ORIGINAL COMMUNICATION

Development and validation of a classification algorithm to diagnose and differentiate spontaneous episodic vertigo syndromes: results from the DizzyReg patient registry

Michael Groezinger[1] · Doreen Huppert[2,3] · Ralf Strobl[1] · Eva Grill[1,2]

Abstract

Background Spontaneous episodic vertigo syndromes, namely vestibular migraine (VM) and Menière's disease (MD), are difficult to differentiate, even for an experienced clinician. In the presence of complex diagnostic information, automated systems can support human decision making. Recent developments in machine learning might facilitate bedside diagnosis of VM and MD.

Methods Data of this study originate from the prospective patient registry of the German Centre for Vertigo and Balance Disorders, a specialized tertiary treatment center at the University Hospital Munich. The classification task was to differentiate cases of VM, MD from other vestibular disease entities. Deep Neural Networks (DNN) and Boosted Decision Trees (BDT) were used for classification.

Results A total of 1357 patients were included (mean age 52.9, SD 15.9, 54.7% female), 9.9% with MD and 15.6% with VM. DNN models yielded an accuracy of $98.4 \pm 0.5\%$, a precision of $96.3 \pm 3.9\%$, and a sensitivity of $85.4 \pm 3.9\%$ for VM, and an accuracy of $98.0 \pm 1.0\%$, a precision of $90.4 \pm 6.2\%$ and a sensitivity of $89.9 \pm 4.6\%$ for MD. BDT yielded an accuracy of $84.5 \pm 0.5\%$, precision of $51.8 \pm 6.1\%$, sensitivity of $16.9 \pm 1.7\%$ for VM, and an accuracy of $93.3 \pm 0.7\%$, precision $76.0 \pm 6.7\%$, sensitivity $41.7 \pm 2.9\%$ for MD.

Conclusion The correct diagnosis of spontaneous episodic vestibular syndromes is challenging in clinical practice. Modern machine learning methods might be the basis for developing systems that assist practitioners and clinicians in their daily treatment decisions.

Keywords Vestibular migraine · Menière's disease · Vestibular disease · Machine learning · Classification

✉ Eva Grill
  eva.grill@med.uni-muenchen.de

[1] Institute for Medical Information Processing, Biometry and Epidemiology, Ludwig-Maximilians-Universität München (LMU), Marchioninistr. 15, 81377 München, Germany

[2] German Centre for Vertigo and Balance Disorders, University Hospital Munich, Campus Grosshadern, Munich, Germany

[3] Department of Neurology, Ludwig-Maximilians-University, Munich, Germany

Introduction

Spontaneous episodic vertigo syndromes—vestibular migraine (VM) as the most frequent form of episodic vertigo with a lifetime prevalence of 1% [1] and Menière's disease (MD) with a prevalence of about 0.2–0.5%—are difficult to differentiate, even for an experienced clinician after a thorough anamnesis and clinical examination. Both VM and MD can have considerable impact on patients' daily life and functioning, so timely and appropriate therapy is essential. The diagnostic criteria for VM and MD correspond to core symptoms of these diseases from the literature for VM [2–10], and for MD [11]. These diagnostic criteria were formulated by the International Bárány Society for Neuro-Otology in 2012 [12] and 2015 [13] (Table 1) Epidemiological studies revealed a coincidence of both conditions in one individual [14–16] which may pose

Table 1 Current definitions of Vestibular migraine and Menière's disease

Vestibular migraine	At least 5 episodes with vestibular symptoms of moderate or severe intensity, lasting 5 min. to 72 h
	Current or previous history of migraine with or without aura according to the International Classification of Headache Disorders (ICHD)
	One or more migraine features with at least 50% of the vestibular episodes:
	headache with at least two of the following characteristics: one-sided location, pulsating quality, moderate or severe pain intensity, aggravation by routine physical activity
	photophobia and phonophobia
	visual aura
	Not better accounted for by another vestibular or ICHD diagnosis
	Probable: A, B or C, and D
Menière's disease	Two or more spontaneous episodes of vertigo, each lasting 20 min to 12 h
	Audiometrically documented low- to medium-frequency sensorineural hearing loss in one ear, defining the affected ear on at least one occasion before, during or after one of the episodes of vertigo
	Fluctuating aural symptoms (hearing, tinnitus, or fullness) in the affected ear
	Not better accounted for by another vestibular diagnosis
	Probable: A, C, and D

Adapted from Lempert et al. 2012 [12] for vestibular migraine and Lopez-Escamez et al. 2015 [13] for Menière's disease

considerable uncertainties in the evaluation of the response of medical treatment.

