

Vergleich Kopfschütteltest und Drehstuhltest bei unilateralen peripheren vestibulären Defiziten

Ein Review mit Praxisempfehlung

Bachelor-Thesis

Selina Fragner

Matrikel-Nr. 21-259-239

Samira Schwaller

Matrikel-Nr. 21-258-975

Berner Fachhochschule, Departement Gesundheit

Bachelor of Science Physiotherapie, 2021

Referentin

Dörte Watzek, Dr. Prof.

Co-Referent

Stefan Schädler, Dipl. Physiotherapeut

FH

Auftraggeber

Berner Fachhochschule

Bern, 30. August 2024

INHALTSVERZEICHNIS

Glossar	4
Danksagung	5
Abstract	6
1 Einleitung	7
2 Theoretische Grundlagen	9
2.1 Das Gleichgewichtsorgan	9
2.1.1 Die Bogengänge	9
2.1.2 Push-pull-Prinzip	10
2.1.3 Vestibulo-okulärer Reflex (VOR)	11
2.1.4 velocity storage	12
2.2 Periphere unilaterale Vestibulopathie	12
2.2.1 Neuritis vestibularis	13
2.2.2 Morbus Menière	14
2.2.3 Nystagmus	14
2.3 Diagnostische Testungen	15
2.3.1 Kopfimpulstest	15
2.3.2 Video-Kopfimpulstest	16
2.3.3 Kalorische Prüfung	16
2.3.4 Dynamic Visual Acuity Test	17
2.3.5 Kopfschütteltest	17
2.3.6 Drehstuhltestung	18
2.3.7 Andere Testungen	23
2.4 Gütekriterien	23
3 Zielsetzung und Fragestellung	26
4 Methodik	27
4.1 Design	27
4.2 Ein- und Ausschlusskriterien	27
4.3 Suchstrategie	28
4.4 Auswahl und Analyse der Studien	29
5 Ergebnisse	30
5.1 Literatursuche und -auswahl	30
5.2 Methodische Qualität der eingeschlossenen Studien	32
5.2.1 Bewertung der Studien nach COSMIN	32
5.3 Zusammenfassung der Studien	34

5.3.1	Kopfschütteltest	34
5.3.2	Drehstuhltest	39
5.4	Gütekriterien der Studien	44
5.4.1	Gütekriterien Kopfschütteltest	44
5.4.2	Gütekriterien Drehstuhltest	48
6	Diskussion	51
6.1	Stärken dieser Arbeit	57
6.2	Limitationen	57
7	Schlussfolgerung	58
8	Praxisempfehlung	59
9	Literaturverzeichnis	60
10	Abbildungsverzeichnis	69
11	Tabellenverzeichnis	70
12	Abkürzungsverzeichnis	71
13	Anhänge	72
13.1	Bewertung der eingeschlossenen Primärstudien anhand der «COSMIN checklist»	72
13.2	Ausführliche Zusammenfassungen der Studien	100
13.2.1	Kopfschütteltest	100
13.2.2	Drehstuhltest	106
13.2.3	Gütekriterien und Ergebnisse Kopfschütteltest	113
13.2.4	Gütekriterien und Ergebnisse Drehstuhltest	120
13.3	Selbständigkeitserklärung	127

GLOSSAR

Asymmetry

Die Spitzenreaktionen der Augen bei Drehungen im Uhrzeigersinn werden mit denen gegen den Uhrzeigersinn verglichen. Das Verhältnis dieser beiden Reaktionen wird als Symmetry bzw. Asymmetry des VOR definiert (Zalewski et al., 2019).

Elektronystagmographie (ENG)

Misst die Augenbewegungen mittels Elektroden durch Aufzeichnung von elektrischen Potenzialen zwischen der Horn- und Netzhaut (Bojrab et al., 2019).

Elektrookulographie (EOG)

Rund ums Auge werden Elektroden angebracht, welche den horizontalen Nystagmus aufzeichnen (Eza-Nuñez et al., 2016).

Ewald'sche Gesetze

1. Gesetz: Durch eine Endolymphströmung wird ein Nystagmus in der Ebene des stimulierten Bogengangs ausgelöst. Die langsame Bewegung der Augen zeigt die Stromrichtung der Flüssigkeit an. 2. Gesetz: Die Stimulation («pull») eines Bogengangs wirkt stärker als dessen Hemmung («push»). 3. Gesetz: Im anterioren und posterioren Bogengang ist die ampullofugale (weg von der Ampulle) Strömung wirksamer als die ampullopetale (in Richtung der Ampulle). Beim horizontalen Bogengang ist es umgekehrt (Reiß & Reiß, 2015).

Frenzelbrille

Reduziert die visuelle Fixation und bietet durch ihre Vergrößerungsgläser eine bessere Untersuchung der Augenbewegungen (Strupp et al., 2014)

Frequenz

Definiert die Anzahl der Hin- und Herbewegungen in einem bestimmten Zeitraum (meist eine Sekunde) und wird in Hertz (Hz) angegeben (Zalewski, 2018).

Gain

Definiert das Verhältnis zwischen den Maxima der Augen- und Drehgeschwindigkeit des Drehstuhls (Zalewski et al., 2019).

Per-rotatorischer Nystagmus

Nystagmus, welcher während der Drehung zu beobachten ist (Ahmed et al., 2009).

Phase

Zeitliche Differenz zwischen dem Start der Rotation des Drehstuhls und dem Beginn der Augenreaktion (Zalewski et al., 2019).

Post-rotatorischer Nystagmus

Nystagmus, welcher nach der Drehung zu beobachten ist (Ahmed et al., 2009).

Pseudorandom Rotation Test

Drehstuhlprotokoll, welches randomisierte Rotationen beinhaltet (Maes et al., 2011).

Skew Deviation

Die Fehlstellung eines Auges aufgrund einer okulomotorischen Störung. Bsp.: Ein Auge ist in Abduktion bzw. Adduktion oder höhergestellt als das andere. Mögliches Anzeichen für eine zentrale Störung (Schädler et al., 2022).

VEMP

Ein vestibuläres Laborverfahren zur Identifizierung von unilateralen Pathologien und Untersuchung der Funktion der Otholithenorgane. Es sind ipsilaterale hemmende myogene Potenziale, die durch akustische Stimuli hoher Intensität, von grossen Oberflächen Elektroden über dem kontrahierten M. sternocleidomastoideus (SCM) abgeleitet werden (Maes et al., 2011).

Videookulographie (VOG)

Anhand einer Infrarotkamera werden die Augenbewegungen bzw. Okulomotorik aufgezeichnet und untersucht (Teiwes et al., 1991).

Videonystagmographie (VNG)

Die VNG verwendet Infrarotkameras mit speziellen Brillen, um die Augenbewegungen bzw. die Pupillenaktivität genau aufzuzeichnen (Bojrab et al., 2019).

DANKSAGUNG

Wir danken unserer Referentin Dörte Watzek, die uns in unserem Prozess umfangreich unterstützt hat und uns bei Fragen immer zur Verfügung stand. Ebenso danken wir für die fachliche Kompetenz und die konstruktiven Rückmeldungen.

Wir danken unserem Co-Referenten Stefan Schädler für die tatkräftige Unterstützung, Beantwortung von Fragen, regelmässigen Austauschmöglichkeiten und Gedankenanstösse, welche uns zu neuen Ideen verholpen haben und wegweisend in unserem Prozess wirkte.

Ebenfalls geht ein Dank an Raphaëlle Genolet, die uns einen Einblick in die Welt der Drehstuhltestung in ihrer Praxis in Sierre ermöglichte.

Ein besonderer Dank geht an unsere Partner, Eltern, Freundinnen und Freunden für das Dasein und die Unterstützung während des Schreibprozesses.

ABSTRACT

Einleitung: Die unilateralen peripheren vestibulären Defizite gehören zu den häufigsten Erkrankungen des Vestibularsystems, weshalb eine frühzeitige und präzise Diagnose entscheidend ist. Die Bachelorarbeit untersucht die diagnostische Qualität des Kopfschüttel- und Drehstuhltests zur Identifizierung solcher Defizite, mit dem Fokus auf die Gütekriterien Reliabilität und Validität. Diese Arbeit schliesst eine Forschungslücke, da bisher keine systematischen Übersichtsarbeiten zu diesen Tests existieren.

Fragestellung: Welche der beiden Testungen, Kopfschüttel- (HST) oder Drehstuhltest (ROT), zeigt die besseren diagnostischen Qualitäten bezüglich Reliabilität und Validität bei Patientinnen und Patienten ab zehn Jahren mit unilateralen peripheren vestibulären Defiziten auf?

Methodik: Diese Thesis verfolgt den Ansatz einer systematischen Literaturanalyse, um relevante Forschungsergebnisse strukturiert zu identifizieren und zu bewerten. Die Beurteilung der methodischen Qualitäten der Primärstudien erfolgte anhand der «COSMIN checklist». Die Gütekriterien wurden zusammengetragen und evaluiert.

Ergebnisse: Nach Ausschluss durch die «PRISMA» Methode wurden 14 Primärstudien eingeschlossen. Die Analyse dieser Studien ergab eine Durchschnittsbewertung der methodologischen Qualität zwischen «doubtful» bis «adequate». Es wurden nur wenig Daten zu Validität und Reliabilität gefunden. Der Kopfschütteltest zeigt insgesamt eine höhere Spezifität, der Drehstuhltest eine höhere Sensitivität. Beide Tests allein sind nicht ausreichend zur Diagnose von unilateralen peripheren vestibulären Defiziten.

Diskussion: Die beiden Testungen zeigen komplementäre Stärken. Sie lassen sich jedoch nicht vergleichen, da viele Unterschiede wie Testkriterien, Testdurchführung, Parameter und Materialien vorliegen. Zudem limitierten die geringe Datenlage und Qualität der Studien die Aussagekraft der Ergebnisse.

Schlussfolgerung: Im Vergleich weist das Drehstuhlprotokoll Sinusoidal Harmonic Acceleration insgesamt die besseren diagnostischen Qualitäten bei der Identifizierung eines unilateralen peripheren vestibulären Defizits auf als der Kopfschütteltest, der nicht valide zu sein scheint. Die diagnostischen Qualitäten wie Reliabilität, Validität und Objektivität sind jedoch aufgrund der geringen Datenlage schwierig einzuschätzen.

Key words: head shaking test, head shaking nystagmus, rotary chair test, sinusoidal harmonic acceleration, velocity step test, unilateral peripheral vestibular deficit, unilateral vestibular hypofunction, diagnostic criteria

1 EINLEITUNG

Die akute unilaterale periphere vestibuläre Funktionsstörung hat eine jährliche Inzidenz von 3.5-15.5 auf 100'000 Personen (Adamec et al., 2015) und gehört somit zu den häufigsten Erkrankungen im vestibulären System (Strupp, Brandt, et al., 2022). Periphere vestibuläre Störungen beziehen sich auf Krankheiten, welche das Gleichgewichtsorgan im Innenohr (Labyrinth) oder den vestibulären Anteil des N. vestibulocochlearis betreffen (Schädler et al., 2022). Das vestibuläre System setzt sich aus vier Efferenzen zusammen. Erstens die Okulomotorik, diese beinhaltet den Vestibulo-okulären Reflex (VOR). Die Funktion des VOR liegt in der Blickstabilisierung während Kopfbewegungen, wobei sich die Augen in die entgegengesetzte Richtung zur Kopfdrehung bewegen. Eine Störung des VOR resultiert in Augenbewegungsstörungen oder Nystagmen. Die zweite Efferenz wird über die Vestibulo-spinalen Bahnen weitergeleitet, welche für den Erhalt des Gleichgewichts zuständig sind. Drittens gelangen Efferenzen in den vestibulären Kortex und das Kleinhirn, welche verantwortlich für die Wahrnehmung der Position im Raum sowie Bewegungen des Kopfes sind. Komplettiert wird die Zusammensetzung mit der vierten Efferenz, dem vegetativen Zentrum. Bei einer Störung des vestibulären, visuellen oder somatosensorischen Systems kann demzufolge Schwindel als Symptom auftreten (Bronstein et al., 2015). Es werden widersprüchliche Informationen empfangen, die nicht eingeordnet werden können. Darauf reagiert das System mit Drehschwindel und Nystagmus, welche sich bei Kopf- und Körperbewegungen verstärken (Reiß & Reiß, 2015). Dies kann sich auch in Kombination mit Angst oder Übelkeit äussern (Bronstein et al., 2015). Schwindel geht oft mit Depressivität einher und Betroffene berichten von signifikant schlechterer gesundheitsbezogener Lebensqualität (Sanftenberg et al., 2021). Die Bestimmung eines unilateralen peripheren vestibulären Defizits kann jedoch nicht ausschliesslich anhand der Beschwerden erfolgen, denn die Diagnose kann mit unterschiedlichen Symptomen assoziiert sein (Strupp, Brandt, et al., 2022).

Eine zuverlässige Erfassung der Untersuchungsergebnisse und deren Dokumentation basierend auf verlässlichen Messwerten ist für Physiotherapeutinnen und Physiotherapeuten wesentlich (Mangold, 2013). In der Praxis können zur Diagnosestellung der unilateralen peripheren vestibulären Läsionen und Erfassung der Therapieziele verschiedene Testungen wie der (Video-) Kopfpulstest (KIT), die kalorische Prüfung oder die Kopfschütteltestung durchgeführt werden (Schädler et al., 2022). Zusätzlich kann eine Lagerungsprüfung oder eine Rotationsprüfung mit dem Drehstuhl angewendet werden (Schmäl, 2020). Es scheint jedoch unklar zu sein, wie die diagnostischen Qualitäten in Bezug auf die Erkennung der unilateralen peripheren

vestibulären Funktionsstörungen aussehen. Kenntnisse über die Verlässlichkeit und Genauigkeit einer Testung erleichtern die Auswahl einer bestimmten Testung für Praktikerinnen und Praktiker im Alltag.

Dieses systematische Review wird sich auf die aktive und passive Ausführung des Kopfschütteltests und den Drehstuhltest beschränken. Beide Assessments können bei der Befunderhebung von horizontalen Nystagmen angewendet werden (Helmchen, 2021). Im physiotherapeutischen Denkprozess werden Hypothesen durch diagnostische Tests überprüft. Die Qualität der Testungen bzw. Assessments werden mittels verschiedener Gütekriterien wie die Reliabilität, Validität, Objektivität, Responsivität, Praktikabilität und klinische Relevanz bewertet. Zur Interpretation ist es wichtig, dass die testende Person die diagnostischen Kriterien kennt (Wirz et al., 2014).

Der aktuelle Forschungsstand zeigt, dass unterschiedliche Studien einzelne Gütekriterien zu den beiden Testungen evaluieren. Es existiert bis anhin jedoch keine Übersichtsarbeit, welche den Kopfschütteltest als auch den Drehstuhltest und deren diagnostischen Kriterien zusammenfasst, vergleicht und eine Praxisempfehlung abgibt. Die vorliegende Arbeit legt den Fokus auf den Gütekriterien Reliabilität und Validität bzw. Sensitivität und Spezifität. Eine Übersicht zu den Gütekriterien der beiden Assessments würde eine Beurteilung der Anwendbarkeit und Interpretation der Testungen in der Praxis erleichtern.

Diese systematische Übersichtsarbeit ist wie folgt aufgebaut. Zuerst werden die theoretischen Grundlagen für das Grundverständnis dieser Arbeit dargestellt. Daraus wird die Fragestellung genauer erläutert. Es folgt die Methodik der Literatursuche und der Ergebnisteil. Aus diesen Ergebnissen wird die Diskussion aufgebaut, verschiedene Aspekte beleuchtet und abschliessend mit der Schlussfolgerung und einer Praxisempfehlung, Kapitel 8 abgerundet.

2 THEORETISCHE GRUNDLAGEN

Nachfolgend werden die Grundlagen der Anatomie und Physiologie des Gleichgewichtsorgans, der unilateralen peripheren Vestibulopathie sowie der dazugehörigen diagnostischen Testungen beschrieben. Anschliessend erfolgt eine allgemeine Erklärung der Gütekriterien. Diese Erläuterungen dienen als Grundverständnis für die vorliegende Arbeit.

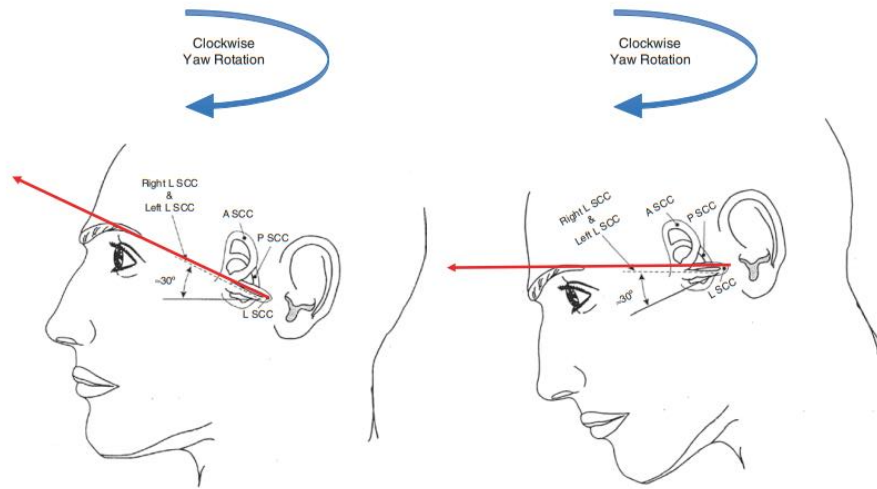
2.1 Das Gleichgewichtsorgan

Im Innenohr befindet sich das Gleichgewichtsorgan, das sich aus dem zentralen Vestibulum mit Utriculus und Sacculus zusammensetzt. Vorne medial ist es mit der Cochlea verbunden und hinten lateral mit den drei Bogengängen (Mackowetzky et al., 2021). Diese Teile sind in einem knöchernen Labyrinth im Felsenbein eingebettet (Reiß & Reiß, 2015). Das vestibuläre System hat komplexe Funktionen, welche sich nicht nur auf das Gleichgewicht beschränken, sondern auch eine wichtige Rolle in Bezug auf die Gehirnfunktion einnehmen. Ausserdem werden dadurch Eigen- und Fremdbewegungen wahrgenommen und die Orientierung sowie Navigation im Raum sind gewährleistet (Day & Fitzpatrick, 2005). Es koordiniert unsere Augenbewegungen, reguliert die Blickstabilisierung, gibt Informationen über lineare und anguläre Beschleunigungskräfte und ist verantwortlich für den Reflexonus sowie für die daraus folgende Stabilisierung der Körperhaltung (Reiß & Reiß, 2015).

2.1.1 Die Bogengänge

Auf beiden Seiten im Innenohr befinden sich drei Bogengänge. Diese werden in die beiden vertikalen Bogengänge, den anterioren und posterioren sowie in den lateralen Bogengang, welcher horizontal verläuft, unterteilt (Walther et al., 2012b). Der laterale Bogengang liegt bei aufrechter Körperhaltung nicht waagrecht, sondern steht nach oben. Erst bei einer Flexion der Halswirbelsäule (HWS) um 30° liegt dieser horizontal (siehe Abbildung 1). Dies muss bei der Ausgangsstellung des Kopfschütteltests oder der Kalorik berücksichtigt werden (Reiß & Reiß, 2015). Die Kanäle melden Rotationsbeschleunigungen des Kopfes in den drei Ebenen (Walther et al., 2012a). Der laterale Bogengang wird bei horizontalen Kopfbewegungen stimuliert. Die vertikalen Bogengänge sind so angeordnet, dass der linke vordere und der rechte hintere Bogengang jeweils in einer Ebene liegen. Diese werden als RALP-Ebene, («right anterior left posterior») und LARP-Ebene («left anterior right posterior») bezeichnet (Walther et al., 2012b).

Abbildung 1: 30° Kopfфлекion für die optimale Lage der horizontalen Bogengänge
(Zalewski et al., 2019, S.78)



Mit fünf Öffnungen münden die Bogengänge in den Utriculus (Walther et al., 2012b). Im Verlauf der Kanäle bilden sich die Ampullen, in denen sich die Cupula, eine gallertartige Masse, befindet. Im Inneren der Cupula sind in Reihen angeordnete mechanorezeptive Haar- und Stützzellen, welche das vestibuläre Sinnesorgan darstellen (Schädler et al., 2022). Die Mechanorezeptoren, sogenannte Stereozilien, welche die Haarzellen bedecken, werden bei Rotationsbeschleunigungen aktiviert (Mackowetzky et al., 2021). Wird der Kopf gedreht, bewegt sich das Innenohr mit. Die Endolymphe, die kaliumreiche Flüssigkeit im Ohr, bleibt dabei inaktiv (Mackowetzky et al., 2021). Dadurch wirkt eine Kraft auf die Cupula, die Stereozilien werden abgelenkt, erfassen die Rotationsgeschwindigkeit, das Aktionspotenzial wird verändert und an das ZNS weitergeleitet (Schädler et al., 2022). Bei einer Kopfdrehung nach rechts, wird die rechte Seite aktiviert und bei einer Linksdrehung, die linke Seite (Mackowetzky et al., 2021).

2.1.2 Push-pull-Prinzip

Ohne Kopfbewegung erzeugen die Haarzellen auf beiden Seiten ein gleich starkes Ruhepotenzial (Fetter, 2016). Auf eine Reizung reagieren die horizontalen Bogengänge mit dem Push-pull-Prinzip. Dieser Mechanismus gewährleistet durch gleichzeitige Inhibition («push») und Exzitation («pull») der Neuronen ein vestibuläres Gesamtsignal. Die durch Kopfbewegung verursachte Beschleunigung wird in dem einen Bogengang erhöht und auf der anderen Seite verringert (Straumann, o.J.). Die Drehrichtung wird als Richtung der Seite mit der höheren Frequenz kodiert, was zu einer langsamen Augenbewegung in die entgegengesetzte Richtung führt (Fetter, 2016). Dieser physiologische Vorgang ist eine Voraussetzung für den Vestibulo-okulären Reflex (Walther et al., 2012a). Die Geschwindigkeit und Richtung von Kopf- und

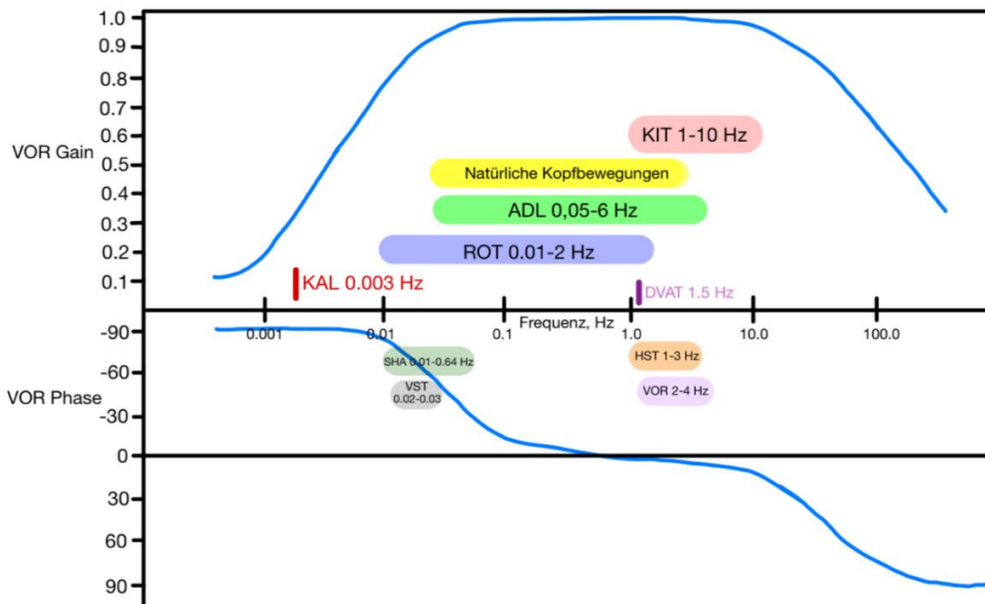
Augenbewegungen hängen eng mit den Bogengangrezeptoren zusammen. Eine Reizung des lateralen Bogengangpaars steuert die Muskeln für horizontale Augenbewegungen. Die vertikalen Push-pull-Paare die für vertikale und rotierende Bewegungen. Dies ist entscheidend für die Erkennung von Störungen der Bogengänge und dient beim Kopfschüttel- sowie beim Drehstuhltest als wichtiger Messwert zur Abgrenzung von zentralen und peripheren Defiziten (Walther et al., 2012a).

2.1.3 Vestibulo-okulärer Reflex (VOR)

Aufgrund vestibulärer Reflexe folgen auf Kopfbewegungen umgehend motorische Reaktionen. Durch den Vestibulo-okulären Reflex (VOR) ist das Auge in der Lage den Bewegungen mit einer Verzögerung von nur 7-10 ms zu folgen und somit die Blickstabilisierung zu gewährleisten. Ebenso realisiert der VOR eine abgestimmte Bewegung beider Augen (Walther et al., 2012b). Als Reaktion auf eine Kopfbewegung bewegt der VOR die Augen reflektorisch und rasch in die Gegenrichtung (Schädler et al., 2022). Durch diesen Mechanismus lösen Kopfbewegungen nach links Augenbewegungen nach rechts aus und umgekehrt (Walther et al., 2012a). Ist der VOR gestört, ist die Blickstabilisierung nicht gewährleistet und die Augen können während einer Kopfbewegung nicht auf dem Ziel gehalten werden. Darauf folgen Rückstellsakkaden. Die Augenbewegung wird somit korrigiert und das Ziel kann wieder fixiert werden. Diese Sakkade tritt meist mit einer Latenz auf und kann deshalb von der Testerin oder dem Tester wahrgenommen werden (Schädler et al., 2022). Sie sind ein Zeichen für einen gestörten VOR (Walther et al., 2012a) und können ebenfalls bei einem positiven Befund des Kopfpulsstests zu erkennen sein (siehe Kapitel 2.3.1) (Schädler et al., 2022).

Neueren Erkenntnissen zufolge, wird der VOR ab einem Frequenzbereich von 2 Hz aktiviert (Verbecque et al., 2018). Abbildung 2 stellt die Frequenzen der jeweiligen Testungen dar, auf welche in der vorliegenden Arbeit näher eingegangen wird. Zudem sind die Aktivitäten des täglichen Lebens (ADLs) und die natürlichen Kopfbewegungen abgebildet. Im Frequenzbereich der natürlichen Kopfbewegungen, zwischen 0.05 Hz und 5 Hz, kann das vestibuläre System (VOR-Gain und VOR-Phase) nahezu perfekt (linear) zusammenarbeiten. Dies verdeutlicht, dass der VOR auf die funktionellen Aktivitäten des täglichen Lebens abgestimmt ist. Liegt die Frequenz ausserhalb der natürlichen Kopfbewegungen, wird die VOR-Antwort deutlich ineffizienter. Dies kann bei der Kalorik Testung mit 0.003 Hz oder dem Kopfpulstest zwischen 1-10 Hz beobachtet werden (Zalewski et al., 2019).

Abbildung 2: Darstellung der Testungen in den jeweiligen Frequenzbereichen (Eigene Darstellung, 2024; nach Zalewski et al., 2019, S.76)



KIT=Kopfpulstest, ADL=Aktivitäten des täglichen Lebens, ROT=Drehstuhl, KAL=Kalorik, DVAT=Dynamic Visual Acuity Test, HST=Kopfschütteltest, VOR=Vestibulo-okular Reflex, SHA und VST=Protokolle des Drehstuhls

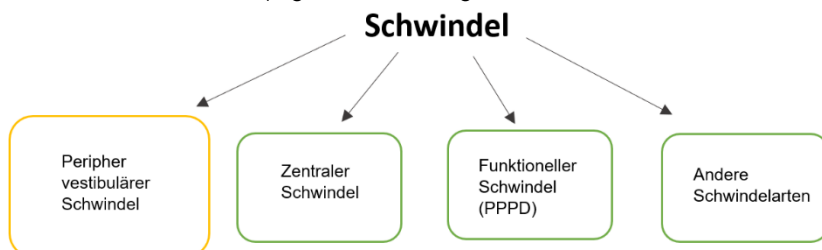
2.1.4 velocity storage

Der «velocity storage» wird als neuronaler Integrator beschrieben, welcher die Leistung des Vestibularsystems bei niedrigen Frequenzen verbessert. Ist eine Seite betroffen, kann dieser Speichermechanismus in seiner Funktion beeinträchtigt sein, was zu einem gestörten VOR im Niederfrequenzbereich führt. Bei Reizen unter 0.1 Hz kann daraus eine Phasen-Voreilung des VOR resultieren. Sind die Vestibulopathien schwergradig oder komplett einseitig, können die VOR-Phasenstörungen auch bei höheren Frequenzen über 1.0 Hz auftreten (Zalewski et al., 2019).

2.2 Periphere unilaterale Vestibulopathie

Fuhry beschreibt vier Diagnosegruppen von Schwindel (2022 in Schädler et al., 2022), welche in Abbildung 3 dargestellt sind. Der periphere vestibuläre Schwindel, der zentrale und funktionelle Schwindel sowie andere Schwindelarten.

Abbildung 3: Arten von Schwindel (Eigene Darstellung, 2024; nach Schädler et al., 2022)



Der zentrale Schwindel besteht aufgrund einer Erkrankung oder Schädigung des zentralen Nervensystems (ZNS). Dabei können Hirnstamm, Kleinhirn oder Großhirn betroffen sein. Der funktionelle Schwindel beinhaltet den «Persistent Postural Perceptual Dizziness» (PPPD) oder den phobischen Schwankschwindel. Es werden auch andere Schwindelarten wie der zervikogene oder der multifaktorielle Schwindel beschrieben, auf welche nicht genauer eingegangen wird. Der periphere vestibuläre Schwindel zeigt sich, wenn das Innenohr und/oder der N. vestibularis betroffen sind. Letzteres umfasst die akute unilaterale Vestibulopathie (AUVP) auch als Neuritis vestibularis bezeichnet, die bilaterale Vestibulopathie, der benigne paroxysmale Lagerungsschwindel (BPLS), Morbus Menière und Vestibularysparoxysmie (Strupp, Brandt, et al., 2022). Studien zeigen, dass die Kopfschüttel- und die Drehstuhltestung häufig bei Neuritis vestibularis und M. Menière angewendet werden. Diese beiden Diagnosen werden im nachfolgenden Unterkapitel genauer definiert.

