Auris Nasus Larynx 49 (2022) 564–570

Contents lists available at ScienceDirect

Auris Nasus Larynx

journal homepage: www.elsevier.com/locate/anl

Machine learning method intervention: Determine proper screening tests for vestibular disorders

Yi Du[a,b,c,d,1], Lili Ren[a,b,c,d,1], Xingjian Liu[a,b,c,d], Ziming Wu[a,b,c,d,*]

[a] College of Otolaryngology Head and Neck Sury, Chinese PLA General Hospital, Chinese PLA Medical School, 28 Fuxing Road, Beijing, China
[b] National Clinical Research Center for Otolaryngologic Diseases, Beijing, China
[c] State Key Lab of Hearing Science, Ministry of Education, Beijing, China
[d] Beijing Key Lab of Hearing Impairment Prevention and Treatment, Beijing, China

ARTICLE INFO

Article history:
Received 23 August 2021
Accepted 13 October 2021
Available online 28 October 2021

Keywords:
Vestibular function screening
Machine learning
Random forest
Video head impulse test

ABSTRACT

Objective: To evaluate the performance of different vestibular indicators in disease classification based on machine learning method.
Methods: This study use retrospective analysis of the vertigo outpatient database from a tertiary care general hospital. 1491 patients with definite clinical diagnoses were enrolled in this study. Spontaneous nystagmus, head-shaking nystagmus, positional nystagmus, unilateral weakness in caloric test, and gain and saccade in video head impulse test (vHIT) were recorded as variables. Diagnoses were mainly reorganized as acute vestibular syndrome, episodic vestibular syndrome, and chronic vestibular syndrome. The trained random forest model was applied based on exploratory data analysis results.
Results: Random forest accuracies on acute, episodic, and chronic vestibular syndrome are 90%, 81.74%, and 91.3%, respectively. The most important features in acute vestibular syndrome are spontaneous nystagmus, and vHIT variables. In episodic vestibular syndrome, unilateral weakness in caloric test, gain and saccades on lateral semicircular canal are the top three parameters. Lateral vHIT gain, head-shaking nystagmus, and unilateral weakness in caloric test are the main parameters on chronic vestibular syndrome. In acute vestibular syndrome, spontaneous nystagmus and vHIT make major contributions in vestibular disorders distinction. When the disease course prolongation, unilateral weakness and head-shaking nystagmus become increasingly important.
Conclusion: Fast clinical test sets including spontaneous nystagmus, head shaking nystagmus, and vHIT should be the first consideration in screening vestibular disorders.

1. Introduction

Vestibular function tests can help determine diagnostic and therapeutic paths on vertigo and dizziness patients. Based on the patient's history and diagnostic criteria, the physician could conduct specialized neuro-otological tests. For instance, if the symptom supports that this patient has the possibility of benign paroxysmal positional vertigo (BPPV), the physician should conduct the Dix-Hallpike and head-roll tests to

* Corresponding author: College of Otolaryngology Head and Neck Surgery, Chinese PLA General Hospital, 28 Fuxing Road, Haidian District, Beijing, 100853, China
E-mail address: zimingwu@126.com (Z. Wu).
[1] These authors contributed equally to this work

https://doi.org/10.1016/j.anl.2021.10.003
0385-8146/

determine which type of BPPV and then provide appropriate repositioning maneuver. Additional examination is not necessary for vertigo management in this case. If the clinician is more aware of the conditions, the patient can undergo fewer tests, thus reducing their suffering time and financial burden.

Determining the appropriate clinical test batteries has become the most essential but complicated issue for physicians. The caloric test (CT) is the most widely used vestibular examination for horizontal semicircular canal (SCC) asymmetry. This method was first proposed by Robert Barany in 1906 [1], and modified and promoted by Fitzgerald and Hallpike in 1942 [2]. The head-shaking test (HST) is a classic but straightforward bedside technique to assess dynamic vestibular asymmetry. Transient nystagmus could be triggered with unilateral vestibular dysfunction after spontaneous nystagmus subsides [3,4]. The head impulse test was first proposed in 1988 [5] by observing corrective saccade to identify vestibular loss. After 27 years, the video head impulse test (vHIT) can automatically calculate VOR gain and saccades (both covert and overt) on each SCC [6,7].

