber, K. A., Blitz, A., Ying, S. H., Hsieh, Y. H., Rothman, R. E., Hanley, D. F., Zee, D. S., & Kattah, J. C. (2013b). Quantitative video-oculography to help diagnose stroke in acute vertigo and dizziness: Toward an ECG for the eyes. Stroke, 44, 1158–1161.

Schneider, E., Villigran, T., Vockeroth, J., Bartl, K., Kohlbecher, S., Bardins, S., Ulbrich, H., & Brandt, T. (2009). EyeSeeCam: An eye movement-driven head camera for the examination of normal visual exploration. Ann NY Acad Sci, 1164, 461–467.

Tjernström, F., Nyström, A., & Magnusson, M. (2012). How to uncover the covert saccade during the head impulse test. Otol Neurotol, 33, 1583–1585.

Weber, K. P., Aw, S. T., Todd, M. J., McGarvie, L. A., Curthoys, I. S., & Halmagyi, G. M. (2008). Head impulse test in unilateral vestibular loss: Vestibulo-ocular reflex and catch-up saccades. Neurology, 70, 454–463.

Yip, C. W., Glaser, M., Frenzel, C., Bayer, O., & Strupp, M. (2016). Comparison of the bedside head-impulse test with the video head-impulse test in a clinical practice setting: A prospective study of 500 outpatients. Front Neurol, 7, 58.