2.2.1 Neuritis vestibularis

Strupp et al. (2022) berichteten, dass die AUVP nach BPLS und M. Menière die dritthäufigste Ursache für peripheren vestibulären Schwindel ist. Die Prävalenz ist bei beiden Geschlechtern gleich (Strupp, Brandt, et al., 2022). Eine Differenzierung vom akuten zentralen vestibulären Syndrom (AZVS) ist entscheidend und lässt sich durch klinische Untersuchungen und der Anamnese zentraler Zeichen gut durchführen (Strupp, Brandt, et al., 2022). Drei einfache klinische Testungen, zusammengefasst im HINTS, bestehend aus dem Kopfpulstest (HI), Blickrichtungsnystagmus (N) und Skew Deviation Test (TS), unterstützen die Differenzialdiagnose. Bei einem positivem HINTS-Test und akutem Schwindel ist eine sofortige Notfalleinweisung erforderlich (Schädler et al., 2022). Die wahrscheinlichste Ursache einer AUVP ist eine durch Reaktivierung einer latenten Herpes-simplex-1 Infektion verursachte Entzündung (Arbusow et al., 2010). Eine Entzündung des N. vestibularis superior führt zu einer Schwellung sowie Eigenkompression und weiter zu einer Ischämie des Gleichgewichtsorgans als auch des Nervs. Betroffen sind primär der horizontale und anteriore Bogengang (Schädler et al., 2022). In der akuten Phase verstärken leichte Kopf- und Körperbewegungen (niedriger Frequenzbereich) die Symptome (Schädler et al., 2022). Drehschwindelattacken können kurz anhaltend sein, klingen fortlaufend ab (Lee et al., 2009), allerdings oftmals unvollständig (Brandt et al., 2010). Durch die zentrale Kompensation, versucht der Körper die Beschwerden zu minimieren und Anpassungen vorzunehmen. In späteren Phasen treten die Symptome verstärkt bei höheren Frequenzbereichen auf, etwa beim Drehen zur betroffenen Seite (Schädler et al., 2022). Die diagnostischen Kriterien umfassen akuter anhaltender Drehschwindel >24 Stunden, Scheinbewegungen, Stand-

und Gangunsicherheiten mit Falltendenz zur betroffenen Seite (verstärkt bei geschlossenen Augen) und Übelkeit, jedoch ohne Hörstörungen (z. B. Tinnitus). Der Nystagmus ist unter Fixation unterdrückbar und der Spontannystagmus schlägt >24 Stunden zur gesunden Seite. Bei einer unilateralen VOR-Störungen wird auf einen positiven Kopfpulstest (KIT) zur betroffenen Seite und/oder positive Kalorik geachtet (Strupp, Brandt, et al., 2022). Zentrale Störungen, wie zentrale Okulomotorikstörungen, vertikaler oder Blickrichtungsnystagmus entgegen der Richtung des Spontannystagmus müssen ausgeschlossen werden (Strupp, Bisdorff, et al., 2022). Die körperliche Untersuchung umfasst die Testung von Spontannystagmus, den KIT, die Bestimmung der subjektiven visuellen Vertikalen (SVV), den Romberg-Test zur Untersuchung der Okulomotorik und des Hörvermögens, den Video-KIT (vKIT) im Hochfrequenzbereich und die Kalorik im niedrigen Frequenzbereich (Strupp, Brandt, et al., 2022).

2.2.2 Morbus Menière

Die Prävalenz des M. Menière variiert nach ethnischer Gruppe zwischen 35/100'000 in Japan (Shojaku et al., 2005) und 513/100'000 in Finnland (Havia et al., 2005). Frauen werden häufiger dokumentiert als Männer (Strupp, Brandt, et al., 2022). Betroffene erleiden zwei oder mehr Attacken von Schwindel. Die Dauer variiert von 20 Minuten bis 12 Stunden. Ebenfalls können Hörminderung (<2000 Hz, min. -30 dB), wie Tinnitus oder Ohrdruck ohne anderweitig mögliche Erklärung, in Kombination mit Schwindel von +/- 24 Stunden nachgewiesen werden (Strupp, Brandt, et al., 2022). Die körperliche Untersuchung wird meist zwischen den Attacken durchgeführt und besteht aus der Testung des Hörvermögens. Dies stellt die wichtigste Differenzierung zur Neuritis vestibularis dar. Zudem kann der KIT angewendet werden. Es gilt zu beachten, dass die Testung ein normales Resultat aufweisen kann, obwohl eine reduzierte kalorische Erregbarkeit vorhanden ist. Während einer Attacke kann zuerst ein Nystagmus zur betroffenen Seite (Reiznystagmus) und im Verlauf zur nicht-betroffenen Seite (Ausfallnystagmus) entstehen (Strupp, Brandt, et al., 2022).

2.2.3 Nystagmus

Unter einem Nystagmus werden unwillkürliche Augenbewegungen verstanden. Diese bestehen meist aus einer langsamen Bewegung und einer schnellen kompensatorischen Rückstellbewegung. Sie können in drei Kategorien eingeteilt werden: *Physiologische* Nystagmen sind die Folgen einer vestibulären oder optokinetischen Stimulation. Die *pathologischen* entstehen aufgrund Störungen des vestibulären oder okulomotorischen Systems. Die *experimentellen* Nystagmen werden durch vestibuläre Untersuchungen ausgelöst (Strupp et al., 2023). Die schnelle Phase kann klinisch besser evaluiert

werden, weshalb die Schlagrichtung des Nystagmus nach dieser benannt wird (Strupp et al., 2023). Pérez et al. (2004) teilen die Nystagmen in die jeweilige Schlagrichtung ein. Schlägt die schnelle Phase auf die betroffene Seite, wird von einem ipsilateralen Nystagmus gesprochen. Schlägt sie auf die gesunde Seite wird es als kontralateral bezeichnet. Ebenso wird betrachtet, ob die schnelle Phase stets auf die gleiche Seite schlägt (monophasisch) oder beim Abklingen die Seite wechselt (biphasisch). Sie erwähnen, dass kleine Defizite eher ipsilaterale Nystagmen mit niedrigerer Geschwindigkeit und längerer Dauer auslösen. Grössere Defizite korrespondieren mit kontralateralen und biphasischen Nystagmen. Sie zeichnen sich durch höhere Geschwindigkeit und kürzerer Dauer aus. Bei einigen Personen kann der Nystagmus asymptomatisch sein. Bei den meisten verursacht er jedoch Schwindel, verschwommenes Sehen oder Oszillopsien (Sekhon et al., 2024).

2.3 Diagnostische Testungen

In diesem Abschnitt wird genauer auf verschiedene diagnostische Testungen eingegangen, welche bei der Befunderhebung von peripheren vestibulären Defiziten genutzt werden. Zuerst werden die Grundlagentestungen, danach die Haupttestungen der Studie beschrieben.

2.3.1 Kopfimpulstest

Der Kopfimpulstest (KIT) basiert auf dem zweiten Ewald'schen Gesetz und dient zur Untersuchung des VOR des horizontalen Bogengangs oder als Verlaufsparmeter. Er kann klinisch sowie apparativ mit einer Videobrille durchgeführt werden (Schädler et al., 2022). Voraussetzungen sind die Fixation eines Objektes und eine freie Kopfbewegung (Reiß & Reiß, 2015). Kontraindikationen wie Verdacht auf Vertebralisdissektion oder ausgeprägten HWS-Problematiken sollten ausgeschlossen werden (Schädler et al., 2022). Die Patientin oder der Patient fixiert die Nasenspitze der untersuchenden Person, während der Kopf mit 30° Rotation zur Seite eingestellt wird. Danach wird er mit einer ruckartigen passiven Bewegung zurück zur Mittelstellung hinbewegt. Beurteilt wird die Augenreaktion. Bei ruckartiger Linksdrehung wird das linke, bei der Drehung nach rechts das rechte Gleichgewichtsorgan untersucht. Physiologisch bleiben die Augen ohne Abweichungen auf die Nase gerichtet (Schädler et al., 2022). Abweichungen der Augen oder Korrektursakkaden deuten auf eine periphere vestibuläre Funktionsstörung hin (Schädler et al., 2022).

2.3.2 Video-Kopfimpulstest

Der KIT kann nur Defizite mit einem Gain von <0.4 feststellen. Dies erschwert das Erkennen von leichten Defiziten, weshalb der Video-Kopfimpulstest (vKIT) eine geeignete Alternative darstellt (Yip et al., 2016). Der vKIT misst die Leistungsfähigkeit des VOR im Hochfrequenzbereich und liefert ausreichende klinische Informationen für die Diagnosestellung (Strupp, Brandt, et al., 2022). Er wird als Primärtest zur Differenzierung zentraler oder peripherer Problematiken empfohlen (Heim, 2024). Die Durchführung entspricht dem KIT, jedoch trägt die Patientin oder der Patient eine Videokamera mit einer hohen Zeitauflösung (250 Hz). Die Augen- und Kopfbewegungen werden mit einem Sensor gemessen. Deren Geschwindigkeiten werden einander gegenübergestellt und daraus der Gain errechnet (Yip et al., 2016). Somit können kleine Korrektursakkaden erkannt und ausgewertet werden (Reiß & Reiß, 2015).

2.3.3 Kalorische Prüfung

Die Kalorik ist eine der ältesten Methoden zur Untersuchung der Gleichgewichtsorgane und prüft primär den pars superior des N. vestibularis (Reiß & Reiß, 2015). Sie wird meist mit Wasser oder Luft durchgeführt (Fetter, 2010), wobei das Trommelfell und der Gehörgang vorgängig auf Perforationen untersucht werden müssen (Reiß & Reiß, 2015). Durch einen Temperaturgradienten zwischen Temperaturquelle und Bogengang wird ein physiologischer Reiz gesetzt, dieser erzeugt eine Flüssigkeitsrotation, was zu Schwindel oder Nystagmus führen kann (Baloh, 2002). Die Testperson liegt mit 30° angehobenem Kopf, mit einer Frenzelbrille zur Aufhebung der Blickfixation in einem dunklen Raum (Reiß & Reiß, 2015). Der thermische Reiz wird für 30 Sekunden in den äusseren Gehörgang appliziert. Die Spülzeit soll für alle Durchgänge gleich sein und mit dem Wärmereiz begonnen werden (Reiß & Reiß, 2015). Der Wärmereiz von 44°C führt zu einer Erregung des horizontalen Bogengangs mit schneller Nystagmusphase zum gespülten Ohr, während der Kältereiz bei 30°C ein Nystagmus vom gespülten Ohr weg auslöst (Fetter, 2010). Getestet wird in einer Frequenz von 0.003 Hz (Zalewski et al., 2019). Die vestibuläre Reaktion wird quantitativ durch die Nystagmusschrift und den Seitenvergleich erfasst, wobei Asymmetrien auf ein Defizit hinweisen können (Fetter, 2010). Personen ohne Defizit zeigen in der Regel symmetrische Reaktionen innerhalb eines festgelegten Normalbereichs (Gonçalves et al., 2015). Zur Quantifizierung der Ergebnisse werden der «vestibularis paresis score» und der «directional preponderance score» nach Jongkees (1950) verwendet (Fetter, 2010). Wird die Reaktion ausschliesslich visuell beobachtet, kann ein VNG zur Aufzeichnung und Auswertung genutzt werden (Eza-Nuñez et al., 2016).

2.3.4 Dynamic Visual Acuity Test

Der Dynamic Visual Acuity (DVAT), auch bewegungsabhängiger Sehtest, prüft, ob die Sehschärfe bei Kopfbewegungen stabil bleibt (Dannenbaum et al., 2005). Eine verminderte Funktion des VOR kann zu solchen Beeinträchtigungen der dynamischen Sehschärfe führen (Blödow et al., 2012). Kontraindikationen (vaskulär, orthopädisch und neuronal) für die Kopffrotation, besonders im Bereich der HWS, sollten zuvor abgeklärt werden (Dannenbaum et al., 2005). Der Test erfolgt vor einer Sehtafel, wobei der Kopf passiv in einem horizontalen und vertikalen Bewegungsausmass von 20° pro Seite und mit einer Frequenz von 1.5 Hz bewegt (Dannenbaum et al., 2005). Für die horizontale Kopfdrehung wird eine Kopfflexion von 30° eingestellt (Erdinest & London, 2022). Das Ergebnis ist die Differenz zwischen der letzten lesbaren Zeile in Ruhe und während Kopfbewegungen. Bei unilateralen vestibulären Defiziten verschlechtert sich das DVAT-Ergebnis mit steigender Frequenz (Dannenbaum et al., 2005).

2.3.5 Kopfschütteltest

Zur Befunderhebung einer unilateralen vestibulären Hypofunktion kann der Nystagmus auch mittels eines Kopfschütteltests (HST) ausgelöst und ausgewertet werden (Jacobson et al., 1990). Um die Stärke und Richtung des Nystagmus zu evaluieren, wird ein ENG eingesetzt (Asawavichiangianda et al., 1999). Pérez et al. (2004) zeichnete die Augenbewegung mit einer VNG auf. Andere nutzten eine Frenzelbrille, um die Augenbewegung zu vergrössern (Jacobson et al., 1990). Als Kontraindikationen gelten Hypermobilität oder Instabilität der HWS sowie ausgeprägte HWS-Problematiken. Diese sollen zu Beginn abgeklärt werden (Schädler et al., 2022). Es ist möglich, den Test aktiv sowie passiv durchzuführen (Burgio et al., 1991). Gimeno-Vilar et al. (2007) empfehlen den HST aktiv mit einer VNG durchzuführen. Wenn kein Nystagmus ersichtlich ist, kann der passive Test als Alternative genutzt werden. Die Testausführung unterscheidet sich unter den Publikationen. Zur Beschreibung der Durchführung wurde, die am häufigsten genutzte Methode verwendet. Die Patientin oder der Patient sitzt in einem dunklen Raum und flektiert die HWS um 30°. Somit steht der laterale Bogengang horizontal (Schädler et al., 2022). Danach wird der Kopf aktiv kräftig für 10-20 Sekunden rotiert. Das Ausmass der Bewegung beträgt 45° und die Geschwindigkeit 2 Hz. Bei der passiven Ausführung wird das Kopfschütteln von einer untersuchenden Person durchgeführt. Die weiteren Testkonditionen bleiben gleich (Asawavichiangianda et al., 1999). Nach dem Test sollte der Kopf schnellstmöglich ruhiggehalten werden, so wird der Nystagmus erkennbar. Bei wiederholtem Kopfschütteln wird die Spannung im Vestibularsystem erhöht, diese hält nach Beendigung der Kopfbewegung kurzzeitig an und führt zu einem Nystagmus

(Kamei, 1988). Bei einer Person mit einer unilateralen vestibulären Hypofunktion kann nach dem Kopfschütteln ein Nystagmus mit der schnellen Phase zur Seite des intakten Ohres festgestellt werden. Bei Personen ohne Läsion löst das Kopfschütteln keinen Nystagmus aus (Jacobson et al., 1990). Es wird vermutet, dass der Mechanismus des Kopfschüttelnystagmus dem «velocity storage» und dem 2. Ewald'schen Gesetz zu Grunde liegt. Die durch die Kopfdrehung ausgelöste Aktivität ist in Richtung des intakten Ohres grösser als diejenige in Richtung des beeinträchtigten Ohres (Angeli et al., 2011). Schnelle Kopfbewegungen verursachen aufgrund der einseitigen Läsion ein asymmetrisches Aufladen des «velocity storage». Wenn das Kopfschütteln gestoppt wird, entlädt sich der Speicher asymmetrisch, was in einem Kopfschüttelnystagmus mit der schnellen Phase zur gesunden Seite resultiert (Jacobson et al., 1990).

2.3.6 Drehstuhltestung

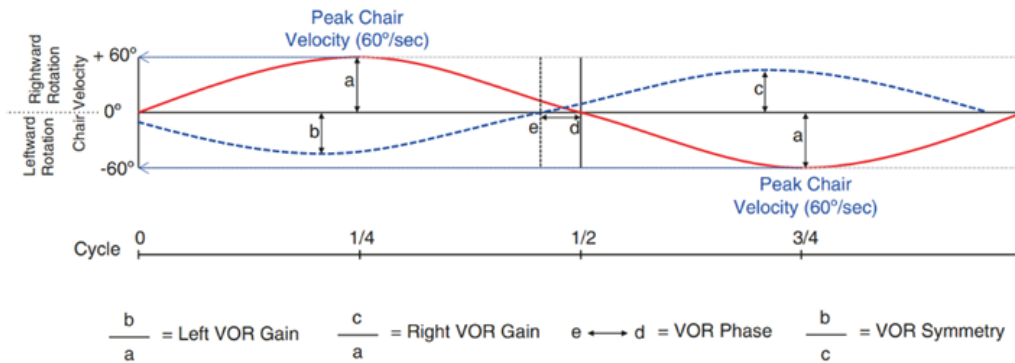
Der Drehstuhltest (ROT) ermöglicht eine umfassende klinische Bewertung des vestibulären Systems (Zalewski et al., 2019). Zum einen hat der ROT die Fähigkeit die Labyrinth gleichzeitig zu stimulieren, zum anderen sind Testungen in höheren Frequenzbereichen möglich (Ahmed, 2015). Es lassen sich somit periphere als auch zentrale vestibuläre Funktionsstörungen identifizieren (Zalewski et al., 2019). Zalewski et al. beschreiben 2019 den ROT als bestes Instrument zur Erfassung und Beurteilung von vestibulären Physiologien und Pathologien. Zu Beginn wurde manuell mit einer Geschwindigkeit von 180 °/s für mindestens 30 Sekunden getestet. Die Zeit des post-rotatorischen Nystagmus wurde mit einer Stoppuhr nach der Drehung nach rechts und links gestoppt und mit der Spitzenintensität und der Dauer des Nystagmus verglichen (Goebel et al., 2000). Der physiologische Nystagmus schlägt während einer Rechtsdrehung nach rechts, stoppt der ROT schlägt die schnelle Phase des Nystagmus nach links. Das Gleiche gilt bei einer Drehung nach links und wird im Protokoll des Velocity Step Test ebenfalls so gehandhabt (siehe Kapitel 2.3.6.2, Abbildung 9). Eine schwache Nystagmusantwort in beide Richtungen weist auf eine bilaterale vestibuläre Funktionsstörung hin, wobei Asymmetrien in Intensität oder Dauer eher auf eine unilaterale Dysfunktion deuten (Goebel et al., 2000). Mit der Einführung der computer- und drehmomentgesteuerten Drehstühlen wurden die manuellen abgelöst, die sich unter anderem auch als nicht reliabel darstellten (Goebel et al., 2000). Bei der Durchführung der ROT wird die Person vertikal auf den Drehstuhl gesetzt mit einer Kopfneigung von 30°, um den lateralen Bogengang optimal für eine maximale Stimulation in die Horizontale zu bringen (siehe Kapitel 2.1.1, Abbildung 1). Es können Frequenzen von 0.003 Hz (kalorischer Reiz) bis 2.0 Hz sowie Geschwindigkeiten von 400-600 °/s bis 1000 °/s appliziert werden (Zalewski et al., 2019). Die Augenbewegungen werden mit

EOG (Eza-Nuñez et al., 2016), ENG (Arriaga et al., 2005) oder Infrarotvideokamera (Maes et al., 2011) aufgezeichnet und elektronisch ausgewertet (Eza-Nuñez et al., 2016). Es werden drei Reaktionsparameter untersucht: Die *Gain* oder auch Verstärkung, die *Phase* und die *Symmetry* (Zalewski et al. 2019, zitiert nach Brey et al., 2008b, und Shepard & Telian, 1996). Grundlegend werden routinemässig zwei Hauptprotokolle angewendet. Der Sinusoidal Harmonic Acceleration Test, auch bekannt als sinusförmiger harmonischer Beschleunigungstest, und der Velocity Step Test, auch Geschwindigkeitsstufentest genannt. Weitere Protokolle werden erwähnt, die jedoch nicht als Standarduntersuchung durchgeführt werden. Wichtig ist, die ROT in die umfassende vestibuläre Untersuchung zu integrieren und somit eine verbesserte Sensitivität und Spezifität für vestibuläre Pathologien zu erlangen (Zalewski et al., 2019).

2.3.6.1 Sinusoidal Harmonic Acceleration

Zalewski et al. (2019) beschreiben, dass der Sinusoidal Harmonic Acceleration Test (SHA) am häufigsten durchgeführt wird. Die Testung gewährleistet eine Untersuchung des Vestibularsystems über einen breiten Frequenzbereich. Die Patientin oder der Patient sitzt auf einem Drehstuhl, welcher sich in der horizontalen Ebene zuerst in eine Richtung dreht und bis auf eine maximale Zielgeschwindigkeit von 50 °/s bis 60 °/s beschleunigt. Nach der Bremsung des Drehstuhls, dreht sich der Stuhl in die entgegengesetzte Richtung bis die gleiche maximale Geschwindigkeit erreicht wird, um dann die Umdrehungen wieder zu verlangsamen. Die am häufigsten applizierten Frequenzen sind 0.01, 0.02, 0.04, 0.08, 0.16, 0.32 und 0.64 Hz (Zalewski et al., 2019). Durch die SHA wird die Reaktion des VOR getestet. Die Rotationsreize werden in einer sinusförmigen Kurve dargestellt (siehe Abbildung 4). Da sich die Geschwindigkeit ständig ändert, wird die Cupula mit allgegenwärtigen Beschleunigungs-, Verzögerungs- und Auslenkungskräften konfrontiert. Durch diese Auslenkungskräfte schwingt die Cupula hin und her, was zu wiederholten Augenbewegungen (Nystagmus) führt. Der VOR-Nystagmus passt sich demnach im Verhältnis der Beschleunigung des Drehstuhls anhand crescendo und decrescendo an. Das Ausmass der maximalen Augenreaktion werden bei Drehungen im Uhrzeigersinn mit denen aus der Gegenuhrzeigerrichtung verglichen (Zalewski et al., 2019). Abbildung 4 ebenfalls zu entnehmen sind die drei Parameter *Gain*, *Phase* und *Asymmetry* (siehe Glossar).

Abbildung 4: Anpassung des VOR-Nystagmus an die Beschleunigung anhand crescendo und decrescendo mit den drei Reaktionsparameter: Gain, Phase, Symmetry visuell dargestellt (Zalewski et al., 2019, S.82)



Blau= Nystagmus, Schwarz= Drehstuhl Be- und Entschleunigung nach rechts und links

Nachfolgend werden in Abbildung 5 bis 8 die Auswertungen der Nystagmusantwort anhand der drei Reaktionsparameter dargestellt. Die Reaktionen werden anhand standardisierter Referenzwerte (weisser Bereich) analysiert und verglichen. Abweichungen (grauer Bereich) deuten auf eine Störung hin (Zalewski et al., 2019). Das heisst bei einer normalen Reizantwort liegen die Ergebnisse innerhalb des weissen Bereichs (siehe Abbildung 5).

Abbildung 5: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei normaler vestibulärer Reizantwort (Zalewski et al., 2019, S.83)

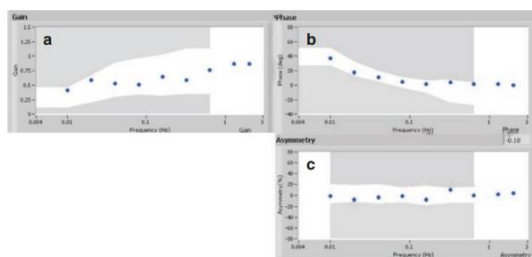
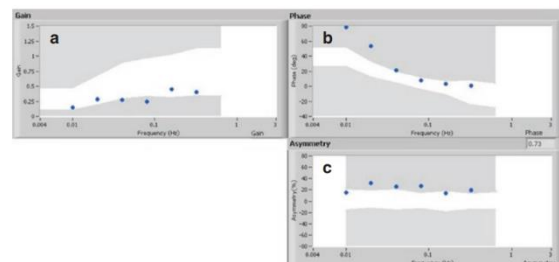


Abbildung 6: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei akuten unilateralen peripheren Störungen (Zalewski et al., 2019, S.83)



Bei unilateralen peripheren Störungen hängt die Untersuchung vom Schweregrad der Läsion sowie vom Zeitpunkt der Störungen ab. Ist die Läsion zu gering, kann ein Ausgleich über die zentrale Kompensation stattfinden und ein normales Resultat liefern. Hingegen bei schweren Läsionen ist meist eine erhöhte Phase (b) bei niedrigen Frequenzen, zudem eine Reduktion des VOR-Gain (a) und eine deutliche Asymmetry (c) zu erkennen, welche in Richtung des Spontanystagmus verzerrt scheint (siehe Abbildung 6). Ist die Asymmetry also rechtsseitig, ist der Nystagmus linksdrehend und vice versa. Bei einer ausgeprägten Asymmetry kann eine zentrale oder periphere Dysfunktion (oft akut oder unkompensiert) zu Grunde liegen (Zalewski et al., 2019).

Liegt eine zentrale Kompensation vor, zeigt sich dies oft im Bereich Gain (a), während Anomalien in der Phase (b) bestehen bleiben (Abbildung 7). Die Phase gilt somit als resistenter Parameter gegen zentrale Kompensationen (Wolfe et al., 1978). Abnorme Phasenverläufe und VOR-Asymmetry bei zentralen Läsionen werden in Abbildung 8 dargestellt (Zalewski et al., 2019).

Abbildung 7: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei unilateralen peripheren Vestibulopathien mit zentraler Kompensation im Gain (Zalewski et al., 2019, S.84)

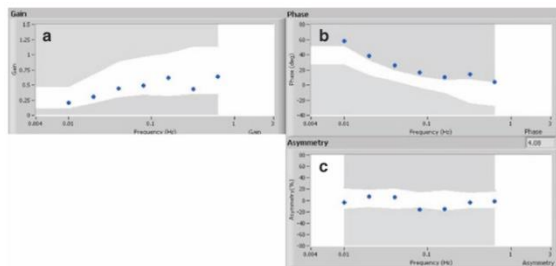
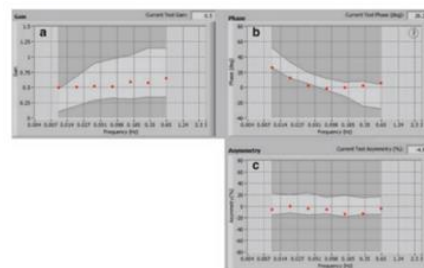


Abbildung 8: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei zentraler Läsion (Zalewski et al., 2019, S.86)



2.3.6.2 Velocity Step Test

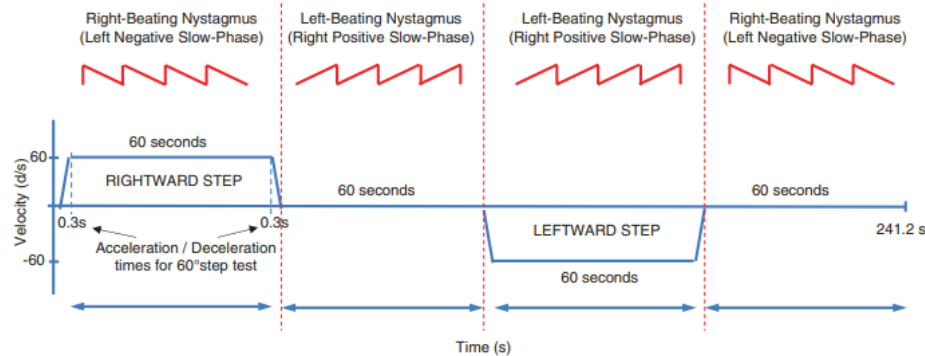
Der Velocity Step Test (VST) ist einer der ältesten klinischen Vestibulartestungen. Er wurde bereits 1907 von Robert Bárány in die Schwindeltherapie eingeführt. Im Gegensatz zum SHA wird der VST mit gleichbleibender Geschwindigkeit durchgeführt, wobei der Drehstuhl am Ende anhand einer abrupten Bremsung zum Stillstand gebracht wird (Zalewski et al., 2019).

Die Patientin oder der Patient sitzt aufrecht auf einem computergesteuerten Drehstuhl, welcher horizontale Beschleunigungen bis zu 120-200 °/s ausübt, bis eine zuvor festgelegte Geschwindigkeit erreicht ist, die anschliessend konstant gehalten wird. Es werden zwei Hauptarten von Geschwindigkeiten am häufigsten angewendet: 60 °/s (niedrig) und 240-300 °/s (hoch) (Zalewski et al., 2019, zitiert nach Brey et al., 2008b, und Brey et al., 2008a).

Je länger die Cupula einer Beschleunigungskraft ausgesetzt ist, desto grösser wird die Auslenkung und somit die neurale Reaktion. Die Beschleunigungsdauer hängt also von der Zielgeschwindigkeit ab (Zalewski et al., 2019, zitiert nach Goulson et al., 2016). Je höher die Zielgeschwindigkeit, desto länger muss der Drehstuhl beschleunigen und resultiert somit in einer höheren vestibulären Augenreflexantwort (Zalewski et al., 2019).

Die Abbildung 9 zeigt ein 60 °/s-VST mit jeweils 60 Sekunden während und nach, rechts- und linksseitiger Drehung mit der entsprechenden Nystagmusrichtung. Die Beschleunigungs- und Verzögerungsreize werden konstant auf 200 °/s für je 60 Sekunden pro Richtung gehalten (Zalewski et al., 2019). Die Entstehung der Nystagmusrichtungen wurde im Kapitel 2.3.6 bereits erläutert.