While in the primary care clinic, usability and efficiency also influence examination selection. The CT takes at least 20 minutes, while the vHIT and HST could be completed within 5 minutes. Besides time-consuming, some issues might restrict CT usability on identifying lesions. Inducing severe dizziness and even vomiting make CT an uncomfortable experience. For the diagnosis of bilateral dysfunction, additional ice irrigation or a further rotatory chair are needed. Poor individual reproducibility can also be observed. Attention, alertness, visual suppression, and abnormal eye movements often lead to a false-positive result. Is CT worth keeping as a vestibular screening test because of its accuracy and irreplaceability? This subject was debated for a long time and the answers are still pretty vague. Morrison and colleagues observed no additional CT needed to confirm acute vertigo patients [8]. Some researchers believed the caloric-vHIT dissociation is a sign of Meniere's Disease (MD) [9,10]. Mezzalira et al. found CT is more sensitive to chronic vestibular hypofunction compared to vHIT [11]. Classic statistical analysis can provide the distinction of CT on different disease groups. However, it might be surpassed by other methods, such as machine learning (ML), as to multiclassification and prediction scenarios.

Due to the capability to produce reliable, repeatable decisions and results, ML has been utilized in various neuro-otological domains [12–14]. However, the application of ML on vestibular results are short and need to be studied. This work aims to use the supervised ML method to identify key test parameters to be used in the model prediction of vestibular disorders. This method could also provide each variable importance to find proper screening methods, making a well-founded decision regarding treatment or referral for further evaluation. This study aims to identify a set of vestibular screening tests that are both effective and in line with expert diagnosis. Those variables selected from those tests can be used as valuable indicators for a more accurate diagnosis of common vestibular disorders.

2. Methods

2.1. Data acquisition

This study retrospectively analyzed 2561 vertigo patients who visited a specialized vertigo clinic at the Chinese PLA General Hospital from July 2019 to May 2021. The ethical committee of the Chinese PLA General Hospital gave exemption on informed consent. Each patient was diagnosed by a careful interview and vestibular function results by an otologist, and another specialist reviewed the clinical notes to confirm the diagnosis. All diagnoses could be divided into MD [15], benign recurrent vertigo (BRV) [16], vestibular neuropathy (VN, the diagnosis was based on the history of acute sustained vertigo or imbalance, positive spontaneous nystagmus or unilateral weakness >25% in vHIT or unilateral VOR gain loss combined with obvious catch-up saccades in vHIT and no additional central lesion signs), BPPV [17], functional and psychiatric vertigo (PV) [18], vestibular migraine (VM) [19], bilateral vestibular hypofunction (BVH) [20], delayed endolymphatic hydrops (DEH) [21], vestibular paroxysmia (VP) [22], acoustic neurinoma (AN, radiologically diagnosed and went through vHIT before surgery), traumatic vertigo (TV, diagnosed by imaging), Ramsay Hunt Syndrome (RHS, diagnosed with an ipsilateral herpetic eruption on the auricle and external ear canal, facial palsy, and vertigo.) and Other (including vascular vertigo, cervicogenic vertigo, tinnitus with vertigo, and complication with MD and VM, complication with MD and BPPV). After ruling out incomplete data, undiagnosed cases, bad examination quality, 1491 patients were finally enrolled in this study. Detailed information and diagnoses were listed in Table 1. All examinations were performed on the same day before any treatment and management and none of them is on the severe acute attack period.

2.2. Videonystagmography, caloric test, and video head impulse test

A video-oculography system (VO425, Interacoustics, Denmark) was used to record spontaneous nystagmus (SN), head-shaking nystagmus(HSN), positional nystagmus(PN) and caloric test results. HSN was recorded right after moving the patient's head in the yaw plane with a frequency of 2Hz. The patient's head was instructed 30 degrees downward to keep the rotation plane is parallel to the axes of lateral canals [3]. PN was tested in different positions (supine, head right, head left) for nystagmus provocation or change. SN (+), HSN(+) and PN(+) were identified as >3°/s of the nystagmus in continuous three periods. Bithermal caloric irrigation (warm air 50°C, cold air 24°C) was performed in all MD patients. Unilateral Weakness (UW) was calculated by Jongkee's formula: UW = |(Right Warm + Right Cold) - (Left Warm + Left Cold)|/ (Right Warm + Left Warm + Right Cold + Left Cold). A UW>25% was considered abnormal.

