The Differential Diagnosis of Ménière's Disease and Vestibular Migraine

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Audiology

Department of Communication Disorders,

University of Canterbury

By

Jonathan Robert Osborne

2017
# Table of Contents

Acknowledgements ........................................................................................................ 6  
Abstract ........................................................................................................................ 7  
Abbreviations ............................................................................................................... 8  
Glossary ....................................................................................................................... 10  
1.0 Introduction ............................................................................................................ 11  
  1.1 The Human Vestibular System ........................................................................... 12  
  1.1.1 Otolith Organs ....................................................................................... 14  
  1.1.2 Semicircular Canals (SCCs) ................................................................. 16  
  1.1.3 Central Vestibular System ................................................................. 18  
  1.2 Ménière's Disease and Vestibular Migraine .................................................... 20  
  1.2.1 Ménière's Disease ........................................................................... 20  
  1.2.2 Pathophysiology of Ménière's Disease ........................................... 25  
  1.2.3 Vestibular Migraine ........................................................................ 27  
  1.2.4 Pathophysiology of Migraine ........................................................... 34  
  1.2.5 Pathophysiology of Vestibular Migraine .......................................... 35  
  1.2.6 Differentiating Ménière's Disease and Vestibular Migraine ......... 37  
  1.2.7 MD and VM – A common pathophysiology? ............................ 39  
  1.3 Diagnostic Accuracy ...................................................................................... 41
1.3.1 Importance of Diagnostic Accuracy ........................................... 42

1.4 Systematic Reviews ................................................................................................................. 43

1.4.1 Systematic Reviews of Diagnostic Accuracy ................................................................. 44

1.4.2 Scoping Reviews .............................................................................................................. 45

1.4.3 Search Filters for Systematic Reviews ........................................................................ 45

1.5 Rationale and Objectives ........................................................................................................ 47

2.0 Methods ......................................................................................................................................... 48

2.1 Protocol and Registration ..................................................................................................... 49

2.2 Eligibility Criteria ................................................................................................................ 50

2.3 Information Sources ............................................................................................................ 50

2.4 Search Strategy .................................................................................................................. 51

2.5 Study Selection and Data Collection Process ..................................................................... 54

2.6 Data Items .......................................................................................................................... 54

2.7 Synthesis of Results ............................................................................................................ 54

2.8 Software .................................................................................................................................... 56

3.0 Results of the Systematic Review – Differential Diagnosis of MD and VM... 57

3.1 Included Studies .................................................................................................................... 58

3.2 Symptoms ............................................................................................................................. 65

3.3 Caloric Testing ..................................................................................................................... 66
3.5 Rotatory Chair Testing ................................................................. 71
3.6 Video Head Impulse Test (vHIT) .................................................. 71
3.7 Head Shaking Nystagmus (HSN) ................................................. 72
3.8 Vibration Induced Nystagmus (VIN) ............................................. 73
3.9 Cervical Vestibular Evoked Myogenic Potentials (cVEMPs) ........... 74
3.10 Ocular Vestibular Evoked Myogenic Potentials (oVEMPs) ............ 75
3.11 Posturography ............................................................................ 76
3.12 Linear Motion Perceptual Thresholds (LMPT) ............................. 77
3.13 Extratympanic click Electrocochleography ................................. 77
3.14 Multivariate Methods .................................................................. 78
3.15 Summary of Evidence ................................................................ 79

4.0 Results of the Scoping Review – Accurate tests for MD or VM ....... 81
4.1 Included Studies ........................................................................... 82
4.2 Caloric Testing ............................................................................. 84
4.10 Vibration Induced Nystagmus (VIN) .......................................... 85
4.3 Gadolinium MRI ........................................................................ 86
4.4 Three-Dimensional Cone Beam Computed Tomography .............. 88
4.5 Electrocochleography (ECoG) ...................................................... 88
4.5.1 Transtympanic click ECoG ....................................................... 89
Acknowledgements

I would like to sincerely thank my supervisors Assoc Prof Greg O’Beirne and Mr Jeremy Hornibrook for their support and guidance throughout the year.

Thank you Dr Rebecca Kelly-Campbell for your advice regarding systematic reviews.

Thank you Mum and Dad for being there for me through the whole thing.

Lastly, thank you Annie for being my rock.
Abstract

The differential diagnosis of Ménière's disease (MD) and vestibular migraine (VM) is difficult, due to a reliance on symptom-based diagnosis despite frequently overlapping symptoms. A systematic review was conducted of studies investigating diagnostic features in patients with MD and VM that may assist differential diagnosis. A scoping review was also conducted of tests with high sensitivity and specificity for MD or VM. The systematic review identified that caloric testing was 54.6% sensitive and 78.9% specific for separating MD from VM (with MD as a positive result). Other potentially useful tests were identified, but more studies are needed. The scoping review identified several tests that not have yet been evaluated for the differential diagnosis of MD and VM (primarily gadolinium magnetic resonance imaging and tone-burst electrocochleography). Several other promising tests were identified, but have not been sufficiently tested with appropriate control groups.

Keywords: Ménière's disease, vestibular migraine, diagnostic accuracy, systematic review, differential diagnosis
Abbreviations

5-HT – 5-hydroxytryptamine (Serotonin)

ABR – auditory brainstem response

AP – action potential

CHAMP – cochlear hydrops analysis masking procedure

CI – confidence interval

cVEMP – cervical vestibular evoked myogenic potential

dB HL – decibels hearing level

ECoG - electrocochleography

EVestG - electrovestibulography

FAR – frequency amplitude ratio

fMRI – functional magnetic resonance imaging

GRADE - grading of recommendations assessment, development, and evaluation

ICHD-3 - International classification of headache disorders 3rd edition

HSN – head shaking nystagmus

LMPT – linear motion perceptual threshold

MD - Ménière's disease

MRI – magnetic resonance imaging
NA – noradrenaline / norepinephrine

oVEMP – ocular vestibular evoked myogenic potential

PIVC - parieto-insular vestibular cortex

PRISMA - preferred reporting items for systematic reviews and meta-analyses

PTA – pure-tone audiometry

QUADAS - quality assessment of diagnostic accuracy studies

SCC – semi-circular canal

SHA – sinusoidal harmonic acceleration

SP – summating potential

STARD - standards for reporting of diagnostic accuracy

SVT – step velocity test

TB – tone burst

VCR – vestibulo-collic reflex

vHIT – video head impulse test

VIN – vibration-induced nystagmus

VM – vestibular migraine

VOR – vestibulo-ocular reflex

VSR – vestibulo-spinal reflex
Dizziness – Sensation of disturbed or impaired spatial orientation without false sense of distorted motion. Distinct from vertigo (Bisdorff, Von Brevern, Lempert, Newman-Toker, & others, 2009).

Ictal – A physiological state or event.

Interictal – The period between physiological events.

Nystagmus – Non-voluntary rhythmic oscillation of the eyes. Nystagmus usually has a fast and slow component. The direction of nystagmus is defined by the fast component (Balog, Honrubia, & Kerber, 2010)

Oscillopsia – False sensation that the visual surround is oscillating (Bisdorff et al., 2009).

Saccade – A rapid movement of the eye between fixation points.

Unsteadiness – The feeling of being unstable while seated, standing, or walking (Bisdorff et al., 2009).

Vertigo – The sensation of self-motion when no self-motion is occurring or the distorted self-motion during otherwise normal head movement (Bisdorff et al., 2009).
1.0 Introduction
The differential diagnosis of Ménière's disease (MD) and vestibular migraine (VM) is difficult, due to a reliance on symptom-based diagnosis despite frequently overlapping symptoms. The intent of this thesis is to conduct a systematic review of the differential diagnosis of MD and VM and a scoping review of accurate diagnostic tests for MD or VM. The first section of the introduction will outline the vestibular system. The next section will summarise MD and VM and the current issues surrounding their differential diagnosis. The third section will briefly summarise the current theory and methodology behind diagnostic accuracy. The fourth section will briefly overview systematic review methodology. The final section of the introduction will summarise the aims of the present study.

1.1 The Human Vestibular System

The peripheral vestibular system is contained within the bony labyrinth in the inner ear (Figure 1). The bony labyrinth is filled with perilymph (having a similar composition to cerebrospinal fluid); (Wangemann & Schacht, 1996). The bony labyrinth includes three semicircular canals (SCCs), the vestibule and the cochlea. Inside the bony labyrinth is the membranous labyrinth. The membranous labyrinth is filled with endolymph (having a similar composition to intracellular fluid); (Wangemann & Schacht, 1996; Hain & Helminski, 2014). Primary afferent neurons from the utricle, anterior saccule, horizontal and superior SCCs project along the superior vestibular nerve, while primary afferent neurons from the remaining saccule and, the horizontal SCC project along the inferior vestibular nerve (McRackan & Brackmann, 2015).
Figure 1. The membranous labyrinth is located within the cavities of the bony labyrinth. The cochlear duct is shown in deep blue, and the remainder of the membranous labyrinth is shown in a light green color. (Reproduced with permission from Siegel, Sapru, & Siegel, 2014).

The superior and inferior vestibular nerves then enter the brainstem at the junction of the pons and medulla, before projecting to the vestibular nuclei (in the brainstem) and the cerebellum. The vestibular nuclei then project to secondary vestibular afferent neurons (McRackan & Brackmann, 2015). The central vestibular system coordinates with the visual and proprioceptive senses to provide stability to the eyes through the vestibulo-ocular reflex (VOR), stability to the head and neck via the vestibulo-collic reflex (VCR), and body posture via the vestibulo-spinal reflex (VSR); (Mudduwa, Kara, Whelan, & Banerjee, 2010).
1.1.1 Otolith Organs

The otolith organs are located within the vestibule. There are two otolith organs, the utricle and the saccule. The utricle is most sensitive to horizontal linear acceleration, whereas the saccule is most sensitive to sagittal linear acceleration. The otoliths contain a sensory epithelium called the macula. The striola is a region that runs along the centre of the macula, and is surrounded by hair cells. Hair cells in the saccule orient away from the striola, whereas hair cells in the utricle orient towards the striola (Figure 2).

There are two types of vestibular hair cells. Type I hair cells typically have bouton endings and are primarily located away from the striola. Type II hair cells have calyx endings and are mainly located close to the striola. Hair cells are embedded in the otolithic membrane, a gelatinous structure containing calcium carbonate crystals (otoconia). Due to inertia, head movement will result in an opposite movement of the otolithic membrane, and therefore hair cells. This causes a change in the firing rate of hair cells in the otoliths (McRackan & Brackmann, 2015).
Figure 2. The otolith organs. a.) Anatomy of the macule b.) Mechanism of hair cell activation c.) Orientation of the utricle and saccule. (Reproduced with permission from Baloh et al., 2010).

Hair cells in the otoliths project to primary vestibular afferent neurons. Vestibular primary afferent neurons are categorised by their interspike interval. Regular afferent neurons primarily connect to extrastriolar type II hair cells, whereas irregular afferent neurons primarily connect to striolar type I hair cells (McRackan & Brackmann, 2015). Regular afferent neurons respond tonically to macula displacement/head position, whereas irregular
afferent neurons respond phasic-tonically to velocity of the macula displacement (Leigh & Zee, 2015). Because of this regular afferents give information regarding head position, while irregular afferents detect sudden movements of the head.

1.1.2 Semicircular Canals (SCCs)

Each labyrinth contains three SCCs, resembling rings, which terminate in the utricle. The horizontal SCCs are positioned 30° from the horizontal plane, while the anterior and posterior SCCs are positioned 45° from the sagittal plane (Figure 3). Each SCC has a dilated compartment called the ampula. The ampula contains the crista ampularis, an assortment of hair cells. The stereocilia and kinocilia of the hair cells are embedded in a gelatinous fluid-filled compartment called the cupula (McRackan & Brackmann, 2015).
Angular acceleration results in movement of the endolymph in the SCCs, which in turn moves the cupula and therefore the hair cells. Ampullopetal movement of endolymph (flow from the utricle to the ampule) is excitatory (in the direction of the kinocilium) in the horizontal SCC and inhibitory (away from the kinocilium) in the anterior and posterior SCCs.
The opposite is true of ampullofugal endolymph movement (flow from the ampule towards the utricle); (McRackan & Brackmann, 2015).

1.1.3 Central Vestibular System

Figure 4 shows a simplified diagram of central vestibular connections. The vestibular nucleus is located in the medulla of the brainstem. The vestibular nucleus receives inputs from the visual, somatosensory, autonomic nervous systems, cerebellum, and both the contralateral and ipsilateral peripheral vestibular systems (McRackan & Brackmann, 2015). From the vestibular nucleus, there are believed to be multiple ascending tracts. Several tracts have been implicated based on animal studies. These include the medial longitudinal fasciculi, the crossed and uncrossed ascending tract of Dieters, the crossed ventral tegmental tract, and the brachium conjunctivum. Projections are sent both ipsilaterally and contralaterally to the midbrain tegmentum, thalamus, and cortex (Dieterich & Brandt, 2015).
The thalamus acts to process vestibular information and as a relay to and between the various cortical regions responsible for vestibular processing. The primary cortical region for vestibular processing is believed to be the parieto-insular vestibular cortex (PIVC). While the exact location of the PIVC is not known in humans, functional magnetic resonance imaging (fMRI) studies have indicated that it may be located at the posterior insula and temporo-parietal junction (Lopez & Blanke, 2011). Neuroimaging studies have implicated several
other cortical regions that are involved in vestibular processing in humans. These regions include the superior temporal gyrus, inferior parietal lobule, somatosensory cortex, precuneus, cingulate gyrus, motor cortex, frontal eye field, and the hippocampus (Lopez & Blanke, 2011). Thalamic and cortical regions mediate the perception of vestibular inputs, while the brainstem mediates sensorimotor reflexes (VOR, VCR, and VSR); (Marianne Dieterich & Brandt, 2015).

1.2 Ménière's disease and Vestibular Migraine

This section will summarise the current understanding of symptomology, pathophysiology, and differential diagnosis of MD and VM.

1.2.1 Ménière's Disease

Ménière's Disease (MD) is characterised by repeated vestibular episodes, fluctuating hearing loss, aural fullness and tinnitus (Lopez-Escamez, Carey, Chung, Goebel, Magnusson, Mandalà, Newman-Toker, Strupp, Suzuki, Trabalzini, & others, 2015). The diagnosis of MD primarily relies on case history. Pure-tone audiometry (PTA) is also used to confirm hearing loss in MD (Gode et al., 2012). In most cases symptoms present unilaterally. However studies indicate that anywhere from 2 to 47% of patients with MD experience bilateral symptoms (Huppert, Strupp, & Brandt, 2010). The reported prevalence of MD varies from 3.5 to 513 per 100,000 (Wladislavosky-Waserman, Facer, Mokri, & Kurland, 1984; Alexander &
Harris, 2010) and accounts for approximately 8.3 to 10.1% of patients in dizziness clinics (Brandt & Strupp, 2006; Bunasuwan, Bunbanjerdsuk, & Nilsuwan, 2011).

Table 1. Current Clinical Definition of Definite MD (Lopez-Escamez, Carey, Chung, Goebel, Magnusson, Mandalà, Newman-Toker, Strupp, Suzuki, Trabalzini, & others, 2015).

A. Two or more spontaneous episodes of vertigo, each lasting 20 minutes to 12 hours.
B. 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.
C. Fluctuating aural symptoms (hearing, tinnitus or fullness) in the affected ear.
D. Not better accounted for by another vestibular diagnosis.

Vestibular symptoms in MD are defined as disabling vertigo which is spontaneous or occurs during normal head movements. While not important for diagnosis, dizziness and unsteadiness are also frequently reported in patients with MD (Lopez-Escamez, Carey, Chung, Goebel, Magnusson, Mandalà, Newman-Toker, Strupp, Suzuki, Trabalzini, & others, 2015). The course of hearing loss is also variable in MD. Most patients with MD show a progressive, sometimes fluctuating, hearing loss in the first 10 years of the condition. Hearing loss typically stabilises at an average level of 50-60 decibels hearing level (dB HL). The configuration of the audiogram is most commonly a flat configuration, but can also be can be rising, sloping, or peaked (Huppert et al., 2010).

Tinnitus in MD is most commonly unilateral (77%; Havia, Kentala, & Pyykkö, 2002). The tinnitus is commonly low in pitch, but its prevalence is reasonably evenly spread across
low, mid and high pitches (Herraiz, Tapia, & Plaza, 2006; Zagólski & Stręk, 2014; Zhang, Liu, Wang, Jia, & Gu, 2016), with an average perceived pitch at 3700 Hz (Zagólski & Stręk, 2014). The perception of the tinnitus in MD is commonly described by patients as roaring, buzzing, ringing, or whistling (Vernon, Johnson, & Schleuning, 1980; Herraiz et al., 2006).