Abbildung 9: 60°/s mit 60 Sekunden vor und nach rechts- und linksseitig gerichteten Stimuli. X-Achse= Zeit (in Sekunden), Y-Achse= Geschwindigkeit (Grad/Sekunde) (Zalewski et al., 2019, S.88)



Der VST mit niedriger Geschwindigkeitsstufe, auch Low-VST, wird aufgrund normativer Referenzdaten für den Nystagmusabfall bei einer Geschwindigkeit von 60° pro Sekunde durchgeführt. Wird zuerst eine Drehung nach rechts untersucht, erfolgt anschliessend ein 60-Sekunden Pausenintervall, bei dem die Nystagmusreaktion bzw. die Abklinggeschwindigkeit in Sekunden aufgezeichnet wird. Nach dieser Minute wird eine Drehung in die linke Richtung initiiert. Die Nystagmusreaktion wird gemessen, bis die Reaktion auf 37% des Maximalwertes erreicht wird (Zalewski et al., 2019, zitiert nach Stockwell & Bojrab, 1997). Bezeichnet wird diese Dauer als VOR-Abklingzeitkonstante oder Zeitkonstante (TC). Der VOR-Nystagmus allein würde aufgrund der neuronalen Reaktion eine Dauer von 6-7 Sekunden aufweisen. Jedoch sorgt der «velocity storage» dafür, dass der Nystagmus weit länger andauert (Zalewski et al., 2019, zitiert nach Shepard et al., 2016). Der normale Referenzbereich für die Abklingzeit liegt zwischen 10 und 30 Sekunden (Zalewski et al., 2019, zitiert nach Baloh & Honrubia, 2001). Bei weniger als 10 Sekunden spricht man von einer Anomalie. Eine verringerte VOR-Zeitkonstante (<10 Sekunden) kann eine periphere, wie aber auch zentrale, Läsion darstellen, da der «velocity storage» bei beiden Störungen beeinträchtigt sein kann. Aufgrund dessen sollte die Diagnosestellung zur definitiven peripheren Vestibulopathie mit Vorsicht erfolgen (Zalewski et al., 2019).

Die hohe Geschwindigkeitsstufe, auch High-VST, ermöglicht die Identifizierung und Quantifizierung der Erregbarkeit des Labyrinths (labyrinthische Symmetry), ähnlich wie die kalorische Testung. Die Durchführung erfolgt mit einer höheren Zielgeschwindigkeit

von 240 °/s Der Unterschied zum Low-VST liegt darin, dass aufgrund der erhöhten Geschwindigkeit eine maximale Cupulaverschiebung des kontralateralen Ohrs stattfindet und somit inhibitorische Reaktionen und anschliessend afferente Reaktionen ausgeschöpft sind. Zeitgleich findet auf der ipsilateralen Seite eine Cupulaverschiebung mit einer maximalen afferenten Erregungsantwort statt. Daraus lässt sich schliessen, dass die VOR-Antwort vom ipsilateralen Ohr stammt, obwohl das kontralaterale Ohr bei abrupten Bremsungen die erregende Reaktion liefert. Somit ist die High-VST geeignet, eine Identifikation und/oder Quantifizierung der erregenden Reaktion aus jedem Labyrinth zu untersuchen (Zalewski et al., 2019). Eine VOR-Asymmetrie >20% wird als signifikant für eine unilaterale Dysfunktion gesehen (Baloh et al., 1979). Wobei die schwächere labyrinthische Reaktion der schwächeren Spitzenreaktion der Augengeschwindigkeit entsprechen. In der Regel gehen eine höhere Sensitivität und Spezifität für die Erkennung von Labyrinth-Asymmetrien oft mit intensiveren Reizen und einer ausgeprägteren Schädigung des Innenohrs einher (Wall, 1990). Folglich lässt sich eine leichte Schädigung nicht ausschliessen, selbst wenn die Stufentestung normale Symmetrien aufweist. Die Notwendigkeit einer gründlichen Beurteilung des vestibulären Systems ist von grosser Bedeutung. Der VST kann in Kombination mit der SHA-Testung eingesetzt werden, um eine periphere Asymmetrie zu identifizieren oder zentrale Kompensationen zu überwachen oder zu validieren (Zalewski et al., 2019).

2.3.7 Andere Testungen

Auch der Romberg- und der Unterberger-Test werden zur Untersuchung des vestibulären Systems verwendet. Jedoch ist zu berücksichtigen, dass der Unterberger-Test in Bezug auf die diagnostische Validität als ungenügend angesehen ist und deshalb in der Praxis nicht empfohlen wird. Auch der Romberg-Test kann aufgrund fehlender Werte zur Validität nicht zur Diagnose einer propriozeptiven Störung empfohlen werden (Schädler et al., 2012).

2.4 Gütekriterien

Priebe (2024) beschreibt Reliabilität, Objektivität und Validität als die drei Hauptgütekriterien. Sind diese drei Kriterien nicht erfüllt, ist ein Test qualitativ nicht hochwertig. Somit ist eine gute Bewertung einer diagnostischen Testung von grosser Wichtigkeit und lässt physiotherapeutische Hypothesen präziser bestätigen oder verwerfen (Wirz et al., 2014). Nachfolgend werden in Tabelle 1 die in dieser Arbeit angewendeten Gütekriterien genauer erläutert.

Tabelle 1: Gütekriterien

Praktikabilität	Die Anwendbarkeit eines Tests wird als Praktikabilität beschrieben. Gemeint ist, inwiefern ein Test sich in der Praxis anwenden lässt. Dazu gehören Kosten, Aufwand und Invasivität (Mangold, 2013) sowie Material, Platzverhältnisse, Lizenzen und Zeitaufwand (Wirz et al., 2014). Die Beurteilung erfolgt aufgrund eigener Beobachtungen der anwendenden Personen (Mangold, 2013).
Reliabilität/ Objektivität	<p>Bleibt das Ergebnis nach mehrmaligen Testungen unverändert, spricht man von einem reliablen Test. Die Reliabilität beschreibt die Zuverlässigkeit einer Testung bei stabilem Zustand der Patientin oder des Patienten. Die Reliabilität setzt sich aus dem Test-Retest und der internen Konsistenz zusammen (Mangold, 2013). Nachfolgend werden verschiedene Arten der Reliabilität bei Testwiederholungen erläutert (Wirz et al., 2014):</p> <ol style="list-style-type: none"> 1. Test-Retest-Reliabilität: Es werden die Messfehler und somit die Zuverlässigkeit des Messinstrumentes bei erneuter Testung beurteilt. Dies unter der Voraussetzung, dass das getestete Merkmal unverändert bleibt. 2. Intra-Rater-Reliabilität/Intra-Tester-Reliabilität: Die gleiche Person führt die Testung sowie den Wiederbefund durch und misst somit die Zuverlässigkeit des Tests. 3. Inter-Rater-Reliabilität/Objektivität: Misst die Zuverlässigkeit, wenn zwei Personen unabhängig voneinander die Testung an der gleichen Person durchführen. Testperson A führt die erste Testung durch, Testperson B die wiederholte Testung. <p>ICC: Bei kontinuierlichen Merkmalen werden die Korrelationen bestimmt, wobei der parametrische Pearson Korrelationskoeffizient (r_p) bei Normalverteilung und die nicht-parametrischen Spearmans Rangkorrelationskoeffizient (r_s) oder Intraclass-Correlation-Coeffizient (ICC) bei nicht normalverteilten Daten angewendet wird. Bei einem Korrelationswert von 1.0 oder -1.0 spricht man von einem perfekten Zusammenhang der ausgewerteten Daten (Wirz et al., 2014). Bei Gruppenvergleichen wird für eine angemessene Reliabilität ein Wert von ICC >0.7 empfohlen (Aaronson et al., 2002).</p> <p>Kappa: Ist kein Referenzwert vorhanden, wird die Reliabilität anhand des Kappawerts (k) angegeben. Dieser korrigiert die zufällige Übereinstimmung. Werte von 0.2-0.39 sind ausreichend, 0.4-0.59 moderat, 0.6-0.8 werden als substantiell, Werte >0.8 als fast perfekt bewertet (Wirz et al., 2014).</p> <p>Interne Konsistenz: Beschreibt, wie vergleichbare Eigenschaften zu vergleichbaren Resultaten führen (Mangold, 2013). Um die Zusammenhänge zu prüfen, wird oft das Cronbach Alpha als Maß genutzt. Ein Wert von 0.7 wird als untere Grenze für eine akzeptable Reliabilität definiert (Mangold, 2013).</p>
Validität	Die Validität, auch Gültigkeit genannt, gibt an, inwiefern ein Instrument das misst, was es zu messen vorgibt (Wirz et al., 2014). Die Validität gilt als wichtigstes Gütekriterium (Kubbe, 2016a). Um sie zu bestimmen, werden die Testungen mit dem Referenztest verglichen. Je besser die Korrelation der Ergebnisse, desto höher die Validität. Für eine hohe Validität muss eine hohe Reliabilität vorausgesetzt sein, jedoch nicht umgekehrt (Kubbe, 2016b). Es werden vier Arten der Validität unterschieden. Die <i>Augenscheinvalidität</i> , die <i>Inhaltsvalidität</i> , die <i>Konstruktvalidität</i> und die <i>Kriteriumsvalidität</i> . Bei Letzterem wird das Ergebnis einer Testung oder eines Assessments mit einem Referenztest (früher Goldstandard) verglichen.

	<p>Diagnostische Testungen haben oft nur zwei Resultate: Positiv oder negativ und werden deshalb dichotom dargestellt. Werden die Ergebnisse mit einem Referenztest verglichen, so kann die Validität anhand der Spezifität und Sensitivität bestimmt werden (Wirz et al., 2014).</p>
Sensitivität	<p>Liegen <i>dichotome</i> Merkmalsausprägungen vor, ist mit der Sensitivität die Anzahl der richtigen positiven Resultate gemeint, die der Test aufweist (Mangold, 2013). Bei <i>kontinuierlichen</i> Daten wird die Sensitivität als Ausprägung der Veränderung einer Patientin oder eines Patienten über einen bestimmten Zeitabschnitt beschrieben.</p>
Spezifität	<p>Die Spezifität beschreibt das Ausmass, welches die Testung als richtig negativ bewertet hat. Die Begrenzung liegt dabei auf dichotomen Merkmalsausprägungen (Mangold, 2013).</p> <p>Die Effektgrößen können prozentual oder dezimal dargestellt werden (Mangold, 2013). Eine Effektgrösse zwischen 0.44 und 0.65 wird als mittlerer Effekt für die Sensitivität bewertet (Cohen, 1988). 0.8-0.89 wird als akzeptabel und 0.9-1.0 gut bis exzellent eingestuft (Plante & Vance, 1994).</p> <p>Positiver prädiktiver Wert (PPV): Die Wahrscheinlichkeit, dass eine kranke Person positiv getestet wird Negativer prädiktiver Wert (NPV): Die Wahrscheinlichkeit, dass eine gesunde Person negativ getestet wird. Beide Werte werden von der Prävalenz einer Erkrankung beeinflusst (Wirz et al., 2014).</p> <p>ROC-Kurve: Als Grenzwert (Engl. cut-off point) wird ein Wert beschrieben, welcher aussagt, ab wann jemand als positiv und wann als negativ bewertet wird. Der Wert beeinflusst die Spezifität und Sensitivität. Wird die Sensitivität besser, sinkt die Spezifität. Dieser Zusammenhang kann in einer ROC-Kurve graphisch dargestellt werden (Wirz et al., 2014).</p>

3 ZIELSETZUNG UND FRAGESTELLUNG

Die theoretischen Grundlagen verdeutlichen die Komplexität der Diagnostik bei einem unilateralen peripheren vestibulären Defizit. Ausserdem wird die Basis für die Identifizierung und Bewertung diagnostischer Methoden gelegt. Aufbauend auf diesem Verständnis besteht das Ziel in der vorliegenden systematischen Literaturarbeit darin, geeignete diagnostische Studien zu identifizieren, analysieren und zu bewerten, wobei der Fokus auf der Untersuchung der diagnostischen Qualität des Kopfschütteltests und der Drehstuhltestung liegt. Die Studien sollen den aktuellen Forschungsstand bezüglich diagnostischer Kriterien für die beiden Testungen bei Patientinnen und Patienten mit unilateralen peripheren vestibulären Defiziten (vor allem Neuritis vestibularis und M. Menière) wiedergeben. Der Fokus liegt insbesondere auf den Gütekriterien Reliabilität und Validität. Basierend auf den Ergebnissen soll eine Praxisempfehlung abgegeben werden.

Auf Basis der Problemstellung, der theoretischen Grundlagen sowie der Zielsetzung lässt sich folgende Fragestellung herleiten:

«Welche der beiden Testungen, Kopfschüttel- oder Drehstuhltest, zeigt die besseren diagnostischen Qualitäten bezüglich Reliabilität und Validität bei Patientinnen und Patienten ab zehn Jahren mit unilateralen peripheren vestibulären Defiziten auf?»

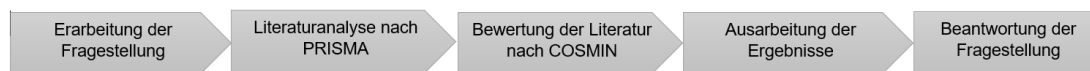
4 METHODIK

Die vorliegende Bachelorthesis stellt ein systematisches Review dar. Diese Methodik dient der kritischen Beurteilungen der wissenschaftlichen Qualität von Studien, zudem soll sie Gemeinsamkeiten und Divergenzen aufzeigen (Mangold, 2013). Nachfolgend wird das Studiendesign genauer erläutert sowie die Ein- und Ausschlusskriterien dargelegt. Zusätzlich wird die Suchstrategie als auch die Auswahl und Analyse der Studien aufgeführt.

4.1 Design

Das Studiendesign orientiert sich an der Durchführung eines systematischen Reviews. Nach der Erarbeitung der Fragestellung wurde anhand des «PRISMA-Flowcharts» eine Literaturanalyse und anschliessend eine Bewertung nach «COSMIN» durchgeführt, um den Einschluss möglichst guter Evidenz zu gewährleisten (Mangold, 2013). Anhand der Ausarbeitung der Ergebnisse wurde abschliessend die Fragestellung beantwortet. Eine Übersicht über die Vorgehensweise der vorliegenden Studie ist in Abbildung 10 dargestellt.

Abbildung 10: Prozess



4.2 Ein- und Ausschlusskriterien

Ein- und Ausschlusskriterien dienen dazu, die Grundgesamtheit der betroffenen Personen zu definieren und anhand dessen die Studien ein oder ausschliessen zu können, um möglichst homogene Studien zur Beantwortung der Fragestellung zu erhalten (Mangold, 2013). In Tabelle 2 werden die angewendeten Ein- und Ausschlusskriterien aufgeführt.

Tabelle 2: Ein- und Ausschlusskriterien zur Studienselektion

	Einschlusskriterien	Ausschlusskriterien
Studiendesign	Diagnostische Studien (min. ein Gütekriterium pro Studie)	Reviews und Interventionsstudien
Sprache	Deutsch und Englisch	Andere Sprachen
Publikationsjahr	Aufgrund der geringen Datenlage wurden keine Einschränkungen bezüglich dem Publikationsjahr gemacht.	
Population	Erwachsene und Jugendliche	Kinder bis zehn Jahre
Diagnose	Unilaterales peripheres vestibuläres Defizit	Andere vestibuläre Erkrankungen
Testungen	Aktiver und passiver HST, SHA und/oder VST	Andere Protokolle der ROT

4.3 Suchstrategie

Die Studiensuche erfolgte am 04.-06. Dezember 2023, 16. Dezember 2023, 23.-26. Dezember 2023, 29. Dezember 2023, 06. und 07. Januar 2024 sowie vom 20.-23. Mai 2024. Es wurden in den Datenbanken «PubMed», «The Cochrane Library», «Shirley Ryan AbilityLab», «ScienceDirect», «Europe PMC», «Cambridge University Press» und «Physiotherapy Evidence Database (PEDro)» nach Literatur gesucht. Für die ROT wurde zusätzlich in den beiden Datenbanken «American Academy of Neurology» und «Physical Therapy Rehabilitation» recherchiert. Auf «MeSH-Terms» wurde verzichtet, um eine breitere Auswahl an Studien durch eine uneingeschränkte Suche mit frei gewählten Schlüsselwörtern zu ermöglichen.

In den aufgeführten Datenbanken wurde nach den folgenden Stichwörtern gesucht: «Unilateral peripheral vestibular lesion», «unilateral peripheral vestibular deficit», «unilateral peripheral vestibular hypofunction», «unilateral vestibular hypofunction», «vestibular neuritis», «vestibular testing», «peripheral», «vestibular», «vestibul», «peripheral vestibulopathy», «lesion». Diese Stichwörter wurden in verschiedenen Variationen mit der OR-Funktion verbunden. Hinzugefügt wurden mit der AND-Funktion die Stichwörter «head shaking test», «head shaking nystagmus», «head shaking nystagmus test», «HST», «cervical rotation test», «nystagmus», «head on body rotation» und «rotary chair», «rotatory chair», «rotating chair», «rotary test», «rotary chair test», «ROTO», «whole body rotation», «manual whole body rotation», «sinusoidal harmonic acceleration», «sinusoidal harmonic acceleration test», «harmonic acceleration», «velocity step test» und verbunden mit AND «validity», «specificity», «sensitivity», «diagnostic», «reliability», «test-retest», «clinical usefulness», «usefulness». Eine Stichwortkombination des HST lautete dementsprechend «unilateral peripheral vestibular deficit» OR «unilateral vestibular hypofunction» AND «head shaking test» AND «reliability». Zusätzlich wurden nach den Stichwörtern «velocity storage», «Vestibulo-ocular reflex», «VOR» gesucht.

Als erweiterte Suchstrategie wurden die Literaturverzeichnisse in den Primärstudien genutzt, um zusätzliche Studien zu identifizieren, welche durch die bisherige Stichwortsuche nicht erfasst wurden. Zudem wurden Empfehlungen von Schwindelexpertinnen und -experten eingeholt und diese ebenfalls in den oben genannten Datenbanken gesucht.

4.4 Auswahl und Analyse der Studien

Anhand des «Preferred Reporting Items for Systematic Reviews und Meta-Analyses (PRISMA) 2009 Flow Diagramm» wurde die Auswahl der Studien durchgeführt (Moher et al., 2009). Aus diesen wurden zuerst die Duplikate entfernt, der Titel und Abstract anhand der Ein- und Ausschlusskriterien (Tabelle 2) geprüft und auf ihre Relevanz zur Beantwortung der Fragestellung beurteilt. Von den bisher eingeschlossenen Studien wurde der Volltext überprüft.

Die Studien wurden mittels der «COnsensus-based Standards for the Selection of health Measurement INstruments (COSMIN) checklist» von beiden Autorinnen unabhängig voneinander bewertet. Die Checkliste ist ein standardisiertes Instrument, um die methodologische Qualität von Studien zu quantifizieren. Anhand der Punktzahl kann die Qualität eines Assessments gemessen werden. Dies dient zur Evaluation von Gütekriterien (Mokkink et al., 2018). Das Manual besteht aus zehn Boxen mit jeweils 3-35 Items, welche mit «very good», «adequate», «doubtful», «inadequate» oder «not applicable» (NA) bewertet werden. Die Bewertung der einzelnen Items lässt Spielraum, weshalb die Interpretationen zu einzelnen Werten vorbesprochen wurde (Mokkink et al., 2018). Zur Bewertung der ausgewählten Studien wurden die passenden Boxen ausgewählt. Beurteilt wurde jede Studie mit Box 6, 7 und 8 (siehe Bewertung im Anhang, Kapitel 13.1) Diese nehmen Bezug auf die Messeigenschaften wie Reliabilität, Messfehler und Konstruktvalidität (Mokkink et al., 2018). Um die Studie insgesamt zu bewerten, schlagen Mokkink et al. (2018) vor, den niedrigsten Score der jeweiligen Box zu betrachten («worst score counts»). Bei der vorliegenden Arbeit wurde jedoch eine alternative Vorgehensweise angewendet, da ansonsten die Mehrheit der Studien mit «inadequate» bewertet worden wären. Es wurde alternativ ein Durchschnittswert der einzelnen Items als Gesamtbewertung festgelegt. Für eine Bewertung «very good» wurden drei Punkte vergeben, für «adequate» zwei Punkte, für «doubtful» ein Punkt und für «inadequate» null Punkte. Der erhaltene Durchschnittswert aller Items wurde auf ganze Zahlen gerundet und zu den ursprünglichen Kategorien «very good», «adequate», «doubtful» oder «inadequate» basierend auf den Punktzahlen umgewandelt (siehe Kapitel 5.2.1, Tabellen 3 und 4). Die Studien, welche einen Durchschnittswert von mindestens «adequate» aufwiesen, wurden mehr gewichtet als die Studien, welche mit «doubtful» bewertet worden sind.

Für die Fragestellung relevante Ergebnisse der Studien wurden in Text- und Tabellenform zusammengetragen. Die Studien wurden unter Berücksichtigung der methodischen Qualität auf ihre Gütekriterien geprüft und zur Beantwortung der Fragestellung herangezogen.

5 ERGEBNISSE

Dieses Kapitel beschreibt die Literatursuche, die Studienbewertung anhand der «COSMIN checklist» und fasst die ausgewählten Studien zusammen. Die Abschnitte werden jeweils in beide Testungen unterteilt.

5.1 Literatursuche und -auswahl

Abbildung 11 zeigt die Suche des HST. Es wurde in den Datenbanken «Pubmed» n=630, «Cochrane» n=57, «Shirley Ryan AbilityLab» n=667, «Science direct» n=794, «Europe PMC» n=178, «Cambridge University Press» n=0, «PEDro» n=1 mittels Stichwortsuche gesucht. Dabei ergaben sich 2327 Treffer für den HST. Es wurden zunächst (n=1632) Duplikate aussortiert, sodass 695 Studien übrigblieben. Diese wurden anhand der Ein- und Ausschlusskriterien (Tabelle 2) auf den Titel geprüft, wodurch weitere 614 Studien ausgeschlossen wurden. Die verbliebenen 81 Studien wurden auf den Abstract geprüft und weitere 65 Studien ausgeschlossen. Gründe für die Ausschluss waren fehlende Gütekriterien, bilaterale oder zentrale Vestibulopathien, kein HST, fehlender Volltext, fremdsprachige Studien ohne englische oder deutsche Übersetzung und die Zielgruppe jünger als zehn Jahre. Durch die erweiterte Suchstrategie wurde eine weitere Primärstudie gefunden (Wei et al., 1989). Die 16 verbliebenen Studien wurden im Volltext gelesen. Nach dem Lesen wurden Studien ohne Relevanz zur Beantwortung der Fragestellung (n=1), fehlende Gütekriterien (n=4), bilaterale oder zentrale Vestibulopathien (n=2) oder Zielgruppe jünger als zehn Jahre ausgeschlossen (n=2). Es konnten sieben Studien zum HST eingeschlossen werden.

In Abbildung 12 ist die Literatursuche der ROT dargestellt, wobei 693 Ergebnisse gefunden wurden («Pubmed» n=442, «Science direct» n=70, «Europe PMC» n=162, «American Academy of Neurology» n=9 und «Physical Therapy Rehabilitation» n=10). Nach Entfernung der Duplikate (n=283) blieben 410 Studien übrig. Es erfolgte eine Titelprüfung anhand der Ein- und Ausschlusskriterien (siehe Tabelle 2), dabei wurden 266 Studien ausgeschlossen. Von 144 Studien wurde der Abstract überprüft und davon 123 ausgeschlossen. Ausschlusskriterien waren folgende: Keine Gütekriterien (n=81), bilaterale Vestibulopathien oder zentrale Störungen (n=24), keine ROT (n=9), fehlende Volltexte (n=5), Sprache (Chinesisch n=2) und Zielgruppe Kinder bis zehn Jahre (n=2). Die restlichen Studien wurden auf den Volltext gescreent und aufgrund der ungenauen Beschreibung der Gütekriterien (n=10), bilateralen Vestibulopathien (n=3) oder zentralen Störungen (n=1) konnten weitere 14 Studien ausgeschlossen werden. Es wurden sieben Studien zum ROT in die Arbeit eingeschlossen.

Abbildung 11 : Eigene Darstellung Studiensuche HST anhand PRISMA-Flowchart (Moher et al., 2009). N=Anzahl

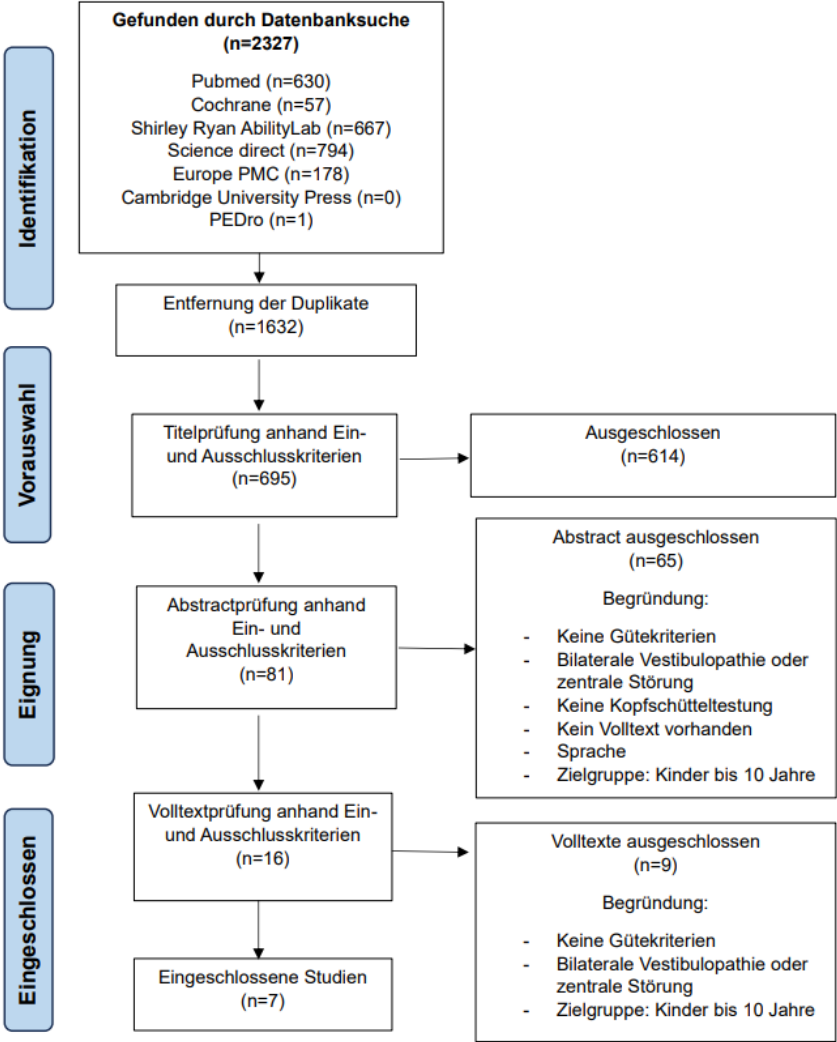
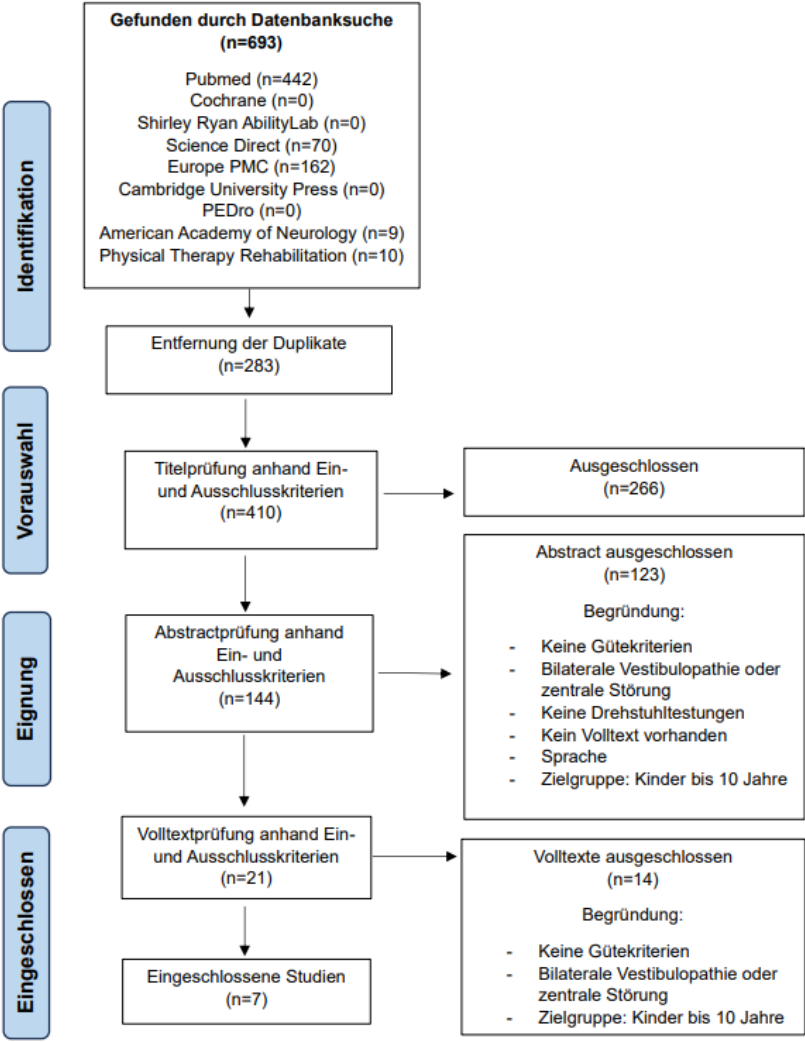


Abbildung 12: Eigene Darstellung Studiensuche ROT anhand PRISMA-Flowchart (Moher et al., 2009). N=Anzahl



5.2 Methodische Qualität der eingeschlossenen Studien

Die verbleibenden 14 Primärstudien wurden von beiden Autorinnen unabhängig voneinander anhand der «COSMIN checklist» bewertet.