Patients were instructed to fixate a target about 1m ahead and performed vHIT with Otometrics (ICS Impulse, GN Otometrics Inc., Denmark). Approximately ten head impulses within small amplitude, fast, passive, and unpredictable im-

Table 1. General characters and clinic tests of different vertigo syndromes diagnosed in a specialized vertigo clinic

Diagnosis	F/M	Age	N	SN	HSN	UW	vHIT_gain	vHIT_saccade incidence
MD	185/218	51.6±12.9	403	19/403	136/403	183/332	281/370	226/370
BRV	234/119	43.7±16.9	353	4/353	53/353	86/306	225/340	122/340
VN	117/126	50.8±14.3	243	39/243	111/243	136/185	199/230	144/230
BPPV	143/73	51.3±14.0	216	4/216	23/216	42/158	80/107	47/107
PV	39/27	47.5±17.2	66	0/66	9/66	13/55	43/65	32/65
VM	28/11	43.7±12.9	39	0/39	3/39	9/32	26/35	12/35
Other	15/20	47.3±16.8	35	4/35	10/35	12/29	27/33	14/33
BVH	17/18	54.9±15.3	35	1/35	11/35	14/22	31/32	31/32
DEH	22/12	49.4±16.9	34	1/34	11/34	10/30	27/32	24/32
VP	11/10	48.7±15.7	21	0/21	4/21	10/16	15/19	7/19
AN	9/9	42.7±15.8	18	2/18	5/18	11/14	14/15	12/15
TV	13/3	44.1±18.5	16	1/16	7/16	7/11	14/16	12/16
RHS	3/9	44.4±14.8	12	4/12	6/12	7/8	12/12	9/12
Total	836/655	48.8±15.3	1491	79/1491	389/1491	538/1198	994/1306	692/1306

pulses were conducted on each patient to each SCC plane [5]. All vHIT gains (ratio of eye movement and head movement) and catch-up saccades selection were recorded automatically by the manufacturer's algorithms. vHIT gain (+) was identified as lateral gain<0.8 or vertical gain<0.7 on any SCC. vHIT saccade (+) was identified as the occurrence of PR score on either SCC.

2.3. Data analysis and visualization

All abnormal rates were presented as proportion, and continuous variables were displayed as mean ± standard deviation. Data visualization was conducted by Python (Python Software Foundation, DE, USA.) 3.7 programming.

2.4. Feature Extraction and Machine-learning Model

In this research, we used the scikit-learn library for model training. Python 3.7 was used for building and training the model. Incomplete data were deleted to minimize the missing effects. All data were labelled by doctors and anonymized. In this research, vestibular results instead of patient history are our first concern. Therefore, feature selection are 9 parameters: SN, HSN, PN, UW, vHIT_gain_lat, vHIT_gain_ant, vHIT_gain_pos, vHIT_saccade_lat, vHIT_saccade_ant, vHIT_saccade_pos.

Diagnoses were regrouped by the intermediate layer of syndromic classification in ICVD (international classification of vestibular disorders) [23,24]: acute vestibular syndrome (AVS, comprising VN, TV, RHS), episodic vestibular syndrome (EVS, comprising MD, VM, VP, DEH), chronic vestibular syndrome (CVS, comprising BVH, PV), positional vestibular syndrome (BPPV). AN, Other, and BRV were excluded.

Supervised algorithm random forests (RF) were applied in our study. RF is an ensemble of decision trees that can generate good predictions and keep off overfitting. This algorithm can perform well without hyperparameter tuning. The vestibular variables were binarized as abnormal or normal. RF classifier can be used for feature selection, which selects the most important features out of the available features. Dataset was firstly split into training and test subset by a ratio of 0.7. The "stratify" parameter was used to ensure the split by the proportion of the raw dataset, avoiding imbalance data bias the tree toward the significant class. The trained RF model was then used to provide the importance of the vestibular variables in distinguishing diagnosis.

3. Results

3.1. General characters and clinical tests

In total, 1941 patients were included in our study population and their general characteristics are outlined in Table 1. The average age of all individuals is 48.8±15.3 and the age distribution between disorders was slightly different. There are more female patients than males, specially 718:564. Most Diagnoses are MD (N=403), BRV(N=353), VN(N=243), BPPV(N=216), and each abnormal rate was quantified and presented as a percent stacked bar chart, enabling characterization of the proportion of negative to positive results (Fig. 1).

3.2. Exploratory data analysis for variables

A percent stacked bar chart was employed to display each of the vestibular variables' abnormal rate. As shown in Fig. 1 d(3) and d(4), vHIT gain (anterior and posterior) has the highest overall abnormal rate, and the proportions display basic unanimously among diseases. vHIT saccades (Fig. 1d(5)) and vHIT gain (Fig. 1d(2)) in lateral SCC, UW (Fig. 1d(1)) and HSN (Fig. 1b) have relatively higher abnormal detection rates, which fluctuate with distinct disorders.