Arguably, clinicians with longstanding expertise and experience will be able to integrate all these diagnostic dimensions. In situations of routine primary care, however, where this expertise is not easily available, diagnostic information will often be too complex to be summarized into one clinical decision. In the presence of seemingly overwhelming information, automated systems can support human decision making. This will be quite straightforward if the logic of the decision-making process is known. However, in the absence of an existing algorithm, diagnosis based on clinical expert experience is a process that is hard to describe.[17].

Several attempts have been made to facilitate bedside diagnosis of vestibular diseases using machine learning (ML) approaches. Some of these approaches were able to classify specific diseases with good accuracy. Interestingly, in these ML studies MD was persistently difficult to predict [18, 19]. As a further development, a combination of several data mining techniques has recently been proposed for 12 vestibular diseases, MD among these, however, without specific solutions for single hard-to-differentiate disease entities [20]. Another study was able to differentiate unilateral canal damage, one potential feature in MD, with 76% accuracy [21]. Combining these promising approaches using recent developments in ML, might be even more successful [22].

Objectives of this study were to develop and test a classification algorithm to differentiate Vestibular Migraine and Menière's disease in clinical practice.

Methods

Our data originates from the DizzyReg patient registry of the German Center for Vertigo and Balance Disorders (DSGZ). Details on purpose and data collection have been published elsewhere [23, 24]. In brief, the registry is a prospective data base including all relevant anamnestic, sociodemographic, diagnostic and therapeutic information of patients that presented at the DSGZ, a specialized tertiary treatment center at the hospital of the LMU Munich with approximately 3000 patients per year since 2016.

A positive vote of the local institutional review board and detailed consulting on data protection issues from the regional data protection officer was obtained for the registry. Informed consent was obtained from each patient or the patient's legal surrogate.

Patients received a thorough neuro-otologic assessment and a validated standardized diagnosis according to current international guidelines [12, 25]. Clinicians with longstanding experience diagnosed all patients. This clinical diagnosis represents the gold standard for our study.

DizzyReg centralizes all data that is collected in electronical health records or medical discharge letters. The data is stored on servers within the hospital firewall and state-of-the-art security techniques are used to protect the data. The data is either retrieved online from the clinical workplace system (CWS) or entered manually via a web-based system. It will only be released fully anonymized

only for predefined purposes after review of an external steering group [23].

Reporting of methods follows the guidelines by Luo et. al.[26].

The classification task was to differentiate cases of VM, MD from other vestibular disease entities. As the number of VM and MD cases was relatively small, we decided to formulate two separate one-vs-all models, one for the prediction of VM and one for the prediction of MD. This approach has been used successfully before [19].

We defined precision (positive predictive value) and sensitivity as criteria to judge the quality of the classification. Sensitivity is the probability of the ML classifier to classify a patient as having MD or VM when this disease is truly present, i.e., the number of correctly classified cases (the true positives TP) divided by the total number of persons with this disease. Precision is the number of correctly classified cases among all cases diagnosed by the algorithm. Additionally, the F-measure [27] was calculated as a combination of precision and sensitivity and as a combination of several thresholds of sensitivity and specificity, respectively.

$$\text{Precision} = \frac{\text{true positives (TP)}}{\text{TP + false positives (FP)}}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP + false negatives (FN)}}$$

$$\text{Specificity} = \frac{\text{true negatives (TN)}}{\text{TN + FP}}$$

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}}$$

$$F\text{ - measure} = \frac{2 *(\text{precision + sensitivity})}{\text{precision*sensitivity}}$$

To build the classification model we first identified candidate predictors that were statistically associated with presence or absence of either VM or MD. Selection was based on VM vs. MD. The initial data set contained 582 variables. Examples of these variables along with their clinical categories is available in the electronic supplement. Variables would be candidates if they had a p-value of 0.2 or below in bivariate chi-square tests (for categorical variables) or Mann–Whitney-Tests (for metric variables). This analysis revealed 105 variables that were then used as input for the subsequent models. Table 1 of the supplementary material lists a summary of all variables used for training.

Variables were then examined for missingness and those with insufficient information (missingness above 90%) were not included in the analyses. Likewise, variables with little variation, and redundant variables were deleted. This reduced the number of variables to 96.

A total of 20 variables had missing values. Overall, 8446 values were missing (6.2%), all patient records had at least one missing value. These remaining 20 variables were imputed.

In line with current recommendations we used multiple imputation techniques for missing data [28], namely Multivariate Imputation by Chained Equations (MICE) [28, 29]. In brief, MICE specify a multivariate imputation model on a variable-by-variable basis by a set of conditional distributions. From those conditional distributions imputed values are drawn with Markov Chain Monte Carlo techniques. To acknowledge uncertainty in the imputation process MICE yields five different imputed data sets that were then used for training the classification models [28].