5.2.1 Bewertung der Studien nach COSMIN

Die Tabellen 3 und 4 zeigen eine zusammenfassende Bewertung von Box 6 bis 8 der eingeschlossenen Primärstudien anhand der «COSMIN checklist». Eine detailliertere Ansicht ist im Anhang, Kapitel 13.1 zu finden.

Die Studien zum HST wurden wie folgt bewertet:

Tabelle 3: Übersicht der Studienbewertung HST anhand der COSMIN Checkliste

Autorinnen/Autoren, Jahr	Box 6: Reliability	Box 7: Measurement error	Box 8: Criterion validity
Angeli et al., 2011			
Pérez et al., 2004			
Wei et al., 1989			
Asawavichiangianda et al., 1999			
Burgio et al., 1991			
Jacobson et al., 1990			
Sun et al., 2022			

Dunkelgrün=very good, Hellgrün=adequate, Orange=doubtful, Rot=inadequate

Die Studien von Asawavichiangianda et al. (1999) und Jacobson et al. (1990) erreichten über alle drei Boxen eine Bewertung von «doubtful». «Doubtful» bis «adequate» wurden Wei et al. (1989) und Burgio et al. (1991) bewertet, welche in Box 8 besser abschnitten als in Box 6 und 7. Wei et al. (1989) erhielt bei der Kriteriumsvalidität ein «very good». Pérez et al. (2004) liegt bei «adequate» und somit kann die Aussagekraft als gut angesehen werden. Doch auch Angeli et al. (2011) erreichte einen Wert von «adequate» bis «very good», wobei das Ergebnis der Kriteriumsvalidität besonders hervorzuheben ist. Sun et al. (2022) wurde in zwei von drei Bereichen mit «very good» und in Box 7 mit «adequate» bewertet. Somit ist diese Studie mit mehr Gewichtung zu beachten. Folglich werden keine Studien zum HST ausgeschlossen.

Die Studien zu der ROT haben folgende Gesamtbewertung erreicht:

Tabelle 4: Übersicht der Studienbewertung ROT anhand der COSMIN Checkliste

Autorinnen/Autoren, Jahr	Box 6: Reliability	Box 7: Measurement error	Box 8: Criterion validity
Palomar-Asenjo et al., 2006			
Ahmed et al., 2009			
Ahmed, 2015			
Maes et al., 2011			
Fattahi et al., 2024			
Eza-Nuñez et al., 2016			
Arriaga et al., 2005			

Dunkelgrün=very good, Hellgrün=adequate, Orange=doubtful, Rot=inadequate

Die Studie von Ahmed et al. (2009) und Maes et al. (2011) werden im Durchschnitt als «adequate» bewertet. Arriaga et al. (2005) und Ahmed (2015) werden mit einem Durchschnittswert von eins als «doubtful» bewertet und weniger gewichtet. Besonders hervorzuheben ist die Studie von Fattahi et al. (2024), die eine Bewertung von «very good» in der Kriteriumsvalidität vorweist. Hingegen die Studie von Palomar-Asenjo et al. (2006) in der Kriteriumsvalidität als «inadequate» bewertet wird und der Durchschnittswert der Studie bei «doubtful» liegt. Eza-Nuñez et al. (2016) wird als einzige Studie über alle drei Boxen mit «adequate» bewertet. Es werden keine Studien zum ROT ausgeschlossen, jedoch werden die Ergebnisse der Studie von Palomar-Asenjo et al. (2006) weniger gewichtet.

5.3 Zusammenfassung der Studien

Es folgt eine Zusammenfassung der Charaktereigenschaften der Studien zum HST und ROT mit Angaben zur Herkunft, Population und Diagnosen, welche in Tabelle 5 und Tabelle 7 als Übersicht zu entnehmen sind. Die Diagnosen sollen einen Überblick der unilateralen peripheren Defizite darstellen.

5.3.1 Kopfschütteltest

Die eingeschlossenen Studien stammen aus dem Zeitraum 1989-2022. Die Stichproben variierten von 53-1364 Personen. Asawavichiangianda et al. (1999) schlossen mit n=1364 die grösste und Angeli et al. (2011) mit n=53 die kleinste Stichprobe ein. Das Alter lag zwischen 10 und 85 Jahren. Die Studie von Burgio et al. (1991) sowie Jacobson et al. (1990) machten keine Angaben zum Alter. Wei et al. (1989) und Sun et al. (2022) gaben ausschliesslich den Mittelwert an. Dieser liegt im Durchschnitt aller Studien bei 49.8 Jahren. Keine der Studien erwähnte, in welcher Krankheitsphase sich die Personen befanden. Einzig Angeli et al. (2011) legten als Einschlusskriterium fest, dass Schwindel oder Gleichgewichtsstörungen seit mindestens zwei Monaten vorliegen müssen. Eingeschlossen wurden Patientinnen und Patienten mit unilateralen peripheren vestibulären Defiziten, darunter sind Neuritis vestibularis und M. Menière am stärksten vertreten. Die Vergleichsgruppen hatten entweder zentralen Schwindel, gemischte Schwindelformen oder keine Beschwerden. Weitere Diagnosen sind der Tabelle 5 zu entnehmen.

Tabelle 5: Charaktereigenschaften Studien HST

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat., Alter (Jahr), Geschlecht	Diagnosen
<p>Angeli et al., 2011</p> <p>Retrospektive casecontrol Studie 2006-2008</p>	<p>USA</p>	<p>n=53</p> <p>32 Frauen 21 Männer</p> <p>Alter: 24-85 Jahre</p> <p>Mittelwert: 57 Jahre</p> <p>Kontrollgruppe (n=10)</p>	<p>Eingeschlossen: 18 Jahre oder älter, Schwindel oder Gleichgewichtsstörung seit min. 2 Monaten, unilaterale oder asymmetrische periphere vestibuläre Hypofunktion festgestellt durch KAL oder VNG</p> <p>Gleichgewichtsstörungen (n=28) Schwindel (n=25)</p> <p>Periphere Vestibulopathie (n=25, 35%) Vestibuläre Neuritis (n=17, 12%) M. Menière (n=7, 14%) Weitere Störungen des Labyrinthes (n=2, 4%) CPA-Tumor (n=2, 4%)</p> <p>Ebenfalls wurden 10 gesunde Erwachsene (Frauen und Männer) erfasst. Diese sollten keine Schwindelsymptome, Gleichgewichts- oder frühere Störungen des Vestibularsystems aufweisen.</p> <p>Ausgeschlossen: BPLS, bilaterale periphere Hypofunktion, Störung des zentralen Nervensystems, Beeinträchtigung des Bewegungsapparates oder Visus.</p>
<p>Pérez et al., 2004</p> <p>Retrospektive Case Studie April 2001-Mai 2002</p>	<p>ESP</p>	<p>n=83</p> <p>Periphere und unilaterale vestibuläre Störungen</p> <p>49 Frauen (59%) 34 Männer (41%)</p> <p>Alter: 14-77 Jahre</p> <p>Mittelwert: 48.6 Jahre</p>	<p>Eingeschlossen: Unilaterale periphere vestibuläre Defizite, keine Anzeichen für zentrale Störungen, ohne Spontannystagmus <2 °/s in der Dunkelheit, keine Einnahme von Schwindelmedikation 24 bis 36 h vor der Studie</p> <p>M. Menière (n=19) Wiederkehrende Vestibulopathie (n=14) Vestibuläre Neuritis (n=13) Akustikusneurinom (n=8) BPLS (n=5) Innenohr Erschütterungssyndrom (n=2) Postoperative Labyrinthitis (n=2) Unilaterale Agenesie des 8. Hirnnervs (n=1) Keine spezifische Diagnose (n=19)</p>

<p>Wei et al., 1989</p>	<p>USA</p>	<p>n=108</p> <p>62 Frauen (57%) 46 Männer (43%)</p> <p>Mittelwert: 49 +/- 18 Jahre</p>	<p>Aufgrund der Hörschärfe wurden die Teilnehmerinnen/Teilnehmer in Gruppen aufgeteilt:</p> <p>Unilaterale sensorineurale Schwäche Bilaterale sensorineurale Schwäche</p> <p>Normal (n=60)</p> <p>Diagnosen: <u>Zentral</u> (n=3)</p> <p><u>CP</u> > +/- 20% (n=30) davon: Perilymphfistel (n=2) N. vestibularis Sektion (n=1) BPLS (n=1) Neuritis vestibularis oder M. Menière (n=26) <u>Undiagnostiziert</u> (n=75)</p>
<p>Asawavi-chiangianda et al., 1999</p> <p>Prospektive Kohorten Studie 1991-1997</p>	<p>THA JAP CAN</p>	<p>1.Phase n=1364 mit Schwindel n=50 ohne Schwindel</p> <p>2.Phase n=30</p> <p>3.Phase n=300</p> <p>173 Frauen (57.7%) 127 Männer (42.3%)</p> <p>Alter: 10-82 Jahre</p> <p>Mittelwert: 48.4 Jahre</p>	<p>Vestibuläre Dysfunktion (n=196, 66%)</p> <p><u>Periphere Dysfunktion</u> (n=190) Periphere vestibuläre Störungen (n=192) BPLS (n=63) M. Menière (n=31) Wiederkehrende Vestibulopathie (n=26) Akustikusneurinom (n=21) Vestibuläre Neuritis (n=8) Anderes (n=16) Unbekanntes peripheres vestibuläres Defizit (n=27)</p> <p><u>Zentrale Dysfunktion</u> (n=4)</p> <p><u>Gemischte Form</u> (n=2, 66%)</p> <p>Psychogener Schwindel (n=58, 19%)</p> <p>Keine spezifische Diagnose (n=46, 15%)</p>

Burgio et al., 1991	USA	n=115 Geschlecht und Alter unklar Kontrollgruppe (n=17)	Unilaterale vestibuläre Störung (n=25), ≥ 25% Bilaterale vestibuläre Störung (n=10) Schwindelpat. mit unauffälligem ENG (n=80)
Jacobson et al., 1990	USA	n=116 Geschlecht und Alter unklar	unklar
Sun et al., 2022 Retrospektive Studie 2019-2021	CHI	n=80 52 Frauen (65%) 28 Männer (35%) Peripherer Schwindel (n=47, 58.8%) Mittelwert: 50.7 Jahre +/- 13.26 Jahre Zentraler Schwindel (n=33, 41.2%) Mittelwert: 47.03 Jahre +/- 16.65 Jahre	Eingeschlossen: Peripherer oder zentraler Schwindel, komplette klinische Daten, vervollständigter HST und HSTST, HSN Periphere Gruppe (n=47) M. Menière (n=22) Neuritis vestibularis (n=9) Plötzlicher Hörverlust mit Schwindel (n=6) Endolymphatischer Hydrops (n=4) Labyrinthitis (n=4) Innenohr (n=1) Malformation (n=1) Akustikusneurinom (n=1) Zentrale Gruppe (n=33) Vestibuläre Migräne (n=16) Mögliche Vestibuläre Migräne (n=12) Transient ischämische Attacke (n=2) Multiple System Atrophie (n=2) Spinocerebellar Ataxie (n=1) Ausgeschlossen: Mischung aus zentralem und peripherem Schwindel

HST=Kopfschütteltest, HSN=Kopfschüttelnystagmus, KAL=Kalorik, CP=Kanalparese, CPA=Kleinhirnbrückenwinkel, BPLS= Benigner paroxysmaler Lagerungsschwindel, VNG=Videonystagmographie, ENG=Elektronystagmographie, HSTST=Head Shaking Tilt Suppression Test, Pat.=Patientinnen/ Patienten

Die Testausführungen des HST werden in folgender Tabelle verglichen.

Tabelle 6: Testausführungen HST

Studie	Kopfbewegung	Licht	Augen	Kopf	Vergleich
Angeli et al., 2011	Passiv, 1-2 Hz, 30-mal	abgedunkelt	unklar, Frenzelbrille	30°	KAL und ROT
Pérez et al., 2004	Passiv, 3 Hz, 30-mal, 45° Radius	dunkel	offen, VNG	30°	KAL und ROT
Wei et al., 1989	Aktiv und Passiv (wenn aktiv n.m.), 1-2 Hz, 10-20-mal, 30° Radius	dunkel	geschlossen, ENG	unklar	KAL
Asawavichiangianda et al., 1999	Aktiv und Passiv, 2 Hz, 15 Sek., 45° Radius	dunkel	geschlossen, ENG	30°	KAL und ENG
Burgio et al., 1991	Passiv, 250-500°/s, 20 Sek.	abgedunkelt	offen, Frenzelbrille	unklar	ENG
Jacobson et al., 1990	Aktiv, 2 Hz, 30 Sek.	dunkel	geschlossen, EOG und Frenzelbrille nach dem Test	unklar	ENG und ROT
Sun et al., 2022	Passiv, 2-3 Hz, 30-mal	unklar	Videobrille und VNG	30°	HSTST

KAL=Kalorik, ROT=Drehstuhl, VNG=Videonystagmographie, ENG=Elektronystagmographie, HSTST=Head Shaking Tilt Suppression Test, n.m.= nicht möglich

Die Studien von Jacobson et al. (1999) und Wei et al. (1989) testeten den HST aktiv. Doch Letztere beschrieben, dass bei HWS-Problematiken auf den passiven Test zurückgegriffen wurde. Die anderen Studien führten ihn ausschliesslich passiv durch. Die Frequenz variierte zwischen 1-3 Hz und die Dauer von 15-30 Sekunden oder 10-30 Drehungen. Mit Ausnahme von Sun et al. (2022) wurde in einem abgedunkelten oder ganz dunklen Raum getestet, um die Blickfixation aufzuheben. Sun et al. (2022) stellten jedoch mit einer Videobrille sicher, dass die Blickfixation verhindert wurde. Vier der sieben Studien hielten die 30° Kopfflexion während der Testung ein, während die drei älteren Studien dies nicht berücksichtigten. Angeli et al. (2011) und Wei et al. (1989) haben den HST mit der Kalorik oder der ROT als alleinige Testung verglichen. Im Rahmen der ENG oder VNG wurden bei Pérez et al. (2004), Asawavichiangianda et al. (1999) und Jacobson et al. (1990) die Kalorik oder ROT als Teil der Testbatterie durchgeführt. Ausschliesslich Burgio et al. (1991) verglich den HST mit der ENG, ohne genauere Angaben dazu zu machen. Als einzige Studie wurde bei Sun et al. (2022) der HST mit dem Head Shaking Tilt Suppression Test (HSTST) verglichen.

5.3.2 Drehstuhltest

Die eingeschlossenen Studien der ROT stammen aus den Jahren 2005 bis 2024. Die Stichprobengrößen variieren zwischen 100 und 989 Personen, wobei die Studie von Arriaga et al. (2005) die grösste Anzahl mit n=989 und Palomar-Asenjo et al. (2006) die kleinste Anzahl mit n=100 aufweist. Das Alter der Testpersonen lag zwischen 13-89 Jahren. Wobei Arriaga et al. (2005) und Palomar-Asenjo et al. (2006) keine Angaben dazu machten. Die Geschlechterverteilung ist über alle Studien gesehen ähnlich. Alle Studien schliessen die beiden Diagnosen Neuritis vestibularis und/oder M. Menière ein. Die retrospektive Case-Studie von Ahmed et al. (2009) vergleicht die Gruppe der unilateralen peripheren Vestibulopathie mit einer Gruppe ohne periphere Vestibulopathie wie beispielsweise migräneinduzierten Schwindel oder zentrale vestibuläre Läsionen. Ähnlich die Studie von Eza-Nuñez et al. (2016), die einen Vergleich zu einer nicht-peripheren Vestibulopathie Gruppe mit Diagnosen wie chronischer Schwindel oder vestibuläre Migräne durchführten. Die Studie von Maes et al. (2011) verglich die betroffene Gruppe mit einer Kontrollgruppe ohne Schwindel. Weiter Diagnosen sind nachfolgend der Tabelle 7 zu entnehmen.

Tabelle 7: Charaktereigenschaften Studien ROT

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen
Palomar-Asenjo et al., 2006 Prospektive Studie	ESP	n=100 Geschlecht und Alter unklar Rechtes Ohr betroffen (n=57) Linkes Ohr betroffen (n=43) Mittelwert Dauer der Symptomatik: 5.4 Jahre	Eingeschlossen: Unilaterales M. Menière Ausgeschlossen: zentrale Störungen, bilaterale Problematiken
Ahmed et al., 2009 Retrospektive Case Studie 2001-2007	USA	n=200 Gruppe 1: 132 Schwindelpat. mit erwarteter peripherer Vestibulopathie 64 Frauen 68 Männer Alter: 20-87 Jahre Mittelwert: 58.56 +/- 15.86 Jahre Gruppe 2: 68 Schwindelpat. ohne erwartete periphere Vestibulopathie 41 Frauen 27 Männer Alter: 17-89 Jahre Mittelwert: 60.19 +/-16.96 Jahre	Eingeschlossen: Gruppe 1: <u>Periphere Vestibulopathie:</u> Definitiver M. Menière (n=46, 34.9%) Neuritis vestibularis (n=41, 31.1%) Nicht-identifiziert (n=20, 15.2%) Bilaterale Vestibulopathie (n=13, 9.9%) Akustikusneurinom (n=6, 4.6%) Labyrinthitis (n=6, 4.6%) Gruppe 2: <u>Ohne periphere Vestibulopathie:</u> Migräne induzierter Schwindel (n=39, 57.5%) Zentrale vestibuläre Läsion (n=11, 16.2%) Posturale Instabilität (n=7, 10.3%) Normaldruck Hydrocephalus (n=4, 5.9%) MS (n=3, 4.5%) Orthostatisch (n=2, 3%) Transient ischämische Attacke (n=1, 1.5%) Angstbedingter Schwindel (n=1, 1.5%) Ausgeschlossen: BPLS, Superiore Bogengang Dehiszenz, Perilymphfistel, mögliche M. Menière oder Pat. mit inkompletten Testungen oder unklarer Interpretation

<p>Ahmed, 2015</p> <p>Retrospektive Studie</p>	<p>EGY</p>	<p>n=119</p> <p>56 Frauen 63 Männer</p> <p>Alter: 22-79 Jahre</p> <p>Mittelwert: 58.23 +/- 15.13 Jahre</p>	<p>Eingeschlossen: Dokumentierte unilaterale periphere vestibuläre Pathologien, dokumentierte unilaterale kalorische Schwäche, die mit der klinischen Diagnose (Anamnese, körperliche Untersuchung neuro-otologischen Untersuchung, audiologischen Beurteilung und radiologischen Untersuchungen) übereinstimmen</p> <p>M. Menière (34.9%) Neuritis vestibularis (31.1%) Unidentifiziert (15.2%) Labyrinthitis (4.6%) Akustikusneurinom (4.6%)</p> <p>Ausgeschlossen: Periphere vestibuläre Läsionen, die nicht durch unilaterale kalorische Schwäche belegt sind, Pat. mit unvollständigen Tests oder nicht-interpretierbaren Daten</p>
<p>Maes et al., 2011</p>	<p>BEL</p>	<p>n=157</p> <p>80 Frauen 77 Männer</p> <p>Gruppe 1: N=77</p> <p>38 Frauen 39 Männer</p> <p>Alter: 19-78 Jahre</p> <p>Mittelwert: 52 Jahre</p> <p>Gruppe 2: N=80</p> <p>42 Frauen 38 Männer</p> <p>Alter: 20-80 Jahre</p> <p>Mittelwert: 48 Jahre</p>	<p>Eingeschlossen:</p> <p>Gruppe 1: <u>Unilaterale periphere vestibuläre Pathologien:</u> M. Menière (n=36) Labyrinthitis (n=12) Neuritis vestibularis (n=10) Akustikusneurinom (n=8) Bogengang-Dehiszenz (n=6) Posttraumatischer Schwindel (n=5)</p> <p>Gruppe 2: Kontrollgruppe ohne Schwindel in der Vergangenheit, keine Gleichgewichtsproblematiken, keine Störungen des Sehens (ausser Brillen- und Kontaktlinsentragende)</p>

<p>Fattahi et al., 2024</p> <p>Retrospektive Analyse 2016-2022</p>	<p>USA</p>	<p>n=150</p> <p>88 Frauen 62 Männer</p> <p>Alter: 13-84 Jahre</p> <p>Mittelwert: 54.10 +/- 15.09 Jahre</p> <p>Periphere Gruppe: 60.9 Jahre +/- 10.8 Jahre</p> <p>Zentrale Gruppe: 48 Jahre +/- 15.8 Jahre</p>	<p>Eingeschlossen</p> <p><u>Periphere Gruppe (n=71)</u> Neuritis vestibularis (n=36, 50.7%) M. Menière (n=22, 30.9%) Labyrinthitis (n=7, 9.86%) Akustikusneurinom (n=4, 5.6%) Ototoxizität (n=2, 2.8%)</p> <p><u>Zentrale Gruppe (n=79)</u> Vestibuläre Migräne (n=51, 64.5%) Zentral Vestibuläre Läsion (n=12, 15.19%) PPPD (n=9, 11.3%) Mal de débarquement (n=3, 3.8%) Schädelhirntrauma (n=4, 5.06%)</p> <p>Ausgeschlossen: Visuelle Störungen, isolierte BPLS, superiore Kanal Dehiszenz</p>
<p>Eza-Nuñez et al., 2016</p> <p>Prospektive Studie</p>	<p>ESP</p>	<p>n=116</p> <p>56 Frauen (48.3%) 60 Männer (51.7%)</p> <p>Alter: 21-80 Jahre</p> <p>Mittelwert: 54 Jahre SD: +/- 14 Jahre</p>	<p>Periphere Vestibulopathie (n=75, 64,6%) M. Menière (n=34) Neuritis vestibularis (n=12) BPLS (n=11) Otosklerose (n=5) Akustikusneurinom (n=3) Iatrogen (n=3) Cochlear Implantat (n=1) Wahrscheinliches M. Menière (n=1) Chronische Otitis media (n=1) Immundefekt- Krankheit (n=1)</p> <p>Nicht-periphere Vestibulopathie (n=41, 35,3%) Vestibuläre Migräne (n=15) Chronischer Schwindel (n=13) Angstbedingter Schwindel (n=3) Post-traumatischer Schwindel (n=2) Undefiniert (n=2) Transient ischämische Attacke (n=2)</p>
<p>Arriaga et al., 2005</p> <p>Retrospektives Chart Review</p>	<p>USA</p>	<p>n=989</p> <p>511 (51,7%) Pat. durchliefen nur die ROT 478 (48.3%) ROT und ENG Batterie.</p> <p>Alter und Geschlecht: Unklar</p>	<p>Labyrinthitis, Neuritis vestibularis, akuter Hörverlust mit Schwindel, Akustikusneurinom, Schläfenbeinbruch und M. Menière</p>

BPLS=Benigner paroxysmaler Lagerungsschwindel, PPPD=Persistierender postural perzeptiver Schwindel, MS=Multiple Sklerose, Pat.=Patientinnen/Patienten, ENG=Elektronystagmographie, ROT=Drehstuhl

Die Testausführungen des ROT werden in der folgenden Tabelle 8 aufgeführt.

Tabelle 8: Testausführungen ROT

Studie	Protokolle (Hz)	Licht	Augen und Aufzeichnung	Kopf	Vergleich	Mentale Aufgaben
Palomar-Asenjo et al., 2006	SHA (50 °/s= 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz) VST (100 °/s)	dunkel	offen EOG	leichte Flexion	KAL	Ja
Ahmed et al., 2009	SHA (60 °/s= 0.025/0.05/0.25/0.5 Hz) VST (100 °/s)	dunkel	offen unklar	30°	KAL	Ja
Ahmed, 2015	SHA (60 °/s= 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz) VST (100 °/s)	dunkel	offen Infrarot Videookulographie	leichte Flexion	KAL	Ja
Maes et al., 2011	SHA (50 °/s= 0.01/0.02/0.05/0.1/0.2 Hz) VST (100 °/s) PRRT (180 °/s)	dunkel	offen Elektroden	gerade	KAL und VEMP	Ja
Fattahi et al., 2024	SHA (40 °/s= 0.01-1.0 Hz)	dunkel	unklar unklar	unklar	KAL und vKIT	unklar
Eza-Nuñez et al., 2016	SHA (50 °/s= 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz) VST (100 °/s)	dunkel	unklar EOG	leichte Flexion	KAL und vKIT	Ja
Arriaga et al., 2005	VOR Tests (60°/s= 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz) VST (100 °/s)	unklar	unklar ENG	unklar	ENG	unklar

KAL=Kalorik, VEMP=Vestibulär Evoked Myogenic Potential, PRRT=Pseudorandom Rotation Test, vKIT=Video-Kopfimpulstest, ENG= Elektronystagmographie, EOG=Elektrookulographie

Alle Studien führten die Testungen in der Dunkelheit durch, mit Ausnahme der Studie von Arriaga et al. (2005) in der diesbezüglich nichts erwähnt wurde. Die Augen waren in vier Studien jeweils offen, bei Fattahi et al. (2024), Eza-Nuñez et al. (2016) und Arriaga et al. (2005) unklar. Die Studie von Ahmed et al. (2009) erwähnt explizit, dass der Kopf in 30° HWS-Flexion gebracht wird. Palomar-Asenjo et al. (2006), Ahmed (2015) und Eza-Nuñez et al. (2016) erwähnen eine leichte Flexion. Hingegen bei Maes et al. (2011) wird der Kopf während der Testung gerade stabilisiert. Bei Fattahi et al. (2024) und Arriaga et al. (2005) ist die Kopfstellung unklar. Sechs von sieben Studien vergleichen die ROT mit der Kalorik, mit Ausnahme von Arriaga et al. (2005), welche die ROT mit dem ENG-Protokoll vergleichen. Zusätzlich vergleicht die Studie von Fattahi et al. (2024) und Eza-Nuñez et al. (2016) die ROT mit dem vKIT. Mentale Aufgaben werden bei fünf von sieben

Studien gestellt, um die Bereitschaft der Patientinnen und Patienten sicherzustellen. Das SHA-Protokoll wird in sechs von sieben Studien durchgeführt, wobei die Geschwindigkeiten und Frequenzen zwischen 40 °/s und 60 °/s respektive 0.01-1.0 Hz variieren. Vier Studien von Palomar-Asenjo et al. (2006), Ahmed (2015), Eza-Nuñez et al. (2016) und Arriaga et al. (2005) verwenden die gleichen Frequenzbereiche. Das VST-Protokoll wird ebenfalls in sechs von sieben Studien durchgeführt, wobei die Geschwindigkeit in allen Studien bei 100 °/s liegt. Einzig Maes et al. (2011) führt zusätzlich das Pseudorandom Protokoll mit einer Geschwindigkeit von 180 °/s. durch. Eine detailliertere Beschreibung der Testungen befindet sich im Kapitel 13.2 im Anhang.

5.4 Gütekriterien der Studien

Nachfolgend werden die Studien zum HST und zum ROT mit ihren Werten zu Sensitivität, Spezifität und anderen Gütekriterien aufgelistet. Die Daten können der Tabelle 9 und Tabelle 10 zur Übersicht entnommen werden.

5.4.1 Gütekriterien Kopfschütteltest

Es zeigt sich, dass der HST in der Spezifität einen höheren Wert erreicht als in der Sensitivität. Basierend auf den ausgewählten Primärstudien liegt die Sensitivität des HST bei peripheren Vestibulopathien und somit die Fähigkeit erkrankte Menschen als krank zu identifizieren im Durchschnitt bei 0.49. Nach Cohen (1988) wird dies als mittlerer Effekt eingestuft. Auffallend ist, dass einzig die Studie von Sun et al. (2022) bei der Sensitivität des HST mit VNG einen hohen Wert von 0.936 aufwies. Zwei Studien gaben die Sensitivität des HST bei M. Menière und Neuritis vestibularis an. Die Studie von Angeli et al. (2011) erhielt 0.31, während Asawavichiangianda et al. (1999) einen höheren Wert von 0.839 aufzeigte. Im Vergleich der Werte beider Studien zu Neuritis vestibularis weist Asawavichiangianda et al. (1999) mit 0.75 einen mittleren bis akzeptablen Effekt auf, hingegen Angeli et al. (2011) mit 0.36 einen tiefen Effekt. Angaben zur Spezifität bei diesen Diagnosen wurden keine gemacht. Angeli et al. (2011) gab Ergebnisse zur Sensitivität des HST und der Kalorik (0.31) an. Es wurde eine Gegenüberstellung des HST mit dem SHA bei normalem VOR-Gain (0.59) und abnormalem VOR-Asymmetry (0.71) aufgelistet. Diese Ergebnisse sind nach Cohen (1988) mittlere bis akzeptable Effekte. Wei et al. (1889) zeigen tiefere Resultate mit tiefen Effekten beim Vergleich des HST mit der Kanalparese als Ergebnis einer kalorischen Testung von 0.4. Ebenso erreichte die Sensitivität des unilateralen Hörverlusts in Bezug auf die Kalorik 0.43 und den HST 0.28 die tiefsten Ergebnisse.

Aus den Studien kann ein Wert der Spezifität von 0.7 für periphere Vestibulopathien entnommen werden. Dies stellt einen mittleren bis akzeptablen Effekt dar. Der HST im Vergleich mit der Kalorik erhielt in der Studie von Angeli et al. (2011) eine gute bis exzellente Spezifität von 0.96. Bei der Gegenüberstellung mit dem SHA konnten hohe Werte von 0.83 und 0.85 erreicht werden. Wei et al. (1989) hingegen erhielten in Bezug auf die Kalorik einen tieferen Wert von 0.6. In der Studie von Sun et al. (2022) wurde beim HST mit VNG eine gute Spezifität von 0.727 erreicht. Die Spezifität vom HST in Bezug auf den Head Shaking Tilt Suppression Test wies ebenfalls einen hohen Wert von 0.939 auf. Einzig bei Sun et al. (2022) liegt der Wert der Spezifität unterhalb der Sensitivität.