Fig.1 c(2) and d(1) describe how are the UW and vHIT gain among SCCs are distributed according to diseases. In UW comparison, the MD, BRV, BPPV, PV, VM, Other, DEH have a median UW<25%, and VN, BVH, VP, AN, TV and RHS have a median UW>25%. Ridgeline plots in Fig. d(1) are overlapping density plots, and the peak of each plot displays where values are concentrated over the samples. BVH, AN, and RHS have abnormal vHIT gain values on three pairs of SCCs. Except for the above three disorders, other vHIT gains on lateral SCC have more values>0.8, but vHIT gains on posterior and anterior have more values <0.7.

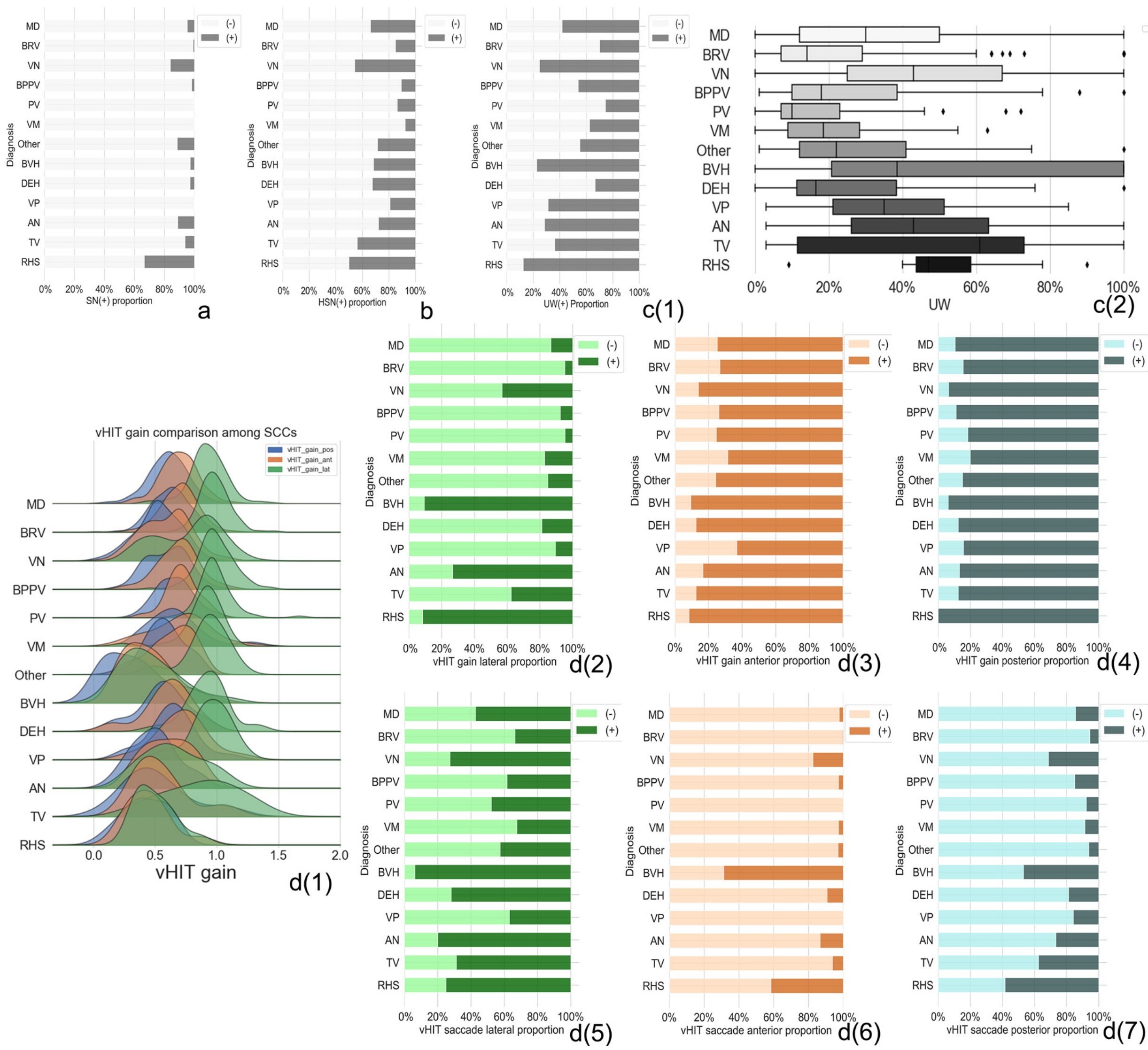

Fig. 1. Exploratory data analysis for different vestibular disorders (a: Percent stacked bar chart of SN; b: Percent stacked bar chart of HSN; c(1): Percent stacked bar chart of UW; c(2): Boxplot of UW; d(1): Ridgeline plot of vHIT gain distribution among SCCs, green, red, blue represent lateral, anterior, and posterior SCCs respectively; d(2), d(3), d(4): Percent stacked bar chart of vHIT gain lateral SCC, vHIT gain anterior SCC, vHIT gain posterior SCC; d(2), d(3), d(4): Percent stacked bar chart of vHIT saccade lateral SCC, anterior SCC, posterior SCC). The prefix of lat, ant, and pos represents anterior SCC, anterior SCC, and posterior SCC individually.