MD is a variable condition, particularly in its early stages. Only 38-40% of patients present with all symptoms necessary for a diagnosis, and symptoms can occur in any order (Havia et al., 2002; Belinchon, Perez-garrigues, & Tenias, 2012; Pyykkö, Nakashima, Yoshida, Zou, & Naganawa, 2013). It may take several years for all symptoms to precipitate. Because of this, diagnosis can be difficult in the early stages of MD. It frequently takes years to reach an official diagnosis of MD (Belinchon et al., 2012; Pyykkö et al., 2013; Hietikko, Sorri, Männikkö, & Kotimäki, 2014).

Results from traditional bedside and oculomotor vestibular tests are also highly variable in MD (Table 2.). During the acute phase, 100% of patients present with spontaneous nystagmus. Spontaneous nystagmus occurs in the horizontal plane, towards or away from the affected ear (Meissner, 1981; McClure, Copp, & Lycett, 1981; Proctor, 2000; Maire & van Melle, 2008; Marques & Perez-Fernandez, 2012; Hirai et al., 2017). The direction of nystagmus can also reverse (Meissner, 1981; Bance, Mai, Tomlinson, & Rutka, 1991). Torsional nystagmus has also been reported in MD (Bance et al., 1991).

However patients with MD are typically seen in the interictal period. This is largely due to the disabling nature of the attacks:- patients cannot typically visit a clinic until the attack is over (Hirai et al., 2017). One study found that 80% of patients with MD have at least one abnormal finding in the interictal period (Shin, Kim, & Park, 2013). While abnormalities are common in these tests, no individual measure is highly sensitive, limiting their usefulness in the diagnosis of MD.
One finding, that has been repeatedly noted, is that caloric hypofunction is common in MD and increases in the early stages of MD (Huppert et al., 2010). This finding is rather unusual because the caloric test has historically been believed to test the horizontal SCC, and patients with MD tend to have a normal video head impulse test (vHIT) measures for the horizontal SCC (Rambold, 2014; Zulueta-Santos, Lujan, Manrique-Huarte, & Perez-Fernandez, 2014; McGarvie, Curthoys, MacDougall, & Halmagyi, 2015). It has been suggested that abnormal caloric results in MD may be due to endolymphatic hydrops of the horizontal semicircular canal (McGarvie et al., 2015).

Table 2. Results of bedside and oculomotor vestibular tests in MD during the interictal period. * = results of tests during ictal period.

<table>
<thead>
<tr>
<th>Test</th>
<th>% Abnormal (SD; Range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous Nystagmus</td>
<td>23.0 (23.6; 0-47.2)</td>
<td>(Mateijsen et al., 2001; Chen &amp; Young, 2006; Marques &amp; Perez-Fernandez, 2012; Faralli, Lapenna, Mandalà, Trabalzini, &amp; Ricci, 2014*; Zulueta-Santos et al., 2014; Maire &amp; van Melle, 2008*; Kumagami, Sainoo, Fujiyama, &amp; Baba, 2009*; Faralli et al., 2014)*</td>
</tr>
<tr>
<td>Spontaneous + Gaze + Static Positional Nystagmus</td>
<td>47</td>
<td>(Marques &amp; Perez-Fernandez, 2012)</td>
</tr>
<tr>
<td>Vibration induced Nystagmus</td>
<td>64.1 (9.9; 53-75)</td>
<td>(Ohki, Matsuzaki, Sugasawa, &amp; Murofushi, 2002; Neff et al., 2012; Marques &amp; Perez-Fernandez, 2012; Shin et al., 2013; Xie et al., 2013)</td>
</tr>
<tr>
<td>Headshake Nystagmus</td>
<td>62.1 (20.6; 33.3-80)</td>
<td>(Neff et al., 2012; Shin et al., 2013; Faralli et al., 2014; Marques &amp; Perez-Fernandez, 2012)</td>
</tr>
<tr>
<td>Random saccade test</td>
<td>2.5 (3.5; 0-5)</td>
<td>(Somefun, Giwa, Bangboye, Okeke-Igbokwe, &amp; Azeez, 2010; Neff et al., 2012)</td>
</tr>
<tr>
<td>Saccadic pursuit</td>
<td>5</td>
<td>(Neff et al., 2012)</td>
</tr>
</tbody>
</table>
Table 2 continued. Results of bedside and oculomotor vestibular tests in MD during the interictal period. * = results of tests during ictal period.

<table>
<thead>
<tr>
<th>Test</th>
<th>% Abnormal (SD; Range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caloric Hypofunction</td>
<td>60.0 (19.2; 33.3-100)</td>
<td>(Kingma, Meulenbroeks, &amp; De Jong, 2000); Mateijsen et al., 2001; Park, Migliaccio, Della Santina, Minor, &amp; Carey, 2005; Chen &amp; Young, 2006; Kumagami et al., 2009; Somefun et al., 2010; Gode et al., 2012 Shin et al., 2013; Blödow et al., 2014; Faralli et al., 2014; Sharon &amp; Hullar, 2014; Marques &amp; Perez-Fernandez, 2012; Satar, Karahatay, Sen, Cekin, &amp; Birkent, 2008; McGarvie et al., 2015; Palomar-Asenjo, Boleas-Aguirre, Sánchez-Ferrándiz, &amp; Perez Fernandez, 2006; Neff et al., 2012; Yetiser, Kertmen, &amp; Yildirim, 2004)</td>
</tr>
<tr>
<td>Caloric Directional Prepoderance</td>
<td>29.6 (4.1; 25.8-34)</td>
<td>(Mateijsen et al., 2001; Palomar-Asenjo et al., 2006; Neff et al., 2012)</td>
</tr>
<tr>
<td>vHIT</td>
<td>37.2 (29.4; 8-66.7)</td>
<td>(Blödow et al., 2014; Rambold, 2014; Zulueta-Santos et al., 2014)</td>
</tr>
<tr>
<td>Rotatory Chair (Sinusoidal harmonic acceleration)</td>
<td>39.3 (28.5; 22.7-72.2)</td>
<td>(Kingma et al., 2000; Palomar-Asenjo et al., 2006; Park, Chen, &amp; Westhofen, 2009)</td>
</tr>
<tr>
<td>Rotatory Chair (Step velocity test)</td>
<td>32.8</td>
<td>(Mateijsen et al., 2001)</td>
</tr>
<tr>
<td>Subjective vertical</td>
<td>19.1 (6.7; 14.3-23.8)</td>
<td>(Kumagami et al., 2009*; Faralli et al., 2014)*</td>
</tr>
<tr>
<td></td>
<td>69.9* (8.8; 63.6-76.1)*</td>
<td></td>
</tr>
</tbody>
</table>
1.2.2 Pathophysiology of Ménière's Disease

Hallpike and Cairns (1938) noted a dilation of the endolymphatic space primarily in the cochlea and saccule in the temporal bones of two patients with MD. They suggested that this dilation may be due to an excess of endolymph. This dilation of the endolymphatic space is now referred to as endolymphatic hydrops, and is associated with MD (Lopez-Escamez, Carey, Chung, Goebel, Magnusson, Mandalà, Newman-Toker, Strupp, Suzuki, Trabalzini, & others, 2015). Endolymphatic hydrops is theorised to cause the symptoms seen in MD due to mechanical deformation of auditory and vestibular structures (Balogh et al., 2010).

Post mortem temporal bone studies are the most accurate way to assess endolymphatic hydrops. A recent meta-analysis of 53 temporal bone studies found that the 1995 criteria for MD had a sensitivity of 31.8% and a specificity of 100% for endolymphatic hydrops (Foster & Breeze, 2013). This indicates that endolymphatic hydrops always occurs in patients that meet the 1995 criteria for MD. However the criteria miss many patients with endolymphatic hydrops, suggesting that the criteria are not sensitive enough for endolymphatic hydrops. One limitation of the study was that they did not clearly define what level of the criteria they considered to be MD.

Foster and Breeze (2013), based on personal experience, estimated that 6% of people have asymptomatic endolymphatic hydrops. In another temporal bone study, 26% of controls had endolymphatic hydrops (Merchant, Adams, & Nadol, 2005. Foster and Breeze (2013) suggested that the existence of asymptomatic endolymphatic hydrops, along with extensive endolymphatic hydrops seen in case studies of patients recently diagnosed with MD (Hallpike & Cairns, 1938; Hallpike & Wright, 1939), supports a causative role of endolymphatic hydrops in MD.
Another meta-analysis of 184 temporal bone specimens with endolymphatic hydrops (93 with MD, 91 without) looked at the location of the hydrops (Pender, 2014). They found endolymphatic hydrops most commonly in the cochlea (100% of cases), followed by the saccule (42%), followed by the utricle (28%), and lastly the SCCs (7%). The endolymphatic hydrops never occurred in the less common temporal bone regions, without simultaneous endolymphatic hydrops in the more common areas. Because of this, Pender (2014) suggested that endolymphatic hydrops starts in the cochlea before spreading to the saccule, the utricle, and lastly the semi-circular canals.

While evidence for the association between endolymphatic hydrops and MD is quite strong, the cause of endolymphatic hydrops is less clear. The classical model suggested that endolymphatic hydrops is due a blockage of the flow of endolymph, causing a build-up of endolymph (Schuknecht & Rüther, 1991). Indeed, blockage of the endolymphatic duct produces endolymphatic hydrops in guinea pigs (Kimura & Schuknecht, 1965). There is also evidence from imaging studies that the reuniting duct, saccular duct, and the endolymphatic sinus are harder to visualise in MD and therefore may be occluded (Yamane et al., 2010; Yamane et al., 2012; Takano, Iguchi, Sakamoto, Yamane, & Anniko, 2013). Yamane and colleagues (2010) have suggested that the blocking material could be detached saccular otoconia (similar to detached otoconia found in benign paroxysmal positional vertigo). It should be noted that studies suggest that there is very little longitudinal flow of endolymph (Salt, Thalmann, Marcus, & Bohne, 1986; Salt, 2001) and radial flow remains to be demonstrated experimentally (Salt & Plontke, 2010). This puts in doubt endolymphatic flow as a potential mechanism of injury.

A more recent, and potentially compatible, theory suggests that dysfunction of spiral fibrocytes could disrupt $K^+$ levels (Nin et al., 2008; Hamid, 2009; Adachi et al., 2013). This change would theoretically result in expansion of the endolymph compartment due to osmotic
pressure, while also disrupting hair cell function. Consistent with this is the fact that in an
guinea pig model of endolymphatic hydrops, spiral ligament fibrocytes are damaged prior to
the formation of endolymphatic hydrops (Shinomori, Kimura, & Adams, 2001; Momin,
Melki, Alagramam, & Megerian, 2009). Other contributing factors to endolymphatic hydrops
have been suggested, including allergies, viruses, genetic abnormalities, diet, and vascular
abnormalities. A “central” theory has been suggested by which multiple mechanisms can lead
to the formation of endolymphatic hydrops (Merchant et al., 2005).

1.2.3 Vestibular Migraine

The primary differential diagnosis of MD is Vestibular Migraine (VM; previously
referred to as migraine-associated vertigo/dizziness, migrainous vertigo and migraine-related
vestibulopathy/dizziness). VM, after benign paroxysmal positional vertigo, is the most
common form of episodic vertigo (Neuhauser, 2007). A large population-based study in
Germany estimated the lifetime prevalence of VM at 0.98% (Neuhauser et al., 2006). VM
also represents between 7-11% of patients at dizziness clinics and 9% of migraine clinics
(Neuhauser, Leopold, von Brevern, Arnold, & Lempert, 2001; Brandt et al., 2005; Maione,
2006; Ahn et al., 2009).

VM is also a highly variable condition. It is characterised by vestibular symptoms
which interfere with routine activities, lasting between seconds and days. Most vestibular
episodes (>50%) are associated with features of migraine, such as headache, transient visual
symptoms (visual aura) and sensitivity to light or sound (photophobia and phonophobia)
(Table 4); (Lempert et al., 2012; Headache Classification Committee of the International
Headache Society, 2013). Migraine typically precedes vestibular symptoms by 8-19 years

Table 3. Current Clinical Definition of VM (Headache Classification Committee of the International Headache Society, 2013).

<table>
<thead>
<tr>
<th>At least five episodes fulfilling criteria B and C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. A current or past history of Migraine without aura or Migraine with aura.</td>
</tr>
<tr>
<td>B. Vestibular symptoms of moderate or severe intensity, lasting between 5 minutes and 72 hours.</td>
</tr>
<tr>
<td>C. At least 50% of episodes are associated with at least one of the following three migrainous features:</td>
</tr>
<tr>
<td>1. headache with at least two of the following four characteristics:</td>
</tr>
<tr>
<td>a) unilateral location</td>
</tr>
<tr>
<td>b) pulsating quality</td>
</tr>
<tr>
<td>c) moderate or severe intensity</td>
</tr>
<tr>
<td>d) aggravation by routine physical activity</td>
</tr>
<tr>
<td>2. photophobia and phonophobia</td>
</tr>
<tr>
<td>3. visual aura</td>
</tr>
</tbody>
</table>

Not better accounted for by another ICHD-3 (The International Classification of Headache Disorders 3rd edition) diagnosis or by another vestibular disorder.

The vestibular symptoms are extremely varied and can include spontaneous, positional, visually-induced, and head motion-induced vertigo or dizziness with nausea (Table 5); (Headache Classification Committee of the International Headache Society, 2013).
These various forms of vestibular symptoms combined with the fact that migraine symptoms are also variable and are not associated with every vertigo episode (Table 4; Neuhauser et al., 2001), can make the diagnosis of VM troublesome. The duration of attacks is also extremely variable, ranging from seconds to days (Neuhauser et al., 2001; Neuhauser et al., 2006; Maione, 2006; Celebisoy, Gokcay, Sirin, & Bicak, 2008; Ahn et al., 2009; Boldingh, Ljøstad, Mygland, & Monstad, 2011; Radtke et al., 2011; Radtke et al., 2012; Boldingh et al., 2013; Shin et al., 2013; Cho et al., 2016).

Table 4. Prevalence of vestibular symptoms in patients with VM.

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% (SD; Range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinning vertigo</td>
<td>60.4 (17.6; 30-81.6)</td>
<td>(Neuhauser &amp; Lempert, 2004; Brevern, Zeise, Neuhauser, Clarke, &amp; Lempert, 2005; Neuhauser et al., 2006; Akdal, 2008; Ahn et al., 2009; Hassan, El-Raouf, &amp; Awad, 2010; Cohen et al., 2011; Radtke et al., 2011; Radtke et al., 2012; Boldingh et al., 2013)</td>
</tr>
<tr>
<td>Unsteadiness</td>
<td>75.2 (15.0; 61-92)</td>
<td>(Neuhauser et al., 2006; Cohen et al., 2011; Eggers et al., 2011; Radtke et al., 2011; Radtke et al., 2012)</td>
</tr>
<tr>
<td>Positional vertigo</td>
<td>33.0 (14.2;13.2-60)</td>
<td>(Neuhauser &amp; Lempert, 2004; Brevern et al., 2005; Neuhauser et al., 2006; Akdal, 2008; Ahn et al., 2009; Hassan et al., 2010; Radtke et al., 2011; Radtke et al., 2012; Boldingh et al., 2013)</td>
</tr>
<tr>
<td>Head motion vertigo</td>
<td>44.7 (18.0; 13.3-65.8)</td>
<td>(Neuhauser &amp; Lempert, 2004; Brevern et al., 2005; Neuhauser et al., 2006; Akdal, 2008; Ahn et al., 2009; Hassan et al., 2010; Radtke et al., 2012; Boldingh et al., 2013)</td>
</tr>
<tr>
<td>Oscillopsia</td>
<td>43 (9.9; 36-50)</td>
<td>(Brevern et al., 2005; Neuhauser et al., 2006)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>39 (42.4; 9-69)</td>
<td>(Neuhauser et al., 2006; Eggers et al., 2011)</td>
</tr>
</tbody>
</table>
Table 5. Migraine symptoms associated with vertigo in VM

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% (SD; range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migrainous headache</td>
<td>76.4 (15.3; 48 - 94.7)</td>
<td>(Neuhauser et al., 2001; Brevern et al., 2005; Neuhauser et al., 2006; Maione, 2006; Ahn et al., 2009; Hassan et al., 2010; Cohen et al., 2011; Radtke et al., 2011; Radtke et al., 2012; Boldingh et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Aura</td>
<td>24.7 (11.6; 7.5-39)</td>
<td>(Neuhauser et al., 2001; Brevern et al., 2005; Neuhauser et al., 2006; Maione, 2006; Ahn et al., 2009; Radtke et al., 2011; Radtke et al., 2012; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Photophobia</td>
<td>66.7 (17.0; 36-89.5)</td>
<td>(Neuhauser et al., 2001; Brevern et al., 2005; Neuhauser et al., 2006; Ahn et al., 2009; Radtke et al., 2012; Boldingh et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Phonophobia</td>
<td>53.1 (21.7; 10-79.8)</td>
<td>(Neuhauser et al., 2001; Neuhauser et al., 2006; Brevern et al., 2005; Ahn et al., 2009; Radtke et al., 2012; Boldingh et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Osmophobia</td>
<td>15</td>
<td>(Brevern et al., 2005)</td>
</tr>
</tbody>
</table>

The variability of diagnostic findings in vestibular migraine is further confounded by the fact that traditional bedside and oculomotor vestibular tests produce variable results in VM (Table 6). As for MD, diagnostic testing in VM typically is performed during the interictal period, presumably for the same reasons (patients are typically bedridden during the course of attacks, and there is typically a lengthy period before a patient can see a specialist).
Studies have shown rates of abnormal bedside and oculomotor test results in interictal VM anywhere between 15% and 78% (Dieterich & Brandt, 1999; Iwasaki et al., 2007; Celebisoy et al., 2008; Teggi et al., 2009; Radtke et al., 2012; Shin et al., 2013; Lee, Jung, Chung, & Suh, 2013; Neugebauer, Adrion, Glaser, & Strupp, 2013), and in ictal VM between 70-100% of patients (Brevern et al., 2005; Hassan et al., 2010; Polensek & Tusa, 2010). The difference in abnormal findings between studies may be partially explained by progression of the disease. Two longitudinal studies of patients with VM found increasing rates of oculomotor abnormalities with increasing duration of VM (Radtke et al., 2012; Neugebauer et al., 2013).