Pérez et al. (2004) vergleicht die unterschiedlichen Arten der Nystagmen untereinander und kommt dabei auf Werte $p < 0.05$. In der Studie von Wei et al. (1989) wurde ein PPV für den HST von 0.32 und ein NPV von 0.75 evaluiert. Angeli et al. (2011) hingegen verglich bei den Angaben zum positiven prädikativen Wert den HST mit der Kalorik bei einem Wert von 0.97, mit der SHA bei normalem VOR-Gain PPV 0.63 und bei abnormalem VOR-Asymmetry mit dem Resultat von PPV 0.66. Einzig Sun et al. (2022) berechneten den Kappawert der Testung und erhielten dabei ein substanzielles Ergebnis von 0.658. Zusätzlich untersuchte diese Studie den Vergleich vom HST mit dem HSTST, wobei ein Resultat von $p < 0.001$ erzielt wurde und evaluierte mit AUC-ROC Kurven die erfolgreichste Methode.

Tabelle 9: Ergebnisse Studien HST

Angeli et al., 2011		
Sensitivität	Spezifität	Andere Gütekriterien
HST + KAL= 0.31 HST + SHA mit normalem VOR-Gain= 0.59 HST + SHA mit abnormalem VOR-Asymmetry= 0.71 Periphere Vestibulopathie= 0.75 Neuritis vestibularis= 0.36 M. Menière= 0.31	HST + KAL= 0.96 HST + SHA mit normalem VOR-Gain= 0.83 HST + SHA mit abnormalem VOR-Asymmetry= 0.85	Validität PPV: HST + KAL= 0.97 PPV: HST + SHA mit normalem VOR-Gain= 0.63 PPV: HST + SHA mit abnormalem VOR-Asymmetry= 0.66
Pérez et al., 2004		
Sensitivität	Spezifität	Andere Gütekriterien
Unklar Unilaterale periphere Defizite= 0.3-0.4	unklar	Validität Negative HSN vs. Cm-HSN=p<0.01 Negative HSN vs. b-HSN=p<0.01 Im-HSN vs. Cm-HSN=p<0.05 Im HSN vs. b-HSN=p<0.01
Wei et al., 1989		
Sensitivität	Spezifität	Andere Gütekriterien
0.32 <u>HSN +:</u> Genereller Hörverlust= 0.38 Unilateraler Hörverlust= 0.5 CP= 0.4 <u>Unilateraler Hörverlust +:</u> CP= 0.43 HSN= 0.28	0.6 <u>HSN +:</u> Genereller Hörverlust= 0.68 Unilateraler Hörverlust= 0.69 CP= 0.6 <u>Unilateraler Hörverlust +:</u> CP= 0.91 HSN= 0.86	Validität PPV= 0.32 NPV= 0.75
Asawavichiangianda et al., 1999		
Sensitivität	Spezifität	Andere Gütekriterien
Vestibuläre Dysfunktion= 0.383 <u>Sensitivität HSN und/oder KAL</u> M. Menière= 0.839 Neuritis vestibularis= 0.75	Seitendifferenz von: >20: 0.732 >40: 0.717 >60: 0.707 >80: 0.701	-

Burgio et al., 1991		
Sensitivität	Spezifität	Andere Gütekriterien
Unilaterale Läsion= 0.44 Bilaterale Läsion=0.4	Unilaterale Läsion= 0.65	-

Jacobson et al., 1990		
Sensitivität	Spezifität	Andere Gütekriterien
0.27 (73% falsch negativ)	0.85 (15% falsch positiv)	-

Sun et al., 2022		
Sensitivität	Spezifität	Andere Gütekriterien
HST mit VNG= 0.936 HST und HSTST= 0.809	HST mit VNG= 0.727 HST und HSTST= 0.939	<p>Reliabilität <i>Kappa</i> HST=0.658 (p<0.001) <i>Kappa</i> HST + HSTST=0.650 (p<0.001) Vergleich von Bedside-Test und Behandlungsraum (Werte übereinstimmend)</p> <p>Validität HST vs. HSTST unidirektional horizontal=p<0.001</p> <p><i>AUC-ROC Kurven:</i> HST=AUC 0.832/HST+HSTST=AUC 0.874 →VNG HST=AUC 0.801/HST+ HSTST=AUC 0.859 → Video- brille</p>

HST=Kopfschütteltest, HSN=Kopfschüttelnystagmus, KAL=Kalorik, CP=Kanalparese, SHA=Sinusoidal Harmonic Acceleration, HSTST=Head Shaking Tilt Suppression Test, Cm=Contralateral monophasisch, b=biphasisch, Im=ipsilateral monophasisch, VOR=Vestibulo-okular Reflex, PPV=Positiver prädiktiver Wert, NPV=Negativer prädiktiver Wert, VNG=Videonystagmographie,

5.4.2 Gütekriterien Drehstuhltest

Die Studie von Palomar-Asenjo et al. (2006) gibt keine genaue Sensitivität und Spezifität an, lediglich wird die Erkenntnis gezogen, dass die ROT eine höhere Sensitivität als die Kalorik darlegt, diese jedoch eine höhere Spezifität. Die Korrelation der beiden Testungen ist mit einem Kappawert von 0.226 ausreichend. Eza-Nuñez et al. (2016) erreichten im Vergleich zum vKIT einen höheren Kappawert von 0.341. Ahmed et al. (2009) stellten die Sensitivität und Spezifität von Gain, Phase und Symmetry in AUC-ROC Kurven dar, diese zeigen jedoch einen begrenzten prädiktiven Wert von weniger als 0.7 an. Ahmed (2015) unterteilte in seiner Studie die Sensitivität der SHA-Testung in Gain, Phase und Symmetry. Diese zeigen gute bis exzellente Ergebnisse bei Gain 0.799, Phase 0.956 und Symmetry 0.748, die Spezifität schneidet schlechter ab mit Resultaten von Gain 0.445, Phase 0.426 und Symmetry 0.282. Ein besonders guter positiver prädiktiver Wert liegt bei 0.913 in der Phase. Als schlechtesten negativen prädiktiven Wert zeigt sich Gain mit 0.342. Die insgesamt besten diagnostischen Qualitäten hinsichtlich Sensitivität und Spezifität stammen aus den beiden Studien von Maes et al. (2011) und Fattahi et al. (2024). Die Studie von Maes et al. (2011) zeigt eine gute Effektgröße der Sensitivität von 0.839 bei der SHA-Testung 0.01 Hz. und eine moderate Spezifität von 0.636 bei gleicher Frequenz. Die Sensitivität des ROT schneidet somit im Vergleich zur Sensitivität der Kalorik bei 0.828 und VEMP 0.684 besser ab. Ähnliche Resultate in der Studie von Fattahi et al. (2024), die eine Sensitivität der ROT von 0.76 und somit besser als die Sensitivität der Kalorik 0.74 und vKIT 0.47 darlegt. Die Spezifität des ROT liegt jedoch tiefer mit 0.785 im Vergleich zu anderen Testungen. Die höhere Sensitivität im Gegensatz zur Spezifität des ROT kann die Studie von Eza-Nuñez et al. (2016) ebenfalls vorweisen bei Effektgrößen der Sensitivität von 0.747 und Spezifität bei 0.634. Eine ähnliche Sensitivität der ROT ist ebenfalls in der Studie von Arriaga et al. (2005) festzustellen mit einem Wert von 0.71. Die Spezifität etwas tiefer bei 0.54. Die Studie beschreibt zudem den positiven prädiktiven Wert von 0.49 und den negativen prädiktiven Wert von 0.75. Zusammenfassend zeigen die Studien, dass die Kombination vom ROT und der Kalorik die genaueste Diagnose von vestibulären Defiziten ermöglicht. Der SHA-Test bietet besonders präzise Ergebnisse in tiefen Frequenzbereichen und ist nützlich für Personen, die den kalorischen Test nicht tolerieren. Der vKIT scheint vor allem in der Praxis von Vorteil zu sein. Eine Kombination aller genannten Tests liefert die umfassendsten Ergebnisse. Diese Resultate sind der Tabelle 10 zu entnehmen. Eine detailliertere Darstellung ist im Anhang im Kapitel 13.2 dargestellt.

Tabelle 10: Ergebnisse Studien ROT

Palomar-Asenjo et al., 2006			
Sensitivität	Spezifität	Andere Gütekriterien	
Unklar ROT höhere Sensitivität als KAL	Unklar KAL höhere Spezifität als ROT	Reliabilität: Die Zeitkonstante des VOR ist stabiler als die Kanalparese des kalorischen Tests in Test-Retest-Studien (Jenkins & Goldberg, 1988). <i>Kappa: 0.226</i> ROT und KAL Validität: Da beide Testungen unterschiedliche Aspekte des vestibulären Systems untersuchen, ist ihre Validität für die Diagnosestellung von vestibulären Dysfunktionen, (wie bei M. Menière), unterschiedlich hoch. <i>PPV KAL=0.74</i>	
Ahmed et al., 2009			
Sensitivität	Spezifität	Andere Gütekriterien	
Unklar Sensitivität KAL=0.55 →bei Cut-off Wert von 29.5%	Unklar Spezifität KAL=0.84 →bei Cut-off Wert von 29.5%	Validität <i>AUC-ROC Kurven:</i> Die einzelnen Parameter der ROT, wie z. B. die Gain bei 0.025 Hz (AUC-ROC von 0.62), die Phase bei 0.025 Hz (AUC-ROC von 0.58) und die TC (AUC-ROC von 0.61)	
Ahmed, 2015			
Sensitivität	Spezifität	Andere Gütekriterien	
SHA (Referenztest VST) Gain= 0.799 Phase= 0.956 Symmetry= 0.748 KAL=0.683	SHA (Referenztest VST) Gain= 0.445 Phase= 0.426 Symmetry= 0.282 KAL=0.286	Validität SHA Gain <i>PPV=0.794</i> <i>NPV=0.342</i> Phase <i>PPV=0.913</i> <i>NPV=0.435</i> Symmetry <i>PPV=0.572</i> <i>NPV=0.692</i>	KAL <i>PPV=0.756</i> <i>NPV=0.567</i>

Maes et al., 2011			
Sensitivität	Spezifität	Andere Gütekriterien	
SHA 0.01 Hz= 0.839 0.05 Hz= 0.790 0.1 Hz= 0.742 KAL=0.828	SHA 0.01 Hz= 0.636 0.05 Hz= 0.606 0.1 Hz= 0.697 KAL=0.786	Validität SHA PPV 0.01 Hz= 0.684 0.05 Hz= 0.653 0.1 Hz= 0.697	NPV 0.01 Hz= 0.808 0.05 Hz= 0.755 0.1 Hz= 0.742
Fattahi et al., 2024			
Sensitivität	Spezifität	Andere Gütekriterien	
0.760 KAL=0.746 vKIT Gain=0.478 Alle Testungen=0.338	0.785 KAL=0.8351 vKIT Gain=0.898 Alle Testungen=0.987	Validität AUC-ROC Kurven AUC isolierte TC=0.802 AUC alle Testungen=0.869 → KI:0.808-0.931 AUC KAL + TC=0.864 AUC vKIT + TC=0.809	PPV ROT= 0.76 NPV ROT= 0.785
Eza-Nuñez et al., 2016			
Sensitivität	Spezifität	Andere Gütekriterien	
0.747 KAL=0.773 vKIT=0.747 Alle drei Testungen=0.547	0.634 KAL=0.658 vKIT=0.585 Alle drei Testungen=0.878	Reliabilität: Spearmans correlations coefficient (p): →KAL vs. ROT= 0.362 →ROT vs. vKIT= 0.341 Cohen`s k coefficient: →CAL vs. ROT= 0.279 →ROT vs. vKIT= 0.341	Validität PPV ROT= 0.788 NPV ROT= 0.558
Arriaga et al., 2005			
Sensitivität	Spezifität	Andere Gütekriterien	
0.71 ENG=0.31	0.54 ENG=0.86	Validität PPV ROT= 0.49 NPV ROT= 0.75	

ROT=Drehstuhl, KAL=Kalorik, SHA=Sinusoidal Harmonic Acceleration, vKIT=Video-Kopfpulstest,, PPV=Positiver prädiktiver Wert, NPV=Negativer prädiktiver Wert, TC=Zeitkonstante (Time Constant), AUC=Area under curve, ROC=Receiver Operating Characteristics, ENG=Elektronystagmographie,

6 DISKUSSION

Eine präzise Diagnose zur Erkennung eines unilateralen peripheren vestibulären Defizits ist für eine erfolgreiche Behandlung entscheidend und unter anderem werden dafür der Kopfschüttel- und Drehstuhltest verwendet. Die vorliegende Arbeit untersucht, welche dieser Testungen hinsichtlich Reliabilität und Validität die besseren diagnostischen Qualitäten bietet. Trotz der weitverbreiteten Anwendung scheint unklar zu sein, welche der beiden Tests bei der Diagnosestellung verlässlichere und genauere Ergebnisse liefert. Diese Unsicherheit erschwert die Entscheidung für die optimale Testung in der Praxis. Das Ziel dieses systematischen Literaturreviews war es, die diagnostische Qualität des HST und der ROT bei der Diagnose unilateraler peripherer vestibulärer Defizite zu vergleichen. Zur Beantwortung der Fragestellung, wurde eine Literaturrecherche durchgeführt, die zu einem Ergebnis von sieben Studien pro Testung führte. Die 14 englischsprachigen Studien von retrospektiver als auch prospektiver Natur stammen aus den Jahren 1989 bis 2024, wobei die Studien zum HST tendenziell älter sind. Trotz häufiger Anwendung in der Praxis (Cohen, 2019), scheint die Anzahl von sieben Studien pro Testung den Autorinnen sehr gering und somit sind auch die diagnostischen Qualitäten beschränkt verfügbar.

Selten hat eine der 14 Studien die Reliabilität genau beschrieben, obwohl diese als wichtigstes Gütekriterium für die Qualität der Testung gilt (Priebe, 2024). Lediglich die Studie von Jenkins & Goldberg, (1988) führte eine Test-Retest Analyse durch, allerdings an gesunden Personen, weshalb sie nicht in die Arbeit eingeschlossen wurde. Diese Studie zeigte eine hohe Reliabilität für die VOR-Phase und VOR-Gain, während die VOR-Asymmetrie die grösste Variabilität aufwies (Jenkins & Goldberg, 1988).

Ältere Studien verglichen die Testungen oftmals mit der Kalorik, welche früher als Referenztest galt (Maes et al., 2011). Die Kalorik erreichte in den eingeschlossenen Studien Werte der Sensitivität zwischen 0.6 und 0.828 und Spezifität von 0.286 bis 0.835. Dies stellt eine hohe Variabilität dar und wertet die Aussagekraft der Testung für die Identifikation von gesunden Personen als gering. Wie bei der ROT untersucht die Kalorik jedes Ohr einzeln (Baloh et al., 1979), zudem wird ein Nystagmus bei der Kalorik und dem ROT als normal gewertet, während er beim HST als pathologisch gilt (Wei et al., 1989). Alle drei Testungen untersuchen unilaterale Hypofunktionen und prüfen dabei den horizontalen Bogengang. Heutzutage wird jedoch der vKIT als Standardtest in der Vestibulariagnostik empfohlen (Heim, 2024). Ein Vergleich zum vKIT wurde bei den Studien zum HST nicht durchgeführt und beim ROT einzig in den Studien von Fattahi et al. (2024) und Eza-Nuñez et al. (2016) angewendet. Der vKIT zeigt gute Werte als alleinige Testung, jedoch ist die Korrelation zwischen ROT und vKIT gering, was sich

durch die Testung unterschiedlicher Frequenzen erklären lässt. Die Kalorik, der vKIT und ROT zeigen zusammen die beste Spezifität von >0.8 und weisen somit einen akzeptablen bis exzellenten Effekt auf.

Der Durchschnittswert der Sensitivität aller HST-Studien liegt unter der Spezifität. Nur die aktuellste Studie von Sun et al. (2022) erreichte eine exzellente Sensitivität. Dies könnte damit begründet werden, dass einzig in dieser Studie mit dem HSTST verglichen sowie die Auswertung mit einer VNG durchgeführt wurde, was eine genauere Erkennung von Nystagmen ermöglichte. Beim Vergleich des HST mit dem HSTST wurde ein statistisch signifikantes Ergebnis von $p < 0.001$ erzielt. Somit gilt der HST als empfindlicher. Es zeigt sich in den ROC-AUC Kurven, dass der HST mit VNG im Untersuchungsraum einen höheren Wert erreicht als die Bedside-Testung mit der Videobrille. Die Testdurchführung mit Videobrille ist weniger aussagekräftig als mit VNG, doch die Differenz ist gering, und auch der Bedside-Test liefert gute Ergebnisse. Die unterschiedlichen Sensitivitätswerte bei M. Menière und Neuritis vestibularis könnten auf die erheblichen Unterschiede in der Stichprobengrösse zurückzuführen sein. Während Asawavichiangianda et al. (1999) mit 1364 die grösste und Angeli et al. (2011) mit 53 die kleinste Population untersuchte, beeinflussen auch unterschiedliche Testdurchführungen und Einschlusskriterien die Ergebnisse. Zudem verwendeten Angeli et al. (2011) eine Frenzelbrille, während Asawavichiangianda et al. (1999) auf ein ENG setzte. Angeli et al. (2011) wird aufgrund der späteren Veröffentlichung und der besseren Bewertung gemäss der «COSMIN checklist» bevorzugt. Pérez et al. (2004) konnte aufzeigen, dass die verschiedenen Arten der Nystagmen bei peripheren vestibulären Läsionen statistische Werte von $p < 0.05$ aufweisen und sich somit grundsätzlich unterscheiden lassen. Die Bewertung und der Vergleich des HST werden durch die Anwendung unterschiedlicher Testkriterien in den Studien erschwert, was auch die Vergleichbarkeit unter den Studien beeinträchtigt. Auffällig ist, dass neuere Studien die passive Testausführung bevorzugen, ohne dies näher zu begründen. Aufgrund der begrenzten Datenlage lässt sich nicht eindeutig bestimmen, welche Methode sich besser eignet. Die eingeschlossenen Studien deuten jedoch darauf hin, dass der passive Test bei HWS-Problemen eher zu empfehlen ist. Allerdings empfehlen Gimeno-Vilar et al. (2007) den aktiven HST mit VNG. Die unterschiedlichen Geschwindigkeiten und Zeitspannen der Durchführung des HST führte in den sieben Studien zu erschwelter Vergleichbarkeit, da nicht eindeutig festgestellt werden kann, ob die beobachteten Effekte auf den Test selbst oder die Ausführung zurückzuführen sind. Dies macht eine Verallgemeinerung schwierig und behindert eine klare Interpretation der Testwirksamkeit.

In den meisten Studien zum ROT zeigte sich eine höhere Sensitivität als Spezifität, wodurch mehr Personen mit unilateralen peripheren Defiziten identifiziert werden können. Die durchschnittliche Sensitivität liegt bei einer mittleren bis akzeptablen Effektgrösse von 0.74, die Spezifität hingegen variiert zwischen 0.2 und 0.78. Die Studie von Ahmed (2015) unterteilte die Sensitivität und Spezifität in die Parameter Phase, Gain und Symmetry. Die Phase erzielte mit einer Sensitivität von 0.959 einen guten bis exzellenten Wert. Das lässt sich dadurch erklären, dass er als besonders resistent gegen zentrale Kompensationen gilt und somit die zuverlässigsten Ergebnisse liefert (Wolfe et al., 1978). Die Studien zur ROT zeigen, dass das Protokoll des SHA eine bessere diagnostische Qualität als der VST aufweist. Dies liegt vermutlich daran, dass der SHA verschiedene Frequenzen testet und somit ein breiteres Spektrum des Vestibularorgans untersucht. Im Gegensatz dazu testet der VST, nur bei einer Geschwindigkeit von 100 °/s und erfasst somit weniger Aspekte des vestibulären Systems. Allerdings ist der VST in der Praxis einfacher anzuwenden. In vier von sieben Studien wurden während der Testung mentale Aufgaben gestellt, was darauf hindeutet, dass die Ablenkung des Kortex durch diese Aufgaben zu einer besseren Erfassung der vestibulären Funktion führt (Bronstein et al., 2015).

Um die Fragestellung zu beantworten ist ein Vergleich der beiden Testungen notwendig. Die Literaturanalyse ergab, dass die Publikationsjahre der Studien stark variieren. Basierend auf den grossen Unterschieden zwischen den Jahren, ist eine Änderung des Wissenstandes über die Physiologie und die Testungen nicht auszuschliessen. Zu sehen ist dies bei der Testausführung, den Materialien und Referenztests. Dies erschwert die Bedingungen für einen optimalen Vergleich. Ebenso liefen teils Studien über sechs Jahre, während andere keine Angaben dazu machten. Auch die Stichprobengrössen variierten erheblich. Kleine Stichproben erschweren es signifikante und präzise Effekte aufzuzeigen und sind anfälliger für Zufallseffekte. Bei grösserer Population könnten mehr Nystagmen erfasst werden, was die Ergebnisse beeinflussen würde. Über alle Studien hinweg sind die Stichproben eines heterogenen Ursprungs. Einige Studien nannten keine Diagnosen und sprachen nur allgemein von peripheren Defiziten, andere schlossen eine Vielzahl an Diagnosen ein. Neuritis vestibularis und M. Menière sind in allen HST-Studien vorzufinden. In den Studien zum ROT ist jeweils eine der beiden Diagnosen vorhanden. Ebenso unterscheiden sich die Kontrollgruppen, die entweder keine Beschwerden, bilaterale oder zentrale Vestibulopathien vorwiesen. Zudem untersuchten die Studien zu beiden Testungen den Nystagmus mit anderen Hilfsmitteln, wobei ältere Studien eher ENG oder die Frenzelbrille und neuere Studien VNG nutzten. Die VNG hat eine sehr gute Aufnahmequalität im Gegensatz zur Frenzelbrille, was die Nystagmuserkennung unterscheiden kann. Diese Heterogenität

zwischen den Gruppen, aber auch innerhalb der Gruppen erschwert einen Vergleich der beiden Testungen sowie die Interpretation der Ergebnisse.

In der Literatur wird der HST als nicht brauchbar beschrieben (Cohen, 2019). Mandalà et al. (2008) empfehlen die Testung jedoch in der Akutphase zur Bedside-Diagnostik. Die ROT hingegen wird eher in der akuten oder chronischen Phase durchgeführt. In der akuten Phase sind die beiden Parameter Gain (Arriaga et al., 2005) und Symmetry (Baloh et al., 1989) oft abgeschwächt, während in der chronischen Phase bereits zentrale Kompensationen auftreten. Daher sind diese Parameter in tieferen Frequenzen zentral kompensiert, während die Asymmetry in höheren Frequenzen weiterhin erkennbar bleibt (Baloh et al., 1989). In der chronischen Zeitspanne bleibt die Phase der aussagekräftigste Parameter (Wolfe et al., 1978). In der subakuten Phase wird die ROT nicht empfohlen, da zu starke Reize die Symptomatik verschlechtern können (Arriaga et al., 2005). Trotz diesen Daten fehlten in den Studien spezifische Hinweise zur Krankheitsphase der Patientinnen und Patienten. Einzig Angeli et al. (2011) legte als Einschlusskriterium fest, dass der Schwindel oder die Gleichgewichtsstörung seit mindestens zwei Monaten bestehen muss, was auf eine akute-chronische Phase hindeutet. Pérez et al. (2004) erwähnten, dass sich die Patientinnen und Patienten in unterschiedlichen Phasen befanden, was den Vergleich der Resultate erschwert. Vergleicht man den HST mit der ROT, zeigt sich, dass der HST eine höhere Spezifität, das SHA-Protokoll jedoch eine bessere Sensitivität aufweist. Nur Angeli et al. (2011) und Pérez et al. (2004) berücksichtigten in ihren Studien den HST und die ROT. Angeli et al. (2011) zeigte, dass bei der Gegenüberstellung des HST mit SHA bei abnormalem VOR-Asymmetry höhere Werte in Sensitivität und Spezifität erreicht wurden als beim Vergleich mit dem SHA bei normalem VOR-Gain. Dies kann darauf zurückgeführt werden, dass der HST ebenfalls auf Asymmetrien ausgelegt ist.

Ein wesentlicher Unterschied zwischen den beiden Testungen liegt in der Standardisierung. Der ROT wird auf einem computergesteuerten Stuhl durchgeführt, wobei die betroffenen Personen immer die gleiche ASTE einnehmen. Hingegen ist die Standardisierung für den HST erschwert, da die Testung nicht immer gleich oder von der gleichen Person durchgeführt werden kann. Ein weiterer Unterschied stellt die Beteiligung der HWS dar. Beim HST wird die HWS aktiv oder passiv belastet, während bei der ROT der gesamte Körper gedreht wird, was andere Afferenzen aus der HWS aktiviert und schonender für die Betroffenen zu sein scheint. Liegen HWS-Problematiken vor, soll dieser Unterschied ebenfalls ein wichtiges Merkmal darlegen und unterstützend im Entscheidungsprozess wirken, ob eine Therapeutin oder ein Therapeut sich für oder gegen eine Testung entscheidet. Dannenbaum et al. (2005) betonten im Kapitel 2.3.4 zum DVAT die Wichtigkeit einer vorherigen Abklärung der Kontraindikationen für die

Kopfrotationen, insbesondere im Bereich der HWS. Keine der 14 Studien hat jedoch spezifische Kontraindikationen für die beiden Testungen erwähnt oder untersucht. Vor allem beim HST können HWS-Pathologien starke negative Auswirkungen haben. Daher wird empfohlen, Kontraindikationen vor den Tests frühzeitig abzuklären.

Das Prinzip der Erkennung der betroffenen Seite erfolgt ebenfalls unterschiedlich, denn Angeli et al. (2011) zeigen, dass der HST-Asymmetrien der Gleichgewichtsorgane aufdeckt, da beide Seiten gleichzeitig getestet werden. Das Push-pull-Prinzip wirkt dabei abwechselnd auf beide Seiten. Im Gegensatz zum ROT, der aufgrund des Push-pull-Prinzips und der Cupulaauslenkung jedes Ohr einzeln prüft, obwohl beide Labyrinth gleichzeitig stimuliert werden.

Zudem untersuchen die beiden Testungen unterschiedliche Frequenzen (siehe Kapitel 2.1.3, Abbildung 2), was eine Aussage nicht vergleichbar macht. Der Drehstuhl ist fähig ein breites Spektrum des Vestibularsystems abzudecken, hingegen der HST in hohen Frequenzen zwischen 1 Hz und 3 Hz testet. Daraus lässt sich schliessen, dass kleine Läsionen mit dem HST nicht erkannt werden. Der SHA-Test kann bis zu einer niedrigen Frequenz von 0.01 Hz prüfen und weist dabei die höchste Sensitivität auf (Maes et al., 2011). Ebenfalls wird die Beurteilung des Nystagmus anders gehandhabt, da bei der ROT drei Parameter verwendet werden (Kapitel 2.3.6), während beim HST hauptsächlich die schnelle Phase des Nystagmus bewertet wird. Bei beiden Testungen wird eine Kopfflexion von 30° zur horizontalen Einstellung des lateralen Bogengangs erwähnt. Die älteren HST-Studien hielten die Kopfflexion nicht ein. Dies deutet darauf hin, dass der Wissensstand über die optimale Ausrichtung des lateralen Bogengangs früher geringer war. Für die optimale Erfassung des horizontalen Bogengangs halten die Autorinnen dieser Arbeit eine Kopfflexion von 30° für sinnvoller (siehe Kapitel 2.1.1). Zur Aufhebung der Blickfixation wurden die Drehstuhl- sowie auch die Kopfschütteltestung in abgedunkelten oder dunklen Räumen getestet. Anpassungen wie zentrale Kompensationen spielen eine wichtige Rolle in der vestibulären Diagnostik. Während die zentrale Kompensation bei ROT, wie von Eza-Nuñez et al. (2016) beschrieben, häufig berücksichtigt wird, findet sie in HST-Studien keine Erwähnung. Sie ist die Fähigkeit des zentralen Nervensystem (ZNS), vestibuläre Dysfunktionen auszugleichen und trotz Schäden am Gleichgewichtsorgan die Körperfunktion zu erhalten (Maes et al., 2011). Diese Mechanismen äussern sich hauptsächlich in Veränderungen der Reflexantwort oder neuronalen Verarbeitung, was darauf hindeutet, dass die Sensitivitäts- und Spezifitätswerte durch Kompensation beeinflusst sein könnten. Die Studie von Fattahi et al. (2024) zeigte keine statistischen signifikanten Unterschiede in der zentralen Diagnostik und führt dies auf zentrale Kompensation und andere Faktoren zurück. Eine Kombination von Tests kann helfen, die Auswirkungen der Kompensation besser zu

verstehen und genauere Diagnosen zu stellen (Fattahi et al., 2024). Ahmed (2015) betont die zentrale Kompensation als entscheidend für den Erholungsprozess. Der SHA-Test kann durch die drei Parameter Anomalien aufzeigen, die auf Kompensationen hindeuten (siehe Kapitel 2.3.6.1). Durch Testungen in verschiedenen Frequenzbereichen wird bewertet, wie gut das ZNS periphere Defizite ausgleicht. Diese Daten können genutzt werden, um das Ausmass der zentralen Kompensation zu beurteilen und mögliche Rehabilitationsstrategien zu entwickeln. Die ROT könnte als Verlaufparameter zur Überwachung der Kompensation dienen. Durch die Identifizierung spezifischer Defizite können Klinkerinnen und Kliniker Rehabilitationsprogramme anpassen, um die zentrale Kompensation zu erhöhen und die Ergebnisse der Patientinnen und Patienten zu verbessern.

Die ROT-Testung wird als objektiver diagnostischer Test beschrieben, der von spezialisierten Fachpersonen durchgeführt werden muss (Cohen, 2019). Hingegen der HST, der ausschliesslich als Screening-Test anzuwenden ist und von ausgebildeten Physiotherapeutinnen und Physiotherapeuten durchgeführt werden kann (Mandalà et al., 2008). Ein direkter Vergleich der Tests lässt sich somit nicht machen. Die beiden Testungen weisen komplementäre Stärken auf. Während der HST viele gesunde Personen identifizieren kann, ist die ROT stärker beim Erkennen der Betroffenen. Ebenfalls gilt Letztere als Referenztest bei bilateralen peripheren vestibulären Läsionen (Baloh et al., 1979).

