Proof of Concept for an "eyePhone" App to Measure Video Head Impulses

T. Maxwell Parker[a]   Nathan Farrell[a]   Jorge Otero-Millan[a, b]   Amir Kheradmand[a]
Ayodele McClenney[c]   David E. Newman-Toker[a, c]

[a]Division of Neuro-Visual and Vestibular Disorders, Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA; [b]Optometry and Vision Science, University of California, Berkeley, Berkeley, CA, USA; [c]Armstrong Institute Center for Diagnostic Excellence, Baltimore, MD, USA

Abstract

Objective: Differentiating benign from dangerous causes of dizziness or vertigo presents a major diagnostic challenge for many clinicians. Bedside presentations of peripheral vestibular disorders and posterior fossa strokes are often indistinguishable other than by a few subtle vestibular eye movements. The most challenging of these to interpret is the head impulse test (HIT) of vestibulo-ocular reflex (VOR) function. There have been major advances in portable video-oculography (VOG) quantification of the video HIT (vHIT), but these specialized devices are not routinely available in most clinical settings. As a first step towards smartphone-based diagnosis of strokes in patients presenting vestibular symptoms, we sought proof of concept that we could use a smartphone application ("app") to accurately record the vHIT. *Methods:* This was a cross-sectional agreement study comparing a novel index test (smartphone-based vHIT app) to an accepted reference standard test (VOG-based vHIT) for measuring VOR function. We recorded passive (examiner-performed) vHIT sequentially with both methods in a convenience sample of patients visiting an otoneurology clinic. We quantitatively correlated VOR gains (ratio of eye to head movements during the HIT) from each side/ear and experts qualitatively assessed the physiologic traces by the two methods. *Results:* We recruited 11 patients; 1 patient's vHIT could not be reliably quantified with either device. The novel and reference test VOR gain measurements for each ear ($n = 20$) were highly correlated (Pearson's $r = 0.9$, $p = 0.0000001$) and, qualitatively, clinically equivalent. *Conclusions:* This preliminary study provides proof of concept that an "eyePhone" app could be used to measure vHIT and eventually developed to diagnose vestibular strokes by smartphone.

© 2020 The Author(s)
Published by S. Karger AG, Basel

Introduction

Stroke is the leading vascular cause of serious misdiagnosis-related harms [1]. Dizziness and vertigo are the symptoms most tightly linked to missed stroke [2]. However, only 3–5% of the 4.4 million patients with acute dizziness or vertigo each year in the USA have stroke as a

T.M. Parker and N. Farrell contributed equally to the project and wish to be listed as co-first authors. University of California, Berkeley, is J. Otero-Millan's primary affiliation.

David E. Newman-Toker
Department of Neurology, The Johns Hopkins University School of Medicine
The Johns Hopkins Hospital, Meyer Building 2-221
600 North Wolfe Street, Baltimore, MD 21287-6921 (USA)
toker@jhu.edu

cause, presenting a major diagnostic challenge for front-line clinicians in the emergency department [3].

Expert bedside examination using a trio of eye movement tests collectively known as "HINTS" (Head Impulse, Nystagmus, Test of Skew) outperforms neuroimaging for diagnosing stroke in acute, continuous dizziness and vertigo with nystagmus known as the "acute vestibular syndrome" [4]. In high-risk patients, one eye movement alone, the head impulse test (HIT) of vestibulo-ocular reflex (VOR) function, which takes less than a minute to assess at the bedside, is vastly more accurate than CT and slightly more accurate than MRI with diffusion-weighted images [5] (the latter is currently considered the gold standard test for acute stroke in a living patient). Unfortunately, assessing this vestibular eye movement is often difficult for frontline clinicians to learn, so, consequently, most providers lack comfort and confidence in using it clinically [6].

It has been roughly a decade since advances in portable video-oculography (VOG) technology made it possible to quantitatively measure VOR responses during the HIT [7]. Quantitative video HIT (vHIT) currently relies on VOG devices worn on the head like goggles that use sensors to measure head movements and a video camera to record and measure eye movements. Despite some successes in deploying such VOG devices to support remote, subspecialty "Tele-Dizzy" consultations in the emergency department [8, 9], these devices are not yet routinely available in most emergency departments, never mind other relevant clinical (or nonclinical) settings. Broad scaling of eye movement diagnosis to ambulances, urgent care clinics, primary care offices, or even patients' homes will almost certainly require more ubiquitous technology, such as mobile phones.

Recent technological advances have now made quantitative vHIT using mobile phones possible. We sought to demonstrate proof of concept that we could use a handheld smartphone and a video-only approach to quantify the vHIT in a similar fashion to commercial VOG devices. Such an approach could ultimately be a "game changer" for remote diagnosis of vestibular strokes.