Both central oculomotor dysfunctions (in 9-63% of patients) and peripheral oculomotor dysfunctions are common during the interictal period (in 12-46% of patients; Dieterich & Brandt, 1999; Celebisoy et al., 2008; Teggi et al., 2009; Hassan et al., 2010; Radtke et al., 2012; Lee et al., 2013; Neugebauer et al., 2013). One of the most common findings in both ictal and interictal VM is static positional nystagmus (Hassan et al., 2010; Polensek & Tusa, 2010; Radtke et al., 2012; Boldingh et al., 2013). Nystagmus in VM can be vertical, horizontal or torsional, and can be direction changing (Brevern et al., 2005; Vitkovic, Paine, & Rance, 2008; Teggi et al., 2009; Hassan et al., 2010; Polensek & Tusa, 2010; Boldingh et al., 2013).
Table 6. Results of bedside and oculomotor vestibular tests in VM during the interictal period. * = during ictal period.

<table>
<thead>
<tr>
<th>Test</th>
<th>% Abnormal (SD; Range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spontaneous Nystagmus</td>
<td>4.8 (5.5; 0-15)</td>
<td>(Dieterich &amp; Brandt, 1999*; Brevern et al., 2005*; Furman, Sparto, Soso, &amp; Marcus, 2005; Maione, 2006; Iwasaki et al., 2007; Vitkovic et al., 2008; Celebisoy et al., 2008; Teggi et al., 2009; Hassan et al., 2010*; Polensek &amp; Tusa, 2010*; Nafie et al., 2011; Radtke et al., 2012; Boldingh et al., 2013; Neugebauer et al., 2013)</td>
</tr>
<tr>
<td>Gaze-evoked Nystagmus</td>
<td>7.3 (11.3; 0-27)</td>
<td>(Dieterich &amp; Brandt, 1999*; Brevern et al., 2005*; Furman et al., 2005; Celebisoy et al., 2008; Radtke et al., 2012; Boldingh et al., 2013; Neugebauer, Adrion, Glaser, &amp; Strupp, 2013)</td>
</tr>
<tr>
<td>Static Positional Nystagmus</td>
<td>29.7 (12.1; 11-42.1)</td>
<td>(Dieterich &amp; Brandt, 1999*; Brevern et al., 2005*; Furman et al., 2005; Maione, 2006; Hassan et al., 2010* Polensek &amp; Tusa, 2010*; Nafie et al., 2011; Radtke et al., 2012; Boldingh et al., 2013)</td>
</tr>
<tr>
<td>Vibration induced nystagmus</td>
<td>22 (9.5; 14.1; 12 32)</td>
<td>(Neff et al., 2012; Shin et al., 2013)</td>
</tr>
<tr>
<td>Headshake Nystagmus</td>
<td>20.9 (15.5;3-50)</td>
<td>(Maione, 2006*; Iwasaki et al., 2007; Akdal, 2008; Polensek &amp; Tusa, 2010*; Nafie et al., 2011; Radtke et al., 2012; Neff et al., 2012; Boldingh et al., 2013; Shin et al., 2013)</td>
</tr>
<tr>
<td>Caloric Hypofunction</td>
<td>21.9 (12.3;0-59)</td>
<td>(Cutrer &amp; Baloh, 1992; Reploeg &amp; Goebel, 2002; Furman et al., 2005; Radtke et al., 2012; Iwasaki et al., 2007; Akdal, 2008; Celebisoy et al., 2008; Teggi et al., 2009; Wang et al., 2009; Dieterich &amp; Brandt, 1999; Vitkovic et al., 2008; Nafie et al., 2011; Gode et al., 2012; Neff et al., 2012; Boldingh et al., 2013; Shin et al., 2013; Sharon &amp; Hullar, 2014; Blödow et al., 2014; Chen, Chang, Chen, &amp; Tseng, 2015; Kang et al., 2016; Morganti et al., 2016; Yoo et al., 2016)</td>
</tr>
</tbody>
</table>
Table 6 continued. Results of bedside and oculomotor vestibular tests in VM during the interictal period. * = during ictal period.

<table>
<thead>
<tr>
<th>Test</th>
<th>% Abnormal (SD; Range)</th>
<th>References</th>
</tr>
</thead>
</table>
| Caloric Directional Preponderance | 15.7 (8.2; 8.1-31.3)  
12.5*                              | (Dieterich & Brandt, 1999*; Vitkovic et al., 2008; Teggi et al., 2009; Neff et al., 2012; Boldingh et al., 2013) |
| vHIT                          | 12.2 (4.5-28.6)        | (Blödow et al., 2014; Rambold, 2014; Kang et al., 2016; Yollu et al., 2016; Yoo et al., 2016) |
| Rotatory Chair (SHA)          | 58                     | (Vitkovic et al., 2008)                                                  |
| Rotatory Chair (SVT)          | 25                     | (Vitkovic et al., 2008)                                                  |
| Romberg Test                  | 14.6 (11.0; 3-25)      
70*                              | (Brevern et al., 2005*; Iwasaki et al., 2007; Boldingh et al., 2013; Radtke et al., 2012) |
| Subjective vertical           | 1.3 (1.8; 0-2.6)       | (Neugebauer et al., 2013; Boldingh et al., 2013)                          |
| Random Saccade Test           | 8.6 (11.1; 0-27.8)     | (Celebisoy et al., 2008; Teggi et al., 2009; Neff et al., 2012; Radtke et al., 2012; Boldingh et al., 2013; Chen, Chang, Chen, & Tseng, 2015) |
| Saccadic Pursuit              | 22.6 (13.8; 8-48)      
25*                              | (Dieterich & Brandt, 1999*; Celebisoy et al., 2008; Teggi et al., 2009; Radtke et al., 2012; Neff et al., 2012 Boldingh et al., 2013; Neugebauer, Adrion, Glaser, & Strupp, 2013; Chen et al., 2015) |

There have also been numerous case studies of hearing loss in patients with migraine (Lipkin, Jenkins, & Coker, 1987; Viirre & Baloh, 1996; Olsson, 1991; Parker, 1991; Lee, Lopez, Ishiyama, & Baloh, 2000; Lee, Whitman, & Lim, 2003; Piovesan, Kowacs, Werneck, & Siow, 2003; Evans & Ishiyama, 2009; Radtke et al., 2012; Chu et al., 2013). However it should be noted that, in VM, hearing loss is typically uncommon and non-fluctuating and occurs primarily in the high frequencies (Brevern et al., 2005; Lempert & Neuhauser, 2009; Radtke et al., 2012).
Because of the variability of findings in vestibular tests and the absence of an objective test for VM, the diagnosis of VM is currently one of exclusion and relies primarily on case history (Neuhauser & Lempert, 2004; Gode et al., 2012). The case history of VM is also subject to significant variability between individual patients and between studies.

### 1.2.4 Pathophysiology of Migraine

To understand the pathophysiology of vestibular migraine, it is useful to look at the current understanding of the pathophysiology of migraine. The sensation of migraine headache pain is believed to be due to the sensitisation of nociceptive trigeminal sensory afferents (Pietrobon & Moskowitz, 2013). Ascending trigeminal sensory afferents project from the meninges to nuclei in the brainstem, hypothalamus, and the ventral posteromedial nucleus of thalamus. The thalamus is then believed to integrate and process nociceptive inputs, in conjunction with various cortical regions (Akerman, Holland, & Goadsby, 2011).

The cause of sensitization of the trigeminal pathway is currently unknown. However, the prevalent theory is that migraine is caused by inflammation of the pia, dura mater, and the cranial blood vessels (Pietrobon & Moskowitz, 2013; Espinosa-Sanchez & Lopez-Escamez, 2015). This would result in release of vasoactive proinflammatory peptides (including calcium gene-related protein, substance P, and neurokinin A) and activation and sensitization of the trigeminal afferents. The activation and sensitization would then spread to the brainstem, thalamus and cortical regions, producing the varied symptoms seen in migraine. It has been noted that due to the heterogeneity of the disorder there may be several mechanisms that induce migraine (Pietrobon & Moskowitz, 2013).
About 15% of patients with migraine will experience a transient focal neurological symptom known as aura (Diener, 2016). In 90% of patients with aura these symptoms are visual, but can also include disturbances to other sensory modalities and speech. Aura typically follows a different time course from other symptoms associated with migraine, in that it typically precedes migraine headache and is of a much shorter duration (5 to 60 minutes for aura compared with 4 to 72 hours for migraine headache); (Headache Classification Committee of the International Headache Society, 2013). Cortical spreading depression is currently believed to be the physiological correlate of aura. Cortical spreading depression is a slowly expanding (2-6mm min\(^{-1}\)) wave of glial and neuronal depolarisation across the cortex. Human studies indicate that it usually begins in the occipital lobe, consistent with visual aura being the most common form (Akerman et al., 2011). It has been suggested that cortical spreading depression may be a mechanism for inducing migraine headache by sustained nociceptive sensitization (Levy, 2012).

1.2.5 Pathophysiology of Vestibular Migraine

In contrast to migraine, very little is known about vestibular migraine. Controversy continues to exist over whether VM is a central or peripheral condition (Millen, Schnurr, & Schnurr, 2011). VM is believed to be caused by interactions in the afferent circuitry of the vestibular system and nociceptive systems (Espinosa-Sanchez & Lopez-Escamez, 2015). Potential central interactions have been suggested such as the thalamus and the brainstem at the parabrachial nucleus (which receives projections from both the vestibular nucleus and the trigeminal nucleus; Balaban, 2011). These central interactions may either facilitate the sensitization of the nociception system in response to vestibular stimulation or enable
migraine to produce vestibular symptoms by sensitizing the vestibular system (Espinosa-Sanchez & Lopez-Escamez, 2015). Cortical spreading depression has also been offered as a potential central cause for vestibular symptoms (Dieterich & Brandt, 1999). This has been suggested as unlikely as vestibular symptoms typically last between seconds to days, which is not consistent with the duration of aura (Liu & Xu, 2016).

Another potential interaction between the vestibular and nociceptive systems is at the inner ear. The inner ear is innervated by the trigeminal nerve via the basilar artery and the anterior inferior cerebellar artery (Espinosa-Sanchez & Lopez-Escamez, 2015). Trigeminal nerve fibres project to the stria vascularis, dark cells in the vestibular labyrinth and the spiral modiolar blood vessels (Vass, Shore, Nuttall, & Miller, 1998). Electrical stimulation of the trigeminal nerve results in increased vascular permeability and leakage of plasma at the basilar and the anterior inferior cerebellar arteries (Vass et al., 2001). Intravenous serotonin (5-HT)-induced vascular permeability results in leakage of plasma into the apical spiral ganglion, the modiolus, and intralabyrinthine segments of the superior and inferior nerves (Koo & Balaban, 2006). This leakage of plasma has been proposed to result in increased inner ear fluid pressure and altered ion homeostasis, therefore leading to peripheral vestibular symptoms in VM (Vass et al., 2004; Koo & Balaban, 2006). Figure 5 shows a current model of VM proposed by Furman, Marcus, and Balaban (2013).
1.2.6 Differentiating MD from VM

As mentioned earlier, both MD and VM are currently diagnosed primarily on case history and exclusion of other conditions. This is not ideal for two reasons. Firstly, patients often find it difficult to articulate their symptoms (Stolte, Holle, Naegel, Diener, & Obermann, 2015). Secondly, there is a significant overlap in the symptoms of VM and MD, as patients with VM can present with auditory symptoms during episodes (Table 7), while migraine symptoms are frequently associated with vestibular symptoms in MD (Table 8).

Figure 5. Proposed vestibular migraine pathways. Serotonin (5-HT), Noradrenaline (NA); (Adapted from Furman, Marcus, & Balaban, 2013).
Table 7. Cochlear symptoms during VM episodes.

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% (SD; Range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hearing loss</td>
<td>23.5 (13.8; 0-44)</td>
<td>(Brevern et al., 2005; Maione, 2006; Neuhauser et al., 2006; Eggers et al., 2011; Radtke et al., 2011; Radtke et al., 2012; Neff et al., 2012; Lopez-Escamez et al., 2014; Lepcha, Tyagi, Ashish, Augustine, &amp; Balraj, 2015; Morganti et al., 2016)</td>
</tr>
<tr>
<td>Tinnitus</td>
<td>31.0 (22.8; 0-69)</td>
<td>(Brevern et al., 2005; Maione, 2006; Neuhauser et al., 2006; Iwasaki et al., 2007; Eggers et al., 2011; Radtke et al., 2011; Radtke et al., 2012; Neff et al., 2012; Lopez-Escamez et al., 2014; Morganti et al., 2016)</td>
</tr>
<tr>
<td>Aural Fullness</td>
<td>32.6 (19.5; 11-70)</td>
<td>(Brevern et al., 2005; Neuhauser et al., 2006; Iwasaki et al., 2007; Radtke et al., 2011; Eggers et al., 2011; Radtke et al., 2012; Neff et al., 2012; Lopez-Escamez et al., 2014; Morganti et al., 2016)</td>
</tr>
</tbody>
</table>
Table 8. Migrainous symptoms during MD episodes.

<table>
<thead>
<tr>
<th>Symptom</th>
<th>% (SD; Range)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Migrainous headache</td>
<td>42.5 (41.0; 8.4-88)</td>
<td>(Neff et al., 2012; Shin et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Photophobia</td>
<td>33.1 (13.1; 18-41.2)</td>
<td>(Neff et al., 2012; Shin et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Phonophobia</td>
<td>48.4 (24.6; 20-63)</td>
<td>(Neff et al., 2012; Shin et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
<tr>
<td>Aura</td>
<td>12.3 (9.1; 4-22)</td>
<td>(Neff et al., 2012; Shin et al., 2013; Lopez-Escamez et al., 2014)</td>
</tr>
</tbody>
</table>

While it has been suggested that the issue of differential diagnosis is difficult in the early stages of MD and VM, evidence suggests that issues persist even after a diagnosis is made. 13% of patients of definite VM also fulfilled the criteria for bilateral MD (Radtke et al., 2011). Other studies have found 17-23% of patients with VM also meet the criteria for MD (Eggers et al., 2011; Neff et al., 2012) and 28-41% of patients with MD also meet the criteria for VM (Neff et al., 2012; Ghavami, Mahboubi, Yau, Maducdoc, & Djalilian, 2016). These findings suggest that there is a significant overlap of the diagnostic criteria for the two conditions, and therefore the diagnostic criteria alone are not sufficiently discriminative.

1.2.7 VM and MD – A common pathophysiology?

One attempt to explain the overlap between VM and MD is that they share a common pathophysiology. Several studies have found that the rate of migraine is higher in patients
with MD (Hinchcliffe, 1967; Morrison, 1981; Radtke et al., 2002; Sen, Georgalas, & Papesch, 2005; Ibekwe et al., 2008; Ghavami et al., 2016). However, another study found the rate of migraine was not different for patients with MD when compared with controls (Rassekh & Harker, 1992). Furthermore, a large population-based study found that the incidence of migraine in MD was similar to that of the general population (Gopen, Viirre, & Anderson, 2009). One explanation is that vertigo may act as a migraine trigger. Indeed, one study found that 49% of patients with migraine experienced a migraine within 24 hours of having caloric/rotatory chair testing (Murdin, Davies, & Bronstein, 2009).