Es sind weitere Forschungen zu den diagnostischen Qualitäten der beiden Testungen erforderlich, um genauere Aussagen tätigen zu können. Primär Reliabilitätsstudien, die Test-Retest Studien bei unilateralen peripheren Defiziten durchführen. Sensitivität und Spezifität allein machen keine verlässliche Aussage darüber, wie zuverlässig die Messinstrumente bei wiederholter Anwendung sind. Eine mögliche Herausforderung ist, dass das getestete Merkmal unverändert bleiben muss (Wirz et al., 2014). Da jedoch mögliche Anpassungen des Vestibularsystems bei oder nach jeder Testung stattfinden könnten, besteht die Möglichkeit einer Veränderung des getesteten Merkmals, was wiederum die Zuverlässigkeit einer Testung beeinträchtigen kann.

Zusammenfassend lässt sich sagen, dass Schwindel ein sehr komplexes und individuelles Symptom darstellt. Die sogenannten diagnostischen Testungen dienen im Schwindelbereich hauptsächlich zur Quantifizierung der komplexen Vestibularfunktionen (Arriaga et al., 2005). Studien legen nahe, dass beide Testungen allein nicht zur Diagnose eines unilateralen peripheren vestibulären Defizits geeignet sind, aber als ergänzende Tests nützlich sein können. Eine Praxisempfehlung zur Durchführung der beiden Testungen sowie den positiven und negativen Aspekten wird nachfolgend in Kapitel 8 genauer erläutert.

6.1 Stärken dieser Arbeit

Im Rahmen dieser Bachelorarbeit wurde eine systematische Literatursuche über mehrere Wochen durchgeführt und somit ein grosses Spektrum an Studien abgedeckt. Die Auflistung in den «PRISMA-Flowcharts» ermöglicht eine transparente Nachvollziehbarkeit der Suche. Durch die Nutzung zahlreicher Datenbanken und Schlüsselwörter konnte ein breites Spektrum an Studien erreicht werden. Die detaillierte Beschreibung der einbezogenen Studien erhöht die Transparenz der Arbeit. Auch wurde eine Vielzahl unterschiedlicher Literaturquellen verwendet. Die Analyse zeigt, dass weiterer Forschungsbedarf in diesem Bereich besteht.

6.2 Limitationen

Viele der zitierten Studien sind veraltet und die heterogenen Gruppen in den Studien erschweren eine pauschale Verallgemeinerung der Ergebnisse, da oft keine Informationen zur Dauer der Symptomatik und zu den Testmethoden gegeben werden. Zudem wurden diagnostische Qualitäten und Gütekriterien häufig nicht erwähnt, was die Zuverlässigkeit beeinträchtigt. Aufgrund zeitlicher Limitationen konnten nicht alle Stichwortkombinationen in allen Datenbanken geprüft werden. Ausserdem sind die meisten Studien nicht ideal für eine COSMIN-Bewertung, da Test-Retest-Studien fehlen, was die Aussagekraft der Bewertung einschränkt.

7 SCHLUSSFOLGERUNG

Das Ziel dieses systematischen Literaturreviews war es herauszufinden, welche der beiden Testungen HST oder ROT die besseren diagnostischen Qualitäten in Bezug auf die Diagnose unilaterale periphere vestibuläre Defizite aufweist. Der HST zeigt moderate Spezifität und somit die Fähigkeit, gesunde Personen als tatsächlich gesund zu identifizieren. Die Testung ist als Screening-Test bzw. Bedside-Test geeignet, um einen groben Überblick über die aktuelle Funktion des Vestibularorgans zu erhalten, jedoch nicht geeignet zur alleinigen Diagnosestellung eines unilateralen peripheren vestibulären Defizits. Hingegen zeigt die ROT insgesamt bessere diagnostische Qualitäten, primär das Protokoll des SHA, in der Sensitivität und somit die Fähigkeit, von unilateralen peripheren vestibulären Defiziten betroffenen Personen, zu identifizieren. Im Gegensatz jedoch schlechtere Werte der Spezifität, was es dem Test erschwert, gesunde Personen als tatsächlich gesund zu identifizieren. Die ROT kann zur diagnostischen Anwendung für ein unilaterales Defizit verwendet werden, obwohl der ROT als Referenztest für bilaterale Vestibulopathien gilt (Baloh et al., 1979). Zudem wird empfohlen, die ROT nicht als alleinstehende Testung anzuwenden (Baloh et al., 1989). Aus folgenden Gründen, lässt sich jedoch kein Vergleich zwischen den beiden Testungen machen: Es werden andere Frequenzen des Vestibularorgans getestet, beim HST liegt eine HWS Beteiligung vor und beim ROT nicht, der HST testet beide Ohren gleichzeitig, der ROT jeweils ein Ohr und dies vermehrt nach dem Push-pull-Prinzip, die Auswertung des Nystagmus wird beim HST anhand der schnellen Phase der Richtung definiert und beim ROT kommen weitere Parameter hinzu wie Gain, Phase, Asymmetry und Zeitkonstante. Die diagnostischen Qualitäten wie Reliabilität, Validität und Objektivität sind aufgrund geringer Studien- und Datenlage schwierig einzuschätzen. Das Protokoll des SHA weist die besseren diagnostischen Qualitäten bei der Identifizierung eines unilateralen peripheren vestibulären Defizits auf als der aktive oder passive HST, der gemäss unserer Forschung, nicht valide zu sein scheint. Es sind somit weitere Untersuchungen notwendig, die genauere Daten über die diagnostischen Qualitäten liefern, insbesondere Test-Retest Studien bei betroffenen Personen mit einem unilateralen peripheren Defizit. Dies scheint wichtig, wenn die beiden Testungen, vor allem jedoch der HST, weiterhin in der Praxis angewendet werden sollen.

8 PRAXISEMPFEHLUNG

Die Autorinnen der vorliegenden Arbeit geben anhand der Zusammenfassung der insgesamt 14 Studien für die Kopfschüttel- und die Drehstuhltestung für Praktikerinnen und Praktiker im Schwindelbereich folgende Empfehlung ab. Eine vorherige Abklärung der Kontraindikationen wird vorausgesetzt (Dannenbaum et al., 2005).

Kopfschütteltest

Aufgrund der geringen Studienlage wird der passive Kopfschütteltest empfohlen, da dieser am häufigsten durchgeführt wurde. Mit einer Kopfflexion von 30° in einem dunklen Raum, einer Frequenz von 1-3 Hz, im Radius von 45° und einer Dauer von 10-20 Sekunden mit einer VNG.



Wenig Aufwand, Material und Kosten und ist von allen Praktikerinnen und Praktikern anwendbar (Angeli et al., 2011). Ermöglicht einen groben Überblick, ob eine mögliche Asymmetrie besteht.



Variationen in der Kopfbewegung machen den Test nicht standardisierbar. Nicht geeignet bei vorbestehenden HWS-Problematiken sowie zur Diagnostik. Es werden keine einheitlichen Durchführungen beschrieben und eine VNG-Brille zur Aufzeichnung der Augenbewegung ist mit höheren Kosten verbunden (Sun et al., 2022).

Drehstuhltestung

Obwohl das Protokoll des Sinusoidal Harmonic Acceleration (SHA) die besseren diagnostischen Qualitäten aufweist, ist die Anwendbarkeit des SHA erschwert. Einfachheitshalber wird deshalb das Protokoll des Velocity Step Test (VST) für die Praxis empfohlen. Zur optimalen Durchführung wird mit 30° Kopfflexion in der Dunkelheit und mit geöffneten Augen getestet. Die VNG oder Elektroden werden zur Augenaufzeichnung empfohlen, zudem müssen während der Drehung mentale Aufgaben gestellt werden.



Eine standardisierte Durchführung mit Frequenzen über ein breites Spektrum, zudem ist er bei Kindern und betroffenen Personen, die vestibulotoxische Medikamente einnehmen, besser tolerierbar als die Kalorik (Ahmed et al., 2009). Der Parameter Phase stellt den aussagekräftigsten Wert dar (Wolfe et al., 1978).



Für die subakute Phase nicht geeignet, da Betroffene sehr starken Reizen ausgesetzt werden. Es wird teures Equipment benötigt und der Zeitaufwand ist gross (Zuniga & Adams, 2021). Es erfordert viel Platz, die nötige Infrastruktur und benötigt zusätzliche Ausbildungen. Der Test ist nicht zur alleinigen Diagnosestellung geeignet (Eza-Nuñez et al., 2016). Mögliches Auftreten von Irritationen nach der Testdurchführung.

9 LITERATURVERZEICHNIS

- Aaronson, N., Alonso, J., Burnam, A., Lohr, K. N., Patrick, D. L., Perrin, E., & Stein, R. E. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 11(3), 193–205. <https://doi.org/10.1023/a:1015291021312>
- Adamec, I., Krbot Skorić, M., Handžić, J., & Habek, M. (2015). Incidence, seasonality and comorbidity in vestibular neuritis. *Neurological Sciences*, 36(1), 91–95. <https://doi.org/10.1007/s10072-014-1912-4>
- Ahmed, M. F. (2015). Rotational chair testing in patients with unilateral peripheral vestibular disorders. *The Egyptian Journal of Otolaryngology*, 31(2), 115–121. <https://doi.org/10.4103/1012-5574.156095>
- Ahmed, M. F., Goebel, J. A., & Sinks, B. C. (2009). Caloric Test Versus Rotational Sinusoidal Harmonic Acceleration and Step-Velocity Tests in Patients With and Without Suspected Peripheral Vestibulopathy. *Otology & Neurotology*, 30(6), 800–805. <https://doi.org/10.1097/MAO.0b013e3181b0d02d>
- Angeli, S. I., Velandia, S., & Snapp, H. (2011). Head-shaking nystagmus predicts greater disability in unilateral peripheral vestibulopathy. *American Journal of Otolaryngology*, 32(6), 522–527. <https://doi.org/10.1016/j.amjoto.2010.11.004>
- Arbusow, V., Derfuss, T., Held, K., Himmelein, S., Strupp, M., Gurkov, R., Brandt, T., & Theil, D. (2010). Latency of herpes simplex virus type-1 in human geniculate and vestibular ganglia is associated with infiltration of CD8+ T cells. *Journal of Medical Virology*, 82(11), 1917–1920. <https://doi.org/10.1002/jmv.21904>
- Arriaga, M. A., Chen, D. A., & Cenci, K. A. (2005). Rotational Chair (ROTO) Instead of Electronystagmography (ENG) as the Primary Vestibular Test. *Otolaryngology–Head and Neck Surgery*, 133(3), 329–333. <https://doi.org/10.1016/j.otohns.2005.05.002>

- Asawavichiangianda, S., Fujimoto, M., Mai, M., Desroches, H., & Rutka, J. (1999). Significance of Head-shaking Nystagmus in the Evaluation of the Dizzy Patient. *Acta Oto-Laryngologica*, 119(540), 27–33. <https://doi.org/10.1080/00016489950181152>
- Baloh, R. W. (2002). Robert Bárány and the controversy surrounding his discovery of the caloric reaction. *Neurology*, 58(7), 1094–1099. <https://doi.org/10.1212/WNL.58.7.1094>
- Baloh, R. W., Jacobson, K. M., Beykirch, K., & Honrubia, V. (1989). Horizontal vestibulo-ocular reflex after acute peripheral lesions. *Acta Oto-Laryngologica Supplementum*, 468, 323–327. <https://doi.org/10.3109/00016488909139069>
- Baloh, R. W., Sills, A. W., & Honrubia, V. (1979). Impulsive and sinusoidal rotatory testing: A comparison with results of caloric testing. *The Laryngoscope*, 89(4), 646–654. <https://doi.org/10.1288/00005537-197904000-00013>
- Blödown, A., Bloching, M., Hörmann, K., & Walther, L. E. (2012). Rezeptorfunktion der Bogengänge: Teil 2: Pathophysiologie, Erkrankungen, klinische Befunde und therapeutische Aspekte. *HNO*, 60(3), 249–262. <https://doi.org/10.1007/s00106-011-2438-1>
- Bojrab, D. I., Lai, W., & Bojrab, D. I. (2019). Electronystagmography and Videonystagmography. In S. Babu, C. A. Schutt, & D. I. Bojrab (Hrsg.), *Diagnosis and Treatment of Vestibular Disorders* (S. 45–65). Springer International Publishing. https://doi.org/10.1007/978-3-319-97858-1_4
- Brandt, T., Huppert, Td., Hübner, K., Zingler, V. C., Dieterich, M., & Strupp, M. (2010). Long-term course and relapses of vestibular and balance disorders. *Restorative Neurology and Neuroscience*, 28(1), 69–82. <https://doi.org/10.3233/RNN-2010-0504>
- Bronstein, A. M., Patel, M., & Arshad, Q. (2015). A brief review of the clinical anatomy of the vestibular-ocular connections—How much do we know? *Eye*, 29(2), 163–170. <https://doi.org/10.1038/eye.2014.262>

- Burgio, D. L., Blakley, B. W., & Myers, S. F. (1991). An Evaluation of the Head-Shaking Nystagmus Test. *Otolaryngology–Head and Neck Surgery*, *105*(5), 708–713. <https://doi.org/10.1177/019459989110500512>
- Cohen, H. S. (2019). A review on screening tests for vestibular disorders. *Journal of Neurophysiology*, *122*(1), 81–92. <https://doi.org/10.1152/jn.00819.2018>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Routledge. <https://doi.org/10.4324/9780203771587>
- Dannenbaum, E., Hakim-Zadeh, R., Paquet, N., & Feldman, A. G. (2005). Optimal Parameters for the Clinical Test of Dynamic Visual Acuity in Patients with a Unilateral Vestibular Deficit. *The Journal of Otolaryngology*, *34*(01), 13. <https://doi.org/10.2310/7070.2005.03105>
- Day, B. L., & Fitzpatrick, R. C. (2005). The vestibular system. *Current Biology*, *15*(15), R583–R586. <https://doi.org/10.1016/j.cub.2005.07.053>
- Erdinest, N., & London, N. (2022). Dynamic visual acuity and methods of measurement. *Journal of Optometry*, *15*(3), 247–248. <https://doi.org/10.1016/j.optom.2021.06.003>
- Eza-Nuñez, P., Fariñas-Alvarez, C., & Fernandez, N. P. (2016). Comparison of three diagnostic tests in detecting vestibular deficit in patients with peripheral vestibulopathy. *The Journal of Laryngology & Otology*, *130*(2), 145–150. <https://doi.org/10.1017/S0022215115003114>
- Fattahi, C. B., Zaro, C., Chung, J. J., Lewis, R. F., & Chari, D. A. (2024). Comparative utility of vestibular function tests in patients with peripheral and central vestibular dysfunction. *Journal of Otology*, *19*(1), 5–9. <https://doi.org/10.1016/j.joto.2023.10.002>
- Fetter, M. (2010). Caloric testing. In *Handbook of Clinical Neurophysiology* (Bd. 9, S. 135–140). Elsevier. [https://doi.org/10.1016/S1567-4231\(10\)09010-6](https://doi.org/10.1016/S1567-4231(10)09010-6)

- Fetter, M. (2016). Acute unilateral loss of vestibular function. In *Handbook of Clinical Neurology* (Bd. 137, S. 219–229). Elsevier. <https://doi.org/10.1016/B978-0-444-63437-5.00015-7>
- Gimeno-Vilar, C., Rey-Martinez, J., & Perez, N. (2007). Active versus passive head-shaking nystagmus. *Acta Oto-Laryngologica*, 127(7), 722–728. <https://doi.org/10.1080/00016480600987800>
- Goebel, J. A., Isipradit, P., & Hanson, J. M. (2000). Manual Rotational Testing of the Vestibulo-ocular Reflex. *The Laryngoscope*, 110(4), 517–535. <https://doi.org/10.1097/00005537-200004000-00004>
- Gonçalves, D. U., Felipe, L., & Lima, T. M. A. (2015). Interpretation and use of caloric testing. *Brazilian Journal of Otorhinolaryngology*, 74(3), 440–446. [https://doi.org/10.1016/S1808-8694\(15\)30580-2](https://doi.org/10.1016/S1808-8694(15)30580-2)
- Havia, M., Kentala, E., & Pyykkö, I. (2005). Prevalence of Menière's disease in general population of Southern Finland. *Otolaryngology--Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, 133(5), 762–768. <https://doi.org/10.1016/j.otohns.2005.06.015>
- Heim, T. (2024). Unilaterale Vestibulopathie sicher diagnostizieren und effektiv behandeln. *InFo Neurologie + Psychiatrie*, 26(4), 60–61. <https://doi.org/10.1007/s15005-024-3940-9>
- Helmchen, C. (2021). *Augenbewegungsstörungen inkl. Nystagmus—Deutsche ... Deutsche Gesellschaft für Neurologie.* <https://de.readkong.com/page/augenbewegungsstorungen-inkl-nystagmus-deutsche-5817924>
- Jacobson, G. P., Newman, C. W., & Safadi, I. (1990). Sensitivity and Specificity of the Head-Shaking Test for Detecting Vestibular System Abnormalities. *Annals of Otolaryngology, Rhinology & Laryngology*, 99(7), 539–542. <https://doi.org/10.1177/000348949009900708>

- Jenkins, H. A., & Goldberg, J. (1988). Test-retest reliability of the rotatory test in normal subjects. *Advances in Oto-Rhino-Laryngology*, 41, 190–195.
<https://doi.org/10.1159/000416055>
- Kamei, T. (1988). Two Types of Head-Shaking Tests in Vestibular Examination. *Acta Oto-Laryngologica*, 105(sup458), 108–112.
<https://doi.org/10.3109/00016488809125112>
- Kubbe, I. (2016a). *Experimente in der Politikwissenschaft*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-09424-9>
- Kubbe, I. (2016b). Gütekriterien experimenteller Forschung. In I. Kubbe (Hrsg.), *Experimente in der Politikwissenschaft: Eine methodische Einführung* (S. 61–76). Springer Fachmedien. https://doi.org/10.1007/978-3-658-09424-9_4
- Lee, H., Kim, B.-K., Park, H.-J., Koo, J.-W., & Kim, J.-S. (2009). Prodromal dizziness in vestibular neuritis: Frequency and clinical implication. *Journal of Neurology, Neurosurgery & Psychiatry*, 80(3), 355–356.
<https://doi.org/10.1136/jnnp.2008.155978>
- Mackowetzky, K., Yoon, K. H., Mackowetzky, E. J., & Waskiewicz, A. J. (2021). Development and evolution of the vestibular apparatuses of the inner ear. *Journal of Anatomy*, 239(4), 801–828. <https://doi.org/10.1111/joa.13459>
- Maes, L., Vinck, B. M., Wuyts, F., D’Haenens, W., Bockstael, A., Keppler, H., Philips, B., Swinnen, F., & Dhooge, I. (2011). Clinical usefulness of the rotatory, caloric, and vestibular evoked myogenic potential test in unilateral peripheral vestibular pathologies. *International Journal of Audiology*, 50(8), 566–576.
<https://doi.org/10.3109/14992027.2011.576706>
- Mandalà, M., Nuti, D., Broman, A. T., & Zee, D. S. (2008). Effectiveness of careful bedside examination in assessment, diagnosis, and prognosis of vestibular neuritis. *Archives of Otolaryngology--Head & Neck Surgery*, 134(2), 164–169.
<https://doi.org/10.1001/archoto.2007.35>

- Mangold, S. (2013). *Evidenzbasiertes Arbeiten in der Physio- und Ergotherapie: Reflektiert - systematisch - wissenschaftlich fundiert*. Springer. <https://doi.org/10.1007/978-3-642-40636-2>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Mokkink, L. B., de Vet, H. C. W., Prinsen, C. a. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(5), 1171–1179. <https://doi.org/10.1007/s11136-017-1765-4>
- Palomar-Asenjo, V., Boleas-Aguirre, Ma. S., Sánchez-Ferrándiz, N., & Perez Fernandez, N. (2006). Caloric and Rotatory Chair Test Results in Patients with Ménière's Disease. *Otology & Neurotology*, 27(7), 945–950. <https://doi.org/10.1097/01.mao.0000231593.03090.23>
- Pérez, P., Llorente, J. L., Gómez, J. R., Campo, A. del, López, A., & Suárez, C. (2004). Functional Significance of Peripheral Head-Shaking Nystagmus. *The Laryngoscope*, 114(6), 1078–1084. <https://doi.org/10.1097/00005537-200406000-00023>
- Plante, E., & Vance, R. (1994). Selection of Preschool Language Tests. *Language, Speech, and Hearing Services in Schools*, 25(1), 15–24. <https://doi.org/10.1044/0161-1461.2501.15>
- Priebe, J. A. (2024). Wie man gute von schlechten Tests unterscheidet. In J. A. Priebe (Hrsg.), *Testkonstruktion – das Herz der psychologischen Diagnostik: Von Gütekriterien, lokaler stochastischer Unabhängigkeit, Normen und dem IQ* (S. 11–32). Springer. https://doi.org/10.1007/978-3-662-67547-2_3

- Reiß, M., & Reiß, G. (2015). *Gleichgewichtsdagnostik: Videonystagmographie und neue Untersuchungsmethoden*. Springer. <https://doi.org/10.1007/978-3-662-45325-4>
- Sanftenberg, L., Bühler, K., Rottenkolber, M., Dreischulte, T., Schelling, J., Gensichen, J., Voigt, K., Bergmann, A., Katzenberger, B., & Grill, E. (2021). Hausärztliche Versorgungsqualität und Lebensqualität aus der Sicht älterer Patienten mit Schwindel: Ergebnisse der Kohortenstudie MobilE-TRA. *MMW - Fortschritte der Medizin*, 163(S4), 11–18. <https://doi.org/10.1007/s15006-021-9703-3>
- Schädler, S., Fuhry, L., Schaaf, H., Marcel, G., & Hauswirth, J. (2022). *Gleichgewicht und Schwindel* (2.). Elsevier GmbH.
- Schädler, S., Kool, J., Lüthi, H., Marks, D., Pfeffer, A., Oesch, P., & Wirz, M. (2012). *Assessments in der Rehabilitation—Band 1: Neurologie* (1. Aufl.). Hans Huber.
- Schmäl, F. (2020). Effektive Diagnostik bei Schwindel und Gleichgewichtsstörungen. *HNO*, 68(9), 703–716. <https://doi.org/10.1007/s00106-020-00910-4>
- Sekhon, R. K., Rocha Cabrero, F., & Deibel, J. P. (2024). Nystagmus Types. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK539711/>
- Shojaku, H., Watanabe, Y., Fujisaka, M., Tsubota, M., Kobayashi, K., Yasumura, S., & Mizukoshi, K. (2005). Epidemiologic characteristics of definite Ménière's disease in Japan. A long-term survey of Toyama and Niigata prefectures. *ORL; Journal for Oto-Rhino-Laryngology and Its Related Specialties*, 67(5), 305–309. <https://doi.org/10.1159/000089413>
- Straumann, D. (o. J.). Fakten und Gesetze bei der Schwindelabklärung. *Schweizer Archiv für Neurologie und Psychiatrie*.
- Strupp, M., Bisdorff, A., Furman, J., Hornibrook, J., Jahn, K., Maire, R., Newman-Toker, D., & Magnusson, M. (2022). Acute unilateral vestibulopathy/vestibular neuritis: Diagnostic criteria. *Journal of Vestibular Research*, 32(5), 389–406. <https://doi.org/10.3233/VES-220201>

- Strupp, M., Brandt, T., & Dieterich, M. (2022). *Vertigo—Leitsymptom Schwindel* (3. Auflage). Springer. <https://doi.org/10.1007/978-3-662-61397-9>
- Strupp, M., Fischer, C., Hanß, L., & Bayer, O. (2014). The takeaway Frenzel goggles: A Fresnel-based device. *Neurology*, *83*(14), 1241–1245. <https://doi.org/10.1212/WNL.0000000000000838>
- Strupp, M., Straumann, D., & Helmchen, C. (2023). Nystagmus und sakkadische Intrusionen. *DNP – Die Neurologie & Psychiatrie*, *24*(2), 56–65. <https://doi.org/10.1007/s15202-023-5651-3>
- Sun, H., Wang, Y., Jiang, H., Gao, Z., & Wu, H. (2022). The clinical application of head-shaking test combined with head-shaking tilt suppression test in distinguishing between peripheral and central vertigo at bedside vs. Examination room. *Brazilian Journal of Otorhinolaryngology*, *88*, S177–S184. <https://doi.org/10.1016/j.bjorl.2022.03.003>
- Teiwes, W., Clarke, A. H., & Scherer, H. (1991). Video-Okulographie—Eine neue Methode zur klinischen Registrierung von dreidimensionalen Augenbewegungen. In K. Fleischer & W. Ey (Hrsg.), *Teil II: Sitzungsbericht* (S. 98–98). Springer. https://doi.org/10.1007/978-3-642-84310-5_86
- Verbecque, E., Van Criekinghe, T., Vanloot, D., Coeckelbergh, T., Van de Heyning, P., Halleman, A., & Vereeck, L. (2018). Dynamic Visual Acuity test while walking or running on treadmill: Reliability and normative data. *Gait & Posture*, *65*, 137–142. <https://doi.org/10.1016/j.gaitpost.2018.07.166>
- Wall, C. (1990). The Sinusoidal Harmonic Acceleration Rotary Chair Test: Theoretical and Clinical Basis. *Neurologic Clinics*, *8*(2), 269–285. [https://doi.org/10.1016/S0733-8619\(18\)30355-4](https://doi.org/10.1016/S0733-8619(18)30355-4)
- Walther, L. E., Hörmann, K., Bloching, M., & Blödown, A. (2012a). Per Funktionsanalyse zur Ursache peripher-vestibulärer Störungen: Rezeptorfunktion der Bogengänge — Teil 1: Diagnostik. *NeuroTransmitter*, *23*(6), 36–50. <https://doi.org/10.1007/s15016-012-0203-7>

- Walther, L. E., Hörmann, K., Bloching, M., & Blödw, A. (2012b). Rezeptorfunktion der Bogengänge: Teil 1: Anatomie, Physiologie, Diagnostik und Normalbefunde. *HNO*, *60*(1), 75–88. <https://doi.org/10.1007/s00106-011-2434-5>
- Wei, D., Hain, T. C., & Proctor, L. R. (1989). Head-shaking nystagmus: Associations with canal paresis and hearing loss. *Acta Oto-Laryngologica*, *108*(5–6), 362–367. <https://doi.org/10.3109/00016488909125540>
- Wirz, M., Köhler, B., Marks, D., Kool, J., Sattelmayer, M., Oesch, P., Hilfiker, R., Rogan, S., Schädler, S., Verra, M., & Lüthi, H. (2014). *Lehrbuch Assessments in der Rehabilitation* (1. Auflage). Verlag Hans Huber.
- Wolfe, J. W., Engelken, E. J., & Kos, C. M. (1978). Low-frequency harmonic acceleration as a test of labyrinthine function: Basic methods and illustrative cases. *Otolaryngology*, *86*(1), ORL-130-142. <https://doi.org/10.1177/019459987808600132>
- Yip, C. W., Glaser, M., Frenzel, C., Bayer, O., & Strupp, M. (2016). Comparison of the Bedside Head-Impulse Test with the Video Head-Impulse Test in a Clinical Practice Setting: A Prospective Study of 500 Outpatients. *Frontiers in Neurology*, *7*, 58. <https://doi.org/10.3389/fneur.2016.00058>
- Zalewski, C. K. (2018). *Rotational vestibular assessment*. Plural Publishing, Inc.
- Zalewski, C. K., McCaslin, D. L., & Carlson, M. L. (2019). Rotary Chair Testing. In S. Babu, C. A. Schutt, & D. I. Bojrab (Hrsg.), *Diagnosis and Treatment of Vestibular Disorders* (S. 75–98). Springer International Publishing. https://doi.org/10.1007/978-3-319-97858-1_6

10 ABBILDUNGSVERZEICHNIS

Abbildung 1: 30° Kopfflexion für die optimale Lage der horizontalen Bogengänge (Zalewski et al., 2019, S.78)	10
Abbildung 2: Darstellung der Testungen in den jeweiligen Frequenzbereichen (Eigene Darstellung, 2024; nach Zalewski et al., 2019, S.76)	12
Abbildung 3: Arten von Schwindel (Eigene Darstellung, 2024; nach Schädler et al., 2022)	12
Abbildung 4: Anpassung des VOR-Nystagmus an die Beschleunigung anhand crescendo und decrescendo mit den drei Reaktionsparameter: Gain, Phase, Symmetry visuell dargestellt (Zalewski et al., 2019, S.82)	20
Abbildung 5: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei normaler vestibulärer Reizantwort (Zalewski et al., 2019, S.83)	20
Abbildung 6: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei akuten unilateralen peripheren Störungen (Zalewski et al., 2019, S.83)	20
Abbildung 7: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei unilateralen peripheren Vestibulopathien mit zentraler Kompensation in Gain (Zalewski et al., 2019, S.84).....	21
Abbildung 8: SHA-Ergebnisse für VOR-Gain (a), Phase (b) und Symmetry (c) Frequenzbereich 0.01 bis 2 Hz bei zentraler Läsion (Zalewski et al., 2019, S.86).....	21
Abbildung 9: 60 °/s mit 60 Sekunden vor und nach rechts- und linksseitig gerichteten Stimuli. X-Achse=Zeit (in Sekunden), Y-Achse=Geschwindigkeit (Grad/Sekunde) (Zalewski et al., 2019, S.88).....	22
Abbildung 10: Prozess.....	27
Abbildung 11: Eigene Darstellung Studiensusuche HST anhand PRISMA-Flowchart (Moher et al., 2009). N=Anzahl.....	31
Abbildung 12: Eigene Darstellung Studiensusuche ROT anhand PRISMA-Flowchart (Moher et al., 2009). N=Anzahl.....	31