Deep Neural Networks (DNN) were used for classification. In brief, neural networks are computational graphs which perform compositions of simpler functions to provide a more complex function used for separation in classification tasks. This type of model has proven to very successful in classification tasks in medical imaging.

DNN provide a wide range of meta parameters which then need to be tuned to optimize the classification results. To limit the computing effort, we relied on recommended standards for meta-parameters and varied breadth (number of nodes per hidden layer) and depth (number of hidden layers) to examine their impact on classification behavior. For a shallow network we trained one hidden layer with 10 different numbers of nodes, each 10 times. The best configuration was selected based on the F-measure. To test the impact of a second hidden layer, we added another hidden layer (10 different number of nodes and 10 training runs). The best configuration for the two layers was again selected by F-measure. For the deep configuration we started with 4 hidden layers and added four additional ones and changed the number of nodes similarly to the shallow configuration.

To identify the best model configuration in terms of breadth vs. depth of the DNN we first used a shallow configuration of the network (a maximum of two hidden layers). After optimizing breadth for the first layer, a second hidden layer was added and optimized. This yielded the best shallow configuration. Then we extended the depth of the network further into a deep configuration with a maximum of eight hidden layers. Deep networks generally do not need the same breadth as shallow configurations to achieve adequate training results, we still kept the breadth in the deep configuration to 1000 nodes per hidden layer. To obtain unbiased estimation of the quality criteria an independent validation set was defined by setting aside 20% of the participants. This validation set was used to estimate accuracy, precision, specificity

and F-measure. This was repeated ten times with a different randomly drawn validation set.

As sample size was relatively small and we wanted to improve training results we then repeatedly trained the whole network again, similar to the work by Erhan [30–33]. We conducted an initial training run, kept the obtained weight parameters, and trained the network again but with a new shuffled random train/test data split. This was used for the best shallow and the best deep network configuration from the initial DNN. The pre-training workflow is described in Fig. 1 of the supplementary material.

Additionally, boosted decision tree (BDT) models were applied [34]. In contrast to DNN, these models yield information on the relative importance of single variables for the model [35]. To identify the best model configuration, we applied sensitivity analyses for the maximum number of trees in the model while keeping all other model parameters fixed (see Appendix II). Training and evaluation was carried out for all five imputed data sets and for both VM and MD as binary outcomes, i.e., prediction of VM or MD vs. all other diagnoses.

We implemented the models in the Python programming language, supported by the TensorFlow library v1.15.

Results

A total of 1357 patients were included (mean age 52.9, SD 15.9, 54.7% female), among these 9.9% with MD and 15.6% with VM. The most frequent other diagnoses were benign paroxysmal positional vertigo (10.4%), chronic unilateral vestibular failure (7.0%), bilateral vestibular failure (5.1%), and polyneuropathy (4.5%).

Shallow Deep Neural Networks (DNN) classified MD with a mean F-measure $55.5 \pm 2.7\%$ across all five imputed data sets (accuracy $91.4 \pm 0.6\%$, precision $55.4 \pm 0.6\%$, sensitivity $54.7 \pm 2.8\%$). Using pre-training on the best shallow network configuration improved F-measure to $90.0 \pm 4.6\%$, accuracy to $98.0 \pm 1.0\%$, precision to $90.4 \pm 6.2\%$ and sensitivity to $89.9 \pm 4.6\%$. Figure 1 shows results averaged over all imputed data sets. Figure 3 depicts the results for pre-training for MD and their values over the number of training runs.

The shallow configuration classified VM yielding a mean F-measure of $36.8 \pm 0.8\%$ (accuracy $81.8 \pm 0.6\%$, precision $39.5 \pm 0.6\%$, sensitivity $35.1 \pm 2.7\%$). Pre-training using the best deep configuration (4 layers 50 nodes) improved classification considerably yielding a F-measure of $90.5 \pm 3.0\%$ (accuracy $98.4 \pm 0.5\%$, precision $96.3 \pm 3.9\%$, sensitivity $85.4 \pm 3.9\%$,). Results of shallow and deep configurations as means of the five imputed data sets are shown in Fig. 2. Pre-training results per run VM are shown in Fig. 3.