One of the mechanisms first suggested for a shared pathology between VM and MD is an ischemic event resulting from migraine-induced vasospasm (Atkinson, 1962 as cited by Evans & Ishiyama, 2009; Parker, 1991; Viirre & Baloh, 1996). Three cases of sudden sensori-neural hearing loss following migraine have been reported that are consistent with an ischemic insult (Lee et al., 2000; Lee et al., 2003). However, the vasospasm theory of migraine has fallen out of favour due to a lack of evidence (Tfelt-Hansen, 2010).

It has been suggested that VM and MD may share a genetic cause (Baloh, 1997; Oliveira, Bezerra, Araujo, Almeida, & Messias, 1997; Oliveira, Messias, & Ferrari, 2002; Cha, Kane, & Baloh, 2008). Studies have identified families in which migraine and MD were highly associated, suggesting a shared heritability (Oliveira et al., 1997; Oliveira et al., 2002; Cha et al., 2008). In particular, it has been suggested that both conditions may be channelopathies (Radtke et al., 2002). Multiple instances of ion channel mutations have been identified in families who show high instances of migraine and MD (Baloh, 1997). There is also a high rate of mutations in aquaporin 3, a water channel, in patients with MD (Candreia, Schmuziger, & Gürtler, 2010). Mutations in such channels could potentially result in both central and peripheral disorders (Baloh, 1997).
Allergy has also been suggested as a common cause of VM and MD (Sen et al., 2005). Several studies have noted higher rates of allergy in patients with MD (Banks, McGinness, Harvey, & Sacks, 2012) and patients with migraine (Mehle, 2012) when compared to healthy controls. Additionally patients with MD and migraine self-report higher rates of allergy than patients with MD but without migraine (Sen et al., 2005).

Another explanation is vasoactive neuropeptide release from vascular trigeminal nerve fibres (Cutrer & Baloh, 1992; Vass et al., 2004; Koo & Balaban, 2006). There is evidence that trigeminal nerve fibres project to the stria vascularis, dark cells in the vestibular labyrinth and the spiral modiolar blood vessels (Vass et al., 1998). Electrical stimulation of the trigeminal nerve results in increased vascular permeability and leakage of plasma at the basilar and the anterior inferior cerebellar arteries (Vass et al., 2001). 5-HT induced vascular permeability results in leakage of plasma into the apical spiral ganglion, and the modiolus, and intralabyrinthine segments of the superior and inferior nerve (Koo & Balaban, 2006). This leakage of plasma has been proposed to result in increased inner ear fluid pressure, in the form of endolymphatic hydrops, and altered ion homeostasis. These changes could produce the peripheral vestibular and auditory symptoms in vestibular migraine (Vass et al., 2004; Koo & Balaban, 2006).

1.3 Diagnostic Accuracy

The main model for assessing the effectiveness of a diagnostic test is diagnostic accuracy. Diagnostic accuracy refers to the amount of agreement between the results of a
diagnostic test of interest (index test) and the results of the best available diagnostic procedure (reference standard). In studies of diagnostic accuracy, an index test is compared to the reference standard to determine the diagnostic accuracy of the index test (Linnet, Bossuyt, Moons, & Reitsma, 2012). Diagnostic accuracy studies operate on the assumption that differences between the index and reference test are due to errors in the index test (Van Rijkom & Verdonschot, 1995; Biesheuvel, Irwig, & Bossuyt, 2007).

There are several measures of diagnostic accuracy. The fundamental measures include sensitivity, specificity, and accuracy and are calculated using a 2 x 2 table (Figure 6).

Sensitivity (TP/(FN+TP)) is the proportion of positive reference test results that the index test correctly classifies. Specificity (TN/(TN + FP)) is the proportion of negative reference test results that the index test correctly classifies. Accuracy ((TN+TP)/(TN+TP+FP+FN)) is the overall proportion of people that the index test correctly classifies (Linnet et al., 2012).

<table>
<thead>
<tr>
<th>Reference Test</th>
<th>Test Result</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index Test</td>
<td>Positive</td>
<td>True Positive (TP)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>False Negative (FN)</td>
<td>True Negative (TN)</td>
</tr>
</tbody>
</table>

Figure 6. 2-by-2 table for estimating diagnostic accuracy.

1.3.1 Importance of Diagnostic Accuracy

Improved diagnostic accuracy is important as it can improve patient outcomes in several ways. These improvements of outcome can occur by way of earlier diagnosis,
reduced numbers of visits to health care professionals, improved prognostic information, better therapeutic decisions, and reassurance for the patient as to the presence or absence of a condition (Fineberg, 1977; Schünemann et al., 2008; Vassiliou, Vlastarakos, Maragoudakis, Candiloros, & Nikolopoulos, 2011; Bossuyt, Reitsma, Linnet, & Moons, 2012; Lin & Aligene, 2013). Additionally improved diagnostic accuracy would theoretically improve the quality of patient populations in clinical trials, giving a better indication of the efficacy of different therapeutics. While none of these benefits are guaranteed, the endeavour to improve the diagnostic accuracy of MD and VM has many potential advantages.

1.4 Systematic Reviews

Systematic reviews and are considered superior to traditional reviews (such as narrative or literature review) for synthesising evidence in health care. This is due to the innate bias associated with traditional reviews, as the methodology is not reported and is not systematic (McGowan, 2012). Several guidelines on recommended protocols for systematic reviews exist such as the preferred reporting items for systematic reviews and meta-analyses (PRISMA; Moher, Liberati, Tetzlaff, Altman, & Group, 2009), the Joanna Briggs Institute Reviewers Manual (Campbell et al., 2015), and the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011).

Systematic review methodologies, while having some differences, still approximately follow the four steps outlined by Goldschmidt (1986): 1) Develop a well-defined research question and inclusion criteria. 2) Develop a systematic search strategy and collect relevant articles according to the inclusion criteria. 3) Assess the methodological validity (or bias) of
studies in a systematic way. 4) Present the process and the informational outcome of the systematic review.

1.4.1 Systematic Reviews of Diagnostic Accuracy

Systematic reviews of diagnostic accuracy provide a transparent overview of the literature for the accuracy of a diagnostic test or tests. They also enable the pooling of results from several studies to provide a more statistically powerful analysis of diagnostic tests than individual studies (Reitsma, Moons, Bossuyt, & Linnet, 2012).

Special considerations arise when conducting systematic reviews of diagnostic accuracy. Studies of diagnostic accuracy are prone to different forms of bias than other studies (Whiting, Rutjes, Westwood, & Mallett, 2013). Recent initiatives have aimed to improve the quality of diagnostic studies. These include the Standards for Reporting of Diagnostic Accuracy (STARD) initiative, which guides diagnostic accuracy studies on how to report their findings so that sources of bias are better reported (Bossuyt et al., 2003; Ochodo & Bossuyt, 2013), and Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 (Whiting et al., 2011), which was designed to be used as a tool for systematically assessing bias in diagnostic accuracy studies. Both highlight common sources of bias in diagnostic accuracy studies, and therefore provide a valuable resource to researchers conducting systematic reviews and primary diagnostic studies.
1.4.2 Scoping Reviews

A subtype of the systematic review methodology is the scoping review (also referred to as a mapping review). Scoping reviews typically follow the same methodology as systematic reviews. The primary distinction is that scoping reviews do not typically assess the methodological validity or bias of selected studies. Instead scoping reviews aim to give an overview of the current evidence in an area, regardless of quality (Peters et al., 2015). Because of this, scoping reviews have a limited scope for impacting on health care, but are useful for providing an overview of a large body of literature and for providing direction for future research (Arksey & O’Malley, 2005). There are several guidelines for conducting a scoping reviews (Arksey & O’Malley, 2005; Levac, Colquhoun, & O’Brien, 2010; Peters et al., 2015).

1.4.3 Search Filters for Systematic Reviews

Search terms (in systematic reviews) can be described in terms of their sensitivity (the proportion of relevant articles retrieved by a search) and specificity (the proportion of irrelevant articles not retrieved by a search). A priority is placed on sensitivity when choosing search terms, as systematic reviews should endeavour to include as many relevant papers as possible (Higgins & Green, 2011). However specificity can also be important as search terms can often return an unmanageable number of irrelevant results. Search filters are a method of improving the specificity of a search.
The use of search filters for diagnostic accuracy studies is controversial in systematic reviews. The Cochrane Handbook for Systematic Review of Diagnostic Test Accuracy recommends using terms for the index test only in combination with terms for the condition of interest (de Vet, Eisinga, Ripagen, Aertgeerts, & Pewsner, 2008). Several studies also indicate that search filters may lead to the omission of important studies (Doust, Pietrzak, Sanders, & Glasziou, 2005; Leeflang, Scholten, Rutjes, Reitsma, & Bossuyt, 2006; Ritchie, Glenville, & Lefebvre, 2007; Mann, Hewitt, & Gilbody, 2008; Whiting et al., 2011; Beynon et al., 2013). However the same is true of limiting a systematic review to certain databases (which is true of all systematic reviews; Whiting, Westwood, Burke, Sterne, & Glenville, 2008).

The sensitive diagnostic clinical queries filters for MEDLINE (Haynes & Wilczynski, 2004) and Embase (Wilczynski & Haynes, 2005) have been shown to maintain high sensitivity (90-100%) relative to unfiltered searches, while also dramatically improving specificity (70-74%; Haynes & Wilczynski, 2004; Wilczynski & Haynes, 2005; Leeflang et al., 2006; Kastner, Wilczynski, McKibbon, Garg, & Haynes, 2009; Whiting et al., 2011; Wilczynski, McKibbon, Walter, Garg, & Haynes, 2013). It should be noted that one study found the MEDLINE sensitive clinical queries filter was only 69% sensitive (Ritchie et al., 2007). However the study’s methodology was flawed, as the authors used studies that were not originally obtained using MEDLINE in their gold standard. Overall these studies indicate the filter will miss some relevant studies when compared to an unfiltered search. However, as long as the use of the filter and the limitations of using it are acknowledged, the sensitive clinical queries search filter offers a valuable method to make otherwise unmanageable searches possible.
1.5 Rationale and Objectives

As mentioned earlier, the differential diagnosis of MD and VM is difficult, due to a reliance on symptom-based diagnosis. Despite the frequent occurrence of both conditions in clinics, a search of the literature identified only one non-systematic review on the topic of the differential diagnosis of MD and VM (Shepard, 2006), one systematic review on the topic of differential diagnosis in MD that did not mention VM (Vassiliou et al., 2011), and one systematic review on the topic of differential diagnosis in VM (Fasunla, Ibekwe, & Nwaorgu, 2012). Fasunla and colleagues (2012) acknowledged the difficulty in differentiating VM and MD but did not provide any evidence base for diagnostic tests that might separate the two conditions.

The present study aimed to conduct a systematic review of the differential diagnosis of MD and VM, with a focus on diagnostic accuracy. For the purpose of this systematic review, any relevant diagnostic test that had been compared between MD and VM was included. The diagnostic criteria for MD and VM were used as reference tests for the diagnostic tests to be compared to. This was done with the intention of establishing the evidence base for diagnostic separation of MD and VM.

A secondary scoping review was also conducted to examine any measures with high sensitivity and specificity for either MD or VM. This was conducted to identify any diagnostic tests that show high diagnostic accuracy but might be missed from the systematic review because the test was not compared in MD and VM simultaneously. The reasoning behind this scoping review was to identify tests that should theoretically separate MD and VM accurately. As will be described in section 2.4, clinical queries filters were used in the scoping review to make the number of search results manageable.
2.0 Methods
2.1 Protocol and registration

A published review protocol exists (Osborne, 2017). The protocol was created according to the preferred reporting items for systematic reviews and meta-analyses (PRISMA)-protocol guidelines (Moher et al., 2015).

2.2 Eligibility criteria

MD was defined as patients meeting either the 1995 (AAO-NHS, 1995) or 2015 AAO-NHS diagnostic criteria (Lopez-Escamez et al., 2015; Goebel, 2016). VM was defined as patients meeting any of the original Neuhauser diagnostic criteria (Neuhauser, Leopold, von Brevern, Arnold, & Lempert, 2001; Neuhauser & Lempert, 2004; Radtke, Neuhauser, von Brevern, Hottenrott, & Lempert, 2011) or the diagnostic criteria defined by consensus of the Bárány Society and International Headache Society (Lempert et al., 2012; Headache Classification Committee of the International Headache Society, 2013).

*Eligibility criteria for Systematic Review:*

1. Article was a meta-analysis, systematic review, randomised control trial, cohort, or case control (prospective or retrospective).
2. Article contained patients with MD and patients with VM.
3. Article contained a measure that could be used to separate MD and VM.
4. Article was written in English.
Eligibility criteria for Scoping Review:

1. Article was a meta-analysis, systematic review, randomised control trial, cohort, or case control (prospective or retrospective).
2. Article contained patients with MD and patients with VM.
3. Article contained a diagnostic test, for which sensitivity and specificity for MD or VM were available or could be calculated.
4. Article was written in English.

After studies meeting the above criteria for the scoping review were identified, diagnostic tests were identified that showed high sensitivity (>70%) and specificity (>70%) for MD or VM in at least one study. If a diagnostic test showed such high sensitivity and specificity in at least one study, all studies meeting the inclusion criteria for that diagnostic test with sensitivity and specificity were included. This is not a traditional PRISMA method, but is instead a novel approach we created to capture diagnostic tests that were accurate for MD, but without excluding studies that showed poorer accuracy in the identified diagnostic tests.

2.3 Information sources

Final searches were conducted on 27 November 2016. PubMed and Ovid (Embase) were used. PubMed was chosen as it provides access to MEDLINE, is updated more rapidly than other MEDLINE search engines, and can search article text for search terms (Kelly & St Pierre-Hansen, 2008). MEDLINE and Embase were chosen as they have been identified as
highly sensitive for diagnostic accuracy studies (Whiting, Westwood, Burke, Sterne, & Glanville, 2008; van Enst, Scholten, Whiting, Zwinderman, & Hooft, 2014). Grey/unpublished literature was not used in this study.

2.4 Search strategy

A search strategy was conducted according to the 2009 PRISMA statement (Moher et al., 2009). PubMed and Ovid (Embase) were searched until 27 November 2016. Filters for the English language, and articles after 1994 were used (as the relevant populations were defined by criteria set after this point). The sensitive diagnostic accuracy clinical queries filters were used (Haynes & Wilczynski, 2004; Wilczynski & Haynes, 2005). As mentioned in the introduction, the use of these filters reduces the sensitivity of a given search. However without the filter, 7563 studies were returned. This was considered to be an impractical amount and so the filters were used. The search strategy for the systematic review (Table 9) and the scoping review (Table 10) are presented below.
Table 9. Search strategy for the systematic review.

<table>
<thead>
<tr>
<th>System</th>
<th>Search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>PubMed</td>
<td>(‘meniere's’ OR ‘meniere’ OR ‘endolymphatic hydrops’) AND (‘vestibular migraine’ OR ‘migrainous vertigo’ OR ‘migraine associated vertigo’ OR ‘migraine associated dizziness’ OR ‘migraine-related vestibulopathy’ OR ‘migraine-related dizziness’)</td>
</tr>
<tr>
<td>Ovid (Embase)</td>
<td>(‘meniere's’ OR ‘meniere’ OR ‘endolymphatic hydrops’) AND (‘vestibular migraine’ OR ‘migrainous vertigo’ OR ‘migraine associated vertigo’ OR ‘migraine associated dizziness’ OR ‘migraine-related vestibulopathy’ OR ‘migraine-related dizziness’)</td>
</tr>
</tbody>
</table>
Table 10. Search strategy for the scoping review

<table>
<thead>
<tr>
<th>Search terms</th>
<th>PubMed</th>
<th>Ovid (Embase)</th>
</tr>
</thead>
</table>
2.5 Study Selection and Data Collection Process

One author (JO) read abstracts from the initial search results to screen whether the studies were potentially relevant to the current systematic review. The articles of all selected abstracts were then read to determine whether the articles meet the eligibility criteria. Data from studies was tabulated using a Microsoft Excel spreadsheet. QUADAS-2 was used to assess bias in individual studies (Whiting et al., 2011).

