Cluster analysis to identify clinical subtypes of Ménière's disease

Running title: Ménière's disease cluster analysis

Mr John Phillips
Consultant in Otolaryngology[1]

Dr Louisa Murdin
Consultant in Audio-Vestibular Medicine[2]

Mizanur Khondoker
Medical statistician[3]

Kelly Grant[3]
Medical statistician[3]

Prof Lee Shepstone[3]
Medical statistician[3]

Dr Erika Sims[3]
Clinical Trials Unit Operations Manager

Prof Peter Rea
Consultant in Otolaryngology[4]

Mr Jonny Harcourt
Consultant in Otolaryngology[5]

[1] Department of Otolaryngology. Norfolk and Norwich University Hospitals NHS Foundation Trust
[2] Department of Otolaryngology. Guy's and St. Thomas' NHS Foundation Trust
[3] Norwich Medical School. University of East Anglia
[4] Department of Otolaryngology. University Hospitals of Leicester NHS Trust
[5] Department of Otolaryngology. Imperial College Healthcare NHS Trust

Correspondence: Dr John Phillips
Department of Otolaryngology,
Norfolk & Norwich University Hospitals NHS Foundation Trust
Colney Lane, Norwich NR4 7UY, United Kingdom
Tel: +44 1603 286770
Fax : +44 1603 287288
Email: john.phillips@mac.com

Conflicts of Interest and Source of Funding:

This project has been funded by the UK Ménière's Society. JP has received funding for his Ménière's disease research by the UK Medical Research Council (MRC) and the UK National Institute of Health Research (NIHR), but his views herein are his own and do not represent those of the MRC, the NIHR nor the UK Department of Health and Social Care.

Cluster analysis to identify clinical subtypes of Ménière's disease

ABSTRACT

Objective: To identify distinct clinical subtypes of Ménière's disease by analysing data acquired from a UK registry of patients who have been diagnosed with Ménière's disease.

Study Design: Observational study.

Methods: Patients with Ménière's disease were identified at secondary/tertiary care clinics. Cluster analysis was performed by grouping participants sharing similar characteristics and risk factors into groups based on a defined measure of similarity.

Results: 411 participants were recruited into this study. Two main clusters were identified: participants diagnosed with Ear infections (OR=0.30, p<0.014, 95%CI:0.11 to 0.78) were more likely to be allocated in Cluster 1 (C1). Participants reporting tinnitus in both ears (OR=11.89, p<0.001, 95%CI:4.08 to 34.64), low pitched tinnitus (OR=21.09, p<0.001, 95%CI:7.47 to 59.54) and those reporting stress as a trigger for vertigo attacks (OR=14.94, p<0.001, 95%CI:4.54 to 49.10) were significantly more likely to be in cluster 2 (C2). Also, participants diagnosed with Benign Paroxysmal Positional Vertigo (OR=13.14, <0.001, 95%CI:4.35 to 39.74), Autoimmune disease (OR=5.97, p<0.007, 95%CI:1.62 to 22.03), Depression (OR=4.72, p<0.056, 95%CI:0.96 to 23.24), Migraines (OR=3.13, p<0.008, 95%CI:1.34 to 7.26), Drug allergy (OR=3.25, p<0.029, 95%CI:1.13 to 9.34) and Hay fever (OR=3.12, p< 0.009, 95%CI:1.33 to 7.34) were significantly more likely to be clustered in C2.

Conclusions: This study supports the hypothesis that Ménière's disease is a heterogeneous condition with subgroups that may be identifiable by clinical features. Two main clusters were identified with differing putative aetiological factors.

Level of Evidence: 3

Key Words: Meniere's Disease, Intratympanic Steroids, Clinical Subtyping

Introduction

Ménière's disease is a vestibular disorder in which affected individuals experience repeated episodes of spontaneous vertigo, fluctuating hearing loss and tinnitus, often with a feeling of fullness in the ear. The exact aetiology remains unknown. The estimated to prevalence of Ménière's disease is 0.25% in the UK (around 162,000 individuals) (1) and is associated with significant physical, psychological, and socioeconomic morbidity (2, 3). There is a sizable lack of knowledge regarding many aspects of Ménière's disease, including a thorough understanding of epidemiological aspects of the disease, aetiological factors, pathogenesis, clinical course, and treatment outcomes. The James Lind Alliance cites Ménière's disease as the theme for four of their top ten priorities for addressing uncertainties in the field of vesribular disorders (4).

Many consider Ménière's disease to represent the final pathway of a number of individual disease processes. This is reflected by the progressive and varied development of diagnostic criteria (5). It is in the context of these observations that Ménière's disease is likely to represent a heterogeneous clinical condition defined only by small groups of common, but not always mandatory, symptoms. Understanding how Ménière's disease may exist as a cluster of clinical subtypes is key to allowing further research into its underlying pathophysiological mechanisms, and the targeting of specific treatment strategies; as well as allowing a better understanding of the physiology of the inner ear microenvironment (6). Previous work has attempted to identify clusters within Ménière's disease with this premise in mind, identifying factors such as family history, presence of migraine and comorbid autoimmune conditions as relevant (ref our psoriasis paper), with different groupings within unilateral and bilateral disease (7, 8).