Methods

Design

This was a cross-sectional agreement study comparing a novel index test (smartphone-based vHIT) to an accepted reference standard test (VOG-based vHIT) for measuring VOR function. We recruited a convenience sample of adult patients visiting an otoneurology clinic for vertigo, dizziness, or related balance symptoms; no sample size calculation was performed for this initial pilot study. All patients agreed to participate and signed a consent form for eye movement recording research approved by the Johns Hopkins Institutional Review Board (IRB). This report follows EQUATOR/GRRAS guidelines for reporting reliability and agreement studies [10].

vHIT Recordings

In each patient, during a single visit, we sequentially recorded vHIT data first using commercially available VOG "goggle" devices and then using our novel smartphone application ("app"). We initially attempted simultaneous recordings, but there was interference between the goggles and the iPhone while recording eye movements.

We used one of two US Food and Drug Administration (FDA)-approved VOG goggles, both deployed routinely in our subspecialty vestibular clinic and laboratory – either the ICS Impulse (Otometrics, Natus Medical, Inc., Pleasanton, CA, USA) or the EyeSeeCam (Interacoustics, Middelfart, Denmark). Each provides a laptop-based software platform for real-time analysis of the vHIT.

The smartphone app was running on an iPhone Xs (Apple, Inc., Cupertino, CA, USA). We developed the app using Apple's ARKit framework. ARKit is a framework designed to create augmented reality experiences on Apple devices. Within ARKit we used the face-tracking configuration that utilizes the TrueDepth camera to create a 3D mapping of the user's face to track head and eye movements (using a proprietary Apple algorithm that combines natural light and infrared sensors). ARKit provides a continuous recording at 60 samples per second that includes the 3D orientation of the head, relative to the phone or gravity, the 3D position of the head relative to the phone, the horizontal and vertical orientation of the eye, and various measurements of eye closure. The results (e.g., VOR gains) were analyzed offline (see below). Note that the iPhone Xs model used in this study has a front-facing ("selfie") camera allowing for video frame rates surpassing 200 Hz, but the current implementation of ARKit is limited to 60 Hz.

For the HIT, patients were asked to sit in a chair facing a wall approximately 2 m away, and to fixate their gaze on a red dot placed at eye level on the wall. A trained examiner (either a vestibular laboratory technician or experienced investigator [A.K.]) stood behind the patient and, using standard VOG HIT technique, manually rotated the head passively in the horizontal semicircular canal plane (chin tucked forward 30°) at velocities of approximately 100–300°/s. We performed ~10–15 HITs to the right and ~10–15 HITs to the left, first using the VOG device and then repeating the same with the iPhone. In the latter case, a second individual held the phone by hand in front of the patient at a distance of ~25–50 cm (the distance was confirmed in real time by the app) without blocking the patient's view of the distance target.

vHIT Data Processing

Both goggle software platforms provide an automatic postprocessing algorithm that filters each individual vHIT recording (for insufficient head velocity or eye movement artifacts [11, 12]), desaccades the remaining VOR curves (i.e., seeks to remove covert refixation saccades [13] occurring during each HIT), and then calculates vHIT gain for expert interpretation. The software is "off the shelf" and no programming is required to utilize either platform. Though the software permits it, we did not filter the automated
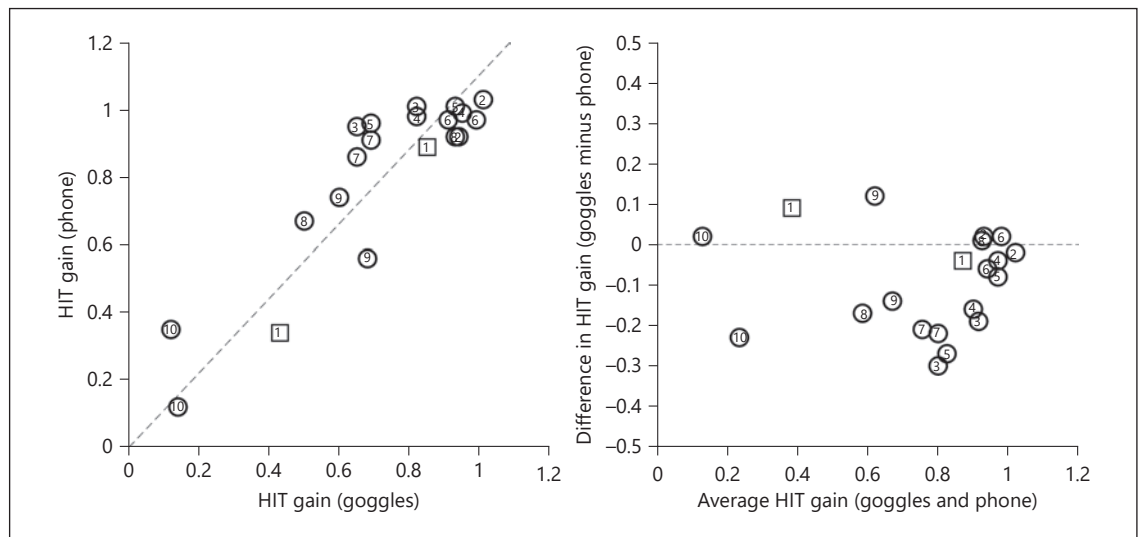
Fig. 1. Correlation between HIT gain measured by smartphone and VOG goggles. The left panel shows a scatter plot of VOR gains for each method on a per-ear (right or left) basis for those 10 of 11 patients undergoing the protocol who had usable data. The right panel shows a Bland-Altman plot of the same data to better visualize possible systematic deviations between the two methods, with positive values indicating a larger gain for the goggles and negative values indicating a larger gain for the smartphone. The results sug- gest that the smartphone app tends to systematically yield higher gain values than the goggles, despite the strong correlation be- tween the results of each measurement method. Circles corre- spond to data collected using the ICS Impulse and squares those collected using the EyeSeeCam. Numbers indicate the number of the participant (online suppl. Table 1). HIT, head impulse test; VOG, video-oculography; VOR, vestibulo-ocular reflex.