11 TABELLENVERZEICHNIS

Tabelle 1: Gütekriterien	24
Tabelle 2: Ein- und Ausschlusskriterien zur Studienselektion	27
Tabelle 3: Übersicht der Studienbewertung HST anhand der COSMIN Checkliste	32
Tabelle 4: Übersicht der Studienbewertung ROT anhand der COSMIN Checkliste	33
Tabelle 5: Charaktereigenschaften Studien HST	35
Tabelle 6: Testausführungen HST	38
Tabelle 7: Charaktereigenschaften Studien ROT	40
Tabelle 8: Testausführungen ROT	43
Tabelle 9: Ergebnisse Studien HST	46
Tabelle 10: Ergebnisse Studien ROT	49

12 ABKÜRZUNGSVERZEICHNIS

AUC-ROC	Area under curve-receiver operating characteristic
ASTE	Ausgangsstellung
CP	Kanalparese
CPA	Kleinhirnbrückenwinkel
DHIS	Dizziness Handicap Inventory Short Form
ENG	Elektronystagmographie
EOG	Elektrookulographie
HSN	Head Shaking Nystagmus/Kopfschüttelnystagmus
HST	Head Shaking Test/Kopfschütteltest
HSTST	Head Shaking Tilt Suppression Test
HWS	Halswirbelsäule
KAL	Kalorik
KIT/vKIT	(Video-)Kopfimpulstest
N.	Nervus
NPV	Negativer Prädiktiver Wert
Pat.	Patientinnen und Patienten
PPV	Positiver Prädiktiver Wert
PRRT	Pseudorandom Rotation Test
ROT	Rotatory Chair/Drehstuhl
SEK	Sekunden
SHA	Sinusoidal Harmonic Acceleration
TC	Time Constant
VEMP	Vestibulär Evoked Myogenic Potential
VNG	Videonystagmographie
VST	Velocity Step Test

13 ANHÄNGE

13.1 Bewertung der eingeschlossenen Primärstudien anhand der «COSMIN checklist»

Angeli et al., 2011

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Angeli et al., 2011

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Pérez et al., 2004

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Pérez et al., 2004

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Wei et al., 1989

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Wei et al., 1989

Box 7. Measurement error		very good	adequate	doubtful	Inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	Inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Asawavichiangianda et al., 1999

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Asawavichiangianda et al., 1999

Box 7. Measurement error						
		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity						
		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Burgio et al., 1991

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Burgio et al., 1991

Box 7. Measurement error

		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity

		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Jacobson et al., 1990

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Jacobson et al., 1990

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Sun et al., 2022

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Sun et al., 2022

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Palomar-Asenjo et al., 2006

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Palomar-Asenjo et al., 2006

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Ahmed et al., 2009

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Ahmed et al., 2009

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Ahmed, 2015

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Ahmed, 2015

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Maes et al., 2011

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Maes et al., 2011

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Fattahi et al., 2024

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Fattahi et al., 2024

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated, Not based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Eza-Nuñez et al., 2016

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Eza-Nuñez et al., 2016

Box 7. Measurement error		very good	adequate	doubtful	Inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Arriaga et al., 2005

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Arriaga et al., 2005

Box 7. Measurement error

		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?			Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another study	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated		% agreement not calculated	Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity

		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

13.2 Ausführliche Zusammenfassungen der Studien

13.2.1 Kopfschütteltest

Charaktereigenschaften Studien zum HST mit detaillierter Beschreibung

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat., Alter (Jahr), Geschlecht	Diagnosen	Testungen
Angeli et al., 2011 Retrospektive casecontrol Studie 2006-2008	USA	N=53 32 Frauen 21 Männer Alter: 24-85 Jahre Mittelwert: 57 Jahre Kontrollgruppe (n=10)	<p>Eingeschlossen: 18 Jahre oder älter, Schwindel oder Gleichgewichtsstörung seit min. 2 Monaten, unilaterale oder asymmetrische periphere vestibuläre Hypofunktion festgestellt durch KAL oder VNG</p> <p>Gleichgewichtsstörungen (n=28) Schwindel (n=25) Periphere Vestibulopathie (n=25, 35%) Vestibuläre Neuritis (n=17, 12%) M. Menière (n=7, 14%) Weitere Störungen des Labyrinthes (n=2, 4%) CPA-Tumor (n=2,4%)</p> <p>Ebenfalls wurden 10 gesunde Erwachsene (Frauen und Männer) erfasst. Diese sollten keine Schwindelsymptome, Gleichgewichts- oder frühere Störungen des Vestibularsystems aufweisen.</p> <p>Ausgeschlossen: BPLS, bilaterale periphere Hypofunktion, Störung des zentralen Nervensystems, Beeinträchtigung des Bewegungsapparates oder Visus</p>	Vergleich HST, KAL und ROT <p>Passiver HST: In einem abgedunkelten Raum mit Frenzelbrille bei einer Geschwindigkeit von 1-2 Hz und einer HWS-Flexion von 30° wurde der Kopf passiv 30-mal gedreht. Als HSN galt, wenn min. 3 Schläge zu sehen waren.</p> <p>VNG-Testbatterie: Spontannystagmus mit und ohne Fixation, blickabhängiger Nystagmus, Okkulomotorik Testung (Sakkaden, Verfolgung, Blick, Optokinetik), Positionstestungen, KAL und SHA</p> <p>Aufzeichnung der Augenbewegungen: Frenzelbrille</p> <p>DHI-S Fragebogen</p>

<p>Pérez et al., 2004</p> <p>Retrospektive Case Studie April 2001- Mai 2002</p>	<p>ESP</p>	<p>N=83</p> <p>Periphere und unilaterale vestibuläre Störungen</p> <p>49 Frauen (59%) 34 Männer (41%)</p> <p>Alter: 14-77 Jahre</p> <p>Mittelwert: 48.6 Jahre</p>	<p>Eingeschlossen: Unilaterale periphere vestibuläre Defizite, keine Anzeichen für zentrale Störungen, ohne Spontannystagmus <2 °/s in der Dunkelheit, keine Einnahme von Schwindelmedikation 24 bis 36 h vor der Studie</p> <p>M. Menière (n=19) Wiederkehrende Vestibulopathie (n=14) Vestibuläre Neuritis (n=13) Akustikusneurinom (n=8) BPLS (n=5) Innenohr Erschütterungssyndrom (n=2) Postoperative Labyrinthitis (n=2) Unilaterale Agenesie des 8. Hirnnervs (n=1) Keine spezifische Diagnose (n=19)</p>	<p>Vergleich HST, VNG, ROT und KAL</p> <p>Passiver HST: Durchführung im Sitzen in einem dunklen Raum. Die Augen durften offengehalten werden. Der Kopf wurde um 30° flektiert und danach passiv in einem 45° Winkel für 30-mal horizontal, mit einer Frequenz von 3 Hz gedreht.</p> <p>Nach der Testung wurde der Nystagmus für 1 Min. mittels einer VNG aufgezeichnet. Als positiv wurde gewertet, wenn min. 5 Schläge von 2 °/s zu sehen waren.</p> <p>Der Nystagmus wurde in monophasisch oder biphasisch und ipsilateral oder kontralateral basierend auf der schnellen Phase eingeteilt.</p> <p>Alle Pat. durchliefen eine VNG-Testbatterie aus den folgenden Testungen: Auftreten von Spontannystagmus und blickabhängiger Nystagmus, Sakkaden, Fähigkeit der Verfolgung mit den Augen, optokinetische Muster, HST, SHA, Positionstestung und der KAL.</p> <p>Aufzeichnung der Augenbewegungen: VNG</p>
---	------------	---	---	---

<p>Wei et al., 1989</p>	<p>USA</p>	<p>N=108 62 Frauen (57%) 46 Männer (43%) Mittelwert: 49 +/- 18 Jahre</p>	<p>Aufgrund der Hörschärfe wurden die Teilnehmerinnen/Teilnehmer in Gruppen aufgeteilt: Unilaterale sensorineurale Schwäche Bilaterale sensorineurale Schwäche Normal (n=60) Diagnosen: <u>Zentral</u> (n=3) CP > +/- 20% (n=30) davon: Perilymphfistel (n=2) Vestibularis Nerv Sektion (n=1) BPLS (n=1) Neuritis vestibularis oder M. Menière (n=26) <u>Undiagnostiziert</u> (n=75)</p>	<p>HST und KAL Vergleich HSN mit Hörverlust und CP Aktiver HST: Die Pat. sollten den Kopf aktiv horizontal für 10-20-mal in einem Winkel von 30° mit 1-2 Hz schütteln. Bei Personen, welche den Kopf nicht aktiv schütteln konnten, wurde der passive HST durchgeführt mit der gleichen ASTE. Der HSN wurde in Dunkelheit mit geschlossenen Augen 30 Sek. vor, während und 100 Sek. nach HST registriert. Eine Augenbewegung wurde als Nystagmus gewertet, wenn min. 3 Schläge in 5 Sek. mit max. Latenz von 5 Sek. und einer Velocity von min. 2.5 °/s sichtbar waren. «Angepasster HSN»: Geschwindigkeitsdifferenz von min. 5 °/s im Vergleich zum Spontannystagmus Aufzeichnung der Augenbewegungen: ENG</p>
-----------------------------	------------	--	--	---

<p>Asawavi-chiangianda et al., 1999</p> <p>Prospektive Kohorten Studie 1991-1997</p>	<p>THA JAP CAN</p>	<p>1.Phase N=1364 mit Schwindel</p> <p>N=50 ohne Schwindel</p> <p>2.Phase N=30</p> <p>3.Phase N=300</p> <p>173 Frauen (57.7%) 127 Männer (42.3%)</p> <p>Alter: 10-82 Jahre</p> <p>Mittelwert: 48.4 Jahre</p>	<p>Vestibuläre Dysfunktion (n=196, 66%)</p> <p><u>Periphere Dysfunktion</u> (n=190) Periphere vestibuläre Störungen (n=192) BPLS (n=63) M. Menière (n=31) Wiederkehrende Vestibulopathie (n=26) Akustikusneurinom (n=21) Vestibuläre Neuritis (n=8) Anderes (n=16) Unbekanntes peripheres vestibuläres Defizit (n=27)</p> <p><u>Zentrale Dysfunktion</u> (n=4)</p> <p><u>Gemischte Form</u> (n=2, 66%)</p> <p>Psychogener Schwindel (n=58, 19%) Keine spezifische Diagnose (n=46, 15%)</p>	<p>Vergleich HST, KAL und ENG</p> <p>Es wurden der aktive und der passive HST durchgeführt.</p> <p>Aktiv: Pat. sitzt mit geschlossenen Augen in einem dunklen Raum mit dem Kopf um 30° flektiert. Danach wird der Kopf für 15 Sekunden horizontal in einem Winkel von 45° und einer Geschwindigkeit von 2 Hz geschüttelt.</p> <p>Passiv: Gleiche ASTE, nur drehte die untersuchende Person den Kopf des Pat.</p> <p>Der Test wurde als positiv gewertet, wenn ein horizontaler Nystagmus mit schneller werdender Frequenz oder mit einer langsamen Phase >5 °/s als beim Spontannystagmus festgestellt werden konnte.</p> <p>Aufzeichnung der Augenbewegungen: ENG</p> <p>Die Studie wurde in drei Phasen unterteilt.</p> <p>1.Phase: Der HST wurde während einer ENG-Testbatterie bei 50 Personen der Kontrollgruppe (ohne Schwindel) und bei 1364 Personen mit Schwindel durchgeführt. Dabei wurde Spontannystagmus, Positionierung (Augen offen/geschlossen), Blicknystagmus, Okkulomotorik, Sakkaden, Blickverfolgung, Optokinetik KAL und HST getestet.</p> <p>2.Phase: Bei 30 Personen wurde die EOG mit der magnetic search (scleral) coil Test Untersuchung in Bezug auf den HSN verglichen.</p> <p>3.Phase: Aus den Pat. wurden 300 Personen mit Schwindel und einer bekannten Diagnose in die dritte Phase eingeschlossen, um die Sensitivität und Spezifität der Testungen zu evaluieren.</p>
--	----------------------------	---	---	--

<p>Burgio et al., 1991</p>	<p>USA</p>	<p>N=115 Geschlecht und Alter unklar Kontrollgruppe (n=17)</p>	<p>Unilaterale vestibuläre Störung (n=25), ≥ 25% Bilaterale vestibuläre Störung (n=10) Schwindelpat. mit unauffälligem ENG: (n=80)</p>	<p>Vergleich HST und ENG Passiver HST: Wurde in einem abgedunkelten Raum durchgeführt. Pat. sitzt, hält die Augen offen und trägt während der Testung eine Frenzelbrille. Der Kopf wird für 20 Sek. gedreht. Max. Geschw. 250-500 °/s Jegliche Ausschläge der Augen wurden als abnormales Resultat gewertet. Aufnahme der Augenbewegungen: Frenzelbrille</p>
<p>Jacobson et al., 1990</p>	<p>USA</p>	<p>N=116 Geschlecht und Alter unklar</p>	<p>unklar</p>	<p>Vergleich HST, ENG und ROT Aktiver HST: Es wurden Elektroden oberhalb und unter den Augen angebracht. Zur Kalibrierung der Pat. wurden 10° horizontal und 10° vertikale Ziele fixiert. Pat. sassen in einem dunklen Raum mit den Augen geschlossen und drehten ihren Kopf mit einer Geschwindigkeit von ungefähr 2 Hz für 30 Sek. hin und her. Erst dann wurde ihnen die Frenzelbrille angezogen. Die Augen sollten danach geöffnet werden und der Kopf nach vorne gehalten. Der Nystagmus wurde anhand der Langsam-Phasen, Augengeschwindigkeit (SPEV), Latenzzeit und Dauer quantifiziert ROT: 0.01/0.02/0.04/0.08/0.16 Hz Max. Geschw. 100 °/s mit max. Beschleunigung von 50 °/s ENG-Testbatterie: Blickabhängiger Nystagmus, Verfolgung, Sakkaden, Optokinetik, Spontannystagmus, Positionierung (Nystagmus) und KAL. Aufzeichnung der Augenbewegungen: EOG</p>

<p>Sun et al., 2022</p> <p>Retrospektive Studie 2019-2021</p>	<p>CHI</p>	<p>N=80</p> <p>52 Frauen (65%) 28 Männer (35%)</p> <p>Peripherer Schwindel (n=47, 58.8%)</p> <p>Mittelwert: 50.7 Jahre +/- 13.26 Jahre</p> <p>Zentraler Schwindel (n=33, 41.2%)</p> <p>Mittelwert: 47.03 Jahre +/- 16.65 Jahre</p>	<p>Eingeschlossen: Peripherer oder zentraler Schwindel, komplette klinische Daten, vervollständigter HST und HSTST, HSN</p> <p>Periphere Gruppe (n=47) M. Menière (n=22) Neuritis vestibularis (n=9) Plötzlicher Hörverlust mit Schwindel (n=6) Endolymphatischer Hydrops (n=4) Labyrinthitis (n=4) Innenohr (n=1) Malformation (n=1) Akustikusneurinom (n=1)</p> <p>Zentrale Gruppe (n=33) Vestibuläre Migräne (n=16) Mögliche Vestibuläre Migräne (n=12) Transient ischämische Attacke (n=2) Multiple System Atrophie (n=2) Spinocerebellar Ataxie (n=1)</p> <p>Ausgeschlossen: Mischung aus zentralem und peripherem Schwindel</p>	<p>Vergleich HST und HSTST</p> <p>Passiver HST: 1. Phase: Bedside-Test mit Videobrille, 30-mal passives Kopfschütteln mit einer Frequenz von 2-3 Hz, Kopfflexion von 30°</p> <p>2. Phase: Gleiche Abfolge im Behandlungsraum mit VNG zur Aufzeichnung von Richtung und Geschwindigkeit der HSN. Wurde von einer anderen Person getestet.</p> <p>HSTST: 1. Phase: Identische Ausführung zum HST (mit Videobrille), jedoch wurde erst nach dem Kopfschütteln der Kopf flektiert (Kinn-Sternum).</p> <p>2. Phase: Gleiche Abfolge im Behandlungsraum mit VNG</p> <p>Aufnahme der Augenbewegungen: 1. Phase Videobrille (Verhindert visuelle Blickfixation) 2. Phase VNG</p>
---	------------	--	---	--

HST=Kopfschütteltest, HSN=Kopfschüttelnystagmus, KAL=Kalorik, CP=Kanalparese, ROT=Drehstuhl, SHA=Sinusoidal Harmonic Acceleration, CPA=Kleinhirnbrückenwinkel, BPLS= Benigner paroxysmaler Lagerungsschwindel, VNG=Videonystagmographie, ENG=Elektronystagmographie, EOG=Elektrookulographie, HSTST=Head Shaking Tilt Suppression Test, Pat.=Patientinnen/ Patienten

13.2.2 Drehstuhltest

Charaktereigenschaften Studien ROT mit detaillierter Beschreibung der Testungen

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
<p>Palomar-Asenjo et al., 2006</p> <p>Prospektive Studie</p>	ESP	<p>N=100</p> <p>Geschlecht und Alter unklar</p> <p>Rechtes Ohr betroffen (n=57) Linkes Ohr betroffen (n=43)</p> <p>Mittelwert Dauer der Symptomatik: 5.4 Jahre</p>	<p>Eingeschlossen: Unilaterales M. Menière</p> <p>Ausgeschlossen: Zentrale Störungen, Bilaterale Problematiken</p>	<p>Vergleich ROT und KAL zur Diagnostik des M. Menière am gleichen Tag</p> <p>ROT: Wurde im Dunkeln mit dem Kopf in leichter Flexion durchgeführt. Die Pat. wurden in Bereitschaft gehalten. SHA: 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz Bei max. Geschwindigkeit 50 °/s (Phase, Gain, Asymmetry) VST: Für 1 Sek. 100 °/s, danach 60 Sek. konstante Geschwindigkeit, Entschleunigung auf 0 °/s für 1 Sek.</p> <p>Resultate unterscheiden zwischen Drehung auf die gesunde Seite (kontralateral) oder betroffene Seite (ipsilateral).</p> <p>Wenn die Zeitkonstante (TC) >10 Sek., liegt ein normaler Befund vor. Ist jedoch der Seitenunterschied <20% liegt eine Asymmetrie vor.</p> <p>Aufzeichnung der Augenbewegungen: EOG</p>

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
<p>Ahmed et al., 2009</p> <p>Retrospektive Case Studie 2001-2007</p>	<p>USA</p>	<p>N=200</p> <p>Gruppe 1: 132 Schwindelpat. mit erwarteter peripherer Vestibulo-pathie</p> <p>64 Frauen 68 Männer</p> <p>Alter: 20-87 Jahre Mittelwert: 58.56 +/- 15.86 Jahre</p> <p>Gruppe 2: 68 Schwindelpat. ohne erwartete periphere Vestibulo-pathie</p> <p>41 Frauen 27 Männer</p> <p>Alter: 17-89 Jahre</p> <p>Mittelwert: 60.19 +/- 16.96 Jahre</p>	<p>Eingeschlossen: Gruppe 1: <u>Periphere Vestibulopathie:</u> Definitiver M. Menière (n=46, 34.9%) Neuritis vestibularis (n=41, 31.1%) Nicht-identifiziert (n=20, 15.2%) Bilaterale Vestibulopathie (n=13, 9.9%) Akustikusneurinom (n=6, 4.6%) Labyrinthitis (n=6, 4.6%)</p> <p>Gruppe 2: <u>Ohne periphere Vestibulopathie:</u> Migräne induzierter Schwindel (n=39, 57.5%) Zentrale vestibuläre Läsion (n=11, 16.2%) Posturale Instabilität (n=7, 10.3%) Normaldruck Hydrocephalus (n=4, 5.9%) MS (n=3, 4.5%) Orthostatisch (n=2, 3%) Transient ischämische Attacke (n=1, 1.5%) Angstbedingter Schwindel (n=1, 1.5%)</p> <p>Ausgeschlossen: BPLS, Superiore Bogengang Dehiszenz, Perilymphfistel, mögliche M. Menière oder Pat. mit inkompletten Testungen oder unklarer Interpretation</p>	<p>Vergleich ROT und KAL</p> <p>ROT: Im Dunklen wurde die Testung mit einer 30° HWS-Flexion, geöffneten Augen und zusätzlichen Aufgaben, ausgeführt. SHA: 0.025/0.05/0.25/0.5 Hz, 60 °/s Geschwindigkeit VST: 100 °/s. rechts, 100 °/s. links</p> <p>Nystagmus Auslösungen pro Drehung</p> <p>TC wurde aufgezeichnet und für jede Testperson wurden vier identifiziert.</p> <p>Aufzeichnung der Augenbewegungen (Per-rotatorisch und Post- rotatorisch): unklar</p>

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
Ahmed, 2015 Retrospektive Studie	EGY	N=119 56 Frauen 63 Männer Alter: 22-79 Jahre Mittelwert: 58.23 +/- 15.13 Jahre	<p>Eingeschlossen: Dokumentierte unilaterale periphere vestibuläre Pathologien, dokumentierte unilaterale kalorische Schwäche, die mit der klinischen Diagnose (Anamnese, körperliche Untersuchung neuro-otologischen Untersuchung, audiologischen Beurteilung und radiologischen Untersuchungen) übereinstimmen.</p> <p>M. Menière (34.9%) Neuritis vestibularis (31.1%) Unidentifiziert (15.2%) Labyrinthitis (4.6%) Akustikusneurinom (4.6%)</p> <p>Ausgeschlossen: Periphere vestibuläre Läsionen, die nicht durch unilaterale kalorische Schwäche belegt sind oder Pat. mit unvollständigen Tests oder nicht interpretierbaren Daten</p>	<p>Vergleich ROT und KAL</p> <p>ROT: In dunklem Raum, auf dem ROT gesichert und gesenktem Kopf. Während der Drehung Augen offen und mentale Aufgaben wurden gestellt. SHA: 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz bei max. Geschwindigkeit von 60 °/s. VST: 100 °/s rechts und links</p> <p>Per-rotatorischer Nystagmus (erste Beschleunigung des ROT) und Post-rotatorischer Nystagmus (beginnt bei komplettem ROT-stopp) Parameter: Gain und TC</p> <p>Aufzeichnung der Augenbewegungen: Infrarot Videokulographie</p>

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
Maes et al., 2011	BEL	<p>N=157</p> <p>80 Frauen 77 Männer</p> <p>Gruppe 1: N=77</p> <p>38 Frauen 39 Männer</p> <p>Alter: 19-78 Jahre</p> <p>Mittelwert: 52 Jahre</p> <p>Gruppe 2: N=80</p> <p>42 Frauen 38 Männer</p> <p>Alter: 20-80 Jahre</p> <p>Mittelwert: 48 Jahre</p>	<p>Eingeschlossen:</p> <p>Gruppe 1: Unilaterale periphere vestibuläre Pathologien: M. Menière (n=36) Labyrinthitis (n=12) Neuritis vestibularis (n=10) Akustikusneurinom (n=8) Bogengang-Dehiszenz (n=6) Posttraumatischer Schwindel (n=5)</p> <p>Gruppe 2: Kontrollgruppe ohne Schwindel in der Vergangenheit, keine Gleichgewichtsproblematiken, keine Störungen des Sehens (ausser Brillen- und Kontaktlinsentragende)</p>	<p>Vergleich ROT, KAL und VEMP</p> <p>ROT: Im Dunkeln, die Testpersonen hatten 10 Min. Anpassungszeit an die Dunkelheit. Gesichert mit Sicherheitsgurt im Stuhl, Kopf gerade stabilisiert (nicht in 30° Flexion), die Augen offen und in Bereitschaft (Kopfrechnen während Drehung).</p> <p>Standardkalibrierung der horizontalen Augenbewegungen wurde durch Fixierung eines visuellen Punktes 10° rechts und links im Vergleich zu vor der Rotation sichergestellt. Pro Seite wurde 10-mal rotiert.</p> <p>SHA: 0.01/0.02/0.05/0.1/0.2 Hz Max. Geschwindigkeit 50°/s (Gain, Phase und Asymmetry)</p> <p>VST: 2x Rotation nach rechts, dann 2x Rotation nach links Unterschwellige Beschleunigung von 2 °/s bis max. 100 °/s. Drehung erfolgte 1 Min., danach die Bremsung von 200 °/s (Aufzeichnung post-rotatorischer Nystagmus) (Gain, Phase, TC, Asymmetry)</p> <p>PRRT: Max. Geschwindigkeit 180 °/s (Gain und Phase)</p> <p>Aufzeichnung der Augenbewegungen und Kopfbewegungen: Elektroden</p>

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
<p>Fattahi et al., 2024</p> <p>Retrospektive Analyse 2016-2022</p>	<p>USA</p>	<p>N=150</p> <p>88 Frauen 62 Männer</p> <p>Alter: 13-84 Jahre</p> <p>Mittelwert: 54.10 +/- 15.09 Jahre</p> <p>Periphere Gruppe: 60.9 Jahre +/- 10.8 Jahre</p> <p>Zentrale Gruppe: 48 Jahre +/- 15.8 Jahre</p>	<p>Eingeschlossen:</p> <p><u>Periphere Gruppe (n=71)</u> Neuritis vestibularis (n=36, 50.7%) M. Menière (n=22, 30.9%) Labyrinthitis (n=7, 9.86%) Akustikusneurinom (n=4, 5.6%) Ototoxizität (n=2, 2.8%)</p> <p><u>Zentrale Gruppe (n=79)</u> Vestibulare Migräne (n=51, 64.5%) Zentral Vestibuläre Läsion (n=12, 15.19%) PPPD (n=9, 11.3%) Mal de débarquement (n=3, 3.8%) Schädelhirntrauma (n=4, 5.06%)</p> <p>Ausgeschlossen: Visuelle Störungen, isolierte BPLS, superiore Kanal Dehiszenz</p>	<p>Vergleich ROT, KAL und vKIT bei peripherer und zentraler vestibulärer Dysfunktion</p> <p>SHA: Wurde in der Dunkelheit mit einer Frequenz zwischen 0.01-1.0 Hz und max. Geschwindigkeit von 40 °/s durchgeführt (Gain, time und bias constant).</p> <p>Altersentsprechende TC <12.7 Sek.= periphere vestibuläre Störung (p=0.05) basierend auf klinischen Vergleichsmassstäben.</p> <p>Aufzeichnung der Augenbewegungen: unklar</p>

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
Eza-Nuñez et al., 2016 Prospektive Studie	ESP	N=116 56 Frauen (48.3%) 60 Männer (51.7%) Alter: 21- 80 Jahre Mittelwert: 54 Jahre SD: +/- 14 Jahre	<p>Periphere Vestibulopathie (n=75, 64,6%)</p> <ul style="list-style-type: none"> M. Menière (n=34) Neuritis vestibularis (n=12) BPLS (n=11) Otosklerose (n=5) Akustikusneurinom (n=3) Iatrogen (n=3) Cochlear Implantat (n=1) Wahrscheinliches M. Menière (n=1) Chronische Otitis media (n=1) Immundefekt- Krankheit (n=1) <p>Nicht-periphere Vestibulopathie (n=41, 35,3%)</p> <ul style="list-style-type: none"> Vestibuläre Migräne (n=15) Chronischer Schwindel (n=13) Angstbedingter Schwindel (n=3) Post-traumatischer Schwindel (n=2) Undefiniert (n=2) Transient ischämische Attacke (n=2) 	Vergleich ROT, KAL und vKIT <p>SHA: In der Dunkelheit, Kopfflexion (horizontalen Bogengänge auf Stimulusebene 30°), durch Konversation wurden Pat. in Bereitschaft gehalten 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz, 50 °/s Geschwindigkeit (Phase, Gain, Symmetry)</p> <p>VST: 100 °/s Geschwindigkeit 60 Sek. rechts, 60 Sek. links →48 min +/- 24 Min. für ROT</p> <p>Aufzeichnung der Augenbewegungen: EOG</p>

Autorinnen/ Autoren, Jahr, Stud. Design	Land	Stichprobe: Anzahl (n) Gruppe der Pat. Alter (Jahr), Geschlecht	Diagnosen	Testungen
Arriaga et al., 2005 Retrospektives Chart Review	USA	N=989 511 (51,7%) Pat. durchliefen nur die ROT 478 (48.3%) ROT und ENG Batterie. Alter und Geschlecht: Unklar	Labyrinthitis, Neuritis vestibularis, akuter Hörverlust mit Schwindel, Akustikusneurinom, Schläfenbeinbruch und M. Menière	Vergleich ROT und ENG VOR Tests: 60 °/s Geschwindigkeit, 0.01/0.02/0.04/0.08/0.16/0.32/0.64 Hz VST: Drehung 100 °/s für 30 Sek. rechts, ROT stopp für 30 Sek., Drehung 100 °/s für 30 Sek. links, ROT stopp für 30 Sek. Aufzeichnung der Augenbewegungen: ENG

ROT=Drehstuhl, KAL=Kalorik, SHA=Sinusoidal Harmonic Acceleration, VST=Velocity Step Test, PRRT=Pseudorandom Rotation Test, vKIT=Video-Kopfimpulstest, VEMP=Vestibulär evoked myogenic potential, VOR=Vestibulo-okular Reflex, BPLS=Benigner paroxysmaler Lagerungsschwindel, PPPD=Persistierender postural-perzeptiver Schwindel, MS=Multiple Sklerose, ENG=Elektronystagmographie, EOG=Elektrookulographie, Pat.=Patientinnen/Patienten

13.2.3 Gütekriterien und Ergebnisse Kopfschütteltest

Ausführliche Darstellung der Ergebnisse und der Werte zur Sensitivität, Spezifität und anderen Gütekriterien zu HST-Studien