Clinical subtyping is a process that has acquired increasing attention over the last decade and is necessary for individualising the practice of medicine (9). This practice is variously termed as precision medicine, personalised medicine, stratified medicine or P4 (Predictive, Preventive, Personalized and Participatory) medicine (10). Clinical subtyping can be based upon many factors, including physical symptoms, clinical course, objective test results, treatment responses, genetics, environment, and lifestyle. The discovery and refinement of disease subtypes can benefit both the practice and science of medicine (11) with a clear associated benefit for patients. Subtyping has benefited the study of a variety of conditions, including cancer, autism, autoimmune diseases, cardiovascular diseases, and Parkinson's disease (12, 13,14). Parkinson's disease provides a good, practical example (15). Like Ménière's disease, Parkinson's disease presents in a variable manner. Ongoing work to define clinical subtypes of Parkinson's has been considered important to identify

homogenous groups of strong clinical, pathological, and genetic coherence (15), leading to a better understanding of the involved biological pathways and ultimately to lead to tailored treatment strategies and prognostic information. To achieve this goal a data-driven approach has been utilised (15).

In many areas of clinical research, the implementation of customised databases and national registries has been demonstrated to be both effective and efficient (16) in answering diverse and complex questions related to the condition(s) being studied. This article outlines work that we have undertaken to identify distinct clinical subtypes of Ménière's disease by analysing data acquired from a UK registry of patients who have been diagnosed with Ménière's disease.

Methodology

In 2020, ethics approval was granted to invite individuals diagnosed with Ménière's disease to have their clinical data entered into a bespoke study database (North West - Liverpool Central Research Ethics Committee, United Kingdom - IRAS ID:275749). Further details regarding the development of the Ménière's disease registry are available elsewhere (17).

Patients with Ménière's disease were identified from eight recruitment sites; four were NHS hospitals, four were independent hospitals. Table 1 lists the hospitals from which patients were identified. Potential participants with a diagnosis of probable or definite unilateral or bilateral Ménière's disease as defined by the 2015 edition of the American Academy of Otolaryngology–Head and Neck Surgery (AAO-HNS) criteria (18) in their medical records were identified at ENT secondary care and private clinics (including services provided by audio-vestibular medicine). All potential participants had received a diagnosis of Ménière's disease within the previous 10 years or had received a new diagnosis during the recruitment window of the study. A full list of inclusion and exclusion criteria for participant recruitment is provided in table 2.

Retrospective participant data was acquired via two routes. Firstly, patients were asked to complete questionnaires related to their condition, together with a number of validated health questionnaires. Secondly, once participants consented to their clinical data being inputted into the study database from their clinical records, the local clinical teams retrieved and uploaded relevant data. Table 3 lists the data collected directly from each study participant. Table 4 lists the data collected by the study team at the participants' hospital site.

Cluster Analysis

Introduction and data

Clustering facilitates grouping participants sharing similar characteristics and risk factors into groups based on a defined measure of similarity (19). The cluster analysis was based on 77 variables of mixed types (categorical, ordinal, and continuous) representing participants' characteristics, indicators of relevant co-morbid conditions, risk factors and various symptoms of Ménière's disease (full list of variables used for clustering is given in the Supplementary Table ST1).

(Dis)similarity measure

We used Gower's distance to measure (dis)similarity between each pair of participants (20), which is a suitable measure for data containing combinations of categorical, ordinal, and quantitative

(continuous) variables. The Gower's distance is always a number between 0 and 1 with lower values indicating more similarity, and the lower and upper bounds representing identical (0) and maximally dis-similar (1) pairs of participants respectively. The calculation was performed using the R function `daisy()` from the `cluster` package. A snapshot of the similarity matrix for a random sample of 10 participants from the sample along with a corresponding visual plot is given in the in the Supplementary Figure SF1.

Clustering algorithm: Partitioning Around Medoids (PAM) clustering

We used the PAM algorithm which is a partition-based clustering approach that allocates participants into a pre-specified number of mutually exclusive clusters based on a given distance matrix (21). The objective of the PAM algorithm is to minimize the average dissimilarity of participants to their closest selected medoid (centre) by identifying a sequence of participants (i.e., medoids) that are centrally located in the clusters (22). The PAM is considered a more robust algorithm when compared to k-means clustering since actual data points (rather than some sort of average) are used as medoids. More detailed overview of PAM and related clustering algorithms can be found in Botyarov *et al.* (22). The PAM algorithm was implemented using the R-function `pam()` from the `cluster` package.

Determining optimal number of clusters

The *Silhouette Width* method (23), which is the recommended method to be used with the PAM clustering algorithm, was used to determine the optimal number of clusters. The Silhouette width measures the quality of separation (partitioning) of the participants allocated to a specified number of clusters. The value of Silhouette width ranges from -1 to 1 with positive values closer to 1 representing better separation of clusters. It is convenient to interpret Silhouette width when visually plotted in a graph (Silhouette width vs. number of clusters), and the number of clusters (k) that corresponds to the largest silhouette width indicates the optimal quantity of clusters (23).

Investigating patterns/characteristics of cluster membership

Following allocation of participants to the identified clusters based on the PAM algorithm, we used multivariable logistic regression (using stepwise forward selection of variables) to identify variables that are associated with the cluster membership.

Association of cluster membership with steroid treatment/effectiveness

The participants were considered to have received steroid treatment if they had indicated having had either oral or intratympanic steroids of any kind. The treatment was defined as effective (coded, effective =1) if at least one form of the treatment was either effective or partially effective and as not effective (coded, effective=0) if none of the treatment forms was effective or partially effective. Logistic regression analysis adjusting for age at onset of Ménière's disease and gender was used to investigate whether cluster membership was associated with participants' receiving steroid treatment(s) and their effectiveness.

Results

Recruitment to this study began in November 2020 and ended in September 2021. Over this ten-month recruitment period, 411 participants were recruited into this study. The 411 participants led to 84,255 possible pairs, and a (dis)similarity measure (Gower's distance) for each pair of participants was calculated based on the 77 variables (listed in Supplementary Table ST1). The estimated distance measures ranged from 0 to 0.86 with a mean of 0.26 and standard deviation of 0.9. A subset of the distance measures for a random sample of 10 participants from the sample along with a corresponding visual plot is given in the in the Supplementary Figure SF1.

Optimal number of clusters

The plot of the Silhouette widths against the number of clusters (Figure 1) shows that two clusters provide optimal separation of the data. The distribution of cluster membership of the 411 participants resulting from the PAM clustering algorithm run with k=2 clusters based on Gower's (dis)similarity was performed. The PAM algorithm allocated 297 (72%) participants to one cluster (C1) and the rest 114 (28%) to the other cluster (C2).

Cluster characteristics

The logistic regression analysis of cluster membership indicator (0=C1, 1=C2) on the 77 variables using forward stepwise selection was used to investigate the characteristics of the two clusters. After exclusions due to missing data, this analysis was based on n=284 participants. The results (presented in Table 5 and in Figure 2) show that participants diagnosed with Ear infections (OR=0.30, p<0.014, 95%CI:0.11 to 0.78) were more likely to be allocated in Cluster 1 (C1). Participants reporting tinnitus in both ears (OR=11.89, p<0.001, 95%CI:4.08 to 34.64), low pitched tinnitus (OR=21.09, p<0.001, 95%CI:7.47 to 59.54) and those reporting stress as a trigger for vertigo attacks (OR=14.94, p<0.001, 95%CI:4.54 to 49.10) were significantly more likely to be in cluster 2 (C2). Also, participants diagnosed with Benign Paroxysmal Positional Vertigo (OR=13.14, <0.001, 95%CI:4.35 to 39.74), Autoimmune disease (OR=5.97, p<0.007, 95%CI:1.62 to 22.03), Depression (OR=4.72, p<0.056, 95%CI:0.96 to 23.24), Migraines (OR=3.13, p<0.008, 95%CI:1.34 to 7.26), Drug allergy (OR=3.25, p<0.029, 95%CI:1.13 to 9.34) and Hay fever (OR=3.12, p< 0.009, 95%CI:1.33 to 7.34) were significantly more likely to be clustered in C2.

Association of cluster membership with steroid treatment/effectiveness

Results of the multivariable logistic regression analysis of steroid treatments (and their effectiveness) on indicator of cluster membership (0=C1, 1=C2) are reported in Table 6. The analysis was adjusted

for age at onset of Ménière's disease and gender. After exclusions due to missing data, this analysis was based on 333 and 138 participants for the steroid treatments and effectiveness respectively. The results show that participants clustered in C2 were more likely to receive steroid treatments (OR= 1.80, p=0.017, 95% CI: 1.11 to 2.92). However, steroid treatment was less likely (OR= 0.38, p=0.043, 95% CI: 0.15 to 0.97) to be effective for participants in C2 compared to those in C1. It should be noted that effectiveness data was available on a smaller number (138, which is less than half of the total sample, n=411) and therefore the association between effectiveness and cluster membership should be interpreted with caution.

Discussion

Since the beginning of the 20th century, endolymphatic hydrops has been upheld as the underlying histologic feature of Ménière's disease. However, there has been an increasingly reported incongruence between the presence of endolymphatic hydrops and a diagnosis of Ménière's disease (24, 25,26), suggesting that endolymphatic hydrops might merely represent an epiphenomenon of what we commonly define as Ménière's disease (6). Furthermore, patients diagnosed with Ménière's disease can demonstrate a wide spectrum of presentations—including hearing loss, which may frequently fluctuate, be relatively static or even regress; hearing loss that is acute, progressive, unilateral, or bilateral; tinnitus of a variety of characters; aural fullness, pressure, or pain or an absence of these features altogether; and vertigo, which often varies significantly with respect to onset, frequency, severity, and time course (6).