Heatmap on Fig. 2 listed all indicators and emphasized the correlation greater than 0.3 as red. vHIT gain on posterior SCC and vHIT gain on anterior SCC have a relatively stronger correlation (0.47). In other variables, the pairwise attributes are not correlated.

3.3. Model performance

From the exploratory data analysis, vHIT_gain_ant and vHIT_gain_pos might not be that efficient on model performance. Therefore, in the next level of model practice, those two parameters were got rid of to improve the accuracy. Classic RF accuracy on AVS is 90%, EVS is 81.74%, CVS is 91.3%, respectively. Fig. 3a listed all variable influence on AVS, and the most important features are SN (19.24%), vHIT_saccade_pos (17.34%), vHIT_gain_lat (15.67%), and vHIT_gain_saccade(14.9%). EVS features importance was shown in Fig. 3b and the top three variables are UW (22.61%), vHIT_saccade_lat (19.50%), and vHIT_gain_lat (16.45%). vHIT_gain_lat (54.14%) HSN (13.3%), and UW (11.2%) were the main parameters on CVS (Fig.3c).

4. Discussion

In the present study, a machine learning method was used to categorize vestibular disorders using vestibular variables. The good model performance and feature importance support that different disorder classification relies on specific vestibular examination. In AVS, SN and vHIT were the main concern when making vestibular disorders identification. When symptom periodically attacks or is long-lasting, the superiority of CT and HST heightens.

Dizziness and vertigo are the most common symptoms in the neuro-otological clinic. Early identification of vestibular