Angeli et al., 2011		
Sensitivität	Spezifität	Andere Gütekriterien
HST + KAL= 0.31 HST + SHA mit normalem VOR-Gain= 0.59 HST + SHA mit abnormalem VOR-Asymmetry= 0.71 Periphere Vestibulopathie= 0.75 Neuritis vestibularis= 0.36 M. Menière= 0.31 Weitere Störungen des Labyrinthes= 0.00 CPA-Tumor= 0.63	HST + KAL= 0.96 HST + SHA mit normalem VOR-Gain= 0.83 HST + SHA mit abnormalem VOR-Asymmetry= 0.85	Validität PPV: HST + KAL= 0.97 PPV: HST + SHA mit normalem VOR-Gain = 0.63 PPV: HST + SHA mit abnormalem VOR-Asymmetry = 0.66
Ergebnisse		
<p>Positiver HST: 16 Pat. (monophasisch)</p> <p>Pat. mit Gleichgewichtsstörungen zeigten öfter einen HSN, als diejenigen mit Schwindel.</p> <p>Positiver HSN und positiver KAL: 14 Pat.</p> <p>Alle 10 Personen der Kontrollgruppe hatten einen negativen HST und negativen KAL.</p> <p>Bei Pat. mit einem positiven HST, war es wahrscheinlicher, dass eine unkompenzierte unilaterale Hypofunktion vorliegt. Die hohe Sensitivität wurde mit den unterschiedlichen Pat. und einer höheren Frequenz der unilateralen Hypofunktion begründet.</p> <p>Der HST kann zur Diagnosestellung einer Vestibulopathie oder Ermittlung der betroffenen Seite hilfreich sein.</p> <p>Bei Pat. mit Schwindel oder Gleichgewichtsstörungen kann ein positiver HST mit grösseren funktionellen Einschränkungen sowie Stürzen verbunden werden.</p> <p>Die Sensitivität des HST ist gross genug, um ihn als Screening-Test für noch unkompenzierte Vestibulopathie dieser Kohorte zu nutzen. Ein negativer Test allein kann ein peripheres vestibuläres Defizit jedoch nicht vollständig ausschliessen.</p>		

Pérez et al., 2004		
Sensitivität	Spezifität	Andere Gütekriterien
unklar Sprechen von 0.3-0.4 bei Pat. mit unilateralen peripheren Defiziten	unklar	Validität Negative HSN vs. Cm-HSN= $p<0.01$ Negative HSN vs. b-HSN= $p<0.01$ Im-HSN vs. Cm-HSN= $p<0.05$ Im HSN vs. b-HSN= $p<0.01$
Ergebnisse		
<p>Von den 83 Pat. hatten 23 (28%) einen positiven HST. 19 (83%) davon monophasisch, davon schlug der Nystagmus bei 9 (47%) auf die gesunde Seite (kontralateral) und bei 10 (53%) auf die Seite mit der Läsion (ipsilateral). 4 (17%) Pat. hatten einen biphasischen Nystagmus.</p> <p>Begründen das seltenere Auftreten vom HSN bei jüngeren Pat. durch die bessere Plastizität. Wird häufiger bei Pat. über 50 Jahre gefunden mit hohem Sturzrisiko. Von 24 Pat. mit einem positiven HSN, wiesen 15 (62%) ein asymmetrisches Resultat bei der SHA. Es zeigte sich eine signifikante Assoziation zwischen einem positiven HST und positiven SHA. Es konnte keine signifikante Korrelation zwischen KAL und HSN festgestellt werden. Ipsilateraler HSN entsprach niedrigeren Defiziten. Kontralaterale und biphasische HSN korrespondieren mit einem grösseren Defizit.</p> <p>Es sei schwierig Pat. mit einem positiven HST zu klassifizieren. Es scheint, dass ein HSN bei unilateralen peripheren vestibulären Defiziten vor allem mit einem funktionellen Defizit assoziiert werden kann. Er zeigt, dass ein vestibuläres Defizit existiert oder existiert hat. Es konnten signifikante Differenzen des HSN zwischen M.Menièrè und wiederkehrende Vestibulopathie festgestellt werden.</p>		

Wei et al., 1989		
Sensitivität	Spezifität	Andere Gütekriterien
<p>0.32</p> <p><u>CP +:</u> Genereller Hörverlust=0.42 Unilateraler Hörverlust=0.65 HSN=0.33</p> <p><u>HSN +:</u> Genereller Hörverlust=0.38 Unilateraler Hörverlust=0.5 CP=0.4</p> <p><u>Genereller Hörverlust +:</u> CP=0.63 HSN=0.47</p> <p><u>Unilateraler Hörverlust +:</u> CP=0.43 HSN=0.28</p>	<p>0.6</p> <p><u>CP +:</u> Genereller Hörverlust=0.82 Unilateraler Hörverlust=0.8 HSN=0.74</p> <p><u>HSN +:</u> Genereller Hörverlust=0.68 Unilateraler Hörverlust=0.69 CP=0.6</p> <p><u>Genereller Hörverlust +:</u> CP=0.65 HSN=0.59</p> <p><u>Unilateraler Hörverlust +:</u> CP=0.91 HSN=0.86</p>	<p>Validität PPV=0.32 NPV=0.75</p>
Ergebnisse		
<p>HSN wurde bei 37 Personen festgestellt: Zentral (n=1, 33%), CP (n=12, 40%) und Undiagnostiziert (n=24, 32%).</p> <p>Der HSN war meist nicht richtungsspezifisch und zeigte keine erkrankte Seite an. Bei 9 von 12 Personen mit HSN und CP schlug er zur Seite mit der stärkeren Antwort bei der KAL. Die KAL zeigte höhere Sensitivität und Spezifität bei Hörausfällen als der HST. Der HSN zeigte meist nicht die erkrankte Seite an, nur bei 4/10 Pat. war dies der Fall.</p> <p>Der HSN war häufiger als CP, dies wurde damit begründet, dass der HSN entweder höhere Sensitivität oder niedrigere Spezifität als KAL hat oder durch das ENG häufiger Augenbewegungen registriert wurden. Ein genauer Vergleich war schwierig, da beim HST eine gesündere Person kein Nystagmus und bei der KAL ein Nystagmus zeigt. Ebenso erwähnen sie, dass der Vergleich durch das Fehlen eines Goldstandards erschwert wird.</p> <p>Die Quantifizierung mittels HST lässt noch Raum für Verbesserung, insbesondere bei den Möglichkeiten zur Aufnahme der Augenbewegungen. Doch als einfacher Bedside-Test kann der passive HST mit Frenzelbrille empfohlen werden.</p>		

Asawavichiangianda et al., 1999		
Sensitivität	Spezifität	Andere Gütekriterien
Vestibuläre Dysfunktion= 0.383 Psychogener Schwindel= 0.207 Keine spezifische Diagnose= 0.152 BPLS= 0.281 <u>Verglichen mit psychogenem Schwindel:</u> M. Menière= 0.581 Akustikusneurinom= 0.476 Neuritis vestibularis= 0.5 Wiederkehrende Vestibulopathie= 0.308 Anderes= 0.314 Unbekanntes peripheres vestibuläres Defizit= 0.444	Seitendifferenz von: >20: 0.732 >40: 0.717 >60: 0.707 >80: 0.701	-
Sensitivität HSN und/oder KAL Akustikusneurinom= 0.905 M. Menière= 0.839 Neuritis vestibularis= 0.75		
Ergebnisse		
<p>1. Phase: 12 (24%) der 50 gesunden Personen und 432 (31.7%) der 1364 Personen mit Schwindel wiesen einen HSN auf. Es konnte keine statistische Differenz zwischen dem aktiven (24%) oder passiven (24.3%) HST ermittelt werden. Der HSN wurde in monophasisch (76.8%), biphasisch (22.7%) und triphasisch (0.5%) unterschieden.</p> <p>2. Phase: Von den 30 untersuchten Personen konnte bei 21 Personen kein HSN festgestellt werden. Bei 4 wurde ein HSN nachgewiesen und bei den restlichen 5 gab es Schwierigkeiten bei der Aufzeichnung.</p> <p>3. Phase: Bei 95 (31.6%) von 300 Personen konnte ein HSN erkannt werden. Dieses Resultat ähnelt der Phase 1. 76 (38.3%) der Pat. mit peripherem vestibulärem Defizit, 12 (20.7%) mit psychogenem Schwindel und 7 (15.2%) ohne Diagnose zeigten einen HSN.</p> <p>Es wurde dann bei 290 Personen eine KAL und ENG durchgeführt. 81 (27.9%) zeigten eine unilaterale kalorische Reduktion und davon hatten 36 (44.4%) einen positiven HST.</p> <p>Zum Vergleich mit der ENG zeigte ein positiver HST und eine Reduktion der KAL eine höhere Sensitivität bei peripheren vestibulären Störungen. Der HST ist ein üblicher Bestandteil der ENG Test-Batterie und dabei wird nicht zwischen der aktiven oder passiven Durchführung unterschieden. Am häufigsten konnte ein HSN bei M. Menière, Neuritis vestibularis oder Akustikusneurinom festgestellt werden. Empfohlen wird eine Kombination aus HST und KAL.</p> <p>Der HST stellt ein geeigneter Bedside-Test dar. Sollte jedoch mit Hilfe einer Frenzelbrille ausgeführt werden. Ein positives Resultat (Nystagmus >5 Sek.) kann sehr wahrscheinlich auf eine Läsion hindeuten, ist jedoch nicht direkt diagnostisch zu interpretieren. Eine Störung kann mit einem negativen Resultat nicht direkt ausgeschlossen, sondern sollte genauer untersucht werden.</p>		

Burgio et al., 1991		
Sensitivität	Spezifität	Andere Gütekriterien
Unilaterale Läsion= 0.44 bilaterale Läsionen=0.4 Schwindelpat. mit normalem ENG=0.51	Unilaterale Läsion= 0.65	-
Ergebnisse		
<p>6 Personen der Kontrollgruppe hatten einen positiven HST. 11/17 Personen der Kontrollgruppe=normal.</p> <p>Die Resultate deuten darauf hin, dass weder die Sensitivität noch die Spezifität des passiven HST ausreicht, um ihn als Screening-Test bei peripheren vestibulären Läsionen zu nutzen, jedoch kann ein HSN eine Asymmetrie des vestibulären Systems aufzeigen.</p> <p>Bezüglich der Sensitivität und Spezifität wurden keine Unterschiede gefunden, welche darauf hinweisen, dass die aktive oder passive Ausführung besser ist. Jedoch gab es Unterschiede der Werte zwischen verschiedenen Studien, welche den Test aktiv oder passiv ausgeführt haben. Dies wird damit begründet, dass sich die physiologische Kontrolle der Augen bei aktiven oder passiven Kopfbewegungen unterscheidet. Bei der aktiven Bewegung kann eine Vorprogrammierung des zentralen Nervs den VOR bereits anpassen. Ebenso wird die Propriozeption der HWS als mögliche Erklärung genannt. Propriozeptorische Signale zum vestibulären System werden bei aktiven Kopfbewegungen inhibiert und haben einen grösseren Effekt bei passiven Bewegungen.</p> <p>In der Studie wird ebenfalls darauf eingegangen, dass der Vergleich von HST (>1.5 Hz) und ENG (0.02 Hz) nicht optimal ist, da in anderen Frequenzen getestet wird.</p> <p>Der HST mit Frenzelbrille und in einem dunklen Raum stellt eine reliable Methode dar, bei der Unterscheidung von physiologischem und pathologischem HSN und kann auf die Seite der unilateralen Läsion hindeuten.</p> <p>Dennoch wird erwähnt, dass die passive Variante des Tests nicht genügend sensitiv ist.</p>		

Jacobson et al., 1990		
Sensitivität	Spezifität	Andere Gütekriterien
0.27 (73% falsch negativ)	0.85 (15% falsch positiv)	Falsch-positiv Rate= 0.15 Falsch-negativ Rate= 0.73
Ergebnisse		
<p>24 von 116 Personen zeigten einen HSN (meist horizontal). Dieses Resultat zeigt eine 21%ige Prävalenz des HSN bei Schwindelapat. Er konnte 14-mal nachgewiesen werden, bei abnormalen ENG und ROT (27% Sensitivität) und 10-mal bei normalen ENG und ROT. Bei 55 Personen war kein HSN zu sehen und ebenfalls ein normales Resultat des ENG und ROT dokumentiert (85% Spezifität).</p> <p>Durchschnittliche Dauer des Nystagmus: 7.03 Sek. und der durchschnittliche SPE 12 °/s.</p> <p>Durch die tiefe Sensitivität des HST kann eine vestibuläre Störung weniger gut festgestellt werden. Es wird vermutet, dass die tiefe Sensitivität auf den geringen peripheren Störungen basiert, die den HSN weniger auslösen als starke periphere Störungen. Ebenso wurden jegliche Grade der Läsion in die Studie miteinbezogen und dies ohne Selektion der Population.</p> <p>Aufgrund der Sensitivität und Spezifität des HST empfiehlt die Studie die Testung nicht, sondern eher die ROT oder KAL. Der HST sei ein schlechter Prädiktor des Vestibularsystems, jedoch ist es schwierig eine genaue Aussage zu machen, da er eine höhere Frequenz testet als ROT und KAL.</p>		

Sun et al., 2022		
Sensitivität	Spezifität	Andere Gütekriterien
HST mit VNG= 0.936 HST und HSTST= 0.809	HST mit VNG= 0.727 HST und HSTST= 0.939	<p>Reliabilität Kappa HST=0.658 (p<0.001) Kappa HST + HSTST=0.650 (p<0.001)</p> <p>Vergleich von Bedside-Test und Behandlungsraum (Werte übereinstimmend)</p> <p>Validität HST vs. HSTST unidirektional horizontal=p<0.001</p> <p><i>AUC-ROC Kurven:</i> HST= AUC 0.832 / HST+HSTST= AUC 0.874 →VNG HST=AUC 0.801 / HST+ HSTST= AUC 0.859 → Videobrille</p>
Ergebnisse		
<p>44 (93.6%) Pat. der peripheren Gruppe zeigten in der 2. Phase mit VNG einen horizontalen HSN Davon 31 (70.5%) unidirektional und 13 bidirektional. 35 (79.5%) Pat. der peripheren Gruppe zeigten beim Bedside-HST einen horizontalen HSN. Der HST und HSTST stellen simple und zeitsparende Testvarianten dar. Es sei möglich, dass der HSN beim HSTST unterdrückt wird und nicht immer sichtbar ist.</p> <p>Durch den HSTST konnten die Mehrheit der unidirektionalen (n=20) und bidirektionalen (n=12) HSN in der 1. Phase unterdrückt werden. In der Phase 2 jedoch nicht mehr ganz so viele.</p> <p>9 Personen der zentralen Gruppe wiesen einen horizontalen HSN auf. Der horizontale HSN weist auch die höchste Sensitivität. Die Resultate des Bedside-Tests sowie VNG waren ähnlich und zeigten eine gute Sensitivität und Spezifität. Die Tests können periphere und zentrale Defizite unterscheiden.</p> <p>Die Kombination der beiden Testungen HST und HSTST konnte zur Diagnose von Schwindel empfohlen werden. Dabei spielt es keine grosse Rolle, ob sie als Bedside-Test oder im Untersuchungsraum ausgeführt werden.</p>		

HST=Kopfschütteltest, HSN=Kopfschüttelnystagmus, KAL=Kalorik, CP=Kanalparese, ROT=Drehstuhl, SHA=Sinusoidal Harmonic Acceleration, HSTST=Head Shaking Tilt Suppression Test, CPA=Kleinhirnbrückenwinkel, BPLS=Benigner Paroxysmaler Lagerungsschwindel, VOR=Vestibulo-okular Reflex, HWS=Halswirbelsäule, PPV=Positiver prädiktiver Wert, NPV=Negativer prädiktiver Wert, VNG=Videonystagmographie, ENG=Elektronystagmographie, Pat.=Patientinnen/Patienten

13.2.4 Gütekriterien und Ergebnisse Drehstuhltest

Ausführliche Darstellung der Ergebnisse und der Werte zur Sensitivität, Spezifität und anderen Gütekriterien zu ROT Studien

Palomar-Asenjo et al., 2006		
Sensitivität	Spezifität	Andere Gütekriterien
<p>unklar ROT höhere Sensitivität als KAL</p>	<p>unklar KAL höhere Spezifität als ROT</p>	<p>Reliabilität: Die Zeitkonstante des VOR ist stabiler als die Kanalparese des kalorischen Tests in Test-Retest-Studien (Jenkins & Goldberg, 1988). <i>Kappa: 0.226</i> ROT und KAL Validität: Da beide Testungen unterschiedliche Aspekte des vestibulären Systems untersuchen, ist ihre Validität für die Diagnosestellung von vestibulären Dysfunktionen, (wie bei M. Menière), unterschiedlich hoch. <i>PPV KAL=0.74</i></p>
Ergebnisse		
<p>Es gibt statistisch signifikante, aber moderate Assoziation zwischen den Ergebnissen des rotatorischen Stuhltests und des kalorischen Tests. Besonders die Richtungspräferenz (Directional Preponderance, DP) weist eine stärkere Übereinstimmung mit dem ROT auf, was auf eine bessere Korrelation dieser beiden Tests hinweist. Ein abnormales Resultat wurde definiert, wenn in 2 oder 3 Frequenzbereichen abnormale Befunde (Phase, Gain und Symmetry) vorlagen. Es wurden weniger abnormale Befunde beim ROT als bei der KAL gefunden. Die Resultate der KAL stimmten oft nicht mit den Resultaten des ROT überein. Der kalorische Test zeigte bei 74% der Patienten abnormale Ergebnisse. Beim ROT waren die häufigsten abnormalen Befunde eine erhöhte Phasenverschiebung bei zwei aufeinanderfolgenden Frequenzen (23% der Patienten) In 45% der Fälle stimmen die Resultate der beiden Testungen überein. Bei 55% dienen die beiden Testungen als Ergänzung, um die vestibulare Funktion genauer zu verstehen. Beide Testungen untersuchen den VOR in unterschiedliche Art und Weise und liefern deshalb oft andere Ergebnisse, vor allem bei leichten Formen der vestibulären Dysfunktion.</p>		

Ahmed et al., 2009		
Sensitivität unklar	Spezifität unklar	Andere Gütekriterien
Sensitivität KAL=0.55 →bei Cut-off Wert von 29.5%	Spezifität KAL=0.84 →bei Cut-off Wert von 29.5%	Validität <i>AUC-ROC Kurven:</i> Die einzelnen Parameter der ROT, wie z. B. die Gain bei 0.025 Hz (AUC-ROC von 0.62), die Phase bei 0.025 Hz (AUC-ROC von 0.58) und die Zeitkonstante (AUC-ROC von 0.61), zeigten einen begrenzten prädiktiven Wert mit AUC-ROC Werten unter 0.7.
Ergebnisse		
<p>Tieffrequente kalorische Stimuli wurden zwischen 0.002-0.004 Hz gesetzt.</p> <p>Tieffrequente ROT lagen zwischen 0.01 und 1.0 Hz.</p> <p>Die wichtigsten Ergebnisse zeigten, dass der kalorische Test mit einer Area Under the Curve (AUC-ROC) von 0,73 bei der Identifizierung peripherer vestibulärer Dysfunktionen am besten abschnitt. Der ROT hingegen hatte eine geringere prädiktive Genauigkeit. Die Kombination aus beiden Tests lieferte jedoch die höchste prädiktive Genauigkeit (AUC-ROC von 0,79) Die höchste c-Statistik von 0,79 wurde erreicht, wenn der kalorische Test mit einem umfassenden ROT kombiniert wurde. Das heisst die ROT kann als Ergänzung zur KAL (primär auch in höheren Frequenzen) genutzt werden.</p> <p>Vergleich von Betroffenen mit Referenzdaten des Herausgebers:</p> <p>Linksseitige Schwäche: Rechtsseitige Drehung → Per-rotatorische TC signifikant $p < 0.05$ / Linksseitige Drehung → Post-rotatorische Gain und TV signifikant < 0.05</p> <p>Rechtsseitige Schwäche: Rechtsseitige Drehung → Per-rotatorischer Gain und TC signifikant < 0.05 / Linksseitige Drehung → Post-rotatorischer Gain und TC signifikant < 0.05.</p>		

Ahmed, 2015		
Sensitivität	Spezifität	Andere Gütekriterien
SHA (Referenztest VST) Gain= 0.799 Phase= 0.956 Symmetry= 0.748 KAL=0.683	SHA (Referenztest VST) Gain= 0.445 Phase= 0.426 Symmetry= 0.282 KAL=0.286	Validität SHA Gain PPV= 0.794 NPV= 0.342 Phase PPV= 0.913 NPV= 0.435 Symmetry PPV= 0.572 NPV= 0.692 KAL PPV=0.756 NPV=0.567
Ergebnisse		
KAL: 72 Pat. zeigten Defizit rechts, 47 Pat. zeigten Defizit links, 18 davon gaben die Richtungsüberlegenheit an Die Studiengruppe weist signifikante Unterschiede in der vestibulären Funktion im Vergleich zu den Normalwerten auf, insbesondere in der per-rotatorischen Phase für die Rotation im Uhrzeigersinn und in der post-rotatorischen Phase für die Rotation gegen den Uhrzeigersinn. SHA (Gain, Phase, Symmetry) zeigten statistische signifikante Unterschiede im VOR-Gain in tiefem Frequenzbereich (0.01-0.08 Hz), bei der VOR-Phase in allen Frequenzen und VOR-Symmetry bei tiefen Frequenzen (0.01-0.04 Hz) ROT ergänzt den kalorischen Test (vor allem in höheren Frequenzen), indem die Testung zusätzliche diagnostische Informationen liefert, insbesondere bei der Unterscheidung zwischen zentralen und peripheren vestibulären Störungen. Positive Faktoren=Stimulus kann in jedem Frequenzbereich angewendet werden, unabhängig von Anatomie und Beschaffenheit des Ohres. Für Kinder gut geeignet, welche die KAL nicht tolerieren und bei Einnahme vestibulotoxischen Medikationen sogar sehr geeignet. Die höchste Sensitivität liefert das Protokoll des VST mit der Phase. Die zentrale Kompensation wird in der Studie berücksichtigt.		

Maes et al., 2011		
Sensitivität	Spezifität	Andere Gütekriterien
SHA 0.01 Hz= 0.839 0.05 Hz= 0.790 0.1 Hz= 0.742 KAL=0.828 VEMP=0.684	SHA 0.01 Hz= 0.636 0.05 Hz= 0.606 0.1 Hz= 0.697 KAL=0.786 VEMP=0.902	Validität SHA PPV 0.01 Hz= 0.684 0.05 Hz= 0.653 0.1 Hz= 0.697 NPV 0.01Hz= 0.808 0.05 Hz= 0.755 0.1 Hz= 0.742
Ergebnisse		
<p>Bei 0,01 Hz sind alle drei Parameter signifikant, wobei <i>Gain</i> den höchsten Wert und die höchste Signifikanz aufweist. Bei 0,1 und 0.5 Hz ist die <i>Phase</i> der einzige signifikante Parameter <0.05.</p> <p>ROT: Zeigt insgesamt die höchste diagnostische Kapazität bei 0.01 Hz mittels SHA. Ideales ROT-Protokoll bei einer Frequenz von 0.1 Hz und 0.05 Hz.</p> <p>Die KAL und VEMP zeigten jedoch bessere diagnostische Genauigkeit. KAL gilt als Goldstandard zur Identifikation von unilateralen Störungen.</p> <p>Alle drei Testungen zusammen ergeben sehr gute Ergebnisse.</p> <p>Der HST wurde unter anderem zur Abklärung einer unilateralen Problematik durchgeführt.</p> <p>Das SHA-Testprotokoll wird dem VST und PRRT vorgezogen (klinisch am relevantesten mit den drei aufgezeigten Frequenzbereichen, nur eine Frequenz ist nicht aussagekräftig genug).</p> <p>Die VEMP-Testung zeigt die höchste Spezifität. Die Kombination aller Tests erhöht somit die Gesamtvalidität und Reliabilität der diagnostischen Beurteilung.</p>		

Fattahi et al., 2024			
Sensitivität	Spezifität	Andere Gütekriterien	
0.760	0.785	Validität	PPV ROT=0.76
KAL=0.746 vKIT Gain=0.478 Alle Testungen=0.338	KAL=0.835I vKIT Gain=0.898 Alle Testungen=0.987	AUC- ROC Kurven AUC isolierte TC=0.802 AUC alle Testungen= 0.869 → KI: 0.808-0.931 AUC KAL + TC=0.864 AUC vKIT + TC=0.809 AUC vKIT + KAL=0.848	NPV ROT=0.785
Ergebnisse			
<p>Die AUC-Kurve aller Testungen, hat die höchste diagnostische Leistung, der hohe Konfidenzintervall spricht für robuste und zuverlässige Ergebnisse.</p> <p>Die Kombination von KAL und ROT zeigt bei der AUC-Kurve höhere diagnostische Leistung mit 0.864 als die Kombination der beiden Goldstandard Testungen vKIT und KAL mit 0.848.</p> <p>Die KAL und ROT waren am aussagekräftigsten in Bezug auf die Erkennung eines peripheren Defizits.</p> <p>Alle Testungen gemeinsam zeigen eine schlechte Sensitivität mit 0.338, dafür ist die Spezifität bei 0.987 und somit werden fast alle gesunden Personen identifiziert.</p> <p>Es kann keine statistisch signifikante Assoziation bei allen drei Testungen, $p > 0.05$ (peripher und zentral) bewiesen werden.</p> <p>Eine Kombination von ROT und KAL verbessert die Genauigkeit der Diagnose, im Gegensatz zu alleiniger ROT.</p>			

Eza-Nuñez et al., 2016			
Sensitivität	Spezifität	Andere Gütekriterien	
0.747	0.634	Reliabilität:	Validität
KAL=0.773 vKIT=0.747 Alle drei Testungen=0.547	KAL=0.658 vKIT=0.585 Alle drei Testungen=0.878	<i>Spearman's correlations coefficient (p):</i> →KAL vs. ROT= 0.362 →ROT vs. vKIT= 0.341 <i>Cohen`s k coefficient:</i> →CAL vs. ROT= 0.279 →ROT vs. vKIT= 0.341	PPV ROT= 0.788 NPV ROT= 0.558
Ergebnisse			
<p>Zentrale Kompensation: Die Studie hat ausreichend Kriterien, um den Test als abnormal zu betrachten, wenn sowohl das Vestibuläre Defizit wie auch die Kompensation berücksichtigt wurden.</p> <p>Die ROT war bei 71 (61,2%) Pat. abnormal. Einige davon hatten zwei abnormale Testungen: 54 (46,6%) Pat. bei der KAL und der ROT.</p> <p>Eine periphere Vestibulopathie wurde diagnostiziert, wenn alle drei Testungen Abweichungen aufzeigten: Sensitivität 0.547, Spezifität 0.878, PPV 0.891 und NPV 0.514.</p> <p>KAL zeigte die besten Ergebnisse an.</p> <p>Insgesamt zeigen die Ergebnisse, dass eine moderate Übereinstimmung zwischen den verschiedenen vestibulären Tests vorliegt. Während die Tests eine gewisse Konsistenz in den Ergebnissen aufweisen, ist die Übereinstimmung nicht perfekt, was darauf hinweist, dass jeder Test unterschiedliche Aspekte der vestibulären Funktion misst.</p> <p>vKIT statistisch als Primärtest zu empfehlen.</p>			

Arriaga et al., 2005		
Sensitivität	Spezifität	Andere Gütekriterien
0.71	0.54	PPV ROT= 0.49 NPV ROT= 0.75
Nur abnormale Phase=0.40 ENG=0.31	Nur abnormale Phase=0.83 ENG=0.86	
Ergebnisse		
<p>ROT wurde als abnormal gewertet, wenn zwei Frequenzen bei der VOR-Testung Abweichungen in Gain, Phase oder Symmetry anzeigen.</p> <p>Resultate ROT: 229 normal und 282 abnormal</p> <p>577 (58.3%) der 989 Personen zeigten Abweichungen entweder in ROT oder ENG.</p> <p>Die ROT wurde als die bessere initiale Testung gewertet, um mögliche Pathologien zu entdecken.</p> <p>Das ENG eher zur Bestätigung dieser Ergebnisse.</p>		

ROT=Drehstuhl, KAL=Kalorik, SHA=Sinusoidal Harmonic Acceleration, VST=Velocity Step Test, PRRT=Pseudorandom Rotation Test, HST=Kopfschütteltest, vKIT=Video-Kopfimpulstest, VEMP=Vestibulär evoked myogenic potential, VOR=Vestibulo-okular Reflex, PPV=Positiver prädiktiver Wert, NPV=Negativer prädiktiver Wert, TC=Zeitkonstante (Time Constant), AUC=Area under curve, ROC=Receiver Operating Characteristics, ENG=Elektronystagographie, Pat.=Patientinnen/Patienten

13.3 Selbständigkeitserklärung

Ich erkläre hiermit,

- dass ich die vorliegende Arbeit selbständig und ohne Anwendung anderer als der angegebenen Quellen und Hilfsmittel verfasst habe;
- dass ich sämtliche Quellen, wörtlich oder sinngemäss übernommene Texte, Grafiken und/oder Gedanken als solche kenntlich gemacht und sie gemäss dem „Leitfaden für schriftliche Arbeiten im Fachbereich Gesundheit“ zitiert habe;
- dass ich diese Arbeit oder Teile davon nicht bereits in gleicher oder ähnlicher Form an einer anderen Bildungsinstitution eingereicht habe;
- dass mir bekannt ist, dass ein Plagiat ein schweres akademisches Fehlverhalten darstellt und entsprechend sanktioniert wird;
- dass mir die Richtlinie über den Umgang mit Plagiaten der Berner Fachhochschule bekannt ist.

Titel der Arbeit **Vergleich Kopfschütteltest und Drehstuhltest bei unilateralen peripheren vestibulären Defiziten**
Ein Review mit Praxisempfehlung

Autorin 1:

Name, Vorname Fragner Selina

Matrikelnummer 21-259-239

Ort, Datum Bern, 29.08.2024

Unterschrift 

Autorin 2:

Name, Vorname Schwaller Samira

Matrikelnummer 21-258-975

Ort, Datum Bern, 29.08.2024

Unterschrift 