Boosted decision trees for Menière's disease yielded a mean F-measure $53.3 \pm 3.0\%$, (accuracy of $93.3 \pm 0.7\%$, precision $76.0 \pm 6.7\%$, sensitivity $41.7 \pm 2.9\%$). Vestibular migraine was classified with a mean F-measure $27.6 \pm 5.1\%$ (accuracy of $84.5 \pm 0.5\%$, precision $51.8 \pm 6.1\%$, sensitivity $16.9 \pm 1.7\%$) (Table 2). After training, BDT provide a listing of features according their importance in the decision
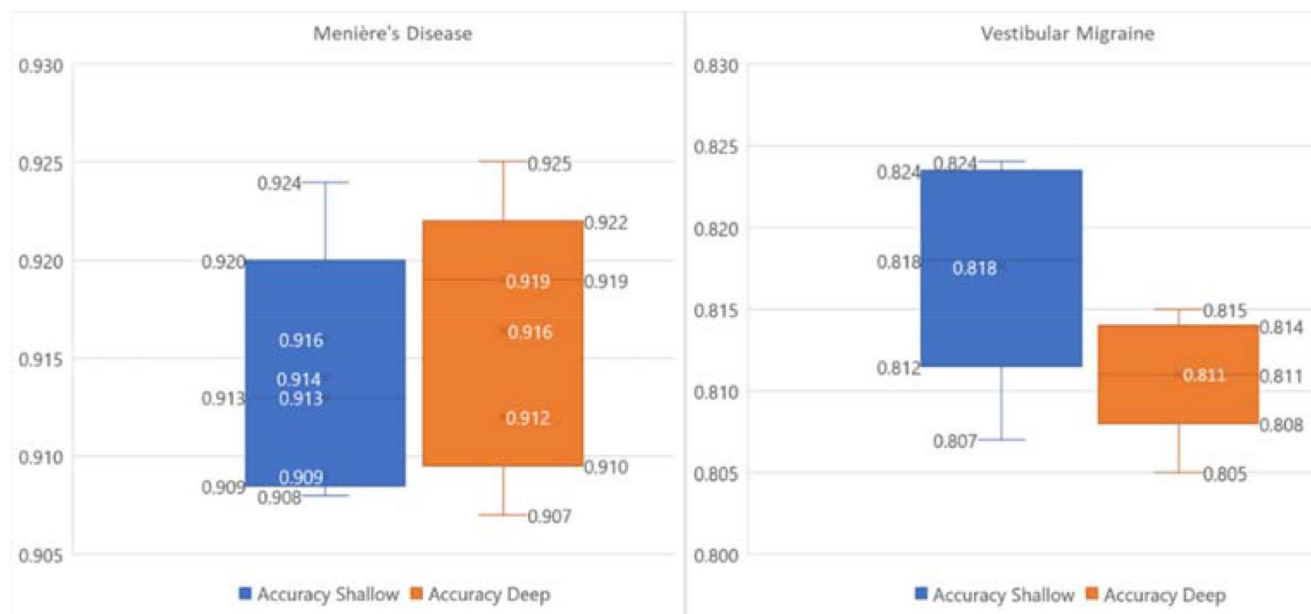