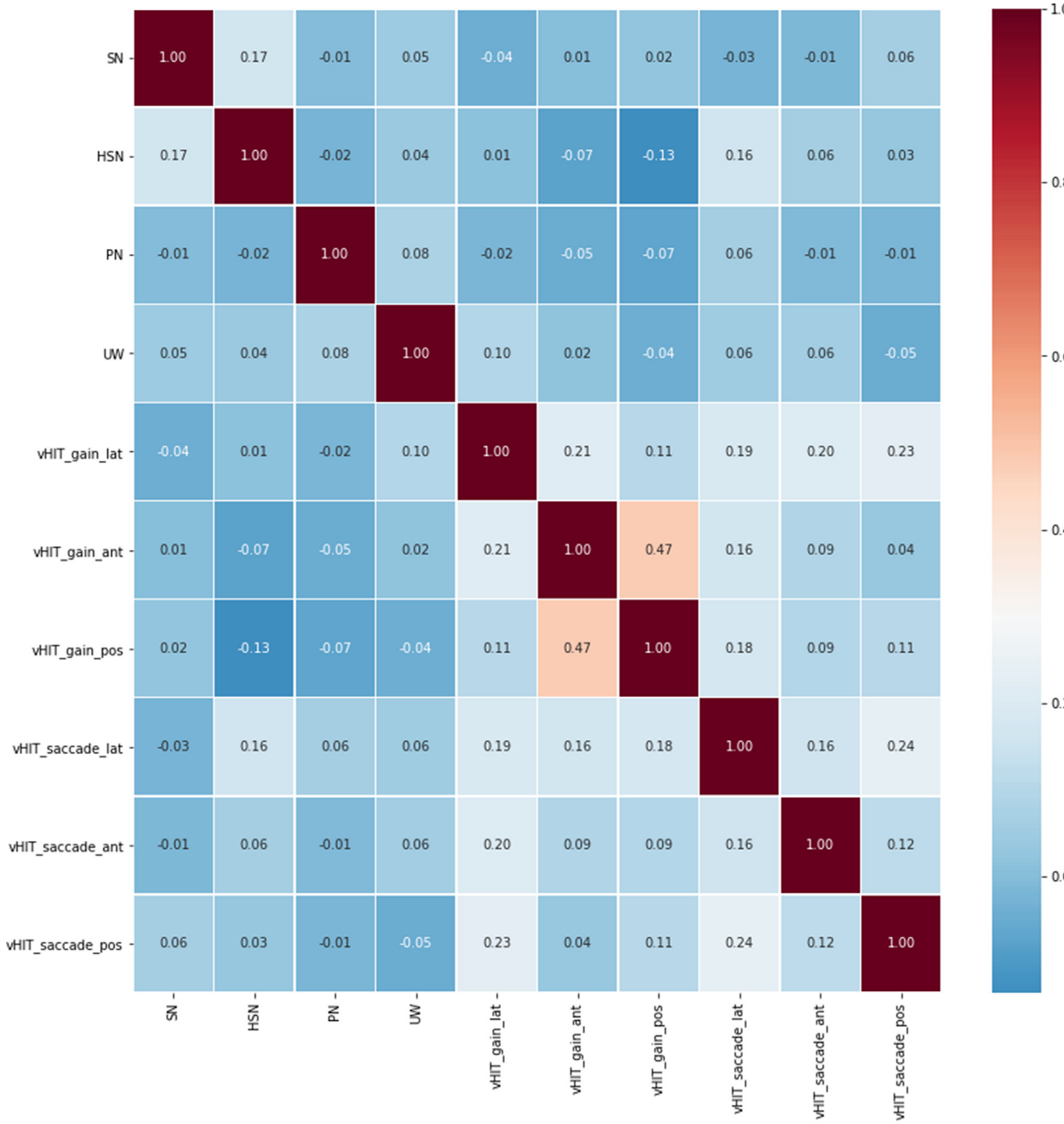

Fig. 2. Correlation heatmap of all vestibular parameters. The correlation of coefficients between every two components was present inside the block, and a correlation greater than 0.3 was shown as red. The prefix of lat, ant, and pos represents anterior SCC, anterior SCC, and posterior SCC individually.

dysfunction is essential for further diagnosis and management and rehabilitation intervention. Although some clinical examinations have been used worldwide for more than 100 years, their application scope and efficiency among various vestibular disorders need to be clarified. Random forest is a highly accurate and robust machine learning method to generate a good prediction, both small samples and high dimensional data. The feature importance could measure the relative importance of each feature on the prediction. In this study, the sequence of feature importances are consistent with clinical expertise choice - SN and vHIT perform well in acute vestibular disorders and UW shows a more reliable performance in EVS and CVS.

As fast screening test sets, SN, HSN combined with vHIT were quick to operate, and with steady and efficient performance in disease differentiation. SN is sensitive enough to perceive the impairment in acute unilateral vestibular disease. HSN is related to afferent stimulus asymmetry and central storage mechanism [3]. Patients with HSN had more self-perceived dizziness and worse feelings than the other [25], helping to distinguish between functional and psychological vertigo [26]. vHIT, the visual acuity test, should be considered the combination of two reflexive tests, vestibular-ocular reflex and saccade. Therefore, the evaluation of vHIT should be based on vHIT gain and saccade independently. Our former research provided evidence that vHIT gain and saccade