results manually or recalculate gains. The complete algorithms are proprietary and not made publicly available, but it is known that the OTOsuite software (ICS Impulse) uses an area under the curve (AUC) method for VOR gain calculations, while the EyeSeeCam software uses an instantaneous velocity method. These two methods of VOR gain calculation yield slightly different values, but the differences are not clinically significant [14].

For each HIT performed with the iPhone, the data from ARKit were recorded and exported out of the phone for visualization and analysis in MATLAB (MathWorks, Inc., Natick, MA, USA) using a custom software developed for postprocessing. To use the software, familiarity with MATLAB programming is ideal, but the user interface to analyze and filter vHIT traces applies a "point-and-click" approach for the end user. After iPhone data export, filtering of HITs with insufficient head velocity or typical eye movement artifacts (blinking, squinting, or noise from eye makeup [11]) was performed manually by one of the investigators (T.M.P.). The identification of eye closure-related artifacts was aided by cross-referencing the video and facial coefficient data provided by the iPhone (i.e., 0-to-1 spectrum value of how much one squinted dur- ing the test). As with the goggle technology, manual filtering of traces with artifacts present requires general familiarity with nor- mal VOR tracings, abnormal VOR tracings, and artifacts. After filtering, we calculated horizontal head and eye velocity as well as VOR gain values for each HIT using an AUC method (in an effort to mirror the OTOsuite calculation). Note that we did not attempt to manually de-saccade the VOR curves in MATLAB, but ensured that saccades did not artificially inflate VOR gain values by trun- cating our AUC calculation at 60 ms after the onset of the HIT (which was defined as head velocity rising and reaching 50°/s). We confirmed that no covert refixation saccades were included in these VOR gain calculations.

VOR Comparisons

We compared the results obtained using the smartphone-based app to those obtained using the standard VOG devices. Each de- vice measured unilateral VOR gain (ratio of eye to head move- ments during the HIT, normally ~0.8–1.0 in healthy individuals) and provided physiologic traces on a per-ear (side) basis. We quantified the statistical correlation using linear regression with- out an intercept between the measured VOR gain from both meth- ods and offer a Bland-Altman plot to identify systematic differ- ences between techniques [15]. Experts qualitatively assessed the physiologic traces by the two methods. Statistical calculations were performed using MATLAB.