Fig. 1 Distribution of the test accuracy predicting Menière's disease and vestibular migraine over 5 imputed data sets for shallow and deep configurations of the Deep Neural Network model
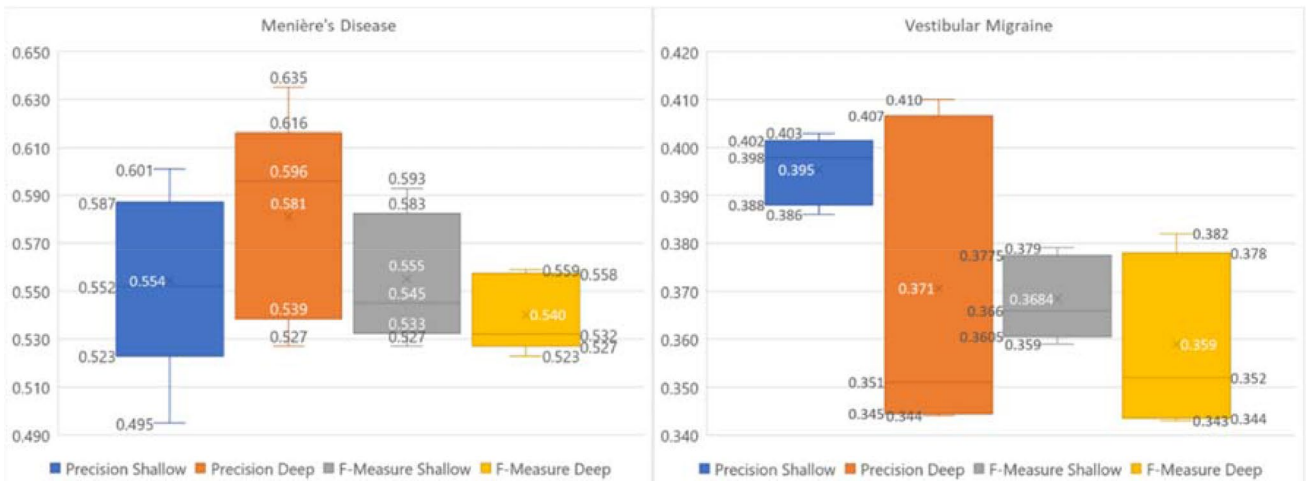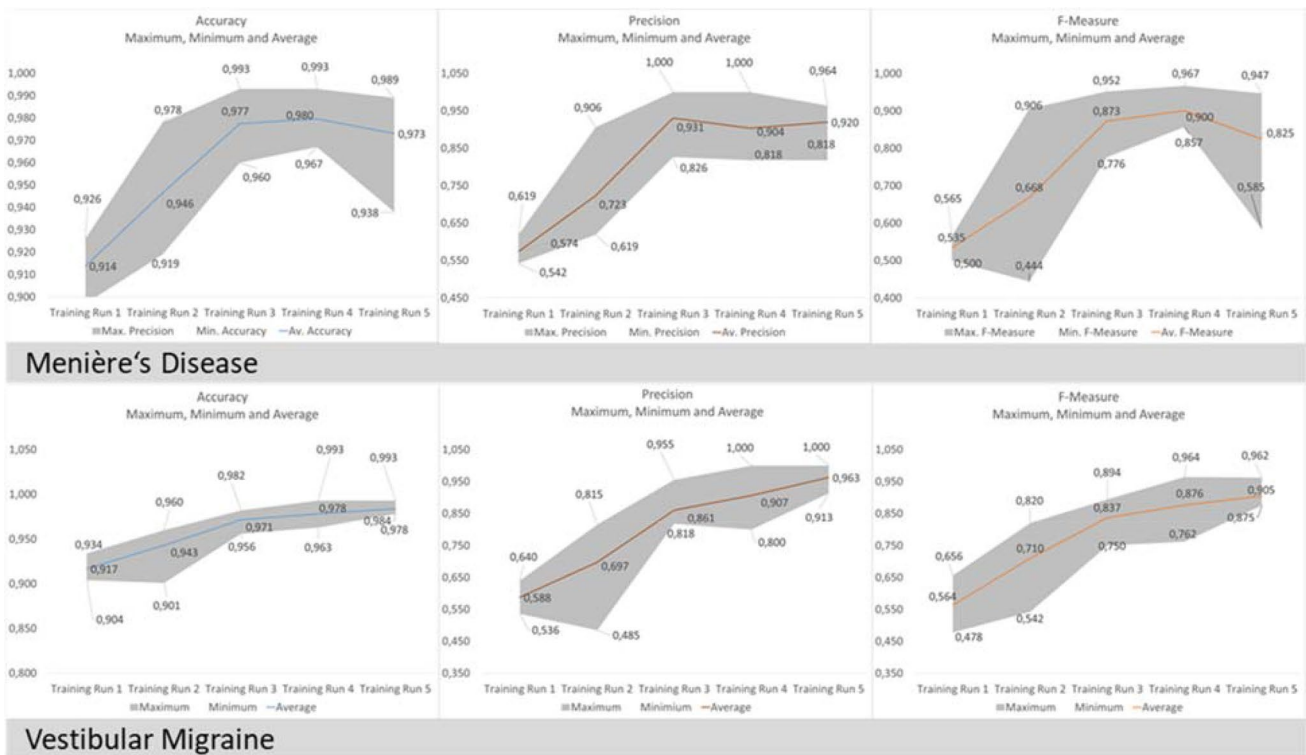
Fig. 2 Distribution of the test accuracy predicting Menière's disease and vestibular migraine over 5 imputed data sets for shallow and deep configurations of the Deep Neural Network model

Fig. 3 Accuracy, precision and F-measure after pre-training for Menière's disease with a shallow network configuration (layer 1/2: 200/50 nodes) and for vestibular migraine with a deep network con-figuration (4 layers/50 nodes each). The lines represent means, the borders of grey areas represent minimum and maximum values over all five imputed data sets

process. Table 2 lists the 10 most important variables for Menière's disease and vestibular migraine, respectively.

Discussion

Differentiating Menière's disease (MD) from vestibular migraine (VM) is a challenge when patients present with episodic vertigo. The Deep Neural Network architecture

Table 2 Results of boosted decision tree models. We show the variables with the highest importance for prediction of Menière's disease and vestibular migraine identified by a boosted decision tree

Menière's disease		Vestibular migraine	
Variable	Importance	Variable	Importance
Age	9,13%	Age	11,94%
Caloric side difference	7,60%	Caloric side difference	7,06%
Patient reported vomiting	7,12%	Gain left (ms)	6,62%
Vertigo lasting several hours	4,92%	Gain right (ms)	5,68%
Gain left (ms)	4,63%	Anamnestic headache	2,62%
Patient reported hearing loss	4,38%	Hours of sleep in 24 h	1,91%
Impaired hearing (audiometry)	2,90%	Visual acuity	1,61%
Hours of sleep in 24 h	2,84%	Patient reported nausea	1,55%
Gain right (ms)	2,56%		
Visual acuity	2,33%		

of this study, augmented by pre-training, showed excellent classification performance for both MD and VM. This result was achieved with a relatively small set of variables, making the system potentially available for applications that run in non-expert settings.

The accuracy of the networks reported here exceeds that of previous models at large. A study using Support Vector Machines and k-Nearest Neighbour reported an overall accuracy of 79.8% [19] for a model differentiating MD from 8 other vestibular syndromes, but this model did not include patients with VM. Likewise, a Support Vector Machines approach to classify unilateral vestibulopathy yielded an accuracy of 76% [21]

The results indicate that DNN perform better to classify MM than to classify VM. Indeed, the differentiation of VM from other episodic vertigo syndromes may be difficult, especially when it comes to differentiating MD and VM. It is important to stress that up to 60% of patients with MD also fulfill some or all of the diagnostic criteria for VM and vice versa [15]. A combination of ear symptoms like a ringing in the ears and dizziness is among the current diagnostic criteria for MD [13]. Another factor that complicates the differentiation of VM and MD is the absence of headache. Interestingly, a history of headache was among the more prominent features that indicated VM in our study. Differentiation will be even more difficult if attacks of MD occur without ear symptoms, which is especially frequent at the beginning of MD [11]. Moreover, both diagnoses (uni- or bilateral Menière's disease and migraine with and without aura) coincided in 56% compared with 25% in an age-matched control group [14–16].

One major disadvantage of DNN is the "black box" approach, because DNN do not yield any information about the relative importance of single variables to the model. Thus, DNN are not particularly useful to create parsimonious models. We, therefore, applied a different machine learning method, boosted decision trees, firstly to compare their performance to that of DNN, but secondly to investigate variable importance that can be useful to concentrate on characteristics with high value for clinical decision-making.

For both MD and VM, caloric side difference and gain of the head impulse test were important predictors derived from boosted decision trees. Most patients with VM have mild central ocular motor disorders in the form of gaze-evoked nystagmus, saccadic smooth pursuit eye movements, or a positional nystagmus even in the attack-free interval [5, 36–39]. Regarding predictors that do not rely on instrumental tests, patient reported hearing loss, vomiting and a duration of attacks of several hours were the most striking characteristics for MD, while a history of headache and patient reported nausea were indicative for VM. These results are not completely surprising, given the typical features of hearing loss for MD and headache for VM. Nevertheless, they can be used for further refinement of diagnostic algorithms.

We acknowledge several limitations. First, identification of MD and VM relied on the clinical diagnoses that were made in a tertiary care center. While these diagnoses are certainly accurate, there might be complex cases, where a definite diagnostic decision can only be made after follow-up. On the other hand, this uncertainty is unavoidable in a cross-sectional study, and we can be sure that diagnoses were generally correct, because these were based on the established Bárány criteria. Second, although boosted decision tree models in contrast to DNN have the advantage to indicate variable importance we found that decision trees had a slightly better accuracy, better precision but lower sensitivity than the DNN models, i.e., these models are better to predict absence of disease than presence. Arguably, this inferior result is due to the relatively low prevalence of MD and VM for this one-vs-all approach. Certainly, the multi-class approach, i.e., building models that can predict multiple classes of diagnoses at one time, is of superior clinical

utility. Future approaches should strive for multicenter data collection to have a sufficient number of cases to address the problem. A multi-class approach for boosted decision trees seems to be an advisable solution for future models.

The correct diagnosis of spontaneous episodic vestibular syndromes is challenging in clinical practice. Modern machine learning methods might be the basis for developing systems that assist practitioners and clinicians in their daily treatment decisions.

Compliance with ethical standards

Conflicts of interest None declared.

Ethics approval Obtained from the Institutional Review Board of the Medical Faculty, University of Munich.

Consent for publication All authors read and consented to the final version of the manuscript.

Code availability Code is available upon reasonable scientific request.