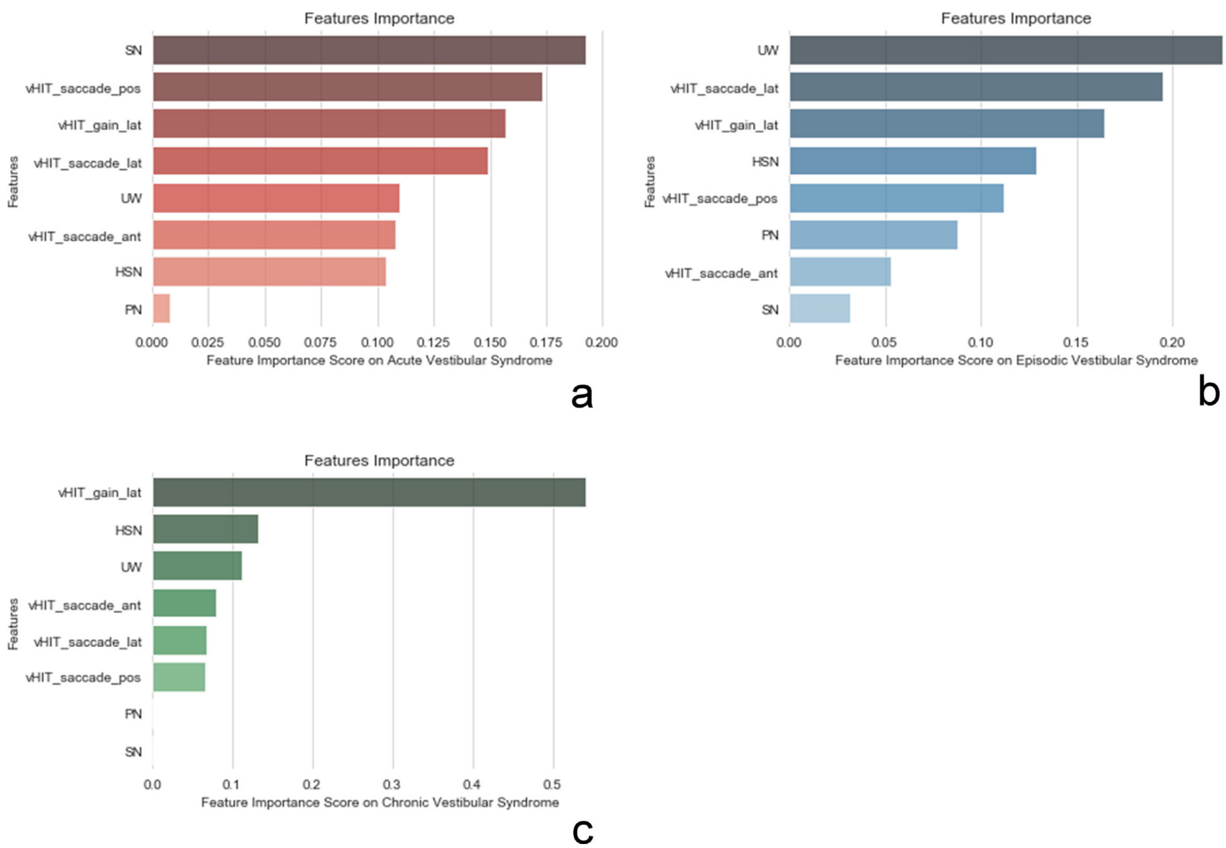

Fig. 3. Features importance on AVS, EVS, and CVS. (a. AVS: acute vestibular syndrome, b. EVS: episodic vestibular syndrome, c. CVS: chronic vestibular syndrome). The prefix of lat, ant, and pos represents anterior SCC, anterior SCC, and posterior SCC individually.

vary among different peripheral vestibular disorders [27]. Van Esch et al. also proposed that the vHIT test be used as the first test to determine vestibular dysfunction, and additional caloric tests should be performed under normal vHIT conditions [28].

Although the ability to record semicircular canal function in a short period is a distinct advantage of vHIT, the accuracy of vHIT gain in vertical SCC is also the main problem. The vertical vHIT gains have a significant downward trend due to angle limitation and elder eyelid problems. A gain lower than 0.7 may not represent a decrease in the corresponding vestibular function. In this situation, either abnormal vHIT gain or detectable saccades is meaningful for diagnosis, rather than rely on the combination results. Besides further investigation on vertical SCCs, more complicated situations should take into concern. Considering the character of “black box” in RF make it hard to interpret, more ML method in the larger dataset should be introduced to make the comparison. The rotatory chair test is another practical vestibular examination that will be introduced in our further study.

5. Conclusion

This research applied a machine learning method to categorize various vestibular disorders. Multiclassification of diseases is based on different vestibular examination variables. Clinical test sets including SN, HSN, and vHIT have been quick to operate, and good performance in disease distinction. CT could be considered when the symptom is not that urgent.

Sponsorship

None.

Funding source

None

Conflicts of interest

This research received no specific grant. The authors declare that they have no conflicts of interest.

References

[1] Lopez C, Blanke O. Nobel prize centenary: Robert Bárány and the vestibular system. J Current Biology 2014;24:R1026–R10R8.

[2] Gerald F, Hallpike CS. Studies in human vestibular function I. Observation on the direct preponderance of caloric nystagmus resulting from cerebral lesions. Brain 1942:115–37.

[3] Hain TC, Fetter M, Zee DS. Head-shaking nystagmus in patients with unilateral peripheral vestibular lesions. Am J Otolaryngol 1987;8:36–47.

[4] Panichi R, Faralli M, Bruni R, Kiriakarely A, Occhigrossi C, Ferraresi A, et al. Asymmetric vestibular stimulation reveals persistent disruption of motion perception in unilateral vestibular lesions. J Neurophysiol 2017;118:2819–32.