Results

We recorded eye movements from 11 patients accord- ing to the protocol described above. Ten had usable re- sults presented here (online suppl. Table 1; see www. karger.com/doi/10.1159/000511287 for all online suppl. material; the eye movements of 9 patients were recorded with ICS Impulse and those of 1 patient with EyeSee-
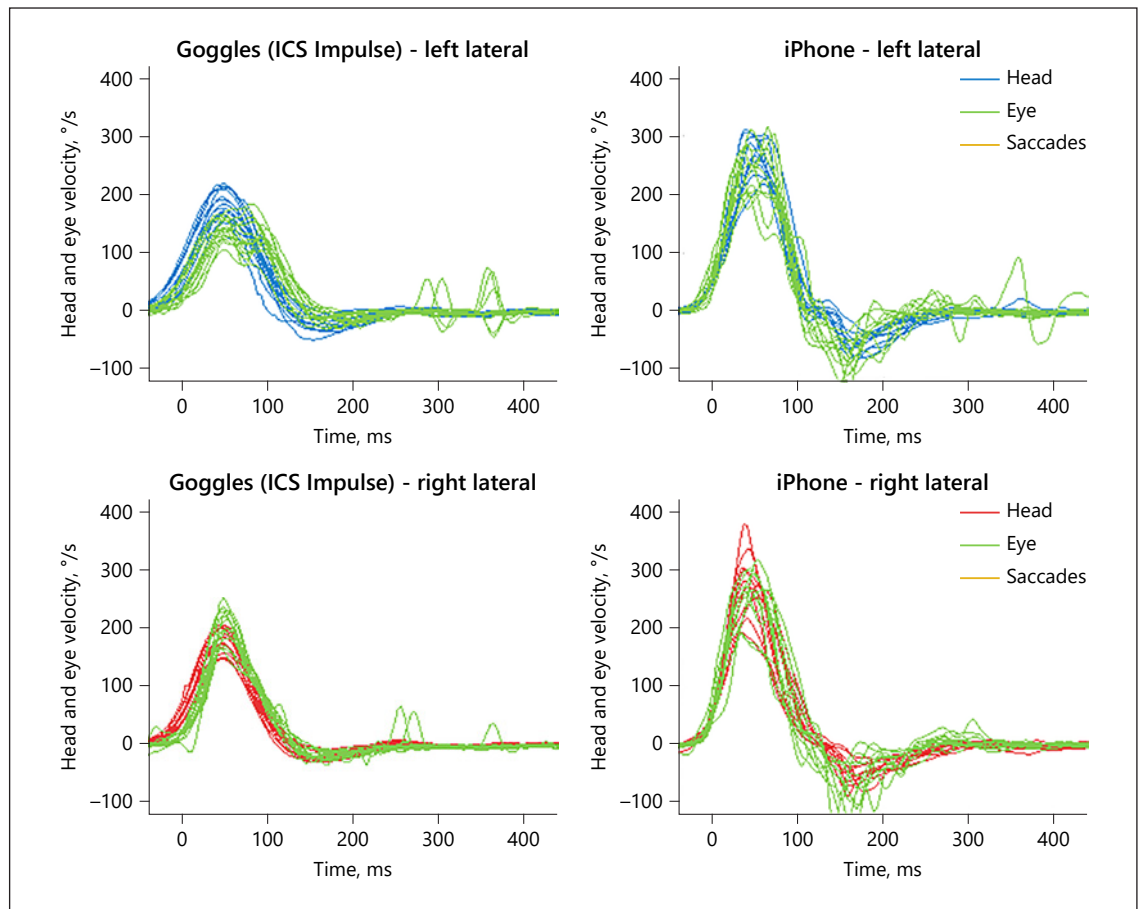
Fig. 2. Patient #4 with normal VOR gains. The left panels show the eye and head velocity during head impulses to the left (top) and right (bottom) recorded with ICS Impulse goggles. The right panels show the equivalent data recorded with the new, custom iPhone application. During the HIT, the eyes and the head move by an equal amount (VOR gain ~1.0), so green traces (reflecting eye movements) are superimposed on head movements (orange and blue). Note that the testing is performed sequentially, so some differences may be real; others, however, may be artifactual (e.g., iPhone "wiggly" eye traces from 100–200 ms, likely from noisy measurement). HIT, head impulse test; VOR, vestibulo-ocular reflex.

Cam); in 1 case results could not be recorded accurately with either VOG goggles or the iPhone due to artifacts, likely due to heavy makeup and shadows around the eyes, which is a known technical problem [11]. Quantitatively, the VOR gains were strongly correlated between the two recording methods (Fig. 1, left) (Pearson's $r = 0.8$, $p < 0.0001$) with a tendency for the smartphone data to report higher gains (Fig. 1, right). Qualitatively, the results were effectively equivalent (Fig. 2–4). Figures 2–4 show side-by-side comparisons of vHIT traces recorded using commercial VOG systems and our custom "eyePhone" app in the same patients; these demonstrate that the physiologic traces are visually similar and interpretively identical (i.e., same clinical conclusion) for both normal (Fig. 2) and abnormal (Fig. 3, 4) VOR responses.