It is in the context of these observations that Ménière's disease is likely to represent a heterogeneous inner ear condition defined by relatively small groups of common but not always mandatory symptoms (6). Appreciating how Ménière's disease may exist as a spectrum of clinical subtypes is key to directing further research into its underlying pathophysiologic mechanisms, to the targeting of specific treatment strategies, and to permitting a better understanding of the physiology of the inner ear microenvironment (6).

In the context of previous work that considered risk factors for the development of bilateral Ménière's disease (27), the dominance of 'Ear infections' in cluster C1, and 'Autoimmune disease' in cluster C2, is worthy of reflection. Both of these variables feature as risk factors for the development of bilateral Ménière's disease, so this provides the opportunity to consider further subtyping of patient groups that progress to bilateral Ménière's disease from unilateral Ménière's disease.

Cluster C2 also yielded a number of other clinical characteristics that merit further exploration including features of the tinnitus symptoms (low pitched, bilateral), complication by benign paroxysmal positional vertigo, and relationship with psychological features including stress as a trigger for vertigo and symptoms of depression.

The findings from this study suggest that those in cluster C2 are more likely to be offered steroid therapy than those in cluster C1, but less likely to report the treatment to be effective. The data presented here do not allow us to be certain of the reason for this but there are a number of potential explanations. For example, this might be because individuals in this group show more of the

clinical features for which steroid therapy is usually administered, such as sudden drops in hearing or frequent vertigo attacks. As well as being features of Ménière's disease, sudden drops in hearing (sudden sensorineural hearing loss) or rapid deteriorations in hearing thresholds are considered also characteristic symptoms of autoimmune inner ear disease, which is considered more steroid sensitive. Although there is a lack of consensus on how to diagnose and manage autoimmune inner ear disease, many clinicians would offer intratympanic or oral steroid therapy to such patients, at least at initial presentation, to establish whether there is any evidence of a steroid responsive condition. Conversely, it may also be that those in cluster C1 are receiving fewer treatments with steroids than those in C2 because their symptoms have been attributed to ear infections that would be more appropriately managed in other ways, such as with antibiotics. Why then would this cluster also report that the steroids were ineffective? We should be cautious in placing too much weight on this observation. Firstly, the absolute numbers of patients in this subgroup for analysis is a relatively small proportion of the total. Secondly, we suspect that cluster C2 might actually be a mixed group of more than one aetiological subtype, noting that cluster C2 also contains 'migraine', a previously nominated aetiological factor that would not be expected to show strong steroid response effects. This would by a dilution effect reduce the magnitude of any measure of benefit. Thirdly, the data point "lack of efficacy for steroids" is patient reported rather than based on validated outcome measures. The complexity and variety of Ménière's disease symptoms might cause patients to report lack of efficacy because at least some of their symptoms were ongoing or recurred after treatment i.e., lack of efficacy from the patient perspective may mean that some symptoms were still ongoing, or that they were still troublesome, or that the steroids did not address the issue of primary concern for the patient. There is also a potential physiological explanation. Ear infections may cause adhesions within the middle ear, limiting the ability of intratympanic steroids to enter the inner ear. Given that Ménière's disease is a chronic condition, ongoing symptoms of some form would be the expected outcome for the majority of those with Ménière's disease, even when there was clear evidence of steroid responsiveness such as improved hearing thresholds.

This last observation reminds us that a limitation of this work is that the medical history and clinical features of the condition were entered into the database directly by patients and not by clinicians. Whilst this might affect accuracy of reporting in some respects, it is interesting to note that despite this potential for inaccuracy, effects consistent with previous literature and current clinical practice were still observed, meaning that the data is likely to be valid. Furthermore, our approach is designed to be easily scalable to facilitate future work, and to be based on "real world" patient directed approaches.

Implications for practice

There have been three high quality large randomised controlled trials investigating the use of intratympanic steroids for Ménière's disease (28). Only one of these trials produced statistically significant outcomes to support the use of intratympanic steroids (29). It has been considered that a possible explanation for the difference in outcome for these clinical trials is due to the heterogeneity of the populations studied. This study supports the principle that there are likely to be subgroups that have differential treatment responses.

Implications for future research

These findings should form the framework for future research to prospectively evaluate whether intratympanic steroid therapy is likely to be more effective if provided to those patients who are members of a 'steroid sensitive' cluster. In addition to this, further basic science work should take place to investigate the complex mechanisms at play that might form the basis of disease development for identified subtypes of Ménière's disease. As a minimum, this should include work into identifying novel autoimmune biomarkers, and to considering how middle ear and inner ear inflammatory modulators might be responsible for disease progression with a view to targeting these agents to provide more effective treatments for patients.

Conclusions

This study supports the hypothesis that Ménière's disease is a heterogeneous condition with subgroups that may be identifiable by clinical features. Two main clusters were identified with differing putative aetiological factors and were also differentiated by treatment related factors. Further work should be carried out to characterise the clusters and to evaluate outcomes of current and future therapies according to cluster or subgroup memberships.

Legends for figures and tables

Figure 1: Plot of plot of the Silhouette widths against number of clusters suggest two clusters gives optimal separation of the data.

Figure 2: Factors associated with cluster membership (C2 vs. C1) based on logistic regression (using stepwise forward selection of variables).

Table 1: Hospital recruitment sites

Table 2: Eligibility criteria

Table 3: Dataset provided by the participant

Table 4: Dataset provided by the participant's clinical team

Table 5: Factors associated with PAM cluster membership (C1 vs. C2) based on logistic regression (using stepwise forward selection of variables).

Table 6: Results of the multivariable logistic regression analysis of steroid treatments (1=Yes, 0=No) and effectiveness (1=Yes, 0=No) on indicator of cluster membership (0=C1, 1=C2).

References

1. Tyrrell J, Whinney DJ, Taylor T. The Cost of Ménière's Disease: A Novel Multisource Approach. Ear Hear. 2016;37(3):e202-9.

2. Anderson JP, Harris JP. Impact of Ménière's disease on quality of life. Otol Neurotol. 2001;22(6):888-94.

3. Tyrrell JS, Whinney DJ, Ukoumunne OC, Fleming LE, Osborne NJ. Prevalence, associated factors, and comorbid conditions for Ménière's disease. Ear Hear. 2014;35(4):e162-9.

4. James Lind Alliance – Priority Setting Partnership. http://www.jla.nihr.ac.uk/priority-setting-partnerships/ear-nose-and-throat-aspects-of-balance/top-10-priorities/ Accessed 6th September 2023.

5. Gürkov R, Pyykö I, Zou J, Kentala E. What is Menière's disease? A contemporary re-evaluation of endolymphatic hydrops. J Neurol. 2016;263:71–81

6. Phillips JS, Murdin L, Rea P, Sutton L. Clinical Subtyping of Ménière's Disease. Otolaryngol Head Neck Surg. 2018;159(3):407-409.

7. Frejo L, Martin-Sanz E, Teggi R, Trinidad G, Soto-Varela A, Santos-Perez S, Manrique R, Perez N, Aran I, Almeida-Branco MS, Batuecas-Caletrio A, Fraile J, Espinosa-Sanchez JM, Perez-Guillen V, Perez-Garrigues H, Oliva-Dominguez M, Aleman O, Benitez J, Perez P, Lopez-Escamez JA; Meniere's disease Consortium (MeDiC). Extended phenotype and clinical subgroups in unilateral Meniere disease: A cross-sectional study with cluster analysis. Clin Otolaryngol. 2017 Dec;42(6):1172-1180.

8. Frejo L, Soto-Varela A, Santos-Perez S, Aran I, Batuecas-Caletrio A, Perez-Guillen V, Perez-Garrigues H, Fraile J, Martin-Sanz E, Tapia MC, Trinidad G, García-Arumi AM, González-Aguado R, Espinosa-Sanchez JM, Marques P, Perez P, Benitez J, Lopez-Escamez JA. Clinical Subgroups in Bilateral Meniere Disease. Front Neurol. 2016 Oct 24;7:182.

9. Mirnezami R, Nicholson J, Darzi A. Preparing for Precision Medicine. New England J. Medicine 2012;366(6):489–491

10. Hood L, Friend SH. Predictive, Personalized, Preventive, Participatory (P4) Cancer Medicine. Nature Rev. Clinical Oncology 2011;8(3):184–187

11. Saria S, Goldenberg A. Subtyping: What it is and its role in precision medicine. Intelligent Systems, IEEE. 2015 Jul;30(4):70-5.

12. Schulam P, Wigley F, Saria S. Clustering Longitudinal Clinical Marker Trajectories from Electronic Health Data: Applications to Phenotyping and Endotype Discovery. Am. Assoc. for Artificial Intelligence 2015 http://pschulam.com/papers /schulam+wigley+saria_aaai_2015.pdf.

13. Verhaak RG, et al. Integrated Genomic Analysis Identities Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 2010;17(1):98–110

14. Lewis SJ, et al. Heterogeneity of Parkinson's Disease in the Early Clinical Stages using a Data Driven Approach. J. Neurology, Neurosurgery, and Psychiatry 2005;76(3):343–348

15. van Rooden SM, Colas F, Martínez-Martín P, Visser M, Verbaan D, Marinus J, Chaudhuri RK, Kok JN, van Hilten JJ. Clinical subtypes of Parkinson's disease. Mov Disord. 2011;26(1):51-8

16. Mandavia R, Knight A, Phillips J, Mossialos E, Littlejohns P, Schilder A. What are the essential features of a successful surgical registry? a systematic review. BMJ Open. 2017;7(9):e017373

17. Phillips J, Murdin L, Rea P, Harcourt J, Shepstone L, Sims E, Bion V, Brunton R, Tetteh A, Duaz S, Daskalakis D, Rea W. Development of a national Ménière's disease registry: a feasibility study. Journal of International Advanced Otology.

18. Goebel JA. 2015 Equilibrium Committee Amendment to the 1995 AAO-HNS Guidelines for the Definition of Ménière's Disease. Otolaryngol Head Neck Surg. 2016;154(3):403-4.

19. Guha S, Meyerson A, Mishra N, Motwani R, O'Callaghan L. Clustering data streams: Theory and practice. IEEE transactions on knowledge and data engineering. 2003;15(3):515-528.

20. Gower JC. A general coefficient of similarity and some of its properties. Biometrics, 1971;857-871.

21. Van der Laan M, Pollard K, Bryan J. A new partitioning around medoids algorithm. Journal of Statistical Computation and Simulation, 2003;73(8):575-584.

22. Botyarov M, Miller EE. Partitioning around medoids as a systematic approach to generative design solution space reduction. Results in Engineering. 2022;15:100544.

23. Batool F, Hennig, C. Clustering with the average silhouette width. Computational Statistics & Data Analysis. 2021;158:107190.

24. Merchant SN, Adams JC, Nadol JB Jr. Pathophysiology of Ménière's syndrome: are symptoms caused by endolymphatic hydrops? Otol Neurotol. 2005;26:74-81.

25. Hallpike CS, Cairns H. Observation on the pathology of Ménière's syndrome. J Laryngol Otol. 1938;53:625-655.

26. Rauch SD, Merchant SN, Thedinger BA. Ménière's syndrome and endolymphatic hydrops: double-blind temporal bone study. Ann Otol Rhinol Laryngol. 1989;98:873-883.

27. Phillips J, Murdin L, Grant K, Shepstone L, Sims E, Rea P, Harcourt J. Risk Factors for the Development of Bilateral Ménière's Disease. Otol Neurotol. In Press

28. Phillips J, Mikulec AA, Robinson JM, Skarinsky D, Anderson JJ. Efficacy of Intratympanic OTO-104 for the Treatment of Ménière's Disease: The Outcome of Three Randomized, Double-Blind, Placebo-Controlled Studies. Otol Neurotol. 2023;44(6):584-592.

29. Patel M, Agarwal K, Arshad Q, Hariri M, Rea P, Seemungal BM, Golding JF, Harcourt JP, Bronstein AM. Intratympanic methylprednisolone versus gentamicin in patients with unilateral Ménière's disease: a randomised, double-blind, comparative effectiveness trial. Lancet. 2016;388(10061):2753-2762.

Table 1: Hospital recruitment sites

NHS centres
Norfolk and Norwich University Hospitals NHS Foundation Trust, Norwich
Leicester Royal Infirmary, University Hospitals of Leicester, Leicester
Charing Cross Hospital, London
Guy's Hospital and St Thomas' Hospital, London
Private centres
Spire Norwich Hospital, Norwich
The London Road Clinic, Leicester
London Hearing and Balance Centre, London
The London Clinic, London

Table 2: Eligibility criteria

Inclusion criteria
- Individuals aged 18 years or over
- Definite or probable diagnosis of unilateral or bilateral Ménière's disease as defined by the 2015 edition of the American Academy of Otolaryngology–Head and Neck Surgery (Goebel 2016)
- Potential participants must have received a diagnosis of Ménière's Disease within the previous 10 years or have received a new diagnosis during the recruitment window of the study.
- Willingness to provide consent for data from health records to be used for research purposes.
Exclusion criteria
- Unable to provide consent
- Unable/unwilling to complete questionnaires.

Table 3: Dataset provided by the participant

Self-reported data themes
Demographics (including ethnicity and occupation)
CoVID diagnosis and implications for symptoms
Circumstances of diagnosis (onset, duration, and laterality)
Nature of vertigo and triggers
Tinnitus experience
Past medical history
Drug history
Family history
Treatment of Ménière's disease (including an indication of efficacy)
Validated questionnaires
Disability Rating Scale (DRS). (Shepard 1990)
Generalized Anxiety Disorder Assessment (GAD-7) (Spitzer 1999)
Patient Health Questionnaire (PHQ-9) (Spitzer 1999)
Tinnitus Handicap Index (THI) (Newman 1996)
Dizziness Handicap Inventory (DHI) (Jacobson 1990)

Migraine Disability Assessment Test (MIDAS)
The Social Life and Work Impact of Dizziness (SWID) (Bronstein 2010)
Situational Vertigo Questionnaire (SVQ) (Jacob 1989)
Self-Administered Comorbidity Questionnaire (SCQ) (Sangha 2003)

Table 4: Dataset provided by the participant's clinical team

Audiometric Data
Pure tone audiometry (Air conduction thresholds for at the following frequencies: 250Hz, 500Hz, 1000Hz, 2000Hz, 3000Hz, 4000Hz, 6000Hz and 8000Hz)
Tympanometry
Vestibular Testing Data
Caloric testing
vHIT (video Head Thrust Test)
cVEMP (cervical Vestibular Evoked Myogenic Potentials)
Posturography
Radiology
Radiological examinations of the Internal Auditory Meati (IAMs) using magnetic resonance imaging (MRI) or computerised tomography (CT) scans. Outcome of specialist specialist scans performed for the purpose of identifying hydrops.

Table 5: *The cluster membership distribution of the two clusters based on PAM algorithm and Gower's distances.*

Clusters	*Number*	*%*
C1	297	72
C2	114	28
Total	**411**	**100**

Table 6: Odds Ratios, standard errors (SE), p-values and 95% confidence Intervals (CI) of factors associated with PAM cluster membership (C1 vs. C2) based on logistic regression (using stepwise forward selection of variables). After exclusions due to missing data, this analysis is based on n=284 participants.

Variables	Odds Ratio (SE)	p-value	95% CI
Tinnitus location (reference: Left ear)			
Right ear	2.31 (1.24)	0.117	(0.81, 6.61)
Both ears	11.89 (6.49)	<0.001	(4.08, 34.64)
Tinnitus type 2 (low pitched?) (reference: No)			
Yes	21.09 (11.17)	<0.001	(7.47, 59.54)
Stress triggers a vertigo attack? (reference: No)			
Yes	14.94 (9.07)	<0.001	(4.54, 49.10)
Diagnosed with Drug allergy? (reference: No)			
Yes	3.25 (1.75)	0.029	(1.13, 9.34)
Diagnosed with Hayfever? (reference: No)			
Yes	3.12 (1.36)	0.009	1.33, 7.34)
Diagnosed with Autoimmune disease? (reference: No)			
Yes	5.97 (3.98)	0.007	(1.62, 22.03)
Diagnosed with Depression? (reference: No)			
Yes	4.72 (3.84)	0.056	(0.96, 23.24)
Diagnosed with Benign Paroxymal Positional Vertigo (BPPV)? (reference: No)			
Yes	13.14 (7.42)	<0.001	(4.35, 39.74)
Diagnosed with Ear infections? (reference: No)			
Yes	0.30 (0.15)	0.014	(0.11, 0.78)
Diagnosed with Migraines? (reference: No)			
Yes	3.13 (1.34)	0.008	(1.34, 7.26)
Do you have Depression? (reference: No)			
Yes	6.48 (5.49)	0.027	(1.23, 34.05)

Table 7: Results of the multivariable logistic regression analysis of steroid treatments (1=Yes, 0=No) and effectiveness (1=Yes, 0=No) on indicator of cluster membership (0=C1, 1=C2): Odds ratios (OR), p-value and 95% confidence intervals.

Outcomes	Predictor variables	Odds Ratio (SE)	p-value	95% CI
Steroid treatments	Cluster (ref: C1)			
	C2	1.80 (0.44)	0.017	(1.11, 2.92)
	Age at onset	1.01 (0.01)	0.490	(0.99, 1.02)
	Gender (male)	0.71 (0.16)	0.140	(0.45, 1.12)
Effectiveness	Cluster (ref: C1)			
	C2	0.38 (0.18)	0.043	(0.15, 0.97)
	Age at onset	1.00 (0.02)	0.951	(0.97, 1.04)
	Gender (male)	1.75 (0.84)	0.247	(0.68, 4.48)

Figures:

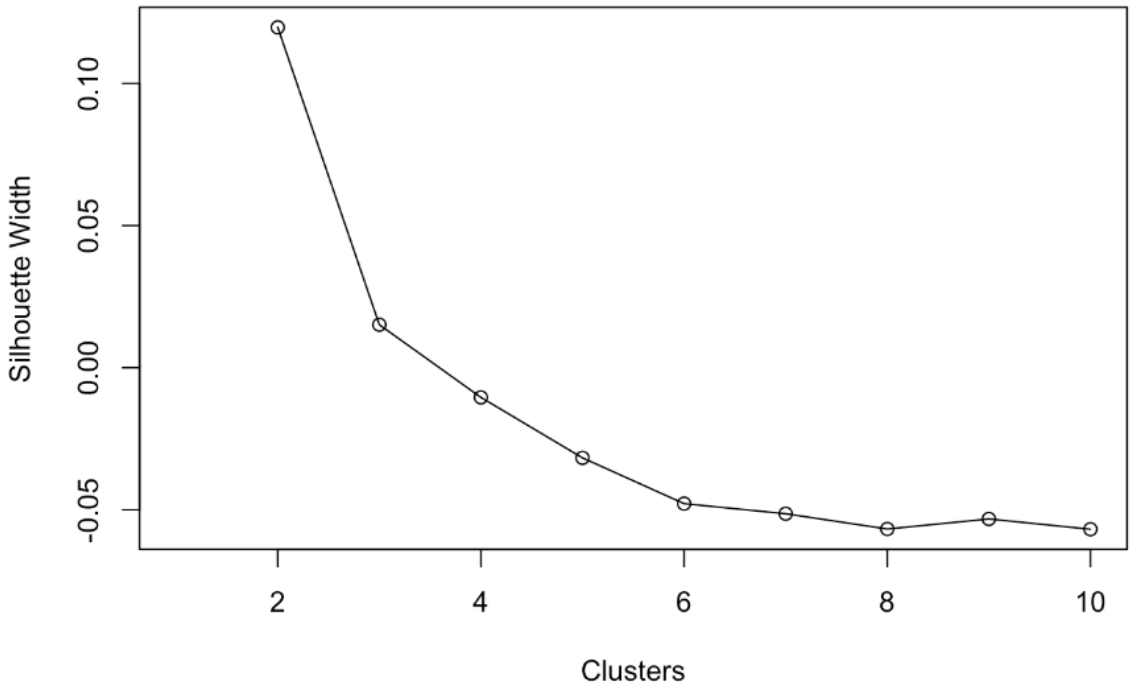

Figure 1: *Plot of plot of the Silhouette widths against number of clusters suggest two clusters gives optimal separation of the data.*

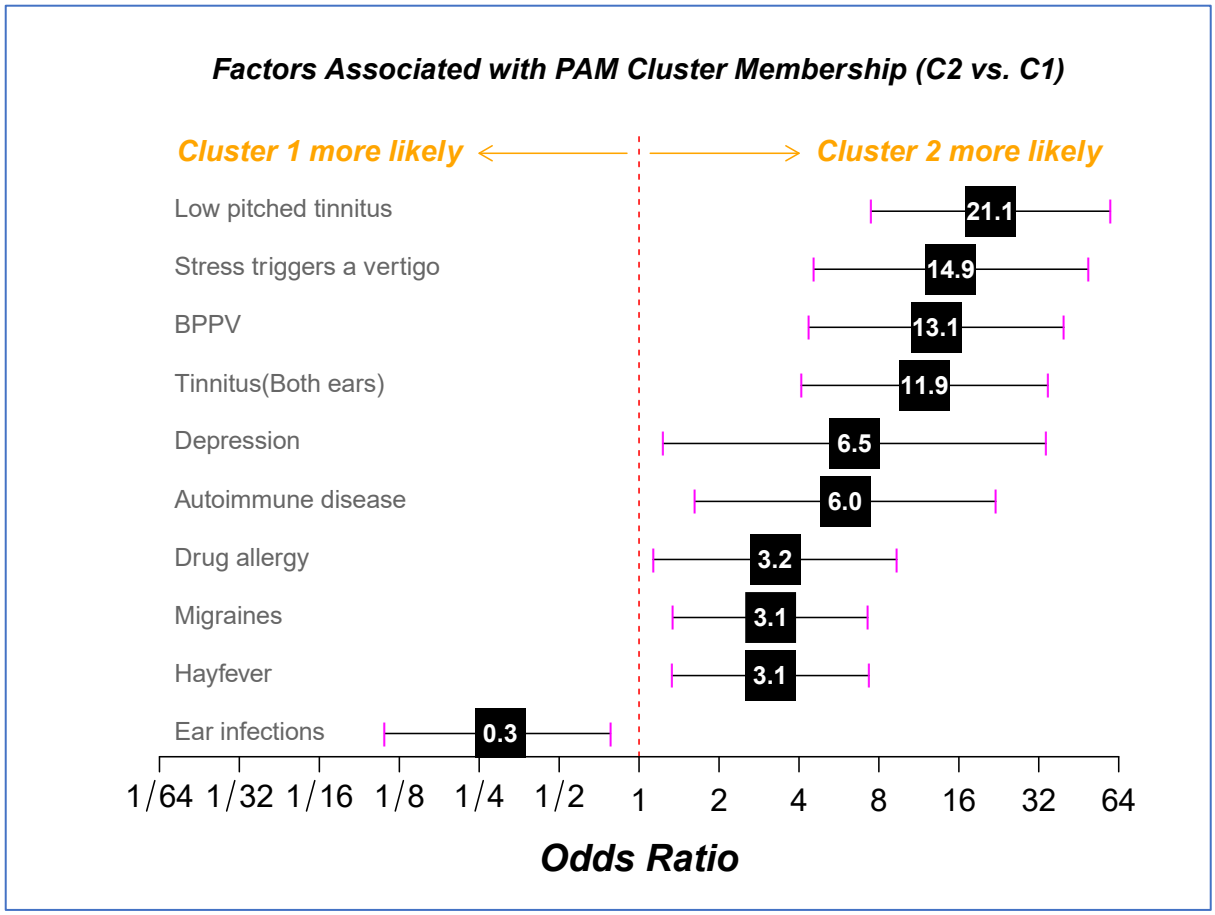

Figure 2: Factors associated with cluster membership (C2 vs. C1) based on logistic regression (using stepwise forward selection of variables).