2.6 Data Items

Measures included the country in which the articles were conducted, the diagnostic criteria used to diagnose MD and VM, the diagnostic tests used in each study, measures of diagnostic accuracy, cut-off values, and bias using the QUADAS-2 items (Whiting et al., 2011). Where a measure of diagnostic accuracy was not available, a 2 x 2 table was constructed based on available information if possible. Where a 95% confidence interval (CI) of the control group was used to define the cut-off value, and specificity was not reported, the n value that gave the closest value to 95% specificity was used in the 2 x 2 table.

2.7 Synthesis of Results

For the systematic review and scoping review, forest plots of sensitivity and specificity were created. Forest plots show a point estimate (in the form of a square) for each study with a horizontal line which typically indicates the 95% CI (Lewis & Clarke, 2001). A
narrative summary was also produced. For the scoping review no further analysis was performed.

Heterogeneity was assessed using $I^2$ (Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003). Hierarchical summary receiver operating characteristics (HSROC); (Rutter & Gatsonis, 2001) were calculated for each test where appropriate. This is currently the recommended method for summary statistic for meta-analyses of diagnostic accuracy (over traditional summary receiver operating characteristics), as it takes into account the relationship between sensitivity and specificity, while also accounting for heterogeneity by using study- and patient-level covariates (Lee, Kim, Choi, Huh, & Park, 2015; Dinnes, Mallett, Hopewell, Roderick, & Deeks, 2016).

Meta-bias was intended to be assessed using the arcsine square-root transformed risk difference model including random effects (Rücker, Schwarzer, Carpenter, & Olkin, 2009), as this model was recommended for dichotomous outcome measures in a recent consensus document on funnel plot asymmetry (Sterne et al., 2011).

The quality of evidence was evaluated using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system (Guyatt et al., 2011). Case-control and cohort studies were considered to start as high quality evidence, before being modified, as recommended when reviewing diagnostic accuracy studies (Schünemann et al., 2008).
2.8 Software

Microsoft Excel, R x64 3.3.2, Meta-DiSc version 1.4 and Revman 5.3.5 were used. Forest plots were produced using Revman. $I^2$ was calculated using Meta-DiSc. HSROC figures were produced using the HSROC package for R (Schiller & Dendukuri, 2015). Funnel plots and bias measures were produced using the meta4diag package for R. These software were all freely available for non-commercial use (with the exception of Microsoft Excel).
3.0 Results of the Systematic Review
3.1 Included studies

Eighteen studies met the inclusion criteria and were included in the systematic review (Figure 7). Seven studies were prospective case-control in design, two were retrospective case-control, two were prospective cohort, and seven were retrospective cohort (Table 11).

Measures that were covered by articles in the systematic review included symptoms, caloric testing, head shaking nystagmus (HSN), vibration-induced nystagmus (VIN), vHIT, cervical vestibular evoked myogenic potentials (cVEMPs), ocular vestibular evoked myogenic potentials (oVEMPs), posturography, linear motion perception thresholds (LMPT), rotatory chair testing, and extratympanic click electrocochleography (ECoG).

All papers contained multiple sources of bias (Table 12). Common sources of bias included inappropriate exclusions in the patient group (such as patients with migraine, bilateral symptoms or hearing loss), case-control design, lack of blinding, use of probable/possible diagnostic criteria (inappropriate reference standard), and use of cut-off points that were determined post-hoc.

One study met the inclusion criteria, but was excluded (Mahringer & Rambold, 2014). The excluded study was a retrospective cohort that looked at vHIT in MD and VM, among other conditions (Mahringer & Rambold, 2014). The study was excluded because only patients with abnormal caloric results were included in their analysis.
Records identified from PubMed (n= 110) and Ovid (Embase) (n=216)

Records after duplicates removed (n = 212)

Records screened (n = 212)

Records excluded (n = 127)

Full text articles assessed for eligibility (n = 85)

Studies included in qualitative analysis (n = 18)

Studies included in quantitative analysis
Caloric (n = 6)

Full text articles excluded, with reasons (n = 67)
26 = not relevant study design
23 = MD and VM not compared
9 = no relevant outcome measures
9 = not relevant patients

Figure 7. PRISMA flowchart for the systematic review.
<table>
<thead>
<tr>
<th>Country</th>
<th>Study Design</th>
<th>Study Period</th>
<th>MD Criteria</th>
<th>VM Criteria</th>
<th>Other Groups</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korea</td>
<td>Retrospective cohort</td>
<td>2014</td>
<td>Neuhauser et al., 1995</td>
<td>AAO-NHS, 1995</td>
<td>None</td>
<td>Design</td>
</tr>
<tr>
<td>Japan</td>
<td>Retrospective cohort</td>
<td>2016</td>
<td>Lempert et al., 2012</td>
<td>AAO-NHS, 1995</td>
<td>Healthy</td>
<td>Study</td>
</tr>
<tr>
<td>Korea</td>
<td>Retrospective cohort</td>
<td>2013</td>
<td>Neuhauser et al., 1995</td>
<td>AAO-NHS, 1995</td>
<td>Other vestibular conditions</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Retrospective cohort</td>
<td>2014</td>
<td>Lempert et al., 2012</td>
<td>AAO-NHS, 1995</td>
<td>Controls with vestibular neuritis</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Prospective case-control</td>
<td>2014</td>
<td>Neuhauser et al., 1995</td>
<td>AAO-NHS, 1995</td>
<td>Healthy</td>
<td></td>
</tr>
<tr>
<td>Taiwan</td>
<td>Prospective case-control</td>
<td>2014</td>
<td>Lopez-Escamell et al., 2015</td>
<td>AAO-NHS, 1995</td>
<td>Controls with caloric and HIT</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Prospective case-control</td>
<td>2016</td>
<td>Bremaoja et al., 2016</td>
<td>AAO-NHS, 1995</td>
<td>Healthy</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Prospective case-control</td>
<td>2014</td>
<td>Biedow et al., 2014</td>
<td>AAO-NHS, 1995</td>
<td>Healthy</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Prospective case-control</td>
<td>2009</td>
<td>Bier &amp; Deibich, 2009</td>
<td>AAO-NHS, 1995</td>
<td>Healthy</td>
<td></td>
</tr>
</tbody>
</table>

Table 11: Summary of included studies.
<table>
<thead>
<tr>
<th>Country</th>
<th>Design</th>
<th>VM Criteria</th>
<th>Other Groups</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>Prospective case-control</td>
<td>AAO-NHS, 1995</td>
<td>Healthy controls</td>
<td>Other Groups, 2012, 2013</td>
</tr>
<tr>
<td>Turkey</td>
<td>Prospective case-control</td>
<td>AAO-NHS, 1995</td>
<td>Healthy controls</td>
<td>Other Groups, 2012, 2013</td>
</tr>
<tr>
<td>Germany</td>
<td>Retrospective cohort</td>
<td>AAO-NHS, 1995</td>
<td>Other conditions</td>
<td>Symptoms, PTA, Caloric, cVEMP, vHIT, Caloric</td>
</tr>
<tr>
<td>Japan</td>
<td>Retrospective cohort</td>
<td>AAO-NHS, 1995</td>
<td>None</td>
<td>Symptoms, PTA, Caloric, cVEMP, vHIT, Caloric</td>
</tr>
<tr>
<td>Spain</td>
<td>Retrospective cohort</td>
<td>AAO-NHS, 1995</td>
<td>None</td>
<td>Symptoms, PTA, Caloric, cVEMP, vHIT, Caloric</td>
</tr>
<tr>
<td>Germany, Spain</td>
<td>Retrospective cohort</td>
<td>AAO-NHS, 1995</td>
<td>None</td>
<td>Symptoms, PTA, Caloric, cVEMP, vHIT, Caloric</td>
</tr>
<tr>
<td>Italy, Luxembourg, Spain</td>
<td>Propective cohort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>USA</td>
<td>Prospective case-control</td>
<td>Healthy controls</td>
<td>cVEMP, oVEMP</td>
<td>Radke et al.</td>
</tr>
<tr>
<td>USA</td>
<td>Prospective cohort</td>
<td>Healthy controls</td>
<td>cVEMP</td>
<td>Lempert et al.</td>
</tr>
<tr>
<td>Australia</td>
<td>Prospective case-control</td>
<td>Healthy controls</td>
<td>cVEMP, oVEMP</td>
<td>Neuhauser et al.</td>
</tr>
<tr>
<td>Korea</td>
<td>Retrospective case-control</td>
<td>None</td>
<td>Caloric, VIN</td>
<td>Neuhauser et al.</td>
</tr>
</tbody>
</table>

Table 11 continued. Summary of included studies.
Table 12. QUADAS-2 assessment.

<table>
<thead>
<tr>
<th>STUDY</th>
<th>RISK OF BIAS</th>
<th>APPLICABILITY CONCERNS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient Selection</td>
<td>Index Test</td>
</tr>
<tr>
<td>Baier &amp; Dieterich, 2009</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Blödow et al., 2014</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Bremova et al., 2016</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Chang &amp; Hsu, 2014</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Heuberger et al., 2014</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hong et al., 2013</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Inoue et al., 2016</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Kim et al., 2014</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Lopez-Escamez et al., 2014</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>
Table 12 continued. QUADAS-2 assessment continued.

<table>
<thead>
<tr>
<th>STUDY</th>
<th>RISK OF BIAS</th>
<th></th>
<th></th>
<th>APPLICABILITY CONCERNS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient Selection</td>
<td>Index Test</td>
<td>Reference Standard</td>
<td>Flow and Timing</td>
</tr>
<tr>
<td>Martin-Sanz et al., 2014</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Neff et al., 2012</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Rambold, 2015</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Salviz et al., 2016</td>
<td>X</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sharon &amp; Hullar, 2014</td>
<td>X</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Shin et al., 2013</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Taylor et al., 2012</td>
<td>X</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>van Tilburg et al., 2016</td>
<td>X</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Zuniga et al., 2012</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>?</td>
</tr>
</tbody>
</table>
3.2 Symptoms

Five of the identified studies compared symptoms between patients with MD and VM (Neff et al., 2012; Shin et al., 2013; Lopez-Escamez et al., 2014; Chang & Hsu, 2014; Sharon & Hullar, 2014).

Features associated with the diagnostic criteria (auditory symptoms, migraine features, duration of attacks) were all significantly different between patients with MD and VM (Neff et al., 2012; Lopez-Escamez et al., 2014). Other studies also noted differences between groups with respect to these, but did not compare them statistically (Shin et al., 2013). Despite differences, most symptoms showed a large degree of overlap, particularly between MD and probable VM (Lopez-Escamez et al., 2014) or when patients meeting the criteria for both definite MD and definite VM were factored in (Neff et al., 2012).

A history of progressive hearing loss (93% sensitive and 78% specific) and having a pure tone average greater than 25 dB HL (100% sensitive and 91% specific) were best for distinguishing patients with MD from patients with VM (Neff et al., 2012). However, these measures did not separate patients with MD from patients with both VM and MD (Neff et al., 2012). It is important to note that these differences are likely to be due to applying the diagnostic criteria, and therefore any association with the criteria is expected.

Demographic features are different between patients with MD and VM. Patients with MD tend to be older at disease onset than patients with VM (Neff et al., 2012; Lopez-Escamez et al., 2014), and are more likely to be male (Neff et al., 2012). Attacks were more frequent in VM (Neff et al., 2012).
Car sickness and motion sickness appear to provide some separation. Car sickness was experienced by 78% of patients with VM in their lifetime, compared with only 18.2% of patients with MD (Chang & Hsu, 2014). Another study also found motion sickness was more common in patients with VM (51%) than in patients with MD (20%); (Neff et al., 2012). Sharon & Hullar (2014) found that scores on a motion sensitivity questionnaire were significantly higher for patients with definite VM when compared to patients with MD. However the scores did not differ significantly when patients with probable VM were included in the VM group.

Comorbidity of chronic subjective dizziness was more common in patients with VM (41%) than patients with MD (4%); (Neff et al., 2012). Interestingly family history was more common in patients who met the criteria for both MD and VM (Neff et al., 2012). This is consistent with the suggestion that the comorbid presentation of MD and VM may actually represent a genetic disorder distinct from the individual conditions (Cha & Baloh, 2007).

3.3 Caloric testing

In caloric testing the ear is irrigated using warm and/or cold water or air, and the resulting nystagmus is recorded. Caloric nystagmus is believed to be due to temperature change of fluid in the lateral canal resulting in activation of the VOR (for a systematic review of caloric methodology see Gonçalves, Felipe, & Lima, 2008). Ten of the identified studies compared caloric results between patients with MD and VM (Neff et al., 2012; Taylor et al., 2012; Hong et al., 2013; Shin et al., 2013; Blödow et al., 2014; Martin-Sanz et al., 2014; Kim et al., 2014; Sharon & Hullar, 2014; Rambold, 2015; Inoue et al., 2016). Kim and
colleagues (2014) used the same VM group as Shin and colleagues (2013), and were therefore excluded from the summary. The rate of caloric asymmetry was consistently higher in MD (43-75%) than VM (7-25%). Patients with MD had significantly higher asymmetry than patients with VM (Sharon & Hullar, 2014). Figure 8 shows the sensitivity and specificity of caloric asymmetry for MD when using VM as a control group. Neff and colleagues (2012) reported abnormal directional preponderance as having 29% sensitivity and 85% specificity for MD, with VM as a control group; however they did not report a cut-off value.

As several studies using caloric in MD and VM were identified (Taylor et al., 2012; Hong et al., 2013; Shin et al., 2013; Blödow et al., 2014; Rambold, 2015; Inoue et al., 2016), an HSROC was produced (Figure 9). Three studies were excluded from further analysis due to lack of a cut-off point (Neff et al., 2012; Martin-Sanz et al., 2014), or lack of 2 x 2 data (Sharon & Hullar, 2014). There was a significant amount of heterogeneity between studies for sensitivity (Chi$^2 = 16.89$, df = 5, $p = 0.005$; $I^2 = 70.4\%$), but not for specificity (Chi$^2 = 4.00$, df = 5, $p = 0.549$; $I^2 = 0.0\%$). The point estimate of sensitivity was 54.6% (95% CI: 38.7-69.0%) and point estimate of specificity 78.9% (95% CI: 70.7-87.4%), indicating that caloric testing would correctly identify about 55% of patients with MD and misclassify 21% of patients with VM.

A contour assisted funnel plot was also produced (Figure 10). The test for funnel plot asymmetry was negative ($t = 0.4489$, $p = 0.6767$). However funnel plot asymmetry tests typically have low power (Sterne et al., 2011). This combined with the fact that no studies were identified in the white region (area of non-significance) indicates that publication bias cannot be excluded as a possibility.
**Figure 8.** Forest plot of sensitivities and specificities for caloric testing when treating MD as a positive result and VM as a negative result.
Figure 9. HSROC of caloric testing, when MD is treated as a positive result and VM is treated as a negative result. Open circles indicates studies. The size of the circle indicates weighting. The blue line indicates the 95% CI of sensitivity and specificity. The red circle indicates the point estimate of sensitivity and specificity. The red line indicates the 95% CI of the point estimate.
Figure 10. Contour enhanced funnel plot. Black circles indicate included studies. No missing studies were identified using the trim and fill method.

<table>
<thead>
<tr>
<th>Arcsine Risk Transformed Difference</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.139</td>
<td>0.104</td>
</tr>
<tr>
<td>0.069</td>
<td>0.058</td>
</tr>
<tr>
<td>0.035</td>
<td>0.025</td>
</tr>
<tr>
<td>0.015</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Legend:
- 10' > d > 30' 0
- 30' 0 > d > 10' 0
- 10' 0 > d
3.4 Rotatory chair testing

Rotatory chair testing involves recording eye movements in response to rotation, and is believed to measure the VOR of the horizontal SCC (Ahmed, Goebel, & Sinks, 2009). One study used rotatory chair testing (Neff et al., 2012). Neff and colleagues (2012) found that patients with MD had more abnormal results than patients with VM for rotatory chair testing. The rotatory chair summary measure was also identified as one of the few tests that could separate patients with MD from patients that met the criteria for both MD and VM. However Neff and colleagues (2012) did not adequately describe their methods or provide cut-off values.

3.5 Video head impulse test (vHIT)

The vHIT involves recording eye movements in response to sudden head impulses in the planes of the SCCs. vHIT enables individual testing of the six SCCs with high sensitivity and specificity for peripheral vestibular disorders (MacDougall, Weber, McGarvie, Halmagyi, & Curthoys, 2009). Three of the identified studies compared vHIT results in patients with MD and VM (Blödow et al., 2014; Rambold, 2015; Heuberger et al., 2014).