References

1. Neuhauser HK, Radtke A, von Brevern M, Feldmann M, Lezius F, Ziese T, Lempert T (2006) Migrainous vertigo: prevalence and impact on quality of life. Neurology 67(6):1028–1033. https://doi.org/10.1212/01.wnl.0000237539.09942.06

2. Brandt T, Dieterich M (2017) The dizzy patient: don't forget disorders of the central vestibular system. Nat Rev Neurol 13(6):352–362. https://doi.org/10.1038/nrneurol.2017.58

3. Cohen JM, Bigal ME, Newman LC (2011) Migraine and vestibular symptoms–identifying clinical features that predict "vestibular migraine". Headache 51(9):1393–1397. https://doi.org/10.1111/j.1526-4610.2011.01934.x

4. Cutrer FM, Baloh RW (1992) Migraine-associated dizziness. Headache 32(6):300–304. https://doi.org/10.1111/j.1526-4610.1992.hed3206300.x

5. Dieterich M, Brandt T (1999) Episodic vertigo related to migraine (90 cases): vestibular migraine? J Neurol 246(10):883–892. https://doi.org/10.1007/s004150050478

6. Furman JM, Balaban CD (2015) Vestibular migraine. Ann NY Acad Sci 1343:90–96. https://doi.org/10.1111/nyas.12645

7. Furman JM, Marcus DA, Balaban CD (2013) Vestibular migraine: clinical aspects and pathophysiology. Lancet Neurol 12(7):706–715. https://doi.org/10.1016/S1474-4422(13)70107-8

8. Kayan A, Hood JD (1984) Neuro-otological manifestations of migraine. Brain 107(Pt 4):1123–1142. https://doi.org/10.1093/brain/107.4.1123

9. O'Connell Ferster AP, Priesol AJ, Isildak H (2017) The clinical manifestations of vestibular migraine: a review. Auris Nasus Larynx 44(3):249–252. https://doi.org/10.1016/j.anl.2017.01.014

10. Strupp M, Versino M, Brandt T (2010) Vestibular migraine. Handb Clin Neurol 97:755–771. https://doi.org/10.1016/S0072-9752(10)97062-0

11. Brandt T, Dieterich M, Strupp M (2013) Vertigo and dizziness: common complaints, 2nd edn. Springer, London

12. Lempert T, Olesen J, Furman J, Waterston J, Seemungal B, Carey J, Bisdorff A, Versino M, Evers S, Newman-Toker D (2012) Vestibular migraine: diagnostic criteria. J Vestib Res 22(4):167–172. https://doi.org/10.3233/VES-2012-0453

13. Lopez-Escamez JA, Carey J, Chung WH, Goebel JA, Magnusson M, Mandala M, Newman-Toker DE, Strupp M, Suzuki M, Trabalzini F, Bisdorff A, Classification Committee of the Barany S, Japan Society for Equilibrium R, European Academy of O, Neurotology, Equilibrium Committee of the American Academy of O-H, Neck S, Korean Balance S (2015) Diagnostic criteria for Meniere's disease. J Vestib Res 25(1):1–7

14. Lopez-Escamez JA, Dlugaiczyk J, Jacobs J, Lempert T, Teggi R, von Brevern M, Bisdorff A (2014) Accompanying symptoms overlap during attacks in Meniere's disease and vestibular migraine. Front Neurol 5:265. https://doi.org/10.3389/fneur.2014.00265

15. Neff BA, Staab JP, Eggers SD, Carlson ML, Schmitt WR, Van Abel KM, Worthington DK, Beatty CW, Driscoll CL, Shepard NT (2012) Auditory and vestibular symptoms and chronic subjective dizziness in patients with Meniere's disease, vestibular migraine, and Meniere's disease with concomitant vestibular migraine. Otol Neurotol 33(7):1235–1244. https://doi.org/10.1097/MAO.0b013e31825d644a

16. Radtke A, Lempert T, Gresty MA, Brookes GB, Bronstein AM, Neuhauser H (2002) Migraine and Meniere's disease: is there a link? Neurology 59(11):1700–1704. https://doi.org/10.1212/01.wnl.0000036903.22461.39

17. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. The MIT Press, Cambridge, Massachusetts, Adaptive computation and machine learning

18. Juhola M (2008) On machine learning classification of otoneurological data. Stud Health Technol Inform 136:211–216

19. Varpa K, Joutsijoki H, Iltanen K, Juhola M (2011) Applying one-vs-one and one-vs-all classifiers in k-nearest neighbour method and support vector machines to an otoneurological multi-class problem. Stud Health Technol Inform 169:579–583

20. Exarchos TP, Rigas G, Bibas A, Kikidis D, Nikitas C, Wuyts FL, Ihtijarevic B, Maes L, Cenciarini M, Maurer C, Macdonald N, Bamiou DE, Luxon L, Prasinos M, Spanoudakis G, Koutsouris DD, Fotiadis DI (2016) Mining balance disorders' data for the development of diagnostic decision support systems. Comput Biol Med 77:240–248. https://doi.org/10.1016/j.compbiomed.2016.08.016

21. Priesol AJ, Cao M, Brodley CE, Lewis RF (2015) Clinical vestibular testing assessed with machine-learning algorithms.

JAMA Otolaryngol Head Neck Surg 141(4):364–372. https://doi.org/10.1001/jamaoto.2014.3519

22. Aggarwal CC (2018) Neural networks and deep learning. Textbook. https://doi.org/10.1007/978-3-319-94463-0

23. Grill E, Muller T, Becker-Bense S, Gurkov R, Heinen F, Huppert D, Zwergal A, Strobl R (2017) DizzyReg: the prospective patient registry of the German center for vertigo and balance disorders. J Neurol 264(Suppl 1):34–36. https://doi.org/10.1007/s00415-017-8438-7

24. Grill E, Akdal G, Becker-Bense S, Hubinger S, Huppert D, Kentala E, Strobl R, Zwergal A, Celebisoy N (2018) Multicenter data banking in management of dizzy patients: first results from the dizzynet registry project. J Neurol 265(Suppl 1):3–8. https://doi.org/10.1007/s00415-018-8864-1

25. Lopez-Escamez JA, Carey J, Chung WH, Goebel JA, Magnusson M, Mandala M, Newman-Toker DE, Strupp M, Suzuki M, Trabalzini F, Bisdorff A (2016) [Diagnostic criteria for Meniere's disease. Consensus document of the Barany Society, the Japan Society for Equilibrium Research, the European Academy of Otology and Neurotology (EAONO), the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) and the Korean Balance Society]. Acta Otorrinolaringol Esp 67(1):1–7. 10.1016/j.otorri.2015.05.005

26. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, Shilton A, Yearwood J, Dimitrova N, Ho TB, Venkatesh S, Berk M (2016) Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res 18(12):e323. https://doi.org/10.2196/jmir.5870

27. Rousseau R (2018) The F-measure for research priority. J Data Infor Sci 3(1):1–18. https://doi.org/10.2478/jdis-2018-0001

28. Buuren Sv (2018) Flexible imputation of missing data. Chapman and Hall/CRC interdisciplinary statistics series, Second edition. edn. CRC Press, Taylor and Francis Group, Boca Raton

29. van Buuren S, Groothuis-Oudshoorn K (2011) Mice: multivariate imputation by chained equations in R. J Stat Softw 1(3):2011

30. SrivastavaGeoffrey NH, Krizhevsky, Alex, Sutskever, Ilya, Salakhutdinov, Ruslan (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learning Res 15:1929–1958

31. Erhan D, Courville A, Bengio Y, Vincent P (2010) Why does unsupervised pre-training help deep learning? In: Yee Whye T, Mike T (eds) Proceedings of the thirteenth international conference on artificial intelligence and statistics. Proceedings of Machine Learning Research, PMLR, vol 9, pp 201–208

32. Erhan D, Manzagol PA, Bengio Y, Bengio S, Vincent P (2009) The difficulty of training deep architectures and the effect of unsupervised pre-training. In: David van D, Max W (eds) Proceedings of the twelfth international conference on artificial intelligence and statistics. Proceedings of Machine Learning Research, PMLR, vol 5, pp 153–160

33. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: Yee Whye T, Mike T (eds) Proceedings of the thirteenth international conference on artificial intelligence and statistics. Proceedings of Machine Learning Research, PMLR vol 9, pp 249–256

34. Hastie T, Tibshirani R, Friedman JH (2009) The elements of statistical learning: data mining, inference, and prediction. Springer series in statistics, 2nd edn. Springer, New York, NY

35. Flach PA (2012) Machine learning : the art and science of algorithms that make sense of data. Cambridge University Press, Cambridge, New York

36. Neugebauer H, Adrion C, Glaser M, Strupp M (2013) Long-term changes of central ocular motor signs in patients with vestibular migraine. Eur Neurol 69(2):102–107. https://doi.org/10.1159/000343814

37. Radtke A, Neuhauser H, von Brevern M, Hottenrott T, Lempert T (2011) Vestibular migraine–validity of clinical diagnostic criteria. Cephalalgia 31(8):906–913. https://doi.org/10.1177/0333102411405228

38. Radtke A, von Brevern M, Neuhauser H, Hottenrott T, Lempert T (2012) Vestibular migraine: long-term follow-up of clinical symptoms and vestibulo-cochlear findings. Neurology 79(15):1607–1614. https://doi.org/10.1212/WNL.0b013e31826e264f

39. von Brevern M, Zeise D, Neuhauser H, Clarke AH, Lempert T (2005) Acute migrainous vertigo: clinical and oculographic findings. Brain 128(Pt 2):365–374. https://doi.org/10.1093/brain/awh351