[5] Halmagyi GM, Curthoys IS. A Clinical Sign of Canal Paresis. Arch Neurol 1988;45:737–9.
[6] MacDougall HG, Weber KP, McGarvie LA, Halmagyi GM. Curthoys IS: The video head impulse test: diagnostic accuracy in peripheral vestibulopathy. Neurology 2009;73:1134–41.
[7] MacDougall HG, McGarvie LA, Halmagyi GM, Curthoys IS, Weber KP. Application of the video head impulse test to detect vertical semicircular canal dysfunction. Otol Neurotol 2013;34:974–9.
[8] Morrison M, Korda A, Zamaro E, Wagner F, Caversaccio MD, Sauter TC, et al. Mantokoudis G: Paradigm shift in acute dizziness: is caloric testing obsolete? J Neurol 2021.
[9] Hannigan IP, Welgampola MS, Watson SRD. Dissociation of caloric and head impulse tests: a marker of Meniere's disease. J Neurol 2021;268:431–9.
[10] McCaslin D, Rivas A, Jacobson G, Bennett M. The dissociation of video head impulse test (vHIT) and bithermal caloric test results provide topological localization of vestibular system impairment in patients with "definite" Ménière's disease. J Am J Audiol 2015;24:1–10.
[11] Mezzalira R, Bittar RSM, do Carmo, Bilecki-Stipsky MM, Brugnera C, Grasel SS. Sensitivity of caloric test and video head impulse as screening test for chronic vestibular complaints. Clinics (Sao Paulo) 2017;72:469–73.
[12] Kamogashira T, Fujimoto C, Kinoshita M, Kikkawa Y, Yamasoba T, Iwasaki S. Prediction of vestibular dysfunction by applying machine learning algorithms to postural instability. Front Neurol 2020;11:7.
[13] Priesol AJ, Cao M, Brodley CE, Lewis RF. Clinical vestibular testing assessed with machine-learning algorithms. JAMA Otolaryngol Head Neck Surg 2015;141:364–72.
[14] Strobl R, Grozinger M, Zwergal A, Huppert D, Filippopulos F, Grill E. A set of eight key questions helps to classify common vestibular disorders-results from the DizzyReg patient registry. Front Neurol 2021;12:670944.
[15] Committee on Hearing and Equilibrium guidelines for the diagnosis and evaluation of therapy in Meniere's disease. American Academy of Otolaryngology-Head and Neck Foundation. Inc. Otolaryngol Head Neck Surg 1995;113:181–5.
[16] Robert Slater. Benign recurrent vertigo. J Neurol Neurosurg Psychiatry 1979.
[17] von Brevern M, Bertholon P, Brandt T, Fife T, Imai T, Nuti D, et al., orientation: Benign paroxysmal positional vertigo: Diagnostic criteria. 2015, 25:105-17.
[18] Staab JP, Eckhardt-Henn A, Horii A, Jacob R, Strupp M, Brandt T, et al. Diagnostic criteria for persistent postural-perceptual dizziness (PPPD): Consensus document of the committee for the Classification of Vestibular Disorders of the Barany Society. J Vestib Res 2017;27:191–208.
[19] Lempert T, Olesen J, Furman J, Waterston J, Seemungal B, Carey J, et al. Vestibular migraine: diagnostic criteria. J Vestib Res 2012;22:167–72.
[20] Strupp M, Kim JS, Murofushi T, Straumann D, Jen JC, Rosengren SM, et al. Bilateral vestibulopathy: Diagnostic criteria Consensus document of the Classification Committee of the Barany Society. J Vestib Res 2017;27:177–89.
[21] Kamei T. Delayed endolymphatic hydrops as a clinical entity. Int Tinnitus J 2004;10:137–43.
[22] Strupp M, Lopez-Escamez JA, Kim JS, Straumann D, Jen JC, Carey J, et al. Vestibular paroxysmia: Diagnostic criteria. J Vestib Res 2016;26:409–15.
[23] Bisdorff AR, Staab JP, Newman-Toker DE. Overview of the international classification of vestibular disorders. Neurol Clin 2015;33:541–50 vii.
[24] Bronstein A. Oxford textbook of vertigo and imbalance. J Biomed Instrumentation Technol 2013.
[25] Angeli SI, Velandia S, Snapp H. Head-shaking nystagmus predicts greater disability in unilateral peripheral vestibulopathy. Am J Otolaryngol 2011;32:522–7.
[26] Asawavichiangianda S, Fujimoto M, Mai M, Desroches H, Rutka J. Significance of head-shaking nystagmus in the evaluation of the dizzy patient. Acta Otolaryngol Suppl 1999;540:27–33.
[27] Du Y, Ren L, Liu X, Guo W, Wu Z, Yang S. The characteristics of vHIT gain and PR score in peripheral vestibular disorders. Acta Otolaryngol 2021;141:43–9.
[28] van Esch BF, Nobel-Hoff GE, van Benthem PP, van der Zaag-Loonen HJ, Bruintjes TD. Determining vestibular hypofunction: start with the video-head impulse test. Eur Arch Otorhinolaryngol 2016;273:3733–9.