The sensitivities and specificities of horizontal vHIT are presented below (Figure 11). Blödow and colleagues (2014) found that abnormal horizontal vHIT gain was more common in patients with MD (37%) than VM (9%), resulting in a sensitivity of 37% and a specificity of 91%. Rambold (2015) found abnormal gains of 8% in of patients with MD and 4.5% of patients with VM. Heuberger and colleagues (2014) found that horizontal vHIT gain was not
significantly different between MD (17%) and VM (7%). Heuberger and colleagues (2014) found that the presence of covert saccades was more common in MD (46%) than VM (18%).

![Figure 11](image)

*Figure 11.* Forest plot of sensitivities and specificities of horizontal vHIT when treating MD as a positive result and VM as a negative result.

### 3.6 Head Shaking Nystagmus (HSN)

HSN is a latent nystagmus provoked by rapid passive head shaking. The nystagmus is believed to be due to asymmetry in the central velocity storage mechanism (Boniver, 2008). Three papers looked at HSN; (Neff et al., 2012; Shin et al., 2013; Kim et al., 2014). Kim and colleagues (2014) had the same VM group as Shin and colleagues (2013). Because of this the duplicate data is not discussed.

The sensitivities and specificities of HSN are presented below (*Figure 12*). Neff and colleagues (2012) found abnormal HSN results in 62% of patients with definite/probable MD, and 15% of patients with definite/probable VM. Shin and colleagues (2013) found that 74% of patients with definite unilateral MD and 50% of patients with definite VM had abnormal HSN. Kim and colleagues (2014) went on to characterise the direction of HSN in MD and VM. They found that the most common type was mixed (46% for MD, 57% for...
VM), followed by horizontal (42% for MD, 30% for VM), and most rarely vertical nystagmus (5% for MD, 4% for VM).

**Figure 12.** Forest plot of sensitivities and specificities for caloric testing when treating MD as a positive result and VM as a negative result.

### 3.7 Vibration-induced nystagmus testing (VIN)

VIN testing involves vibrating the skull of a patient, and recording the resultant nystagmus (for a summary of VIN methodology see Dumas, Perrin, Ouedraogo, & Schmerber, 2016). Two studies compared VIN in patients with VM and MD (Neff et al., 2012; Shin et al., 2013). The sensitivities and specificities of HSN are presented below (Figure 13). Neff and colleagues (2012) found that 60% of patients with definite/probable MD had abnormal VIN, compared to 12% of patients with definite/probable VM. Shin and colleagues (2013) found abnormal VIN results in 53% of patients with definite unilateral MD and 32% of patients with definite VM.

**Figure 14.** Forest plot of sensitivities and specificities of VIN when treating MD as a positive result and VM as a negative result.
3.8 Cervical Vestibular Evoked Myogenic Potentials (cVEMP)

The cVEMP is a response from the sternocleidomastoid muscle of the neck produced by loud sounds, vibration, or tap stimuli. The cVEMP is believed to reflect the VCR, and it is thought to originate from the saccule. (For reviews see: Rosengren, Welgampola, & Colebatch, 2010 or Murofushi, 2014). Seven studies compared cVEMP results between patients with MD and VM (Baier & Dieterich, 2009; Taylor et al., 2012; Zuniga et al., 2012; Bremova et al., 2016; van Tilburg et al., 2016). One study also used cVEMPs (Neff et al., 2012). However they did not adequately describe their methods or the definition of an abnormal cVEMP, so their data was omitted.

Absent cVEMPs at 500 Hz TB are more common in MD (12.5-37%) than in VM (0-6%); (Baier & Dieterich, 2009; Taylor et al., 2012). Thresholds were significantly higher for the affected ear of patients with MD than for patients with VM (Taylor et al., 2012). Latencies were not significantly different between patients with MD and VM (Baier & Dieterich, 2009; Zuniga et al., 2012; Bremova et al., 2016), but there was a trend for patients with VM to have shorter p13 latencies in one study (Bremova et al., 2016).

P13-n23 amplitudes were significantly different between MD and VM for cVEMPs in one study (Taylor et al., 2012), but were not significantly different in other studies (Baier & Dieterich, 2009; Zuniga et al., 2012; Bremova et al., 2016). Baier & Dieterich (2009) reported a much lower sensitivity of 69% and specificity of 32% for cVEMP amplitudes, indicating that cVEMP amplitudes do not separate MD from VM.

Bremova and colleagues (2016) reported that p13-n23 amplitudes could separate MD and VM with 83% sensitivity and 80% specificity, despite being not significantly different. However, this appears to be an error as the mean amplitudes for MD and VM were within
one standard deviation of each other, indicating a large degree of overlap. Additionally the mean amplitude of the MD group was also above the cut-off value and it is highly unlikely that 80% of patients were below the mean.

Taylor and colleagues (2012) found the 0.5/1 kHz frequency amplitude ratio (FAR); (<-20.4%) separated patients with MD and VM with 35% sensitivity and 90% specificity. Murofushi, Ozeki, Inoue, and Sakata (2009) also compared the cVEMP 0.5/1 kHz FAR between patients with MD and VM but their definition of VM did not meet the inclusion criteria.

Taylor and colleagues (2012) found that the inter-aural asymmetry ratio (>29.2%) could separate patients with definite unilateral MD from patients with definite VM with 45% sensitivity and 85% specificity. Conversely Martin-Sanz and colleagues (2014) found a higher rate of inter-aural asymmetry in patients with VM than in patients with MD. This is likely because Martin-Sanz and colleagues (2014) performed testing before diagnosis, and bilateral MD, probable, and possible diagnoses were included. However it suggests that cVEMP inter-aural asymmetry may not be useful in the early stages of MD and VM.

Overall these findings support some ability of cVEMP asymmetry ratios and FAR may be useful in separating patients with MD and VM. However there are not currently enough studies to support this.

3.9 Ocular Vestibular Evoked Myogenic Potentials (oVEMP)

The oVEMP is a response from the inferior oblique muscle of the eyes caused by loud sounds, vibration, or tap stimuli. The oVEMP is believed to reflect the VOR, and its origin is
believed to be from the utricle (for a review see Rosengren et al., 2010). Three studies used oVEMPs in both patients with MD and VM (Taylor et al., 2012; Zuniga et al., 2012; Bremova et al., 2016). Zuniga and colleagues (2012) found that the click and 500 Hz TB oVEMP latency was longer for patients with MD than VM. However, Bremova and colleagues (2016) found that oVEMP latencies did not separate MD and VM.

Zuniga and colleagues (2012) found that, for 500 Hz oVEMPs, patients with MD had significantly smaller n10 amplitude than patients with VM and healthy controls. However there were no significant differences between groups for bone conducted or tap oVEMP n10 amplitudes. Bremova and colleagues (2016) found that oVEMP amplitudes did not help to differentiate MD and VM.

Despite having oVEMP amplitude and latency data for MD and VM, Taylor and colleagues (2012) did not statistically compare these measures between MD and VM. They noted that oVEMPs showed a similar trend as cVEMPs in that healthy controls and patients with VM had largest oVEMPs at 0.5 and 1 kHz, while patients with MD had largest responses at 1 kHz. Taylor and colleagues (2012) were unable to analyse the 0.5/1 kHz frequency ratio for oVEMPs due to a large number of absent oVEMPs. Overall there is not sufficient evidence that oVEMPs can separate patients with MD and VM, but 500 Hz TB oVEMPs may be useful.

3.10 Posturography

Posturography involves measuring posture changes in response to experimentally induced balance perturbations. (For a review see Visser, Carpenter, van der Kooij, & Bloem,
2008). Two studies used posturography in patients with MD and VM (Neff et al., 2012; Hong et al., 2013). Neff and colleagues (2012) found no significant difference in the sensory organisation test composite score between patients with VM and MD. Hong and colleagues (2013) found that patients with VM have a higher rate of abnormalities somatosensory score for the sensory organisation test (19%) than patients with MD (0%). However no other posturography measure was significantly different.

3.11 Linear Motion Perceptual Threshold (LMPT)

LMPTs are obtained by placing a patient in a seat. The seat then moves for 1 second in a linear direction, and the patient makes a judgement of the direction that they moved. One study used LMPT (Bremova et al., 2016). Bremova and colleagues (2016) found that patients with MD had elevated thresholds for naso-occipital and inter-aural movement relative to healthy controls and patients with VM. They also found that patients with VM had elevated thresholds for head-vertical movements relative to MD and controls. The sensitivities and specificities for separating MD from VM using LMPT were 84% and 70% for inter-aural, 72% and 83% for naso-occipital and 88% and 63% for head-vertical movements. The study groups were not age-matched.

3.12 Extratympanic click Electrocochleography (ECoG)

Extratympanic click ECoG is a technique that involves recording bioelectrical responses from the cochlea, via the tympanic membrane or ear canal, in response to loud
clicks (for a review see: Wuyts, Van De Heyning, Van Spaendonck, & Molenberghs, 1997). One study used extratympanic click ECoG (Martin-Sanz et al., 2014). Martin-Sanz and colleagues (2014) found that 43% of patients with MD had abnormal ECoG (SP/AP > 50%), compared with 24% of patients with VM. While this is only one study, it indicates that extratympanic click ECoG is not particularly useful in separating the two conditions.

3.13 Multivariate Methods

Neff and colleagues (2012) found that moderate/severe headache, caloric (directional preponderance), rotatory chair test, and initial PTA results could separate MD, VM, and patients with both MD and VM (Area under curve = 0.996-1.00). However they did not adequately describe the diagnostic tests or provide cut-off values.

Taylor and colleagues (2012) found that a combination of 0.5/1 kHz TB cVEMP FAR (< -20.4%), 0.5 kHz TB cVEMP amplitude asymmetry (> 29.2%) and caloric asymmetry (>25%) separated MD and VM with 90% sensitivity and 70% specificity. However their study only included unilateral MD (and not bilateral), and therefore asymmetry ratios are likely to overestimate sensitivity for MD.

It is important to note that, as with cut-off values for individual diagnostic tests, the use of combinations of individual diagnostic tests should ideally be set a priori. As all of these studies determined the best combination of test for separating MD and VM post-hoc, it is likely that these sensitivities and specificities are inflated. Despite this, these findings are promising and warrant further research.
3.14 Summary of Evidence

A table summarising the evidence base is presented below (Table 13). There was a moderate level of evidence supporting that caloric testing will correctly diagnose MD in approximately 50% of patients, while misdiagnosing approximately 20% of patients with VM as MD (54.6% sensitivity and 78.9% specificity). This indicates that the caloric test does provide some diagnostic value in the differential diagnosis of MD and VM. However a caloric result is insufficient to either confirm, or exclude a diagnosis of MD or VM.

Several other diagnostic tests have been compared between patients with MD and VM, and show some potential for assisting the differential diagnosis of the two conditions. These tests included VIN, HSN, rotatory chair testing, cVEMP asymmetry, 0.5/1 kHz cVEMP FAR, and LMPT. However these diagnostic tests had sensitivity and specificity data for only one to three studies and therefore meta-analysis of these tests was not considered to be appropriate.
<table>
<thead>
<tr>
<th>Evidence</th>
<th>87.4% (95% CI: 70.7–78.9%)</th>
<th>69.0% (95% CI: 38.7–54.6%)</th>
<th>No concern</th>
<th>No concern</th>
<th>No concern</th>
<th>No concern</th>
<th>No concern</th>
<th>No concern</th>
<th>Design (case-control and cohort)</th>
<th>No concern (lack of blinding)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality</td>
<td>WM as a control</td>
<td>MDD as a control</td>
<td>Bias</td>
<td>Imprecision</td>
<td>Indirectness</td>
<td>Inconsistency</td>
<td>Limitations</td>
<td>Limitations</td>
<td>(Design of studies)</td>
<td>No of studies</td>
</tr>
<tr>
<td></td>
<td>WM as a control</td>
<td>MDD as a control</td>
<td>Bias</td>
<td>Imprecision</td>
<td>Indirectness</td>
<td>Inconsistency</td>
<td>Limitations</td>
<td>Limitations</td>
<td>(Design of studies)</td>
<td>No of studies</td>
</tr>
</tbody>
</table>

Table 13: GRADE Summary of Evidence
4.0 Results of the Scoping Review
4.1 Included Studies

Tests that had greater than 70% sensitivity and specificity in at least one study included: caloric, VIN, gadolinium magnetic resonance imaging (MRI), 3-dimensional cone beam computed tomography, extraympanic click ECoG, transtympanic click ECoG, transtympanic TB ECoG, electrovestibulography (EVestG), cochlear hydrops analysis masking procedure (CHAMP), and distortion product otoacoustic emission (DPOAE) phase shift. Sixty-five studies meeting the inclusion criteria were identified that used one of the above tests (Figure 15). Studies that could not be accessed can be found in the Appendix.
Figure 15. PRISMA flowchart for the scoping review.

Records identified from PubMed (n= 1686) and Ovid (Embase) (n=1496)

Records after duplicates removed (n = 2327)

Records screened (n = 2327)

Records excluded (n = 1511)

Full text articles assessed for eligibility (n = 831)

Studies that met the first stage of inclusion criteria (n = 184)

Studies with measures that had >70% sensitivity/specificity in at least one study (n = 65)

Full text articles excluded, with reasons (n = 647)

321 = not relevant study design
165 = not relevant outcome measures
147 = not relevant patient group
14 = could not access
4.2 Caloric testing

Twenty studies using caloric testing were identified that met the inclusion criteria for the scoping review (Jouko, Martti, & Arto, 2003; Yetiser et al., 2004; Furman et al., 2005; Chen & Young, 2006; Celebisoy et al., 2008; Vitkovic et al., 2008; Baier & Dieterich, 2009; Teggi et al., 2009; Neff et al., 2012; Taylor et al., 2012; Hong et al., 2013; Shin et al., 2013; Blödow et al., 2014; Martin-Sanz et al., 2014; Kim et al., 2014; Sanyelbhaa Talaat & Sanyelbhaa Talaat, 2014; Sharon & Hullar, 2014; Zhu et al., 2014; Rambold, 2015; Inoue et al., 2016).

Ten studies had data comparing MD and VM (Neff et al., 2012; Taylor et al., 2012; Hong et al., 2013; Shin et al., 2013; Blödow et al., 2014; Martin-Sanz et al., 2014; Kim et al., 2014; Sharon & Hullar, 2014; Rambold, 2015; Inoue et al., 2016). However, as these were covered in the systematic review, they will not be covered here.

Only two studies compared caloric in patients with MD with a group of healthy controls (Yetiser et al., 2004; Sharon & Hullar, 2014). Due to the limited number of studies that met the inclusion criteria, not much can be said about the sensitivity and specificity of caloric testing for MD based on the studies that met the inclusion criteria (Figure 16). Overall these results indicate that there are large variations in the sensitivity of caloric for MD. Additionally caloric testing differentiates some patients VM from healthy controls (Figure 17). As would be expected, specificity was poorer (36 - 79.1%) when MD and VM are compared with other vestibular conditions.

A meta-analysis of normative data in the caloric test indicated that the 95% CIs for healthy controls was below 25% canal paresis in all identified studies, with an average level of 22% (Wuyts & Boniver, 2008). Because of this it can be assumed that
caloric testing using 25% as a cut-off value is highly specific (Over 95% of healthy controls should have negative caloric results). This is consistent with studies that used caloric testing in patients with healthy controls (Figure 16 and 17).

Figure 16. Sensitivity and specificity of caloric testing from studies, when a positive result is treated as MD.

<table>
<thead>
<tr>
<th>Study</th>
<th>Control group</th>
<th>Cutoff point</th>
<th>MD Group</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharon and Hurler, 2014</td>
<td>Healthy controls</td>
<td>&gt;25%</td>
<td>Not Clear Prob/Pos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellander et al., 2004</td>
<td>Healthy controls</td>
<td>&gt;25%</td>
<td>Prob/Pos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen and Young, 2005</td>
<td>SSNHL</td>
<td>&gt;25%</td>
<td>Prob/Def</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jouke et al., 2003</td>
<td>Vestibular patients</td>
<td>&gt;15%</td>
<td>Def/Prob</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhu et al., 2014</td>
<td>Vestibular patients</td>
<td>&gt;25%</td>
<td>Def/Prob</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 17. Sensitivity and specificity of caloric testing from studies, when a positive result is treated as VM.

4.3 Vibration Induced Nystagmus (VIN)

Four studies using VIN met the inclusion criteria for the scoping review (Ohki, Murofushi, Nakahara, & Sugasawa, 2003; Neff et al., 2012; Shin et al., 2013; Xie et al., 2013). Data comparing MD and VM was covered in the systematic review and will not be covered here (Neff et al., 2012; Shin et al., 2013). The sensitivity and specificity of VIN is presented below (Figure 18). One study identified VIN as having 74% sensitivity and 100%
specificity when distinguishing definite MD from healthy controls (Xie et al., 2013).