Discussion

These preliminary results demonstrate proof of concept that it should eventually be possible to accurately measure vHITs using smartphones. This is important because (1) the vHIT is the single most useful test in distinguishing inner ear disease from stroke in the acute vestibular syndrome [7], (2) it is a very challenging test for frontline clinicians to perform and interpret correctly [6], and (3) VOG devices are rarely available. Because we were able to do this without special phone attachments or equipment, only with the phone held by hand in front of the patient, the potential for broad scalability of this approach is established. However, readers should note that these pilot study results will need to be validated, and the
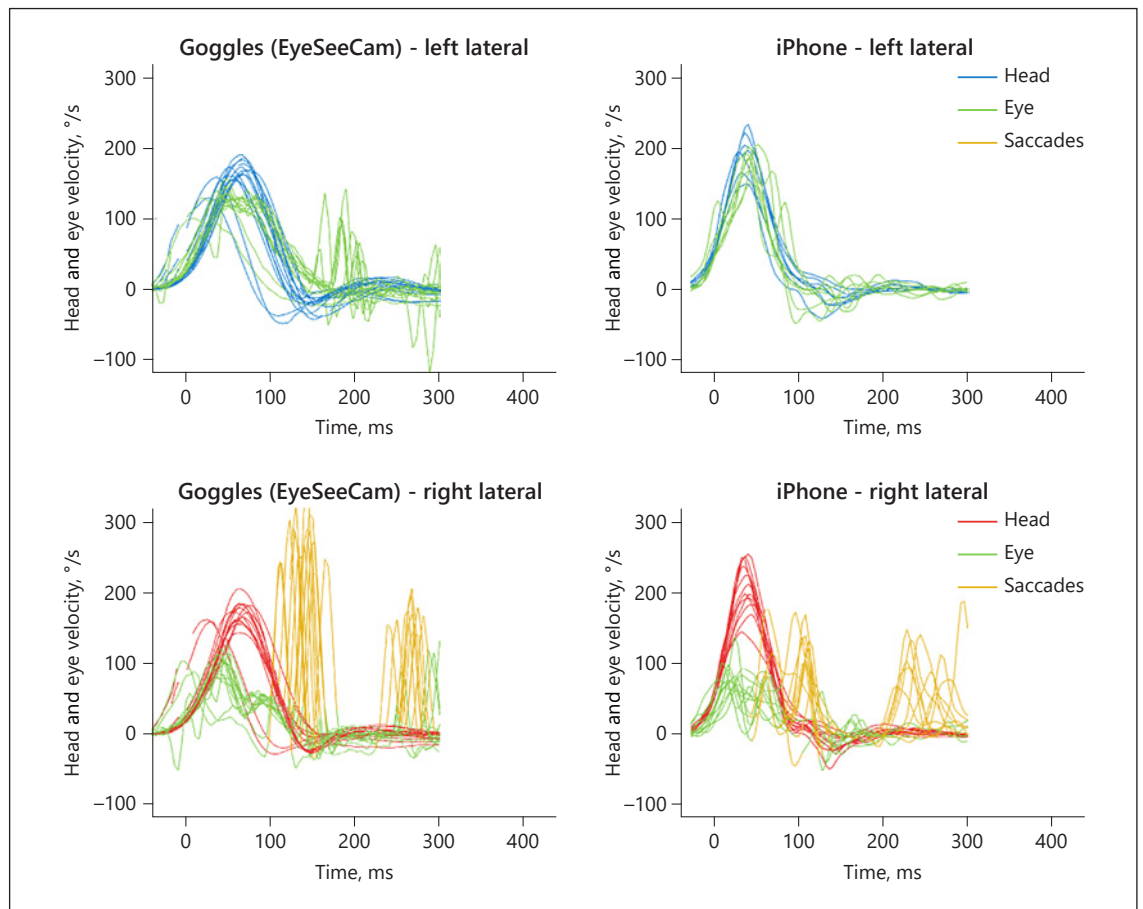
Fig. 3. Patient #1 with unilateral VOR loss on the right side. The left panels show the eye and head velocity during head impulses to the left (top) and right (bottom) recorded with EyeSeeCam goggles. During head impulses towards the right, the eyes do not move as quickly as the head does during the first 100 ms. Subsequent corrective eye movements (catch-up saccades) compensate for the deficient VOR response during the HIT. These corrective eye movements appear as spikes that occur shortly after the head has finished moving ("overt" saccades). Note that the testing is performed sequentially, so some differences may be real; others, however, may be artifactual (e.g., the iPhone-measured velocities of overt saccades are systematically low). It is unclear whether the differences in the leftward responses (top row) are real or artifactual. HIT, head impulse test; VOR, vestibulo-ocular reflex.

app will need to be further developed to incorporate real-time, autonomous quantification of results. Thus, widespread clinical dissemination is still likely years away.

When fully developed, this approach could initially help support remote teleconsultation by specialists [8]. and eventually be combined with real-time diagnostic decision support, as is currently under development for traditional VOG devices [16, 17]. We envision a future in which such tools will be routinely used as a method of initial screening and triage for patients with acute dizziness and vertigo for out-of-hospital care, even in low-resource settings. This could include immediate diagnosis and on-site treatment of common, benign vestibular dis-

orders or care escalation for suspected stroke, which is often missed in patients with dizziness [2].

Furthermore, the vHIT approach has the potential for even broader applications in the bedside assessment of patients with inner ear disease (e.g., vestibular neuritis, Ménière's disease, and bilateral vestibulopathy), whether for diagnosis and management in the clinic, or for research. Finally, when paired with phone-based video recordings or nystagmography for intermittent vertigo with nystagmus [18], this approach has the potential to transform care for all patients with dizziness and vertigo, making in-home telediagnosis and treatment a real possibility. Because minority populations more often use
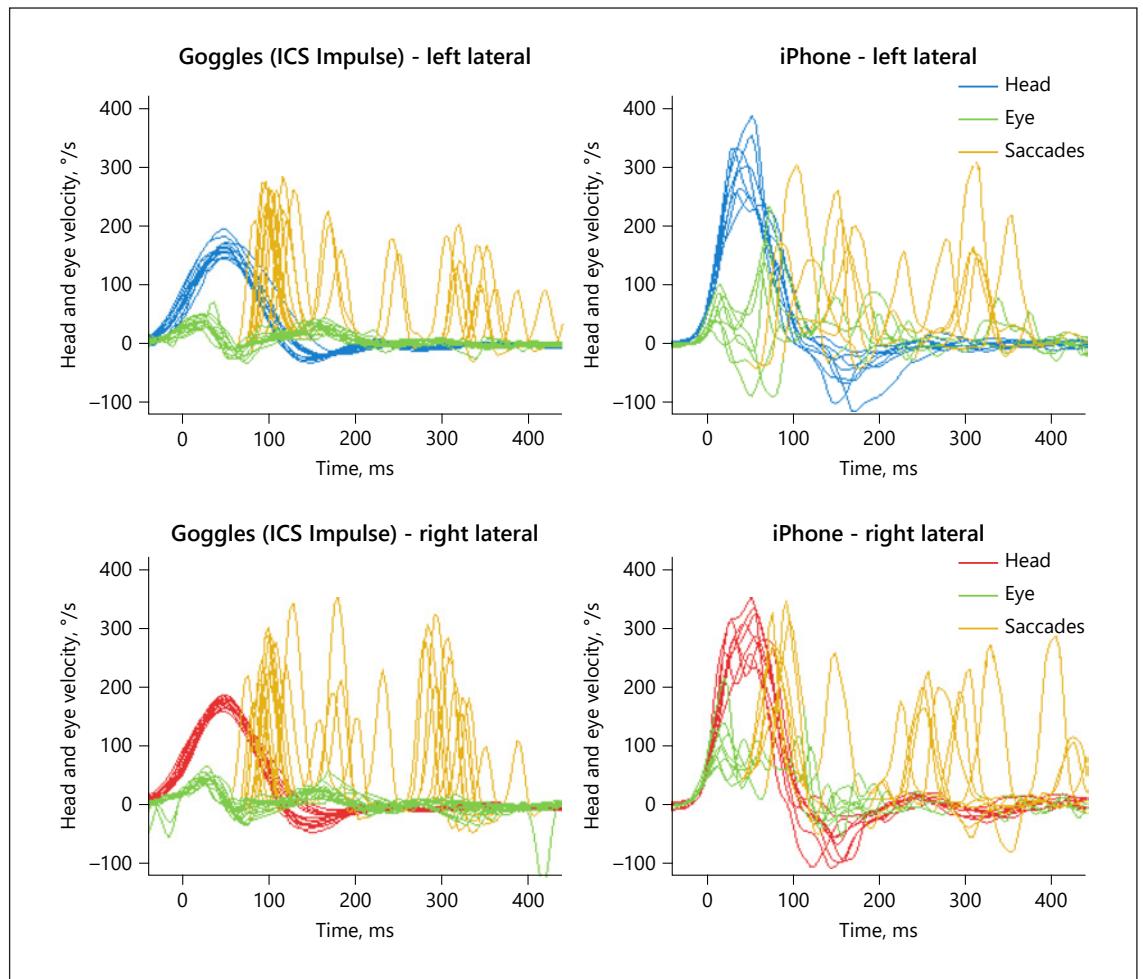
Fig. 4. Patient #10 with bilateral VOR loss. The left panels show the eye and head velocity during head impulses to the left (top) and right (bottom) recorded with ICS Impulse goggles. Note how during head impulses towards either side the eyes do not move as quickly as the head. Then, subsequent corrective eye movements (catch-up saccades) compensate for the deficient VOR response during the HIT. These corrective eye movements appear as spikes that occur during the head movement ("covert" saccades) and afterwards ("overt" saccades). Note that the testing is performed sequentially, so some differences may be real; others, however, may be artifactual (e.g., iPhone "wiggly" eye traces from 100–200 ms, likely from noisy measurement). HIT, head impulse test; VOR, vestibulo-ocular reflex.

smartphones as their primary tool for engaging in health-related activities [19], this is a way to place equity first in stroke and vestibular care.

Limitations

This study is limited by the small sample of patients, the gestalt method of comparing physiologic VOR traces during vHITs, and the lack of gold standard scleral search coil eye movement recordings to assess whether the observed differences were errors of VOG versus smartphone-based testing. Additional app development is needed to address technical challenges (see below). The extent to which these methods will be smartphone (brand/camera) specific is unknown. It is also unknown whether this technique can be completed alone without a second person (i.e., with the patient using a tripod or hand-held phone and active [20] [self-initiated] head impulses). Future research is required to validate these preliminary findings and confirm that smartphone-based approaches can accurately and reliably record eye movements in frontline clinical care settings, in the hands of untrained examiners, and in the context of intrusive eye movements, such as nystagmus.

In addition to the study limitations above, specific technical issues remain for an app-based approach. Cur-

rent ARKit recordings are limited to 60-Hz sampling, but higher frame rates might permit more accurate vHIT quantification and assessment. Processing algorithms to filter, de-saccade, and calculate VOR gains automatically are still needed; these will likely require additional testing and analysis of artifacts (see Fig. 2–4). For example, some smartphone vHIT traces had head velocities that deviated from those of the goggles (e.g., Fig. 4). Because we could not do simultaneous recordings, these may be real differences (i.e., the examiner used different head velocities), but they could also be artifactual; if so, it could be because the iPhone's face-tracking mechanism lacks a special calibration mechanism (Apple discloses little about their FaceID algorithm). Despite our choice to use a similar method of calculation to that of the OTOsuite software (used for 9 of the 10 patients reported on), our AUC methods of calculating VOR gain may not have exactly matched those from the goggles, whose algorithms are proprietary. More testing is needed to determine which method of calculating VOG gains on the iPhone provide the most accurate results (e.g., when compared to the gold standard of scleral search coils). Furthermore, additional work is needed to confirm the test-retest reliability of the method, estimate the number of head impulses required for a stable, valid measurement of VOR gain, and assess its relationship to manual filtering of artifacts; for goggles, there is diminishing marginal utility over 20 HITs [12].

Conclusions

This small pilot study used a video-only smartphone-based approach to measure the vHIT, a key bedside diagnostic test whose use is still mostly restricted to vestibular specialists. Technical challenges remain for ongoing "eyePhone" app development, and more research is needed to validate these preliminary findings. Nevertheless, this small study provides encouraging early evidence that this line of inquiry might ultimately help reduce the important public health problem of missed posterior circulation strokes among patients with acute dizziness and vertigo [3].

Acknowledgement

Statement of Ethics

All patients agreed to participate and signed a consent form for eye movement recording research approved by the Johns Hopkins Institutional Review Board (IRB).

Conflict of Interest Statement

D.E.N.-T. conducts research related to dizziness diagnosis and diagnostic error, including serving as the principal investigator for multiple grants on this topic. He serves as an unpaid member of the Board of Directors of the Society to Improve Diagnosis in Medicine and as its current President. He serves as a medico-legal consultant for both plaintiff and defense in cases related to diagnostic error. The Johns Hopkins University has filed a provisional patent (US 62883373) related to the technology described in this manuscript; the authors D.E.N.-T., J.O.-M., T.M.P., and N.F. are listed as co-inventors. There are no other conflicts of interest. None of the authors have any financial or personal relationships with other people or organizations that could inappropriately influence (bias) their work.

Funding Sources

Author Contributions

T.M.P.: "I declare that I helped conduct the research, helped design and code the 'app,' and co-authored the first draft of the manuscript; edited the manuscript for scientific content; and have seen and approved the final version."

N.F.: "I declare that I helped conduct the research, led design and coding of the 'app,' and co-authored the first draft of the manuscript; edited the manuscript for scientific content; and have seen and approved the final version."

J.O.-M.: "I declare that I supervised design and development of the 'app' and conducted the primary statistical analyses; edited the manuscript for scientific content; and have seen and approved the final version."

A.K.: "I declare that I helped conduct the research; edited the manuscript for scientific content; and have seen and approved the final version."

A.M.: "I declare that I helped conduct the research; edited the manuscript for scientific content; and have seen and approved the final version."

D.E.N.-T.: "I declare that I developed the research plan; authored the first and subsequent major revisions; edited the manuscript for scientific content; and have seen and approved the final version."

Data Availability Statement

Anonymized VOR gain data will be shared by request from any qualified investigator.

References

1 Newman-Toker DE, Schaffer AC, Yu-Moe CW, Nassery N, Saber Tehrani AS, Clemens GD, et al. Serious misdiagnosis-related harms in malpractice claims: the "Big Three" – vascular events, infections, and cancers. Diagnosis (Berl). 2019;6(3):227–40.

2 Tarnutzer AA, Lee SH, Robinson KA, Wang Z, Edlow JA, Newman-Toker DE. ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: a meta-analysis. Neurology. 2017;88(15):1468–77.

3 Newman-Toker DE. Missed stroke in acute vertigo and dizziness: it is time for action, not debate. Ann Neurol. 2016;79(1):27–31.

4 Tarnutzer AA, Berkowitz AL, Robinson KA, Hsieh YH, Newman-Toker DE. Does my dizzy patient have a stroke? A systematic review of bedside diagnosis in acute vestibular syndrome. CMAJ. 2011;183(9):E571–92.

5 Newman-Toker DE, Kerber KA, Hsieh YH, Pula JH, Omron R, Saber Tehrani AS, et al. HINTS outperforms ABCD2 to screen for stroke in acute continuous vertigo and dizziness. Acad Emerg Med. 2013;20(10):986–96.

6 Kene MV, Ballard DW, Vinson DR, Rauchwerger AS, Iskin HR, Kim AS. Emergency physician attitudes, preferences, and risk tolerance for stroke as a potential cause of dizziness symptoms. West J Emerg Med. 2015; 16(5):768–76.

7 Newman-Toker DE, Curthoys IS, Halmagyi GM. Diagnosing stroke in acute vertigo: the HINTS family of eye movement tests and the future of the "Eye ECG". Semin Neurol. 2015; 35(5):506–21.

8 Gold D, Peterson S, McClenney A, Tourkevich R, Brune A, Choi W, et al. Diagnostic impact of a device-enabled remote "Tele-Dizzy" consultation service [abstract]. Diagnostic Error in Medicine, 12th Annual Conference, November 10–13, 2019. Washington, DC; 2019.

9 Muller-Barna P, Hubert ND, Bergner C, Schütt-Becker N, Rambold H, Haberl RL, et al. TeleVertigo: diagnosing stroke in acute dizziness: a telemedicine-supported approach. Stroke. 2019;50(11):3293–8.

10 Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. 2011;64(1):96–106.

11 Mantokoudis G, Saber Tehrani AS, Kattah JC, Eibenberger K, Guede CI, Zee DS, et al. Quantifying the vestibulo-ocular reflex with video-oculography: nature and frequency of artifacts. Audiol Neurotol. 2015;20(1):39–50.

12 Mantokoudis G, Saber Tehrani AS, Wozniak A, Eibenberger K, Kattah JC, Guede CI, et al. Impact of artifacts on VOR gain measures by video-oculography in the acute vestibular syndrome. J Vestib Res. 2016;26(4):375–85.

13 Lee SH, Newman-Toker DE, Zee DS, Schubert MC. Compensatory saccade differences between outward versus inward head impulses in chronic unilateral vestibular hypofunction. J Clin Neurosci. 2014;21(10):1744–9.

14 Janky KL, Patterson JN, Shepard NT, Thomas ML, Honaker JA. Effects of Device on Video Head Impulse Test (vHIT) Gain. J Am Acad Audiol. 2017;28(9):778–85.

15 Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb). 2015;25(2): 141–51.

16 Newman-Toker DE, Saber Tehrani AS, Mantokoudis G, Pula JH, Guede CI, Kerber KA, et al. Quantitative video-oculography to help diagnose stroke in acute vertigo and dizziness: toward an ECG for the eyes. Stroke. 2013; 44(4):1158–61.

17 Newman-Toker DE. AVERT Phase II Trial: Acute Video-Oculography for Vertigo in Emergency Rooms for Rapid Triage. Johns Hopkins University: National Institute on Deafness and Other Communication Disorders (NIDCD 1U01DC013778-01A1); 2014–2019. Clinical Trials.gov ID NCT02483429.

18 Young AS, Lechner C, Bradshaw AP, MacDougall HG, Black DA, Halmagyi GM, et al. Capturing acute vertigo: a vestibular event monitor. Neurology. 2019;92(24):e2743–53.

19 Perrin A, Turner E. Smartphones help blacks, Hispanics bridge some – but not all – digital gaps with whites. Pew Research Center. Available from: https://www.pewresearch.org/fact-tank/2019/08/20/smartphones-help-blacks-hispanics-bridge-some-but-not-all-digital-gaps-with-whites/. Last updated August 20, 2019. Accessed April 10, 2020.

20 Black RA, Halmagyi GM, Thurtell MJ, Todd MJ, Curthoys IS. The active head-impulse test in unilateral peripheral vestibulopathy. Arch Neurol. 2005;62(2):290–3.