However as would be expected the specificity is much lower when compared with other vestibular conditions (Ohki et al., 2003; Xie et al., 2013). Interestingly, nystagmus towards the affected side appears to be pathognomonic to MD (Ohki et al., 2003; Xie et al., 2013).

![Figure 18. Sensitivity and specificity of VIN when a positive result is considered to indicate MD.](image)

### 4.4 Gadolinium MRI

Nine studies were identified that used gadolinium MRI (Fiorino, Pizzini, Beltramello, Mattellini, & Barbieri, 2011; Tanigawa et al., 2011; Grieve, Obholzer, Malitz, Gibson, & Parker, 2012; Baráth et al., 2014; Hagiwara et al., 2014; Homann et al., 2015; Hornibrook et al., 2015; Sepahdari et al., 2015; Ito et al., 2016). Typical MRI cannot clearly visualise the inner ear. However gadolinium, a contrast agent, when injected intratympanically or intravenously, enters the inner ear and allows imaging of the perilymph (and therefore the endolymph in regions without perilymph); (Nakashima et al., 2007). This enables calculation of the relative endolymph and perilymph, and therefore allows visualisation of endolymphatic hydrops. (For a review see Naganawa & Nakashima, 2014).

One systematic review was identified on the topic of gadolinium MRI in MD (Ziylan, Smeeing, Stegeman, & Thomeer, 2016). However they limited their search to studies with
both gadolinium MRI and click ECoG, and therefore did not sufficiently cover the literature on gadolinium MRI.

The sensitivity and specificity of gadolinium MRI for MD is presented below (Figure 19). Most studies used the grading system that was proposed by Nakashima and colleagues (2009). Gadolinium MRI showed 45-64% sensitivity for MD in the vestibule or cochlea, with high specificity (Homann et al., 2015; Sepahdari et al., 2015). Sensitivity was higher in studies that defined a positive result as hydrops from either the cochlea or vestibule (64-90%), while maintaining high specificity (Grieve et al., 2012; Baráth et al., 2014; Hornibrook et al., 2015; Ito et al., 2016). Several studies found that the vestibule had a higher rate of hydrops than the cochlea (Baráth et al., 2014; Homann et al., 2015; Ito et al., 2016). This contradicts histological studies that suggest a higher rate of cochlear hydrops (Pender, 2014).

Figure 19. Sensitivity and specificity of gadolinium MRI from studies, when a positive result is treated as MD.

One study was identified using gadolinium MRI in patients with MD and VM that met the inclusion criteria. Nakada and colleagues (2014) used the Nakashima grading system (Nakashima et al., 2009) in patients with VM and vestibular MD (Alford, 1972). While vestibular MD is not within the inclusion criteria, it can be treated as a group for comparison with VM. They found that significant hydrops in the vestibule was present in the vestibule of
100% (7/7) of patients with vestibular MD, whereas only 14% (1/7) patients with VM had hydrops in the vestibule.

One study using gadolinium MRI in patients with VM did not meet the inclusion criteria for the scoping review due to lack of a control group (Gürkov et al., 2014). However it will be summarised here due to its extreme relevance to the topic. Gürkov and colleagues (2014) used the Nakashima grading system (Nakashima et al., 2009) in patients with VM who also had auditory symptoms. They found that gadolinium MRI identified hydrops in only 21% of patients. This indicates that gadolinium MRI would have relatively high specificity for excluding patients with VM.

### 4.5 Three-Dimensional Cone Beam Computed Tomography (CT)

One study found that three-dimensional cone beam CT of the membranous labyrinth could separate patients with definite MD from healthy controls with 84% sensitivity and 100% specificity (Yamane et al., 2014).

### 4.6 Electrocochleography (ECoG)

ECoG met the inclusion criteria for the scoping review. ECoG is a technique that involves recording bioelectrical responses from the cochlea in response to acoustic stimuli (loud clicks or tone bursts [TB]). The responses are recorded by placing an electrode through the tympanic membrane and onto the promontory of the cochlea (transtympanic) or in the ear canal or on the tympanic membrane (extratympanic); (For a review see Wuyts et al., 1997).
In healthy controls, ECoG results in a small summating potential (SP; a direct current shift due to cochlea hair cell function) and a large compound action potential (AP; an analogue current shift due to primary cochlea afferents of the auditory nerve (Ferraro & Durrant, 2006). Gibson, Moffat, and Ramsden (1977) found that 65% of patients with MD had enlarged SP in transtympanic click ECoG. Gibson and colleagues (1977) suggested that this enlarged SP could be due to displacement of the basilar membrane due to endolymphatic hydrops or a change in the ionic composition of endolymph.

One group conducted a general systematic review of ECoG (Lamounier, Gobbo, de Souza, de Oliveira, & Bahmad, 2014). They did not produce a meta-analysis, but they concluded that TB ECoG was more sensitive than click ECoG. Another group conducted a systematic review of click and TB ECoG (Qi & Nunez, 2013). They noted that click ECoG had a sensitivity of 53-70% and specificity of 80-90%, and that the combination of click and TB measures improved sensitivity to 84-92% and specificity to 85-90%. However this systematic review was a poster and was therefore excluded. The studies using ECoG from the scoping review were divided into transtympanic click ECoG, transtympanic TB ECoG and extratympanic ECoG and are presented below.

### 4.6.1 Transtympanic click ECoG

Ten studies using transtympanic click ECoG met the criteria (Orchik, Ge, & Shea, 1998; Sass, 1998; Ikino & de Almeida, 2006; Gibson, 2009; Baba et al., 2009; Ohashi, Nishino, Arai, Hyodo, & Takatsu, 2009; Iseli & Gibson, 2010; Claes, De Valck, Van de Heyning, & Wuyts, 2011; de Carvalho Lopes, Munhoz, Santos, Moraes, & Chaves, 2011; Hornibrook et al., 2015). Transtympanic click ECoG showed varied sensitivities between studies (Figure)
20). Various SP/AP ratio cut-off values have been used between different studies, ranging from 28-50%. There is a notable trend that transtympanic click ECoG had lower sensitivity in studies that used higher cut-off values. Problematically, Gibson (2009) found that transtympanic ECoG did not separate MD and controls, when they were matched for hearing loss. This indicates that transtympanic click ECoG does not have diagnostic value in MD.

### Figure 20. Sensitivity and specificity of transtympanic click ECoG from studies, when a positive result is treated as MD.

#### 4.6.2 Transtympanic TB ECoG

There were six studies that met the inclusion criteria using transtympanic TB ECoG (Sass, 1998; Sass, Densert, Magnusson, & Whitaker, 1998; Gibson, 2009; Iseli & Gibson, 2010; Claes et al., 2011; Hornibrook et al., 2015). Transtympanic TB ECoG consistently had high sensitivity for MD (61.7-91.2%) and specificity (70.9-91.1%) for separating MD from patients with vestibular conditions (Iseli & Gibson, 2010; Claes et al., 2011) and patients with hearing loss (Gibson, 2009; Hornibrook et al., 2015); (Figure 21).
Using a 30 Hz tone to increase the size of the SP, Iseli and Gibson (2010) found that patients with MD had a reduced increase in SP. A bias ratio (between biased and unbiased SP amplitude) of < 1.4 could separate patients with MD from patients with hearing loss or vestibular conditions with 85% sensitivity and 80.6% specificity. When this bias measure was combined with the 1 kHz TB SP (<-6µV) the tests had 95% sensitivity and 79.1% specificity.

Combining click (>41% SP/AP) and 1 kHz TB ECoG measures (>3.1µV) improved sensitivity (82-83%) and specificity (95-100%) for separating patients from MD and patients with hearing loss (Sass, 1998; Sass et al., 1998). The addition of click action potential shift between rarefaction and condensation to click and TB measures further improved sensitivity to 87% while maintaining specificity at 100%.

Claes and colleagues (2011) found that a combination of 4 kHz TB ECoG and air conduction pure-tone audiometry thresholds at 150 Hz, and 8 kHz could separate patients with MD from patients with vestibular conditions with 94% sensitivity and 98% specificity. While impressive, it is obviously expected that this measure would be much less effective if patients were matched for hearing loss.

Figure 21. Sensitivity and specificity of transtympanic TB ECoG from studies, when a positive result is treated as MD.
4.6.3 Extratympanic ECoG

Eleven studies using extratympanic click ECoG met the inclusion criteria (Chung, Cho, Choi, & Hong, 2004; Noguchi, Nishida, Kawashima, Tokano, & Kitamura, 2004; Hwang, Ho, Hsu, Yang, & Liu, 2008; Satar et al., 2008; Vitkovic et al., 2008; Lee et al., 2011; Oh et al., 2014; Lamounier et al., 2014; Martin-Sanz et al., 2014; Gerenton et al., 2015; Yollu et al., 2016). Data directly comparing MD and VM was covered in the systematic review, and will not be covered here (Martin-Sanz et al., 2014). The sensitivity of extratympanic click ECoG was highly variable between studies (Figure 22), and this heterogeneity did not appear to be consistent with differences in patient groups or cut-off values.

Figure 22. Sensitivity and specificity of extratympanic click ECoG from studies, when a positive result is treated as MD. Data from Oh and colleagues (2014) was estimated from a receiver operating characteristic, at the point which gave 95% specificity.

None of the studies used groups of patients with hearing loss. Noguchi and colleagues (2004) found that 50% of patients with acute low-tone hearing loss had endolymphatic hydrops according to extratympanic click ECoG. They suggested that acute-low-tone hearing loss may represent a form of endolymphatic hydrops. It should be noted that only a minority of patients with acute low-tone hearing loss progress to MD (Yamasoba, Kikuchi, Sugasawa,
Yagi, & Harada, 1994). Because of this it is important to consider that the patients may have had positive ECoG results due to their hearing loss rather than endolymphatic hydrops, as was indicated for transtympanic click ECoG (Gibson, 2009).

One study found that a combination of the SP/AP ratio with the SP/AP area ratio was highly sensitive (90%) and specific (95%) for MD (Ferraro & Tibbils, 1999). However another study found no significant difference between MD and health controls using the SP/AP area ratio (Oh et al., 2014). Three other studies were identified that used the SP/AP area ratio, but were excluded as they did not meet the inclusion criteria (Devaiah, Dawson, Ferraro, & Ator, 2003; Al-momani, Ferraro, Gajewski, & Ator, 2009; Baba et al., 2009).

Two studies meeting the inclusion criteria looked at extratympanic ECoG in VM. Vitkovic, Paine, & Rance, (2008) found that 0% of patients with definite VM had positive ECoG (SP/AP >50%). However using the same cut-off point, Yollu and colleagues (2016) found that 38.1% of patients with definite VM had positive ECoGs. With a less conservative cut-off value (SP/AP >40%), Yollu and colleagues (2016) found that 71.4% of patients with definite VM had positive ECoGs.

Overall there is a large amount of inconsistency between studies for extratympanic ECoG. It is possible that the inconsistencies are due to the higher noise floor of extratympanic ECoG.

### 4.7 Electrovestibulography (EVestG)

Two studies using EVestG met the inclusion criteria (Blakley, Dastgheib, Lithgow, & Moussavi, 2014; Dastgheib, Lithgow, Blakely, & Moussavi, 2016). EVestG uses
extratympanic ECoG, but tilts or rotations of the body are used as a stimulus, instead of an acoustic stimulus (Lithgow, Garrett, & Heibert, 2008; Dastgheib et al., 2016). EVestG using a sideways body tilt as the stimulus shows both high sensitivity and specificity (Figure 23). The specificity of EVestG using sideways tilt remains high when vestibular patients are included (94%). However unclassified responses were common in these patients (33%); (Dastgheib et al., 2016). All studies using EVestG have been conducted by one research group with small sample sizes. It will be important for findings to be replicated by other labs, and in larger populations. It will also be important to determine whether hearing loss influences the EVestG.

Figure 23. Sensitivity and specificity for sideways body tilt EVestG in included studies, when a positive result is treated as MD.

4.8 Cochlear Hydrops Analysis Masking Procedure (CHAMP)

Seven studies using CHAMP met the inclusion criteria (Don, Kwong, & Tanaka, 2005; De Valck, Claes, Wuyts, & Van de Heyning, 2007; Don, Kwong, & Tanaka, 2007; Ordóñez-Ordóñez et al., 2009; Kingma & Wit, 2010; Lee et al., 2011; Shang et al., 2012). CHAMP involves measurement of auditory brainstem response (ABR) latencies in response to clicks and high-pass masking noise. CHAMP is performed by calculating the difference between latencies for wave Vs at clicks with high pass noise at 500 Hz and 8 kHz. This effectively calculates the time that sound takes to travel from the basal cochlear (8 kHz) to
the apical cochlear (500 Hz); (Don et al., 2005). When high pass masking noise is used to estimate the speed of sound along the basilar membrane, patients with MD have been shown to have higher travelling wave velocity than patients with noise-induced hearing loss (Thornton & Farrell, 1991; Donaldson & Ruth, 1996). A physiological explanation is that increased stiffness of the basilar membrane, from increased pressure in MD, would result in faster movement of sound along the basilar membrane, and therefore shorter latencies (Don et al., 2005).

CHAMP is performed by calculating the difference between latency and amplitude for wave Vs at 500 Hz TB and clicks. This effectively calculates the time that sound takes to travel from the basal cochlear (click) to the apical cochlear (500 TB). The sensitivity and specificity of CHAMP latency is presented below (Figure 24). When a standardised criteria is used such as <0.3 ms (Don et al., 2005), there is a huge variability of the sensitivity and specificity of CHAMP latency measures. Fewer studies have investigated CHAMP amplitude (Figure 25), but these studies indicate that CHAMP amplitude is highly sensitive and specific for separating MD from healthy controls (Don et al., 2007; Shang et al., 2012) and patients with vestibular conditions (Lee et al., 2011). Shang and colleagues (2012) found that the sensitivity of CHAMP could be improved by combining amplitude and latency measures.

![Figure 24](image1.png) Sensitivity and specificity of CHAMP latency in response to sideways tilts for studies, when treating MD as a positive result.
An issue arises in the fact that all studies compared CHAMP in MD patients with significant hearing loss (mean pure tone average: 37.8-47.8 dB HL) to controls with mild to no hearing loss (mean pure tone average: < 10-28.3 dB HL). This is not ideal as the wave V latency and amplitude are influenced by hearing loss (Stapells & Oates, 1997; Maloff & Hood, 2014). An excluded study found that hearing loss had a significant influence on CHAMP results (Kim, Jung, Lee, Jung, and Suh, 2015). When they compared CHAMP results from patients with MD with controls who were matched for hearing loss they found that the means for latency difference were abnormal for both groups when using the <2 ms criteria proposed by Kingma and Wit (2010). Unfortunately the study did not contain sensitivity and specificity data and was therefore excluded. However the study highlights the need for more investigations of the influence of hearing loss on the diagnostic accuracy of CHAMP.

Additional problems occur with CHAMP, as 16-50% patients with MD do not have CHAMP latency responses (Ordóñez-Ordóñez et al., 2009; De Valck et al., 2007; Shang et al., 2012). This is usually due to an absent wave V at 500 Hz (De Valck et al., 2007). For instance CHAMP is not recordable in patients with ABR thresholds over 60 dB nHL as the wave V is not present in the ABR trace (Shang et al., 2012).
4.9 Distortion product otoacoustic emissions (DPOAE) phase shift

Two studies using DPOAE phase shift met the inclusion criteria for the scoping review (Avan, Giraudet, Chauveau, Gilain, & Mom, 2011; Gerenton et al., 2015). DPOAE phase in response to a body tilt enabled separation between patients with definite MD, who had experienced an attack in the past week, and healthy controls with high sensitivity and specificity (Figure 26). However, only 9-28% of patients with MD had abnormal results when tested in the interictal period. This indicates that DPOAE phase shift may be a sensitive test for acute MD. However all studies looking at DPOAE phase shift have been conducted by one lab. It will be important for these findings to be replicated by other labs.

![Figure 26](image)

Figure 26. Sensitivity and specificity of DPOAE phase shift when a positive result is considered to indicate MD.

4.10 oVEMP and cVEMP 0.5/1 kHz frequency amplitude ratio (FAR)

The oVEMP and cVEMP 0.5/1 kHz FAR met the inclusion criteria. Three studies used cVEMP 0.5/1 kHz FARs (Kim-Lee, Ahn, Kim, & Yoon, 2009; Murofushi et al., 2009; Taylor et al., 2012); (Figure 27). The specificity of the cVEMP FAR was consistently high (90-100%). However the sensitivity was inconsistent between studies, with one study showing 93% sensitivity (Kim-Lee et al., 2009), while other studies showed much lower sensitivities of 35-56% (Murofushi et al., 2009; Taylor et al., 2012). No clear methodological
reason for this discrepancy was identified, although several differences existed between the studies, such as stimulus properties and patient selection. Overall these findings support that cVEMP 0.5/1 kHz FAR has some diagnostic value for separating MD from healthy controls, but further studies are needed to determine its sensitivity for MD.

<table>
<thead>
<tr>
<th>Study</th>
<th>Control group</th>
<th>MD Group</th>
<th>Cut-off (%)</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim-Lee et al., 2009</td>
<td>Healthy controls</td>
<td>Def / Prob</td>
<td>≤ -17.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murcufshi et al., 2009</td>
<td>Healthy controls</td>
<td>Unilateral Def</td>
<td>≤ -19.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor et al., 2012</td>
<td>Healthy controls</td>
<td>Unilateral Def</td>
<td>≤ -20.4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 27. Sensitivity and specificity of 0.5/1 kHz cVEMP FAR when a positive result is considered to indicate MD.*

One study found that oVEMP 0.5/1 and 0.5/0.75 kHz FARs separated definite unilateral MD from healthy controls with 88.9% sensitivity and 100% specificity (Singh & Barman, 2016).

Another study used a combination of cVEMP and oVEMP 0.5/1 kHz FAR and cVEMP asymmetry ratio (Maxwell, Jerin, & Gürkov, 2016). They found that the combination of the techniques only had had 64% sensitivity and 93% specificity. This suggests that the high sensitivities seen in studies using the individual measures (Kim-Lee et al., 2009; Singh & Barman, 2016) may not be representative.
5.0 Discussion
5.1 Summary of findings from the Systematic Review

Several diagnostic tests have been compared between patients with MD and VM, and show some potential for assisting the differential diagnosis of the two conditions. These tests included VIN, rotatory chair testing, vHIT, cVEMP asymmetry, 0.5/1 kHz cVEMP FAR, and LMPT. However these diagnostic tests have only been compared in one or two studies and so meta-analysis of these tests was not considered to be appropriate.

There was a moderate level of evidence supporting that caloric testing will correctly diagnose MD in 55% of patients, while misdiagnosing approximately 21% of patients with VM as MD (54.6% sensitivity and 78.9% specificity). This indicates that the caloric test has some diagnostic value in the differential diagnosis of MD and VM. However a caloric result is insufficient to either confirm or exclude a diagnosis of MD or VM.

5.2 Sources of Bias

Notably all of the studies in the systematic review component of this thesis had multiple sources of bias. This is consistent with a previous study that identified that most diagnostic accuracy studies have at least one methodological flaw (Reid, Lachs, & Feinstein, 1995).

One of the primary sources of bias was due to the use of the “probable” and “possible” category of diagnosis as a reference test. Diagnostic accuracy studies operate on the assumption that the reference test is 100% accurate at diagnosing a condition, and that differences between the index and reference test are due to errors in the index test (Van
Rijkom & Verdonchot, 1995; Biesheuvel, Irwig, & Bossuyt, 2007). A “probable” or “possible” diagnosis is not highly accurate. This creates a problem when using these criteria to evaluate diagnostic tests.

Misclassification due to an inappropriate reference test (non-definite diagnosis) results in errors. In the diagnostic accuracy paradigm these errors are assumed to be due to the index test, resulting in an underestimation of the accuracy of the index test (Biesheuvel et al., 2007). Because of this, studies should ideally use pure groups of patients with definite MD and VM when evaluating diagnostic tests. Probable and possible diagnoses have value also, as these indicate earlier stages of the illness. However the data for these groups should be separate from the definite group, so as to minimise misclassification of the reference test.

Lack of blinding was another systemic problem. Only one study used blinding (van Tilburg et al., 2016). This is understandable as most studies used the index test to assist diagnosis in a clinical setting. However it introduces significant bias as a patient’s diagnosis or predicted diagnosis of MD or VM (reference test) may influence how a clinician conducts or interprets the index test. Alternatively the results of an index test may influence the diagnosis of MD or VM. For instance Bremova and colleagues (2016) excluded patients with VM if they had a positive caloric result. While other studies did not state doing this, it is very likely that a positive caloric result could lead a clinician to avoid making a diagnosis of VM. If either of these biases is present, the results of the index test and reference test are not truly independent. Because of this, blinding is an important factor in the quality of diagnostic accuracy studies.
5.3 Summary of findings from the Scoping Review

Several tests with high accuracy for MD were identified. The most well-established tests with high sensitivity and specificity were gadolinium MRI and transtympanic TB ECoG. TB ECoG is believed to detect hydrops of the cochlea, whereas gadolinium MRI is believed to be able to detect hydrops of both the cochlea and the vestibule. It is therefore understandable that these techniques were found to separate patients with MD from patients with hearing loss or other vestibular conditions. Both tests have limitations in that they require invasive procedures (transtympanic electrode or transtympanic or intravenous injection of gadolinium). Gadolinium MRI is further limited due to the cost of obtaining an MRI. MRI is standard procedure for unilateral/asymmetrical hearing loss (to exclude vestibular schwannoma). However this cost may not be justifiable in the earlier stages of MD where asymmetrical hearing loss may not be present.

The CHAMP and Click ECoG were also identified as sensitive and specific for MD. However Click ECoG is influenced by hearing levels, and does not sufficiently separate patients with MD from controls when matched for hearing loss (Gibson, 2009). Similar concerns have been raised regarding the diagnostic accuracy of CHAMP (Kim et al., 2015). No CHAMP studies that met the inclusion criteria for the scoping review matched controls for hearing loss. Studies evaluating the sensitivity and specificity of CHAMP under these conditions are needed before it is accepted as a diagnostic test for MD.

Other promising techniques for the diagnosis of MD included EVestG, VIN, 0.5/1 kHz cVEMP/oVEMP FAR, and DPOAE phase shift. However more studies are needed before these techniques are established as diagnostic tests for MD. Additionally VIN and the 0.5/1 kHz cVEMP FAR were compared between patients with MD and VM, and both
indicated some degree of separation of the two conditions, supporting that these techniques may be valuable in MD diagnostics. These techniques are also all non-invasive.

While DPOAE phase shift could only detect acute MD (within one week of an attack), it is still promising. This is because it is a non-invasive procedure and would be able to be performed by a trained individual in either private or public settings. If DPOAE phase shift is confirmed to be highly sensitive and specific by future studies, it would be possible to have patients tested close to an attack, as a specialist would not need to be involved. This would make seeing a patient close to attack more viable.

It is important to note that studies using a control group of healthy individuals tend to over-estimate diagnostic accuracy (Lijmer et al., 1999; Whiting et al., 2013). This is because we do not typically use diagnostic tests on healthy individuals. Studies evaluating diagnostic tools for MD ideally need to be evaluated against hearing loss matched individuals and patients with other vestibular conditions, as these are the patients that will likely undergo these diagnostic tests.

5.4 Inconsistencies between Gadolinium MRI and Histology

One interesting finding was that typically gadolinium MRI identifies hydrops of the vestibule more commonly than cochlear hydrops (Fiorino et al., 2011; Homann et al., 2015; Ito et al., 2016). This may be contrasted with the recent meta-analysis of endolymphatic hydrops in temporal bones which indicated that saccular and utricular hydrops did not occur in the absence of cochlear hydrops (Pender, 2014).
This may be for several reasons. Firstly the visualisation of the cochlea is more difficult with gadolinium MRI. This is particularly true with the apical cochlea (Tanigawa et al., 2011), which is responsible for low frequency hearing. This is problematic because the apical cochlea is believed to be the most affected part of the cochlea in MD.

Secondly it may be due to a high rate of false positive results in gadolinium MRI (detection of hydrops in the vestibule when there is none). This is also possible as the current definition of mild hydrops in the vestibule is >33% which was the mean for normal patients (Nakashima et al., 2009). By this definition, approximately 50% of healthy controls are expected to have “mild hydrops” of the vestibule.

A third reason is that the criteria for hydrops in histological studies are typically qualitative. (For examples of qualitative criteria for endolymphatic hydrops see Rauch, Merchant, & Thedinger, 1989 or Lin et al., 2006). A recent temporal bone study constructed 3D images of temporal bones (Morita et al., 2009). They found that saccular hydrops was 100% sensitive in temporal bones of patients with MD (more sensitive than cochlear hydrops) when using the 95% CI of healthy controls as a cut-off value. It is possible that quantitative methods, that are now being utilised (Morita et al., 2009; Nakashima et al., 2009), are more sensitive for detecting endolymphatic hydrops of the vestibule. If this is true, a significant re-evaluation of temporal bone studies for endolymphatic hydrops needs to be conducted.

Lastly this may due to processes that occur in active disease, that are not observable in temporal bone studies. Gadolinium MRI is the first technique that has enabled in vivo measurement of vestibular hydrops, and therefore these inconsistencies may be simply due to processes that are unique to active disease.
5.5 Lack of Diagnostic Accuracy Data

For studies systematic review using diagnostic tests with continuous variables (such as VEMPs), only a minority of studies (Taylor et al., 2012; Bremova et al., 2016) produced cut-off values and data to enable the calculation of sensitivity/specificity. This was also encountered in the scoping review. It appeared to be particularly common in studies where a significant difference was not found, but also occurred in many studies where findings were significantly different. This creates an issue as meta-analyses of diagnostic accuracy cannot account for these studies because they do not contain the necessary metrics. This effect would be expected to create a source of bias in meta-analyses by making diagnostic tests appear more accurate than they actually are. It is therefore important that if a diagnostic test has poor sensitivity and specificity in a study, that those aspects are reported. Good options for selecting a cut-off value exist in the form of a 95% CI of the control group, or Youden’s J statistic (Youden, 1950). Alternatively a receiver operating characteristic could be used.

5.6 Limitations of the Present Study

It should be noted that diagnostic accuracy alone is not sufficient when determining the value of a diagnostic test (Fineberg, 1977). The utility of improved accuracy should also be considered. The primary purpose of health care is to improve patient outcomes. A test may have an impressive diagnostic accuracy, but if it does not improve patient outcomes its value is limited (Schünemann et al., 2008). Unfortunately studies of these outcomes, while ideal, are often impractical (Bossuyt, Lijmer, & Mol, 2000; Bossuyt, Reitsma, Linnet, & Moons,
Because of this, systematic reviews of diagnostic tests typically rely on diagnostic accuracy as a surrogate measure of patient outcome (Schünemann et al., 2008).

The current study may have missed some relevant studies. Only two databases were searched, the search was limited to studies in English. Grey literature and unpublished literature were also not considered. Additionally the scoping review used diagnostic accuracy filters. It is likely that the sensitivity of the systematic and scoping reviews for relevant articles was reduced by these factors. One article meeting the criteria for the systematic review which was identified but was not available on PubMed or Embase that compared VEMPs between patients with MD and patients with VM (Utkur, Durankaya, Idiman, Serbetcioglu, & Guneri, 2013). Additionally, one significant paper was identified which was missed by the diagnostic accuracy filter, despite being available through PubMed (Pyykkö et al., 2013). While the study did not meet the inclusion criteria for the scoping review, it appears to be the largest published paper to date looking at gadolinium MRI and its sensitivity for MD. While these studies were identified effectively through a grey literature search (not formally conducted in this study), it highlights the facts that the present study is likely to have missed some relevant studies due to methodological limitations.

The quality of studies was not assessed in the scoping review, and no meta-analysis was performed. While this limits conclusions from the scoping review, the intent of the scoping review was to briefly overview the literature, rather than to be a formal systematic review. It is likely that several of the measures identified in the scoping review would benefit from formal systematic reviews.
5.7 Recommendations for Future Research

Gadolinium MRI and transtympanic TB ECoG were identified as highly sensitive and specific tests for MD. However neither test has been properly evaluated for the differential diagnosis of patients meeting the criteria for MD or VM (with the exception of Nakada and colleagues [2014] who compared patients with VM and auditory symptoms with patients with vestibular MD). For this reason studies are needed to evaluate the efficacy of these diagnostic tests for separating the two conditions. Because they are not proven to separate the two conditions yet (particularly because patients with VM may actually have endolymphatic hydrops) any use of these techniques for this purpose should be performed cautiously.

The value of diagnostic tests is not limited to diagnosis. Our current understanding of the endolymphatic hydrops in patients is primarily based on symptomology and cross-sectional studies of post mortem histology. Gadolinium MRI and TB ECoG provide important tools to be able to monitor endolymphatic hydrops in a living patient, and longitudinal studies may give us important insights into the progression of endolymphatic hydrops. For instance a recent study has shown a reduction of endolymphatic space following acetazolamide treatment (a carbonic anhydrase inhibitor diuretic) in patients with MD. The reduction in endolymphatic space was correlated with an improvement of symptoms (Sepahdari, Vorasubin, Ishiyama, & Ishiyama, 2016).

CHAMP was identified as an accurate test for MD. However future studies should attempt to compare CHAMP results between patients with MD and hearing loss matched patients. This will be needed to determine whether the test can separate MD from patients with patients that are matched for hearing loss.
To determine whether there is an overlapping pathophysiology between MD and VM, post-mortem temporal bone studies of patients with definite VM will be needed to establish the incidence of endolymphatic hydrops in these patients (and if possible) any association with vestibular and auditory symptoms. It is understandable that no studies of this nature were identified as VM is a relatively new entity. Because of this it may take some time before temporal bones in these patients become available. However studies of this nature will be critical to understanding the potential pathophysiological overlap of MD and VM.

5.8 Recommendations for Clinical Practice

The differential diagnosis of MD and VM should continue to be based on the current diagnostic criteria (IHS & others, 2013; Lopez-Escamez et al., 2015). There was not sufficient evidence supporting any measure that could separate the two conditions adequately. A moderate level of evidence supported some diagnostic value of caloric testing with an estimated sensitivity of 54.6% and specificity of 78.9%. However caloric testing neither confirms neither the presence nor absence of either condition.

Problematically both conditions rely on a diagnosis of exclusion. A reasonable approach in the absence of an objective method of differentiation may be to exclude a diagnosis of definite MD before attempting a diagnosis of VM. This is supported by the fact that systematic reviews show that endolymphatic hydrops occurs in 100% of patients with a diagnosis of MD, indicating that a diagnosis of MD is almost certain to have endolymphatic hydrops (Foster & Breeze, 2013; Pender, 2014). However, as it is not known with confidence whether endolymphatic hydrops occurs in VM, patients meeting the diagnostic criteria for both conditions may represent an overlap of the two conditions.
While not directly assessed in the systematic or scoping review, the literature (cited in the introduction) supports that spontaneous nystagmus is essential in the acute phase of MD but not VM. Confirmation of nystagmus can be incredibly difficult due to the transient nature of attacks. Our lab has had some success with requesting that patients record their eyes using their phones during attacks. This has provided definitive proof of spontaneous nystagmus on multiple occasions, and we suggest this as a valuable strategy for both specialists and supporting clinicians.

5.9 Concluding statements

The differential diagnosis of MD and VM continues to be troublesome. There was a moderate level of evidence that caloric hypofunction is more common in MD. However caloric testing does not reliably confirm or exclude either diagnosis. Temporal bone studies are needed to determine whether endolymphatic hydrops occurs in VM. Additionally studies are needed to examine whether gadolinium MRI and TB ECoG can differentiate MD from VM, as these tests have been demonstrated to be sensitive and specific for MD. Several other promising techniques were identified but there was currently insufficient evidence for these. Further research is needed to establish the value of these tests.

5.10 Funding

This project did not receive funding, and the author has no conflict of interest.
6.0 References


https://doi.org/10.1097/MAO.0b013e31820e7614

https://doi.org/10.1097/MAO.0b013e31829e83df

https://doi.org/http://dx.doi.org.ezproxy.canterbury.ac.nz/10.1016/S1474-4422(13)70107-8


https://doi.org/10.1016/j.heares.2015.07.016


https://doi.org/10.1080/00016480902729843


https://doi.org/http://dx.doi.org.ezproxy.canterbury.ac.nz/10.1136/jnnp.2009.176602


**Otology and Neurotology, 36**(6), 1109–1114.
https://doi.org/10.1097/MAO.0000000000000782

https://doi.org/10.3109/00016480903382808


https://doi.org/10.1097/01.mlg.0000227976.43352.61


https://doi.org/10.1097/MAO.0000000000000688


https://doi.org/10.1111/j.1468-2982.2009.01860.x


https://doi.org/10.3109/1651386X.2011.607248

https://doi.org/http://dx.doi.org.ezproxy.canterbury.ac.nz/10.1007/s11604-014-0290-4


https://doi.org/10.1080/00016480902729827


https://doi.org/http://dx.doi.org.ezproxy.canterbury.ac.nz/10.1016/j.clinimag.2014.09.014


APPENDIX

Papers that could not be accessed:


