ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SALUTE, SICUREZZA E SISTEMI DEL VERDE

Ciclo 38

Gruppo Scientifico Disciplinare: 06/MEDS-26

Settore Scientifico Disciplinare: MEDS-26/C

METHODOLOGICAL CONSIDERATIONS AND NOVEL PERSPECTIVES IN VESTIBULAR PHYSIOTHERAPY

Presentata da: Nicola Ferri

Coordinatore Dottorato

Daniele Torreggiani

Supervisore

Paolo Pillastrini

Co-supervisori

Marco Tramontano

Andrea Turolla

Esame finale anno 2026

Table of Contents

Abstract

The foundation of modern evidence-based healthcare rests upon two critical pillars: the trustworthy synthesis of global research evidence and the precise application of this evidence to individual patient care. This thesis investigates a crucial disconnect between these two pillars, exploring both the systemic crisis in the reliability of synthesized evidence and the challenges in translating robust evidence into precise, instrument-guided rehabilitation practices, using vestibular rehabilitation in neurology as a paradigmatic case study.

This dissertation is structured around two complementary research lines that, together, map a pathway from evidence creation to its clinical implementation.

Line 1: Meta-research. The exponential growth of systematic reviews (SRs) has paradoxically created a challenge in discerning reliable from unreliable evidence. This line of research aimed to: (1) quantify the methodological quality and risk of bias in contemporary SRs, particularly in physiotherapy; (2) critically appraise the tools used (AMSTAR-2, ROBIS); and (3) identify and propose technological solutions to help clinicians and policymakers navigate this complex evidence landscape efficiently.

Line 2: Clinical research. Despite strong evidence for its efficacy, vestibular physical therapy (VPT) is poorly integrated into neurorehabilitation. This line of research aimed to: (1) objectively quantify the prevalence and characteristics of vestibular dysfunction in key neurological populations (Parkinson's Disease, Severe Traumatic Brain Injury, Multiple Sclerosis, Cerebral Palsy, Stroke) using the video Head Impulse Test (vHIT); (2) identify the barriers preventing the routine use of instrumental assessment and VPT; and (3) advocate for a new standard of personalized, assessment-driven rehabilitation.

A multi-method approach was employed, combining meta-research, survey-based design, evidence synthesis, and clinical observational studies.

For Line 1, we conducted cross-sectional analyses of large samples of SRs to assess methodological quality (using AMSTAR-2) and risk of bias (using ROBIS). We also performed a meta-research evaluation of the Jadad algorithm and surveyed healthcare decision-makers to assess their needs and interest in AI-assisted evidence appraisal tools.

For Line 2, we conducted instrument-based cross-sectional studies on patients with specific neurological conditions and two systematic reviews with meta-analysis of accuracy and prevalence, respectively. All participants underwent a comprehensive vestibular assessment using vHIT. Correlation analyses between vestibular parameters and clinical scales were performed. Additionally, surveys were distributed to rehabilitation professionals to identify barriers to VPT implementation.

The results from both lines of research reveal parallel and interconnected challenges. From Line 1, a pervasive crisis in the trustworthiness of SRs was confirmed. Analyses revealed critically low confidence in the results

of 90% of musculoskeletal physiotherapy SRs, with broader healthcare analyses showing low-quality or high-risk-of-bias rates of 73-81%. Primary deficiencies were found in search strategies, reporting of excluded studies, and funding source information. Crucially, Cochrane reviews consistently demonstrated superior quality, proving that high standards are achievable. The inadequacy of existing manual solutions for selecting the best evidence among many SRs on the same topic was highlighted, with the Jadad algorithm shown to be non-reproducible. In response, a strong need was identified, with 87.0% of decision-makers expressing interest in an automated, AI-informed tool (e.g., the proposed WISEST AI) to appraise and compare SRs rapidly.

From Line 2, a significant and underestimated burden of vestibular dysfunction was objectively documented. vHIT revealed a considerable prevalence of VOR gain dysfunctions in all the populations investigated. Crucially, reliance solely on mean VOR gain was found to mask isolated canal dysfunctions, which significantly correlated with balance and gait deficits in specific samples analyzed. Despite vHIT's proven diagnostic utility (84% sensitivity, 85% specificity) in ruling out stroke in acute vestibular syndrome, nearly half of rehabilitation professionals do not use instrumental assessment, citing a lack of skills and cost as key barriers. Furthermore, we demonstrated that the current diagnostic criteria for bilateral vestibulopathy may have a low sensitivity of 46%, suggesting a possible revision that includes vertical canal and otolithic function as well.

This thesis has shown that the path from evidence to effective practice is obstructed by a dual challenge: a *production crisis* at the level of evidence synthesis and an *application crisis* at the point of care. The way forward requires a synergistic, two-pronged approach. First, the evidence synthesis ecosystem must shift from quantity to quality, supported by AI-driven tools that empower users to efficiently identify trustworthy evidence. Second, clinical practice in neurorehabilitation must embrace a new paradigm of precision, moving beyond subjective assessment to integrate objective vestibular profiling for truly individualized therapy. This will necessitate educational initiatives, revised diagnostic standards, and innovative care models.

Ultimately, this work argues that strengthening healthcare decision-making requires closing the loop between these two crises: by ensuring the production of rigorously appraised, high-quality evidence and by equipping clinicians with the skills and technologies to apply this evidence with precision and confidence at the bedside.

Constructing the reliability of evidence: methodological considerations

Over the past two decades, the exponential growth of systematic reviews (SRs) has transformed evidence-based healthcare. This offers clinicians and policymakers a wealth of synthesized evidence but also creates new challenges in discerning reliability and methodological rigor. Despite their central role in clinical decision-making, many SRs differ substantially in quality, transparency, and interpretability, often leading to discordant results and conclusions. The first two studies explored the practical and cognitive challenges that arise when clinicians, researchers, and policymakers are confronted with multiple or conflicting SRs, emphasizing the need for structured frameworks and user-centered strategies to identify the most trustworthy evidence. The subsequent two works advanced this discussion by critically meta-researching the tools used to appraise SRs, specifically AMSTAR-2 and ROBIS, and by quantifying trends in the methodological quality of physiotherapy SRs over time. Collectively, these studies delineate a coherent research trajectory aimed at improving how SRs are produced, appraised, and implemented in healthcare. They ultimately advocate for a shift from quantity-driven publication practices toward a culture of methodological excellence and decision-making confidence.

Literature Inflation and Clinical Impact: Why Methods Matter?

Study #1. How can clinicians choose between conflicting and discordant systematic reviews? A replication study of the Jadad algorithm

Published as: *Lunny C, Thirugnanasampanthar SS, Kanji S, Ferri N, Pieper D, Whitelaw S, Tasnim S, Nelson H, Reid EK, Zhang JHJ, Kalkat B, Chi Y, Abdoulrezzak R, Zheng DW, Pangka LRS, Wang DXR, Safavi P, Sooch A, Kang KT, Tricco AC. How can clinicians choose between conflicting and discordant systematic reviews? A replication study of the Jadad algorithm. BMC Med Res Methodol. 2022;22(1):276.*[1]

Introduction

The exponential growth of published SRs presents a significant challenge for knowledge users who strive to base their practice on the best available evidence. When multiple SRs address the same clinical question but yield conflicting or discordant results[2], the task of identifying the most reliable evidence becomes daunting. Indeed, the number of SRs increased to 20 times from 2000 to 2019, leading to 80 SRs published each day in 2019[3]. Paradoxically, the growth in annual publications of SRs is more pronounced than that of secondary studies[4]. *Literature Inflation and Clinical Impact: Why Methods Matter?* is not merely a rhetorical question; it underscores a fundamental crisis.

In 1997, Jadad and colleagues[5] proposed an algorithm to assist healthcare providers in selecting the most trustworthy SR among multiple discordant ones. This tool was designed to navigate inconsistencies arising from differences in the research question, inclusion criteria, data extraction, methodological quality, and statistical methods across reviews (Figure 1). Despite its age and the lack of formal validation, the Jadad algorithm remains in use in contemporary research, underscoring a persistent gap in the methodological toolkit available to clinicians and researchers.

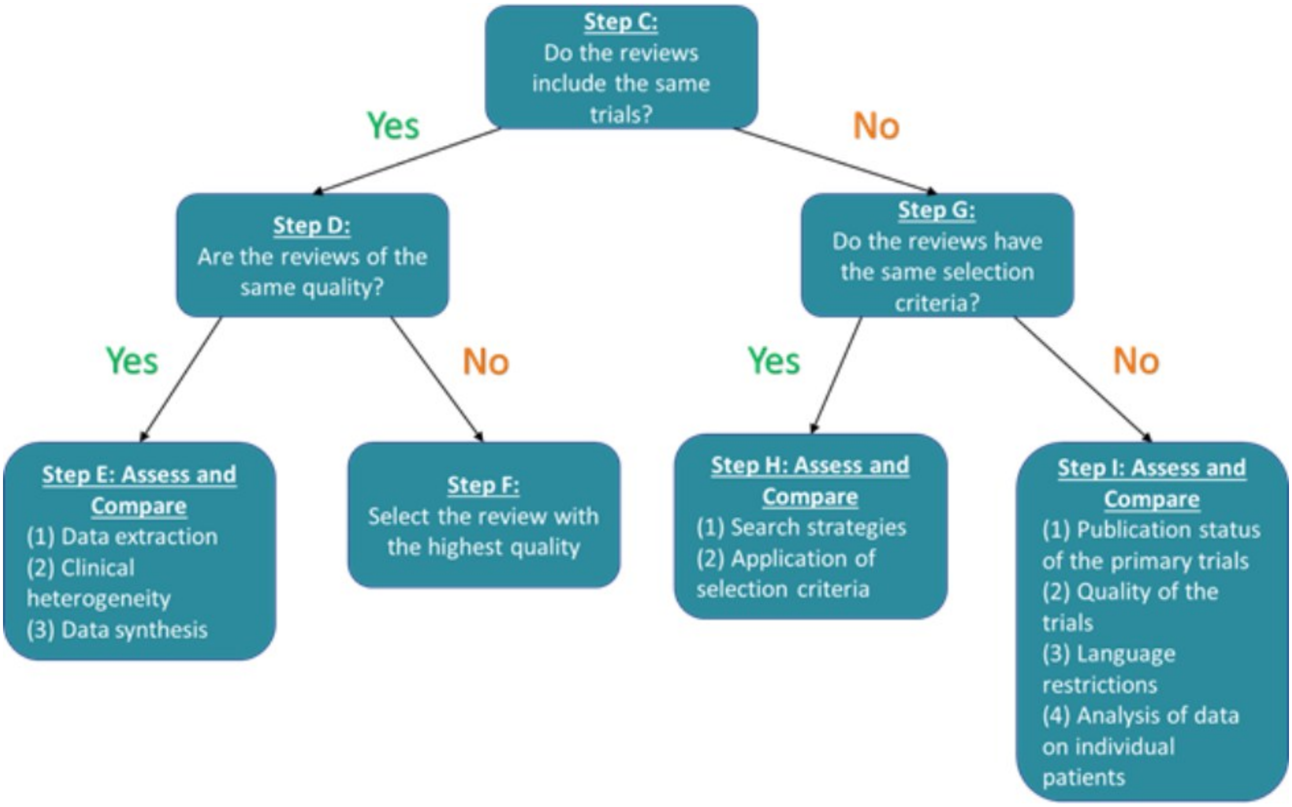

Figure 1. Jadad algorithm (source: Lunny C. et al., 2022)

However, as a stakeholder navigating this complex evidence landscape, it is crucial to understand whether such tools are reliable and reproducible. Thus, this study aimed to perform the first empirical replication of the Jadad algorithm to determine if different users would arrive at the same conclusion when applying it to the same set of conflicting reviews.

Methods

We conducted a replication study, identifying SRs of SRs that explicitly used the Jadad algorithm to choose between conflicting studies on the same clinical question. We refer to a previous bibliometric study[6] that included 1218 SRs, searching from MEDLINE, Epistemonikos, and the Cochrane Database of Systematic Reviews published from 2000 to 2020. Furthermore, we updated the search on MEDLINE in April 2021.

A key challenge we encountered was the lack of instructions in the original Jadad publication. To ensure a consistent and transparent application, our research team developed a consensus-based, standardized interpretation of each step of the algorithm, complete with detailed decision rules.[1] Furthermore, to minimize bias, the results of the original Jadad assessments in the included studies were blinded before we independently reapplied the algorithm.

The primary outcomes measured were whether the replication led to the selection of the same SR as the original authors and whether the overall direction of the findings (e.g., beneficial effect or not) was consistent.

A prospective protocol of this study has been published.[7]

Results

From the original dataset and the updated search, 21 discordant reviews[8–28] that met our eligibility criteria were included (Figure 2).

Over the 21 Jadad assessments, in 8 studies, our judgment led to the same original selection of the most appropriate SRs; instead, in 13 studies, there was disagreement in the best evidence selection between the authors and our replication process. However, of the 13 discordant Jadad assessments, the direction of the effect was consistent with our judgment in 10 studies. Thus, 86% (18/21) of the independent Jadad assessments agreed with the direction of the results, although 62% (13/21) disagreed with the selection of the best evidence.

Author groups interpreted the Jadad algorithm in very different ways. Many (12/21) did not follow its sequential steps, instead using select components. Half (10/21) considered factors outside the algorithm entirely, such as GRADE, specific statistical methods, or funding sources. Furthermore, the most common final step was Step I (12/21), which we defined as assessing publication status, RCT quality, language, and individual patient data. However, other authors reinterpreted it to mean the review's recency or to consider only published literature. Step H was the second most common final step (5/21). We defined it as comparing search strategies and applying inclusion criteria. Others, however, interpreted it as checking which journals or databases were searched, or whether publication status was reported. Some authors used Step H but ultimately selected the review with the most RCTs, a rationale not specified by the algorithm.

Completing Steps H and I of the Jadad algorithm took an average of 60 minutes per review (18/21). Assessments were generally rated as easy (9/10), especially for those who chose Step H as the final decision step, which involved comparing search strategies and inclusion criteria. The SR that selected Step I as the final decision regarding publication status and trial quality was rated as moderately difficult. Easier assessments were generally associated with better reporting and fewer systematic reviews included.

The Jadad algorithm was found to lack several key methodological considerations. For example, tt does not account for the recency of the literature search or publication, the number of primary studies included, or the certainty of the evidence (e.g., as graded by GRADE).

Using the ROBIS tool, we assessed 98 SRs included in the 21 discordant reviews. Of these, 41 SRs were at low risk of bias, and 57 SRs were at high risk. In our replicated Jadad assessments, we finally selected 19 low-risk and 9 high-risk reviews as the best evidence, a result similar to the original authors' selections (17 low-risk and 8 high-risk).

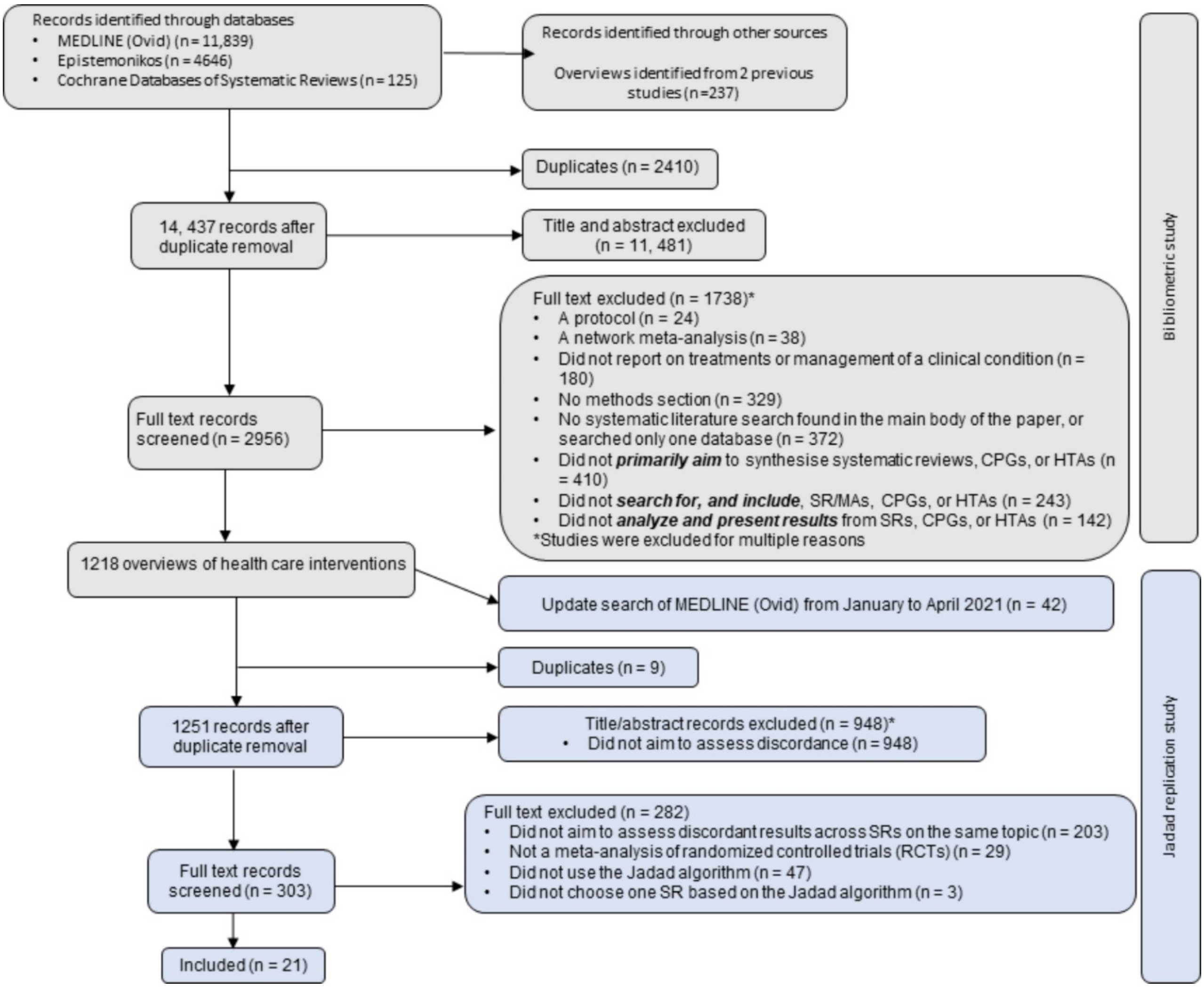

Figure 2. PRISMA flow-chart (source: Lunny C. et al., 2022)

Discussion

Our replication of the Jadad algorithm revealed a 62% disagreement rate with the original authors on which SR constituted "best evidence." This lack of reproducibility stems from the algorithm's vague instructions and inconsistent application by users, who often omit steps or introduce external criteria. While the algorithm was rated as easy to use (taking ~60 minutes for a partial assessment), its process was lengthened when there was a high number of SRs, quality assessments were required, or poor reporting was present.

Crucially, the Jadad algorithm is outdated and lacks comprehensiveness. It fails to account for the recency of a review's literature search, the number of included trials, the certainty of evidence (e.g., GRADE), and most importantly, the methodological quality or risk of bias of the SRs themselves. This is a fundamental design flaw, as quality assessment is only triggered if SRs include identical trials, a scenario that did not occur in our sample.

To date, no prior replications of the Jadad algorithm exist, highlighting a significant evidence gap. Thus, our study was the first to test its reliability empirically. It is clear that, without a reliable tool, clinicians may face substantial challenges when encountering discordant SRs. This uncertainty can directly impact patient care and policy, as seen in cases where reimbursement decisions were based on conflicting reviews. In addition, we think that selecting a single SR simplifies knowledge translation, but it risks losing valuable context and information.

This replication study has several strengths. First, we used consensus-based standardized instructions to operationalize the algorithm. Second, we conducted a comprehensive and updated search strategy without restrictions, in order to reduce publication bias. Lastly, two independent reviewers conducted each judgment process to minimize selection bias, and both screening and assessment procedures were piloted to ensure consistency between the authors' judgments.

However, some limitations must be acknowledged. Even though our research team is international, multicultural, and multidisciplinary, there may be a lack of clinical expertise for all the topics included. Furthermore, a broader search strategy may have identified more discordant reviews. Finally, orthopaedics was the most representative topic in our replication study, thus limiting generalizability.

Given the lack of reliability and robustness in the Jadad algorithm's methodology, there is an urgent need for a new, validated tool that ensures better user agreement, incorporates modern methodological advances (such as ROBIS and GRADE), and is applicable across diverse settings. This first study is actually part of a large international project (WISEST - WhIch Systematic rEview iS best) to develop a new automated AI-driven tool that will support knowledge users in selecting the best evidence when many discordant SRs are available on the same topic.

Conclusion

Our findings suggest that the Jadad algorithm is not reproducible and is methodologically obsolete. Pending the development of an optimal tool, we recommend clinicians choose among conflicting SRs by prioritizing those that are most relevant to their research question, most recent, most comprehensive (in terms of the number of included trials), and of the highest quality (lowest risk of bias).

Study #2. Exploring decision-makers' challenges and strategies when selecting multiple systematic reviews: insights for AI decision support tools in healthcare

Published as: *Lunny C, Whitelaw S, Reid EK, Chi Y, Ferri N, Zhang JHJ, Pieper D, Kanji S, Veroniki AA, Shea B, Dourka J, Ardern C, Pham B, Bagheri E, Tricco AC. Exploring decision-makers' challenges and strategies when selecting multiple systematic reviews: insights for AI decision support tools in healthcare. BMJ Open. 2024;14(7):e084124.*[29]

Introduction

SRs are a cornerstone of evidence-based healthcare.[2,30] However, their rapid proliferation means that decision-makers often encounter multiple SRs on the same topic.[31,32] This can lead to confusion, especially when reviews present conflicting results for the same research question. However, it remains unclear to date whether researchers or clinicians truly feel this difficulty, and what the obstacles and challenges are in translating knowledge into these discordant situations. Thus, this study aimed to understand how healthcare decision-makers (e.g., practitioners, policymakers, researchers) select among multiple discordant SRs and to gauge interest in an artificial intelligence (AI) tool to assist in this process.

Methods

This is a cross-sectional study conducted as a web-based survey, reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines.[33] The prospective protocol has been uploaded at https://osf.io/nbcta.

We developed a 21-question survey in English, mainly consisting of closed-ended questions, using Qualtrics software (Qualtrics Labs, Provo, Utah). The survey was organized into three sections: (a) demographic information, (b) experiences and barriers in selecting from multiple SRs on the same topic, and (c) key factors considered when making that selection. Respondents could skip questions and revise their answers before final submission. The full survey questions, CHERRIES checklist, dissemination list, and responses by decision maker types are available as Supplementary Material of the published article on BMJ Open at https://bmjopen.bmj.com/content/bmjopen/14/7/e084124/DC1/embed/inline-supplementary-material-1.pdf?download=true.

In this research, our primary goal was to identify the specific SR features that decision-makers value most when comparing reviews, as this information is vital for the development of the WISEST tool. To avoid biasing respondents by presenting a pre-determined list of features, we randomly assigned one of two similar question formats to participants: one group (Q11) selected from a multiple-choice list of features, while the other (Q12a-c) analyzed a case study of three similar SRs and then indicated which review(s) they would use and why.

The survey was piloted by eight academics and refined iteratively to enhance its clarity and validity.

A sample size of 440 decision-makers was calculated as necessary to detect a 50% mean difference between groups with 90% power and a 5% significance level. To achieve this, we contacted approximately 1,760 individuals, anticipating a response rate of 25%.

The survey targeted individuals from organizations that produce SRs (e.g., Cochrane Collaboration, Joanna Briggs Institute, US Agency for Healthcare Research & Quality's Evidence-Based Practice Centre, GRADE group, etc.) as well as a broad range of SR users. Indeed, we compiled an email list of major evidence-synthesis groups and also distributed the survey anonymously to 28 additional organizations via professional networks. It was further promoted through a university methods series, e-newsletters, professional healthcare associations, and social media platforms, including LinkedIn and Twitter, utilizing targeted advertisements and hashtags. Thanks to our international authors, dissemination was broad and transdisciplinary. Following the Dillman method[34], we sent three reminder emails and maintained consistent social media promotion. The survey was active from July 19 to August 19, 2022, with reminders sent bi-weekly. The estimated completion time was 10 minutes.

Responses were exported from Qualtrics to Microsoft Excel for analysis. Incomplete responses were included, but entirely blank questionnaires were excluded. Completion times were recorded for all attempts. We used descriptive statistics to analyze closed-ended questions. For open-ended responses, one researcher independently identified and coded emerging themes, and these responses were then summarized using counts and frequencies. A stratified descriptive analysis was performed in relation to the type of decision-maker; if a respondent identified with multiple roles, their answers were counted for each relevant category. Finally, we compared the results from the two randomly assigned question formats (Q11 and Q12).

Results

Overall, 3158 email invitations were distributed to advertise the survey. After 197 emails failed to deliver, 2961 invitations were successfully received. Following the removal of 25 duplicate and 83 blank responses, 684 completed surveys were included in the final analysis (Figure 3). The majority of respondents (n = 450) accessed the survey via an anonymous link on social media or newsletters, while a smaller group (n = 234) used the direct Qualtrics email link. The final response rate to email invitations was 7.9%, falling short of the anticipated 25%.

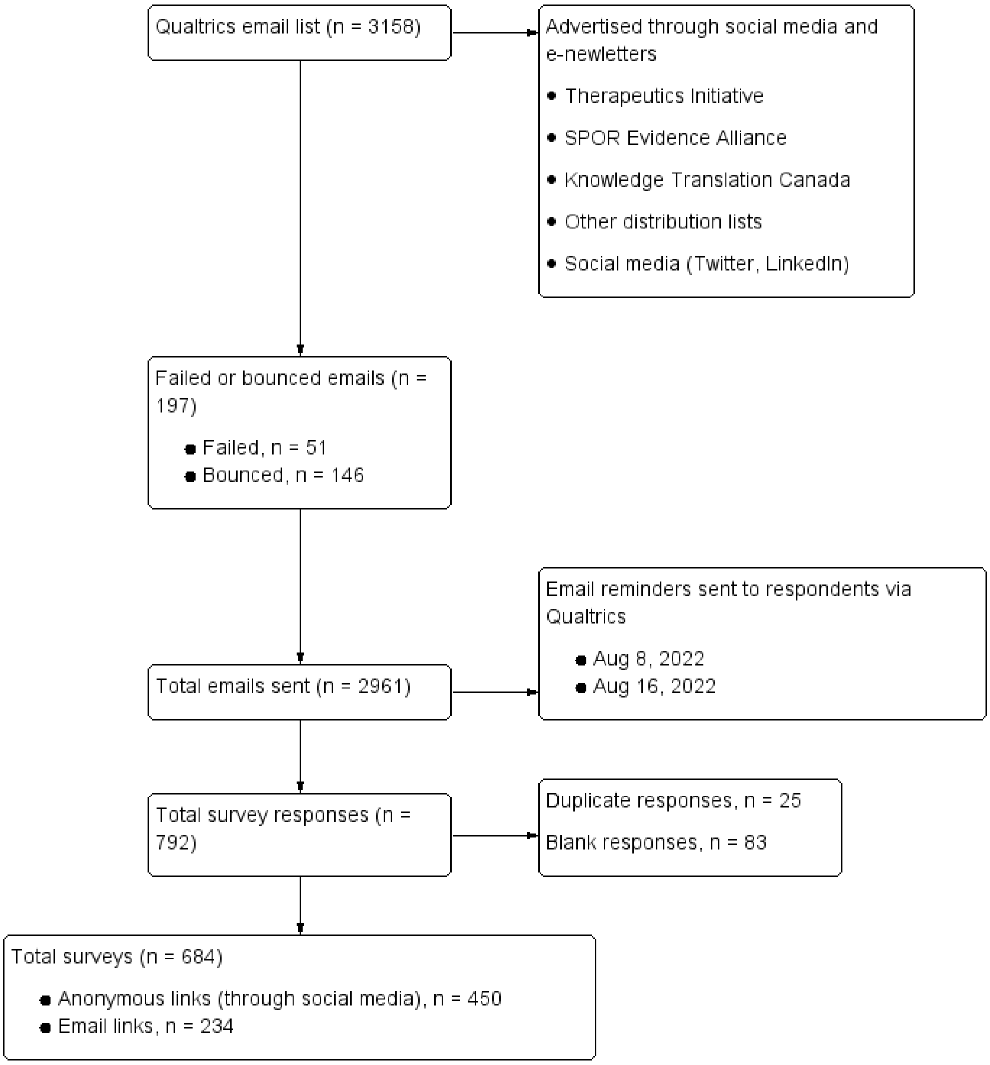

Figure 3. Survey flowchart (source: Lunny C. et al., 2024)

Among the 684 respondents, 462 (67.5%) answered all questions, and 97 (14.2%) completed less than half. The median completion time was 11 minutes.

The respondent pool was predominantly composed of researchers (58.2%), practitioners (37.1%), students/trainees (19.2%), and policymakers (13.5%), with many individuals identifying with multiple roles. For instance, most policymakers (59.7%) also identified as researchers. Geographically, the majority of respondents were from North America (52.8%) and Europe (34.2%).

A significant proportion of respondents (90.8%) reported familiarity with SRs. To objectively assess understanding, two knowledge-testing questions were included: a majority (61.8%) correctly identified that the PRISMA checklist is a reporting guideline, not a tool for assessing methodological quality. Just over half (56.3%) correctly identified that the risk of bias in primary studies was not a limitation at the SR level. A post-hoc analysis revealed that respondents recruited via targeted email lists demonstrated a better understanding of these concepts than those recruited via anonymous social media links.

Nearly all respondents frequently used SRs to inform their decisions (97.1%) and encountered multiple SRs on the same topic (97.9%). Some users (50.8%) reported struggling to choose the most valid and trustworthy SR.

Overall, the primary barriers to selection were:

- Lack of time to read each SR thoroughly (55.2%).
- Variability in the methodological quality of SRs (54.2%).
- Differences in results and conclusions across SRs (49.7%).
- Variation in the primary studies included in the SRs (44.6%).

Approaches to selecting a SR varied by role:

- Practitioners most often selected the most recently published relevant SR (44.1%).
- Policymakers and researchers were more likely to find and assess all relevant SRs (54.8% and 41.7%, respectively).

A substantial majority of respondents (87.0%) expressed interest in a free, automated AI tool to assist in comparing and selecting SRs, with practitioners showing the highest level of interest (90.5%).

When asked about the most important features for choosing an SR, respondents prioritized:

1. Relevance of the SR's research question to their own (83.6%).
2. Methodological quality and reproducibility of the SR (79.2%).
3. Recency of the literature search (74.8%).
4. Comprehensiveness of the search strategy (69.0%).

This was supported by a case study, where respondents most frequently selected a SR based on a hierarchy of multiple features (73.1%) rather than a single criterion. The number of included studies or patients was also a key consideration, particularly for practitioners.

Discussion

This survey of policymakers, practitioners, and researchers revealed that respondents, who possessed a good understanding of SRs, frequently used them to inform decisions and often encountered multiple SRs on a single topic. Nearly half struggled to select the most trustworthy SR, primarily due to time constraints, variability in the SRs' methodological quality and in the included primary studies, and differences in the key findings. The proposed use of an AI tool to assist in this comparison was widely supported.

When comparing SRs, key criteria were the relevance to their specific question, the robustness and recency of the search, and the methodological quality. Selection strategies differed by role: practitioners often prioritized the most recent review, while policymakers and researchers typically appraised all relevant SRs using a hierarchy of criteria. This suggests that practitioners may prefer a tool providing a ranked synthesis, while policymakers and researchers seek distilled data from all reviews to inform their own judgments.

A further theme was that there is rarely a single "best" SR. A review with a high-quality rating might still have significant flaws or omit critical data. This complexity often led respondents, particularly policymakers, to review all available SRs to fully assess their strengths and weaknesses before making a decision.

This study has some key strengths. Indeed, the survey was conducted in adherence to a pre-published protocol. Second, we employed a multimodal strategy (using newsletters, email lists, and social media) to recruit a global range of decision-makers. This approach, which included reminders and repeated social media posts, successfully broadened our reach, although it limited our control over exposure. Lastly, to our knowledge, this is the first study to investigate decision-makers' interest in AI tools for navigating SR evidence.

However, several limitations should be noted too. The email response rate was lower than anticipated (under 10% versus an expected 25%). Furthermore, a precise overall response rate was unattainable as two-thirds of responses came via anonymous social media links. While this multimodal strategy was necessary to maximize participation, it came at the cost of calculable metrics. Moreover, the survey was piloted solely by academics; maybe involving a more diverse group in the piloting process could have improved the applicability of the questions and the response rate. Lastly, our sample also exhibited potential response bias, with an over-representation of decision-makers from high-income countries and those involved in guideline development and policymaking. Consequently, we may have missed other important SR users.

The growing number of SRs on the same topic will likely intensify the challenge of comparing and choosing between them in the future. Our survey indicates that over one-third of decision-makers currently base their choice of SR solely on the title and abstract, rather than the full text. However, titles and abstracts often lack sufficient detail for an informed decision.

While an AI tool can aid in comparing SRs, it cannot compensate for inherent methodological flaws in the reviews themselves. The fundamental responsibility for producing high-quality evidence lies with researchers who design rigorous studies and journals that uphold high methodological standards in publication. This underscores the need for enhanced training for researchers, peer reviewers, and editors in conducting and appraising SRs.

Our findings confirm the potential for an AI tool to help decision-makers efficiently compare SRs. Since user needs vary by role, we are developing the WISEST tool to be suitable for a diverse range of users. WISEST will display the scope, strengths, limitations, and a quality ranking of all relevant SRs on a topic, thereby supporting, not replacing, the user's own clinical and methodological judgment.

The survey also revealed that many SR users lack confidence in their ability to assess methodological quality and often struggle to identify biases at the systematic review level. Therefore, the WISEST tool must be complemented by supporting education and training resources to bridge this knowledge gap effectively.

Conclusion

Policymakers, practitioners, and researchers frequently use SRs for decision-making, but they commonly encounter multiple SRs on a single topic. Slightly more than half of the surveyed individuals reported difficulty in selecting the most valid and trustworthy SR. This challenge was attributed to a lack of time and the complexity of comparing SRs with varying methodological quality and characteristics. Thus, an AI tool designed to quickly outline the features, strengths, and limitations of SRs could effectively address these challenges and support healthcare decision-making.

Study #3. Exploring the methodological quality and risk of bias in 200 systematic reviews: A comparative study of ROBIS and AMSTAR-2 tools

Published as: *in press on Research Synthesis Methods*

Introduction

Critically appraising research is essential for informed healthcare decisions. Policymakers, clinicians, and patients rely on high-quality evidence, with SRs being the most reliable source. A 2022 survey[29] confirmed that 98% of decision-makers use SRs, yet approximately 40% struggle to choose between the numerous options available on a single topic.

The concept of SR quality remains vague in the literature, as it can imply multiple constructs spanning various domains.[35] Indeed, it is crucial to distinguish between a SR methodological quality (how well it was conducted), its reporting quality (how completely it is described), and its risk of bias (are there limitations in the SR process that could have systematically skewed the results away from the truth).

For example, bias in SRs, or meta-bias, can lead to over- or under-estimating an intervention's true effect. Furthermore, SR authors may "spin" their conclusions, presenting results in an overly favorable light, which poses a separate risk of bias in the interpretation of the findings.[36–39]

To evaluate potential bias in a SR, assessors can use the ROBIS tool,[40] while for assessing methodological quality, the AMSTAR-2 tool is available.[41] Both are current, validated tools endorsed by the Cochrane Handbook, with AMSTAR-2 probably most widely used.[42]

AMSTAR-2[41] consists of 16 items to judge how well a SR was conducted. However, its lack of clear guidance on some items can lead to varying interpretations.[43,44] ROBIS assesses the risk of bias across 21 items in four domains and can be applied to various types of reviews.[40] While ROBIS offers detailed instructions, it has a steeper learning curve.[45,46]

Using both tools requires significant training, and their application is highly dependent on the assessor's expertise and experience.[43] The time required for assessment is substantial (14-60 minutes per tool, excluding reading time), and proper practice necessitates two independent assessors, further increasing the time commitment.[46]

Thus, this study applies both AMSTAR-2 and ROBIS to a sample of 200 SRs to: (a) map and compare both tools' items; (b) determine the methodological quality and risk of bias in the SRs; (c) compare assessment times; and (d) calculate inter-assessor agreement.

Methods

This is a meta-research study that was conducted following SR guidance[47] and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)[48] guidelines for cross-sectional studies. A prospective protocol was stored on https://osf.io/nbcta/.

We included all kinds of SRs except for overviews of reviews, SRs including qualitative studies, methodological SRs, and scoping reviews. Two published datasets[45,49] of SRs were leveraged to collect studies that met our eligibility criteria. Then, two assessors independently screened all titles/abstracts and full texts for inclusion.

Since the AMSTAR-2 and ROBIS tools require extensive training and are prone to inconsistent interpretation of items, this ambiguity gives users significant leeway; for instance, defining a 'comprehensive search strategy' as described in item 4 of AMSTAR-2 is subjective and can be interpreted differently. Thus, to ensure standardized application, we developed specific decision rules and clear data extraction instructions.

To assess the synthesis sections, we first identified the primary outcome, defined as the one explicitly stated in key manuscript sections, the one used for power calculations, or the first outcome mentioned. For the primary intervention, we selected the experimental intervention reported in the manuscript versus a control, or the first one highlighted in the title/abstract.

Assessors documented their judgments alongside a verbatim quote from the SR to justify their evaluations. They also rated the clarity of each item assessment as: *clear* (i.e., easy to judge), *weak* (i.e., vague or incomplete), *different* (i.e., contradictory text), or *no information*.

Data were extracted on the SRs characteristics and eligibility criteria, presence of a certainty of evidence assessment (i.e., GRADE [Grading of Recommendations, Assessment, Development, and Evaluation][50]), and authors' interpretation of results (i.e., spin or overestimation bias). We synthesized the results with descriptive statistics.

Results

Among 200 analyzed SRs, 85% were published before 2018. The median number of authors was five, with first authors primarily from Europe, North America, and Asia. Half of the SRs focused on three medical topics. Most (73%) were interventional reviews, and pairwise meta-analysis of RCTs was the most common synthesis method (45%).

The ROBIS (24 items) and AMSTAR-2 (16 items) tools were compared. ROBIS includes a phase for assessing the applicability of findings, which AMSTAR-2 does not. The tools had considerable conceptual overlap, though some items were unique to each. For instance, nine ROBIS items and three AMSTAR-2 items did not sufficiently overlap, covering distinct aspects like relevance, reporting comprehensiveness, and specific biases.

Both tools indicated generally poor quality among the SRs. ROBIS assessments found 81% of SRs had a high risk of bias, while AMSTAR-2 rated 73% as low or critically low quality. Cochrane reviews consistently outperformed non-Cochrane reviews. Over half of Cochrane SRs were rated as low risk/high quality, whereas only a single non-Cochrane SR received this rating from each of the tools. SRs with and without meta-analysis showed similar proportions of low-risk/high-quality ratings.

Item-level analysis revealed that Cochrane SRs, especially those with meta-analysis, achieved high positive rates (often >90%) for most items in both tools. In stark contrast, non-Cochrane SRs had much lower positive rates, sometimes below 15% for critical items like having a pre-defined protocol.

A small proportion of SRs (9%) received discordant overall ratings from the two tools (e.g., high risk but high quality). The unique, non-overlapping items of each tool primarily drove these discrepancies.

Inter-assessor agreement was high for most items, with only 25% falling below 70% agreement. The median time to complete an assessment was longer for ROBIS (64 minutes) than for AMSTAR-2 (51 minutes) when applied first. However, the time per item was slightly faster for ROBIS. Assessors reached a plateau in assessment time much more quickly with AMSTAR-2 (after ~5 reviews) than with ROBIS (after ~12 reviews), suggesting a steeper learning curve for ROBIS.

Discussion

This study assessed 200 SRs using the AMSTAR-2 and ROBIS tools. While significant item overlap exists, their unique components mean ratings can conflict; 9% of SRs received high-quality AMSTAR-2 ratings but high-risk ROBIS ratings. Indeed, ROBIS is superior for assessing bias in conclusions and generalizability, while AMSTAR-2 focuses on methodological quality.

Most SRs were of low quality and high risk of bias, especially non-Cochrane reviews and those without meta-analyses. Cochrane reviews consistently demonstrated higher quality. A significant issue was the inadequate reporting, which hindered complete assessment for over 50% of SRs on key items like protocol registration and excluded studies lists. AMSTAR-2 was faster to complete (median 51 min) than ROBIS (median 64 min). High inter-assessor agreement was achieved, likely due to extensive training and piloting. The choice of tool depends on the user's goal: ROBIS provides a fuller assessment of bias in conclusions, while AMSTAR-2 is a faster option for evaluating methodological quality.

Some study limitations have to be acknowledged. First, the sample was not representative, over-including high-quality Cochrane reviews. Second, not all assessments underwent senior verification. Third, AMSTAR-2 was applied to some non-interventional SRs, for which it was not designed. Lastly, agreement was measured by percentage, which overestimates reliability compared to chance-corrected statistics.

Conclusion

In summary, most systematic reviews assessed were of low/critically low quality (AMSTAR-2) and high risk of bias (ROBIS). While the tools share substantial overlap, they are not directly comparable due to unique items in each. ROBIS offers a more comprehensive evaluation of meta-biases, including external validity and bias in review conclusions. In contrast, AMSTAR-2 is faster to apply and uniquely assesses the rationale for study designs and conflicts of interest. Therefore, AMSTAR-2 may be suitable for rapid methodological quality assessment, while ROBIS could be preferable for a thorough bias evaluation.

Quality in Rehabilitation Evidence: A Rude Awakening

Study #4. The confidence in the results of physiotherapy systematic reviews in the musculoskeletal field is not increasing over time: a meta-epidemiological study using AMSTAR 2 tool.

Published as: *Ferri N, Ravizzotti E, Bracci A, Carreras G, Pillastrini P, Di Bari M. The confidence in the results of physiotherapy systematic reviews in the musculoskeletal field is not increasing over time: a meta-epidemiological study using AMSTAR 2 tool. J Clin Epidemiol. 2024;169:111303.*[51]

Introduction

SRs with meta-analysis are a cornerstone of evidence-based medicine, providing the highest level of evidence by synthesizing RCTs for guideline developers and policymakers.[52] Their publication has grown exponentially, with an estimated 80 new SRs published daily.[3]

Physiotherapy is a globally relevant topic, with musculoskeletal disorders the most common indication for treatment and affecting an estimated 1.71 billion people worldwide.[53] Despite this volume of research, the overall quality of SRs and their impact on practice remain unclear. This uncertainty can hinder clinical progress and credibility. Indeed, recent meta-research has consistently found low confidence in SRs across diverse medical fields.[54–60]

This "quantity over quality" issue is partly driven by academic pressure to publish.[61] However, a recent survey indicated that over 50% of decision-makers report difficulty in selecting the best evidence, primarily prioritizing factors such as methodological quality and the reputation of the journal or authors.[29] In response, new methodologies have been proposed to enhance knowledge translation and support clinical reasoning in the face of uncertainty.[62–64]

Given the essential role of physiotherapy in global health and the recent literature inflation, this study aimed to analyze the methodological quality of SRs of RCTs on physiotherapy for musculoskeletal conditions published over the last decade and to explore any related predictors.

Methods

This is an observational, meta-research study in the field of physiotherapy. In the absence of specific reporting guidelines for such type of studies, we followed the Cochrane Handbook[47] for selection and data extraction processes, and adhered to the STROBE[48] guidelines for reporting. The study was prospectively registered on the Open Science Framework (https://osf.io/bc8zw/).

We included SRs of RCTs from the last 10 years, focusing on the clinical effectiveness of physiotherapy for musculoskeletal conditions. All physiotherapy-administered treatments and all comparator interventions were eligible. We excluded economic studies, network meta-analyses, and other non-standard evidence syntheses.

We searched four major databases (MEDLINE, Cochrane Database of Systematic Reviews, CINAHL, and PEDro) from 2012 to 2022. After deduplication, two researchers independently screened records by title and abstract, and then by full text, against the eligibility criteria. From the eligible full-text SRs, the first 100 were selected from a randomized list for inclusion.

A predefined data sheet was used for extraction, including authors, journal, and study details. The methodological quality of each included SR was appraised by two independent reviewers using the AMSTAR 2 tool, which provides an overall confidence rating (high, moderate, low, or critically low) rather than a summary score, according to the recommendations of Shea et al.[41]

Data were analyzed using descriptive statistics. Variables were compared across the AMSTAR 2 confidence categories using appropriate statistical tests. Indeed, an ordinal logistic regression model was built to identify predictors of the overall confidence rating among the extracted variables.

Results

The initial search yielded 2,078 records. After removing 155 duplicates and 2 retracted papers, 1,921 records were screened by title and abstract. This led to a full-text review of 475 articles, from which 395 SRs were ultimately included in the study. A random sample of 100 SRs was selected for detailed analysis from this larger pool. The whole process is reported in Figure 4.

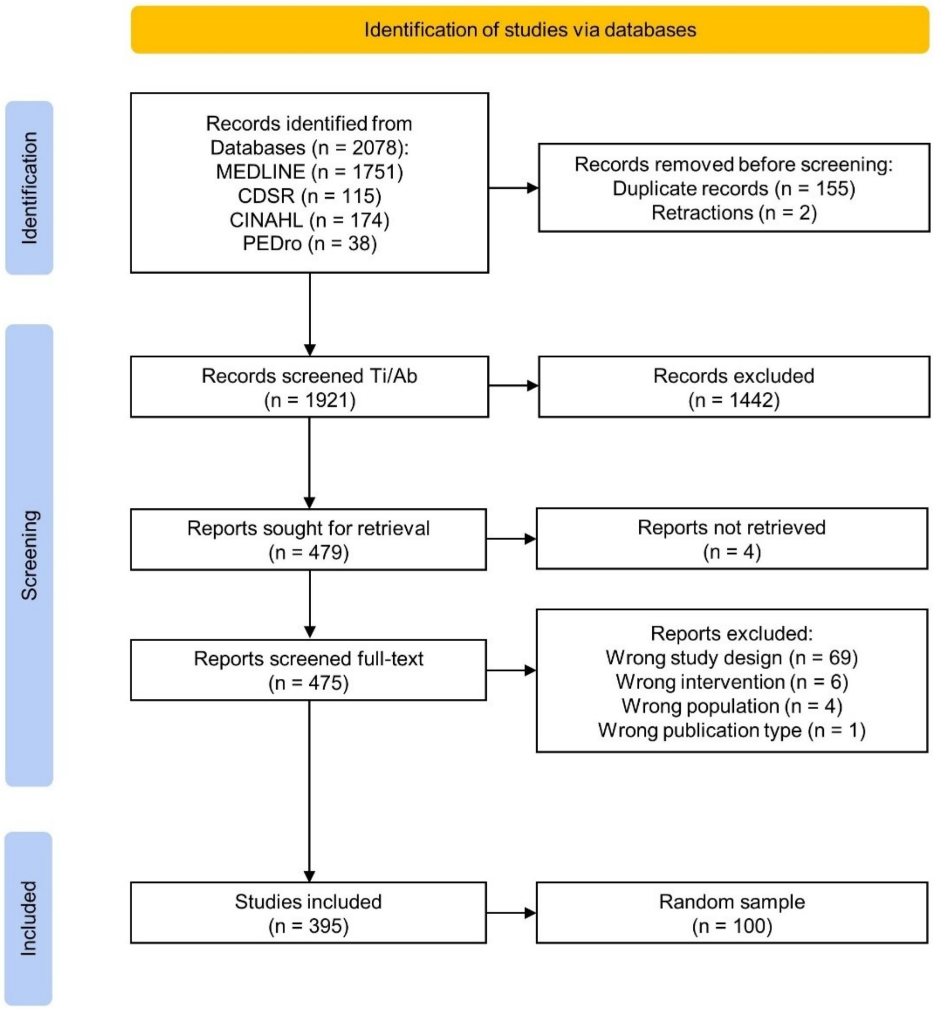

Figure 4. PRISMA flow diagram (source: Ferri et al., 2024)

The sampled SRs covered various interventions commonly used in musculoskeletal physiotherapy, with manual therapy, exercise, physical agents, and their combinations being the most prevalent (Figure 5).

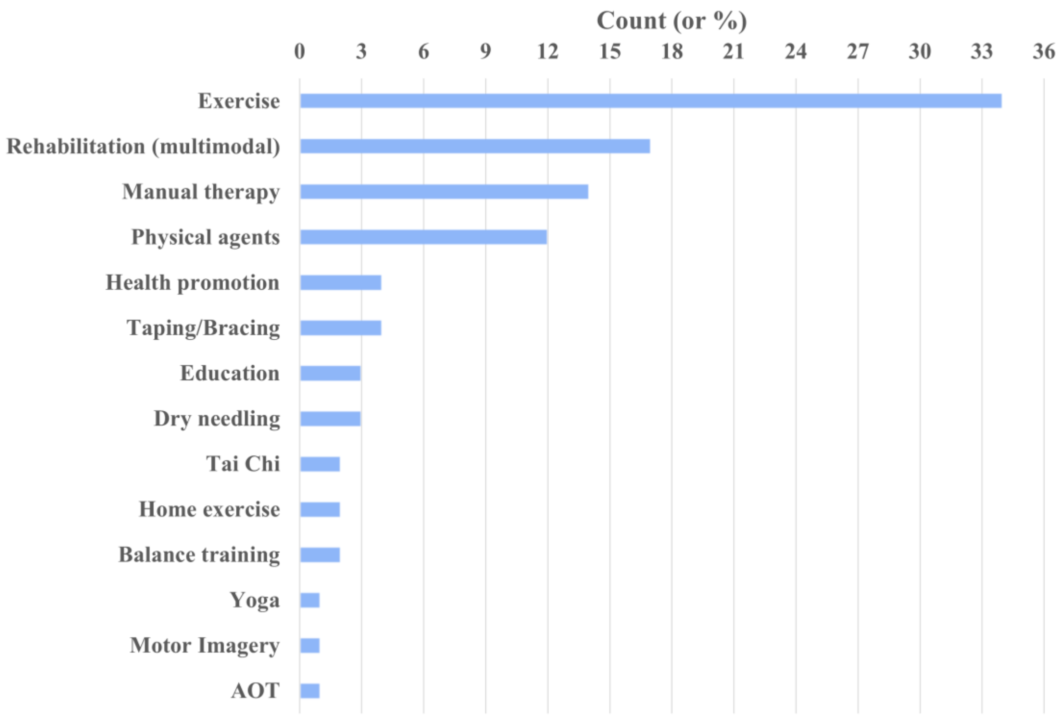

Figure 5. Interventions considered (source: Ferri et al., 2024)

The overall methodological quality of the SRs was very low, with 90% of the sampled reviews rated as having "critically low" confidence in their results according to the AMSTAR 2 tool. This poor quality showed no sign of improvement over the years analyzed (2013-2022).

While a statistically significant association was found between a higher Journal Impact Factor (JIF) and better AMSTAR 2 ratings, this relationship was entirely driven by Cochrane reviews. When Cochrane reviews were excluded from the analysis, the JIF was no longer a predictor of quality.

The SRs originated from all continents. The analysis revealed that the scientific influence (H-index) of the last author was a significant predictor of SR quality, with a one-unit increase in the last author's H-index slightly increasing the odds of a higher AMSTAR 2 rating. Having more authors on the review also showed a borderline significant association with better quality. Crucially, all SRs with "moderate" or "high" confidence ratings were exclusively Cochrane reviews.

An in-depth look at specific methodological items revealed common weaknesses: 93% of SRs failed to justify their chosen study designs, 78% did not provide a list of excluded studies, and 90% did not investigate the funding sources of the primary studies they included. Furthermore, less than 10% employed a truly comprehensive search strategy, which is a critical flaw. Strengths included transparent reporting of the research question (i.e., PICO), declaration of conflicts of interest, and a well-described study selection process.

Finally, our statistical model (Table 1) revealed that the JIF and the year of publication (specifically before or after the introduction of the AMSTAR 2 tool) were not independent predictors of an SR's quality. The most significant predictor was the last author's H-index.

Table 1. Multivariable ordinal logistic regression (n = 100)

Variable	OR (95% CI)
Journal IF	1.00 (0.99-1.01)
H-index last author	1.04 (1.01-1.06)
Number of authors	1.53 (0.99-2.39)
Intercept 1	5.8
Intercept 2	6.44
Intercept 3	6.88

Abbreviations: OR, odds ratio; CI, confidence interval; IF, impact factor.

Discussion

This analysis of SRs in musculoskeletal physiotherapy found that 90% were rated with critically low confidence, a figure that showed no improvement over time. The only factor predicting higher confidence was the senior author's H-index.

This prevalence of low-quality evidence is concerning, as healthcare professionals are ethically and legally bound to use the best available evidence. Our findings align with prior research, which has also identified similar issues in the field.[65] While the AMSTAR 2 tool can be challenging to apply, partly because it assesses methodological quality through reported details, we employed rigorous methods, including extensive training and consensus, to ensure a reliable assessment.

The tool itself may be exceptionally rigorous,[66] potentially contributing to the low scores, even for older Cochrane reviews. For instance, marking a SR down for not providing a list of excluded studies (a PRISMA requirement) is a stringent criterion, as authors may have this list but omit it from the publication. Despite this, the fact that some Cochrane reviews achieved high confidence proves that superior methodological quality is attainable, underscoring that improvement remains a priority.

Furthermore, the predictive power of the senior author's H-index may be logical, as this individual often oversees the study's scientific integrity and offers validation for a metric that is frequently criticized.[67–70]

This study's strengths include a comprehensive selection process and a statistically sound approach. However, some limitations have to be acknowledged. In particular, we used AMSTAR 2 (published in 2017) to assess older studies, although its principles were established in earlier guidelines. Then, for resource optimization, we analyzed a large randomly selected sample rather than the entire population of SRs. Moreover, we chose the H-index, a dynamic metric that can vary significantly even within the same year. We must also acknowledge the statistical challenges due to the imbalanced data, represented by the "rare event" of high-quality reviews. Lastly, we are aware that a simplified "open-access" classification may not reflect more complex publishing models.

Conclusion

Evidence-based medicine integrates patient preferences with the most robust scientific evidence to guide clinical decisions. This study reveals that SRs on musculoskeletal disorders, the primary reason for physiotherapy worldwide, suffer from very low methodological quality. This directly threatens patient care and places more liability on healthcare providers. To improve research quality, stricter adherence to methodological and reporting standards is essential. Strengthening the prepublication peer-review process could be an effective strategy to achieve this.

General discussion and future perspectives

The collective evidence consistently highlights a pervasive crisis in the trustworthiness of systematic reviews. This crisis, characterized by critically low methodological quality and a high risk of bias across various healthcare domains, urgently demands structural and technological solutions to safeguard evidence-based decision-making. The overwhelming majority of reviews analyzed, particularly in musculoskeletal physiotherapy, exhibit critically low confidence in results (90%). This trend is mirrored in broader assessments, where low quality or high risk of bias rates reach 73% to 81%, confirming that this issue is not limited to specific medical specialties or review types. This compromised quality primarily results from inadequate adherence to rigorous methodology and reporting standards. Deficiencies are frequently noted in critical areas such as comprehensively defining the search strategy, listing excluded studies, and reporting the funding sources of primary studies. Conversely, the analyses consistently demonstrate that Cochrane reviews are predominantly represented in the higher AMSTAR 2 confidence levels and are generally found to be at lower risk of bias compared to non-Cochrane reviews, suggesting that high methodological standards are still achievable. For decision-makers, who are constantly challenged by the lack of time and the difficulty comparing methodological quality across multiple SRs, existing manual solutions are inadequate; specifically, the Jadad algorithm has been demonstrated to be non-reproducible and lacking in comprehensiveness, failing to account for critical modern factors like recency and systematic assessment of methodological quality. Looking forward, the pathway to resolving this crisis involves a twofold approach: first, an imperative focus on improving the quality of research itself through greater adherence to methodological guidelines (like PRISMA) and careful assessment during the peer-review process. Second, the future of evidence synthesis navigation lies in leveraging AI tools to efficiently assist users in critical appraisal and comparison. The strong reported interest in an automated, AI-informed tool to rapidly highlight the features, strengths, and weaknesses of SRs (with 87.0% of decision-makers expressing interest) confirms the market need for systems like the proposed WISEST AI tool. Such tools will aim to provide a supporting role, distilling complex information related to relevance, search robustness, recency, and methodological quality, thereby helping users make informed judgments when faced with the necessary trade-offs in clinical, public health, and policy decisions. Ultimately, strengthening the trustworthiness of SRs through enhanced creation and intelligent, expedited appraisal mechanisms is crucial to ensure that the best available evidence truly supports clinical practice and decision-making.

Applying methods in vestibular research. Are we missing something in people with central nervous system disorders?

Professional Practices: Where Do We Stand?

Study #5. Knowledge, Barriers, and Future Directions of Vestibular Rehabilitation Practice in Neurorehabilitation: An Italian Survey.

Published as: *Ferri N, Morone G, Piermaria J, Manzari L, Turolla A, De Tanti A, Ciancarelli I, Pillastrini P, Tramontano M. Knowledge, Barriers, and Future Directions of Vestibular Rehabilitation Practice in Neurorehabilitation: An Italian Survey. Healthcare (Basel). 2024;13(1):22.*[71]

Introduction

Vestibular Physical Therapy (VPT), also known as Vestibular Rehabilitation (VR), is an evidence-based approach that involves specific exercises to promote central compensation for disorders affecting gaze and balance.[72] While initially developed for peripheral vestibular conditions,[73,74] its application has expanded, with growing evidence supporting its effectiveness for central nervous system (CNS) disorders.[75]

Research demonstrates the benefits of VPT across various neurological conditions, including reducing dizziness and improving stability in multiple sclerosis,[76] enhancing gait and reducing fall risk in Parkinson's disease,[77] and addressing persistent dizziness and motor coordination issues after traumatic brain injury.[78–80] In children with cerebral palsy, VPT improves motor function and sensory integration,[81] and in stroke survivors, it enhances gait symmetry and postural control.[82]

Given that neurological disorders are a leading cause of global disability and rehabilitation demand,[53,83] VPT's role is increasingly vital. However, despite the evidence, significant challenges impede its implementation. These may include a lack of clinician training,[84] limited medical referrals,[85] significant discrepancies in clinical practice,[86,87] and the underutilization of the vestibular approach.[88,89]

This study aims to map the knowledge, application, and perceived barriers to VPT among Italian neurorehabilitation professionals, highlighting current practice and future directions.

Methods

This cross-sectional survey, approved by the Bioethics Committee at the University of Bologna, targeted Italian healthcare professionals in neurorehabilitation (e.g., physicians, physiotherapists). The published protocol outlines a 29-item questionnaire covering socio-demographics, professional information, clinical practice, and future perspectives.

The survey was developed through an iterative expert drafting process and piloted by six representatives of the target audience to optimize clarity and validity. It was structured using Qualtrics XM and adhered to Dillman's methodology[34] to enhance response rates. We distributed the survey via national professional associations and used IP address screening to prevent duplicate responses. Participation was voluntary and anonymous, with no incentives offered.

A target sample size of 275 respondents was estimated. For analysis, continuous variables are summarized as the mean (SD) or median (IQR), depending on their distribution, and categorical variables are summarized as counts and percentages. Thematic analysis was performed for open-ended questions, with results visualized in word clouds. All analyses were conducted using STATA SE 18, and the study was reported in accordance with the CROSS guidelines.[90]

Results

The survey was distributed via 20,000 emails sent through major Italian professional organizations. From the 435 individuals who opened the survey link, 290 complete responses were included in the final analysis after excluding those who denied consent or submitted blank surveys. The respondents, who had a median work experience of 10 years, represented the entire Italian peninsula. The demographic characteristics are reported in Table 2.

Table 2. Sample characteristics

Female (%)		60.63%
Age (median, IQR)		39.5 (19)
Job title (%)		
	Physical therapist	77.35%
	Physician	16.03%
	Occupational therapist	4.53%
	Developmental therapist	2.09%
Professional role (%)		
	Clinician only	74.74%
	Clinician and researcher	23.86%
	Researcher only	1.40%
Employment type (%)		
	Independent	39.45%
	Public	36.68%
	Private	23.88%

Source: Ferri N et al., 2024

While most respondents (72.98%) were aware of the role of vestibular rehabilitation in neurology, only 24.22% reported having complete expertise. Physiatrists were the most frequent prescribers of this therapy (41.84%), and stroke was the most common CNS disorder referred for vestibular rehabilitation (46.39%).

In clinical practice, balance training (52.69%) and visual input exercises (26.35%) were the most frequently used strategies. A significant barrier was the underutilization of instrumental assessments, as 46.33% of respondents did not use them at all. Indeed, the main perceived obstacles to practicing vestibular therapy were the cost of equipment and a lack of personal skills (Figure 6).

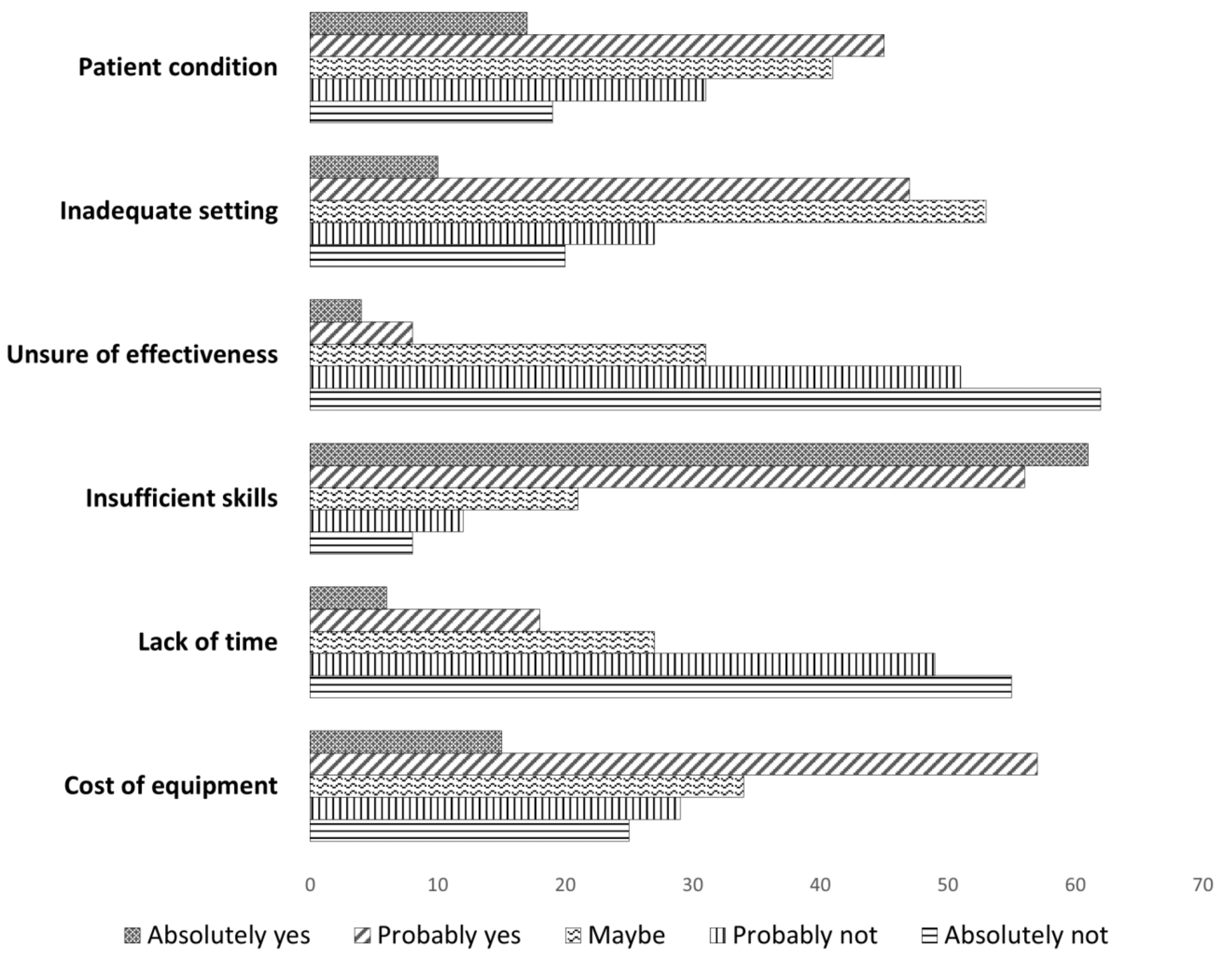

Figure 6. Barriers for using VPT (source: Ferri N et al., 2024)

There was strong consensus on the need for improvement, with 73.01% of respondents completely agreeing that more vestibular rehabilitation education should be provided in introductory academic courses. Additionally, 62.20% completely agreed on the need for specific clinical guidelines for neurological disorders.

Discussion

Neurological disorders significantly impact patients' quality of life, their families, and healthcare costs.[83,91] While scientific evidence for effective treatments grows exponentially,[3] this creates a knowledge translation gap, where the volume of research makes it challenging to select and apply the best evidence in clinical practice.[29]

This challenge is evident in VR. Initially developed in the 1940s,[92,93] VR has evolved considerably, yet our survey reveals a perceived need for updated academic training across multiple rehabilitation professions. This skills gap is further highlighted by the fact that VR, once used for selective peripheral disorders, now applies to a broad spectrum of neurological conditions.[94]

However, our findings highlighted key barriers to the implementation of VR. Firstly, not all patients are deemed suitable candidates; therefore, future guidelines must clearly define functional and cognitive inclusion criteria, tailoring VR exercises to specific pathologies (e.g., Parkinson's disease, stroke). Second, the inadequate setting may not be an actual barrier, as VR is a low-cost approach that can be delivered in standard rehabilitation settings. Moreover, the perception of it requiring a special setting reflects a knowledge-to-practice gap. Third, insufficient skills and equipment costs are barriers that could be addressed by enhancing VR education in professional and academic curricula. Establishing specialized "Hub and Spoke" centers would also provide access to complex assessments for complicated cases, which may eventually require more expensive equipment. It is important to note that practitioners did not report uncertainty about VR's effectiveness or a lack of time as significant barriers. This suggests a belief in its value as an unmet clinical need, reinforcing the educational priority.

However, we must acknowledge some limitations, such as a potentially low response rate, which may indicate a broader lack of expertise in the community, as well as regional variations that might reflect Italy's decentralized healthcare system.

Conclusion

This investigation delineates the current state of vestibular rehabilitation practice within Italy's neurorehabilitation sector, pinpointing substantial knowledge deficits and operational barriers. A key finding is that clinicians perceive a fundamental shortfall in their training for vestibular physical therapy, despite its established efficacy. Addressing this will require strategic initiatives in education and improving access to specialized resources.

Peripheral Vestibular Involvement in Central Nervous System Diseases

Study #6. Abnormal Vestibulo–Ocular Reflex Function Correlates with Balance and Gait Impairment in People with Multiple Sclerosis.

Published as: *Tramontano M, Casagrande Conti L, Orejel Bustos AS, Ferri N, Lelli T, Nocentini U, Grasso MG, Turolla A, Pillastrini P, Manzari L. Abnormal Vestibulo-Ocular Reflex Function Correlates with Balance and Gait Impairment in People with Multiple Sclerosis. Audiol Res. 2024;14(5):799-808.*[95]

Introduction

Multiple Sclerosis (MS) is a chronic, inflammatory, and degenerative disease of the CNS, affecting about 2.8 million people globally.[96,97] Symptoms in people with MS (PwMS) vary with lesion location, but fatigue and walking difficulties are common, significantly impacting daily life and leading to disability. Balance impairments, present in nearly 86% of individuals, are a primary mobility limitation and may be partly linked to dysfunction of the vestibular system.[98]

Rehabilitation is crucial for improving function and activity in PwMS, with moderate-quality evidence supporting its long-term benefits.[99] Vestibular-specific training has been shown to be as effective as other exercises and superior to no intervention in improving balance and dizziness.[76] The role of the vestibular system is further highlighted by its potential correlation with fatigue and non-motor functions.[100,101] A key limitation in existing studies showing positive effects of vestibular rehabilitation, however, is the lack of instrumental assessment of vestibular reflexes before and after training.[102,103]

The video Head Impulse Test (vHIT) is a modern tool that objectively measures the angular vestibulo-ocular reflex (aVOR) gain through the Head Impulse Paradigm (HIMP).[104] Its recent variant, the Suppressed Head Impulse (SHIMP) paradigm, provides a complementary assessment that can give information about possible gain recovery.[105] Using this technological tool, previous studies have found significantly lower aVOR gains in PwMS with brainstem involvement, suggesting its utility in detecting related lesions.[106,107] Worse aVOR function has also been associated with greater MS-related disability.[108]

To our knowledge, no study has yet investigated the correlation between aVOR alterations and balance disorders in PwMS, nor compared the HIMP and SHIMP paradigms for this purpose. We hypothesize that altered semicircular canal function correlates with balance and gait impairment. Therefore, the primary aim of this study is to assess aVOR gains in all six semicircular canals using both HIMP and SHIMP. A secondary objective is to correlate these aVOR gains with balance and gait scale scores.

Methods

This cross-sectional study was conducted at Santa Lucia Foundation Hospital (Rome, Italy) with approval from the Local Independent Ethics Committee (Prot. CE/2022_011). The study adhered to the ethical standards outlined in the Declaration of Helsinki and the STROBE[48] reporting guidelines.

Adults diagnosed with relapsing-remitting or secondary-progressive MS (McDonald criteria)[109] were consecutively recruited between March 2022 and July 2023. Eligibility required an Expanded Disability Status Scale (EDSS)[110] score of 1–6 and the ability to walk 50 meters independently. The exclusion criteria were represented by other significant neurological/psychiatric disorders, by a clinical relapse or steroid use in the previous 3 months, recent lower extremity fractures, or a history of vestibular disorders. All participants provided written informed consent before entering the study.

Clinical and demographic data were collected at the time of enrollment. A researcher, blinded to the other assessments, evaluated balance and gait using the Berg Balance Scale[111] (BBS; score 0-56, higher=better) and the Mini-BESTest[112] (MBT; score 0-28, higher=better). The "Walking" sub-score of the MBT was also analyzed.

The vHIT was used to assess the aVOR gain for all six semicircular canals, following both the HIMP and SHIMP. The protocol followed standardized procedures to ensure reliability,[113] with two trained physiotherapists performing all assessments under expert supervision.

The following tests were performed. For the HIMP assessment, participants fixated on a stationary target during brief, passive, unpredictable head turns. VOR gain was calculated as eye velocity/head velocity, with functional gain ranges defined as 0.76–1.29 for vertical canals and 0.80–1.29 for horizontal canals. For the SHIMP assessment, participants tracked a laser dot moving with their head; gains outside the predefined functional ranges (0.66–1.29 for SHIMP) were classified as hypo- or hyper-gain.

Data were analyzed using SPSS software (IBM Corp. Released 2021. Version 28.0. Armonk, NY, USA: IBM Corp). Continuous data are presented as means ± standard deviations, and categorical data are presented as counts (percentages). Spearman's correlation tests were used to assess relationships between balance scores and VOR gains. A subgroup analysis was performed based on EDSS severity (Class 1: EDSS 1.0–3.5; Class 2: EDSS 4.0–6.0). The dataset was complete with no missing values.

Results

The study sample consisted of 27 PwMS, with a mean age of 47.9 years and a majority being female (n=18). A comprehensive assessment of the aVOR was conducted, evaluating a total of 216 semicircular canals. This included 162 canals tested with the HIMP paradigm and 54 canals with the SHIMP paradigm (Table 2).

Table 3. Sample characteristics

Sex F (%)	66.67
Age (years)	47.93±8.51
Stature (m)	1.70±0.09
Weight (kg)	66.4±15.05
EDSS 1.0-3.5 (%)	25.93
EDSS 4.0-6.5 (%)	74.07
Time from diagnosis (years)	16.22±10.02
Education (years)	13.89±2.98
Mini-BESTest	17.30±5.74

Walking MBT	6.00±2.13	
BBS	44.67±8.85	
HIMP VOR gain (mean ±SD)	Left	Right
Anterior	0.78 (±0.22)	0.86 (±0.14)
Horizontal	0.92 (±0.19)	0.98 (±0.24)
Posterior	0.82 (±0.12)	0.76 (±0.17)
SHIMP VOR gain (mean ±SD)	Left	Right
Horizontal	0.78 (±0.21)	0.87 (±0.23)

EDSS: Expanded Disability Status Scale; MBT: MiniBESTest; BBS: Berg Balance Scale; Walking MBT: refers to the Walking section of MBT; ±=standard deviation

A high prevalence of vestibular dysfunction was observed, with 81% of participants exhibiting an abnormal aVOR gain in at least one semicircular canal. A detailed breakdown revealed that 26.5% of the 162 canals assessed with HIMP and 18.5% of the 54 canals assessed with SHIMP showed abnormal gains. These abnormalities were categorized as either hyper-gain (abnormally high) or hypo-gain (abnormally low), with specific clinical thresholds defined for each paradigm.

A statistically significant correlation was identified between the aVOR gain of the left anterior semicircular canal (measured via HIMP) and patient performance on two clinical balance scales: the MBT and the BBS. This relationship was particularly robust, as the correlation with the BBS remained significant even when analyzing a subgroup of patients with milder disability (EDSS Class 2).

A notable and more nuanced finding emerged from the SHIMP testing. In seven instances involving the horizontal canals, the aVOR gain itself was within a functional range. However, these cases were characterized by a conspicuous absence of anti-compensatory saccades, a key expected response in the SHIMP paradigm. This suggests a specific disruption in higher-level vestibular processing or saccadic suppression, even when the basic reflex arc is intact.

Discussion

This study used the vHIT to evaluate the function of all six semicircular canals in PwMS. The results demonstrated that PwMS often have impaired aVOR, and notably, the reduced function of the vertical canals was significantly correlated with poorer balance scores measured by standard and validated scales. From a rehabilitation perspective, these findings are crucial. They suggest that vestibular assessment could guide the development of personalized rehab programs. Specifically, exercises targeting the aVOR reflex could be a viable strategy to improve dynamic balance and gait in PwMS.[102,114–118] Future research should utilize instrumental vestibular measures to objectively track improvements in vestibular function following rehabilitation. We found that the relationship between aVOR function and clinical scores varied depending on the patients' level of disability (EDSS classification), and this aligns with existing evidence linking greater disability to more severe vestibular dysfunction.[108]

The vestibular deficits observed in PwMS, which often affect isolated canals, differ from those typically seen in peripheral vestibular disorders.[119] This pattern suggests that the cause may be central, likely due to demyelinating plaques within the brainstem vestibular pathways or peripheral connections, which disrupts signal processing.[120] Interestingly, we identified both reduced (hypofunction) and increased (hyperfunction) aVOR gains. While hypofunction is more common, hyperfunction may indicate a loss of inhibitory control from the cerebellum, a phenomenon also observed in other CNS disorders.[121–123] The complex integration of the vestibular reflex arc suggests that these impairments likely stem from central lesions in the brainstem, rather than just the inner ear.

These findings address a significant knowledge gap and highlight the need for a standardized, interdisciplinary approach to diagnose and treat vestibular issues in MS. Indeed, tailored vestibular rehabilitation could be key to improving mobility and quality of life.

However, some limitations must be acknowledged. First, the observational design of the study prevents establishing causation for the aVOR deficits. Second, the small sample size, which is predominantly composed of individuals with moderate disabilities, may limit generalizability. Lastly, the absence of additional vestibular tests (e.g., caloric testing) and the lack of comparison with the clinical Head Impulse Test are also limitations. Thus, future studies with larger cohorts and more comprehensive testing are needed to elucidate these complex vestibular dysfunctions further.

Conclusion

PwMS often exhibit impairments in the aVOR across one or more semicircular canals. Specifically, aVOR gain deficits in the vertical canals are associated with balance and gait disorders measured by clinical scales. Both the HIMP and SHIMP testing paradigms appear to be valid methods for evaluating vestibular canal function in PwMS. However, further research is needed to confirm these findings and establish the role of instrumental VOR assessment in the routine clinical management of PwMS.

Study #7. Video head impulse test in subacute and chronic stroke survivors: new perspectives for implementation of assessment in rehabilitation.

Published as: *Tramontano M, Ferri N, Turolla A, Orejel Bustos AS, Casagrande Conti L, Sorge C, Pillastrini P, Manzari L. Video head impulse test in subacute and chronic stroke survivors: new perspectives for implementation of assessment in rehabilitation. Eur Arch Otorhinolaryngol. 2024;281(10):5129-5134.*[124]

Introduction

The vHIT is a key instrumental test for objectively assessing the aVOR by measuring the gain, which is the ratio of eye movement speed to head movement speed, during brief, unpredictable head turns.[125] While its utility is well-established for diagnosing peripheral vestibular disorders, such as vestibular neuritis,[126,127] its role in CNS pathologies is less defined.

Recent research has begun applying vHIT also to conditions such as stroke,[128,129] MS,[107,108] and Parkinson's disease (PD).[130] Findings have been mixed; for instance, some MS studies show no aVOR deficits,[131] while others link reduced aVOR gain to brainstem lesions.[107] In stroke, vHIT has proven helpful for acute differential diagnosis,[132,133] but its relevance in recovery phases is less clear.[134]

Currently, the literature on vHIT in CNS disorders is still fragmented and conflicting. Furthermore, existing studies are primarily confined to diagnostic settings, with limited application in neurorehabilitation. In this study, we hypothesize that vHIT could identify semicircular canal dysfunctions in persons with stroke (PwS), enabling more personalized rehabilitation to improve gaze and postural stability. Therefore, the objective is to evaluate the VOR in PwS using both the HIMP and SHIMP vHIT paradigms.

Methods

This cross-sectional study was conducted at the Santa Lucia Foundation (Rome, Italy) and received ethical approval from the Local Independent Ethics Committee (Prot. CE/2022_011). All procedures adhered to national and institutional ethical standards, as well as the Declaration of Helsinki and the STROBE guidelines[48]. Written consent was obtained from all participants for the publication of their clinical and instrumental data.

PwS were recruited from the neurorehabilitation services between January and September 2023. Inclusion criteria were: first-ever stroke with unilateral hemiparesis, ability to walk with or without a device, and a Functional Ambulation Classification[135] score > 3. Exclusion criteria were: cognitive impairment (Mini-Mental State Examination[136] > 24), severe unilateral spatial neglect, severe aphasia, and significant neurological, orthopedic, or cardiac comorbidities.

Clinical and demographic data were collected at the time of enrollment. Balance and gait were assessed using the BBS[111] and the MBT[112]. The BBS (0–56 points) and MBT (0–28 points) are standardized tools for evaluating balance and fall risk, with higher scores indicating better function.

The VOR function was evaluated using the vHIT. Both the HIMP and the SHIMP paradigms were used to assess semicircular canal function. The aVOR gain was considered normal between 0.8 and 1.29 for HIMP and 0.66 and 1.29 for SHIMP.[137,138] In particular, based on local norms, gains below 0.80 for horizontal canals and below 0.76 for vertical canals were classified as dysfunctional.

Data were analyzed using STATA 17 (StataCorp. 2023, College Station, TX, USA). Continuous data are presented as mean ± standard deviation, and categorical data as frequencies. Spearman's correlation was used to examine the relationships between balance outcomes and aVOR gains, given the non-normal distributions. Then, a subgroup analysis compared participants in subacute (<6 months) and chronic (>6 months) phases post-stroke.

Results

The study included 36 participants with hemiparesis following a stroke. The sample characteristics are reported in Table 4.

Table 4. Sample characteristics

	Subacute (n = 22)	Chronic (n = 14)
Sex F (%)	22.73	42.86
Age (years)	56.13±15.00	53.50±15.65
Time from diagnosis (days)	87.20±49.95	729.21±663.16
Mini-BESTest	19.14±6.85	17.30±7.18
BBS	44.90±13.29	41.07±16.55
Fallers in last year (%)	5.88	20.00

BBS: Berg Balance Scale. Unless stated otherwise, data are reported as mean ± standard deviation.

Overall, the average aVOR gains for the semicircular canals were predominantly outside the normal physiological range (Table 5).

Table 5. Vestibulo-ocular reflex gains (n = 36)

HIMP VOR gain	95% CI
Left Anterior	0.73-0.87
Left Horizontal	0.92-1.02
Left Posterior	0.86-0.96
Right Anterior	0.74-0.91
Right Horizontal	0.85-0.93
Right Posterior	0.74-0.83
SHIMP VOR gain	
Left Horizontal	0.70-0.83
Right Horizontal	0.81-0.95

CI: confidence interval.

On an individual level, a significant number of participants showed impaired VOR function, with the highest rates observed in the left anterior (38.9%) and right posterior (33.3%) canals. Impairments in the horizontal canals were less frequent (right = 2.8%, left = 11.1%) and more often detected using the SHIMP paradigm (right = 5.6%, left = 22.2%).

No significant differences in aVOR gains, balance, or motor function were found between participants in the sub-acute and chronic phases of recovery.

Discussion

This study evaluated the aVOR gain in PwS using the vHIT with both the HIMP and SHIMP paradigms. Assessments were conducted on 216 and 72 semicircular canals with HIMP and SHIMP, respectively, across 36 participants (22 subacute, 14 chronic). Seventy-five percent of PwS showed altered aVOR gain in at least one canal. A key finding was the presence of isolated, sometimes bilateral, aVOR hypofunction, contrasting with the pattern seen in unilateral peripheral vestibular loss. This indicates that stroke leads to vestibular alterations that do not affect the entire system uniformly. Notably, the mean aVOR gains for some canals, particularly the vertical ones, were near the lower limit of the normal range (0.76), confirming a tendency towards hypofunction. No statistical correlation was found between balance and vestibular function across disease phases, suggesting the aVOR impairment is independent of the chronicity of the stroke in this cohort. The prevalence of vertical canal deficits may explain the results from the SHIMP paradigm, which only tests the horizontal canals.

The utility of vHIT in stroke is well-established for differential diagnosis in acute vestibular syndrome.[139,140] However, its application in neurorehabilitation remains limited. Our findings underscore the necessity of including vestibular assessment in the care of PwS to create tailored rehabilitation programs. Standardizing assessments and adopting an interdisciplinary approach could significantly improve care for post-stroke balance disorders. Future studies should use aVOR metrics as outcomes to measure rehabilitation efficacy. Consequently, rehabilitation should evolve beyond a purely neuromusculoskeletal focus. It must integrate training that addresses continuous sensory integration, combining vestibular, proprioceptive, and visual inputs, which underlie dynamic postural stability and gait.[141–143]

Finally, this study has some limitations to be acknowledged. The observational design prevents establishing a causal link between neurological lesions and abnormal aVOR gain. Moreover, it is difficult to determine if aVOR dysfunctions stem from the stroke itself, reduced physical activity, or other confounding factors. Lastly, the lack of vestibular-evoked myogenic potential testing is another limitation that could have provided more pathophysiological insights.

Conclusion

An altered aVOR gain can be present in stroke patients across subacute and chronic disease phases. Assessing vestibular function instrumentally could assist the rehabilitation team in personalizing balance training according to patient-specific deficits.

Study #8. Vestibulo-Oculomotor Reflex Dysfunction in Children with Cerebral Palsy Correlates with Gross Motor Function Classification System.

Published as: *Casagrande Conti L, Ferri N, Manzari L, Lelli T, Mangeruga M, Dal Piaz M, Manzotti A, Verrecchia L, Tramontano M. Vestibulo-Oculomotor Reflex Dysfunction in Children with Cerebral Palsy Correlates with Gross Motor Function Classification System. Audiol Res. 2025;15(2):21.*[144]

Introduction

Cerebral palsy (CP) is a leading cause of childhood physical disability, affecting 2 to 3 per 1000 live births.[145] It is a non-progressive disorder caused by damage to the developing brain, resulting in movement and postural disturbances.[146] While motor impairments like spasticity are well-known,[147,148] the impact of sensory dysfunctions, particularly vestibular impairment, is less understood.

The vestibular system, crucial for balance, gaze stability, and spatial orientation, provides essential sensory input for navigating the environment.[149] Its dysfunction can lead to poor balance, unstable gaze, and a higher risk of falls.[150–152] For children with CP, who already have motor challenges, vestibular issues can further limit their daily activities and social participation.[148] Despite its significance, vestibular function in CP remains under-investigated, with research traditionally focusing more on motor and visual deficits.[153–157] The vHIT now allows for quantitative assessment of the aVOR in children,[158] where vestibular impairments likely affect crucial multisensory integration for motor planning, impacting activities like walking and playing sports.[148,159]

While vHIT has been used in adults with neurological conditions,[95,124,160] its application in children with CP is novel.[158] We hypothesize that children with perinatal brain injury may have impaired aVOR pathways, even with a developed vestibular system. Thus, this study aims to determine the prevalence of aVOR dysfunction in children with CP, explore its correlation with motor function, and evaluate the feasibility of using vHIT in a neurorehabilitation setting.

Methods

This cross-sectional study, approved by the Local Independent Ethics Committee (Prot. CE/2022_011), was conducted at a pediatric neurorehabilitation center (Santa Lucia Foundation, Rome, Italy) from September 2023 to September 2024, following STROBE guidelines.[144]

Participants were children aged 3-11 years with cerebral palsy (spastic paresis or ataxia), a Level of Sitting Scale[161] score >4, and the ability to understand the visual fixation task. Children with significant spasticity or bone deformities were excluded. Motor function was classified using the Gross Motor Function Classification System (GMFCS).[162,163]

Vestibular function was assessed using the vHIT to evaluate the aVOR gain. Two trained physiotherapists performed the tests under supervision, using both the HIMP and SHIMP paradigms.[105,125,164,165] Standardized protocols were followed for horizontal and vertical canal testing, with controlled lighting and techniques to maintain child engagement.[113]

Statistical analysis was performed using STATA 18 (StataCorp, 2023, College Station, TX, USA). Data normality was assessed with the Shapiro-Wilk test, and the relationship between aVOR gain and GMFCS level was analyzed using Spearman's rank correlation coefficient.

Results

A total of 20 children were initially recruited. However, seven were excluded from the final analysis: four could not follow instructions, one could not be fitted with the goggles, and two discontinued due to discomfort.

Consequently, thirteen children completed the assessment of the horizontal semicircular canals, but only one completed the vertical plane evaluation. The SHIMP paradigm was successfully performed on only three children, as the remaining ten found the procedure too difficult (Table 6).

Table 6. Sample characteristics (n = 13)

Female (n, %)		8 (61.5)
Age (years ± SD)		7.8±2.7
Clinical characteristics (n, %)		
	Ataxia	3 (23.1)
	Diplegia	4 (30.8)
	Hemiplegia	2 (15.4)
	Other	4 (30.8)
GMFCS (n, %)		
	Level I	6 (46.2)
	Level II	4 (30.8)
	Level III	3 (23.1)
	Level IV	0 (0.0)
	Level V	0 (0.0)

GMFCS: Gross Motor Function Classification System.

Of the thirteen children who completed the horizontal canal assessment, six (46.2%) exhibited dysfunction in at least one canal. Overall, eight out of twenty-six (30.8%) assessed horizontal canals were dysfunctional. While the overall sample mean showed a functional aVOR gain (Table 7), subgroup analysis revealed that children with ataxia had significantly reduced gains, with half of the horizontal canals being dysfunctional (Table 8).

Table 7. Vestibulo-ocular reflex gains (n = 13)

HIMP aVOR gain		Total canals (n)	Mean ± SD	95% CIs
	Left Horizontal	13	0.80±0.17	[0.69, 0.90]
	Right Horizontal	13	0.90±0.17	[0.79, 1.01]
	Left Anterior	1	0.86	NA
	Right Anterior	1	1.22	NA
	Left Posterior	1	0.92	NA
	Right Posterior	1	0.60	NA
SHIMP aVOR gain		Total canals (n)	Mean ± SD	95% CIs
	Left Horizontal	3	0.73±0.19	NA
	Right Horizontal	3	0.85±0.07	NA

NA: not applicable.

Table 8. Subgroup analysis (GMFCS levels)

HIMP aVOR Gain		Total Canals (n)	Mean ± SD
GMFCS 1	Left Horizontal	6	0.91 ± 0.09
	Right Horizontal	6	1.02 ± 0.09
GMFCS 2	Left Horizontal	4	0.75 ± 0.18
	Right Horizontal	4	0.86 ± 0.17
GMFCS 3	Left Horizontal	3	0.65 ± 0.18
	Right Horizontal	3	0.73 ± 0.19

GMFCS: Gross Motor Function Classification System.

Furthermore, correlation analysis identified a strong negative association between aVOR gain and the GMFCS level (ρ = -0.73, p-value = 0.0092), indicating that vestibular function declined as motor impairment increased (Figure 7).

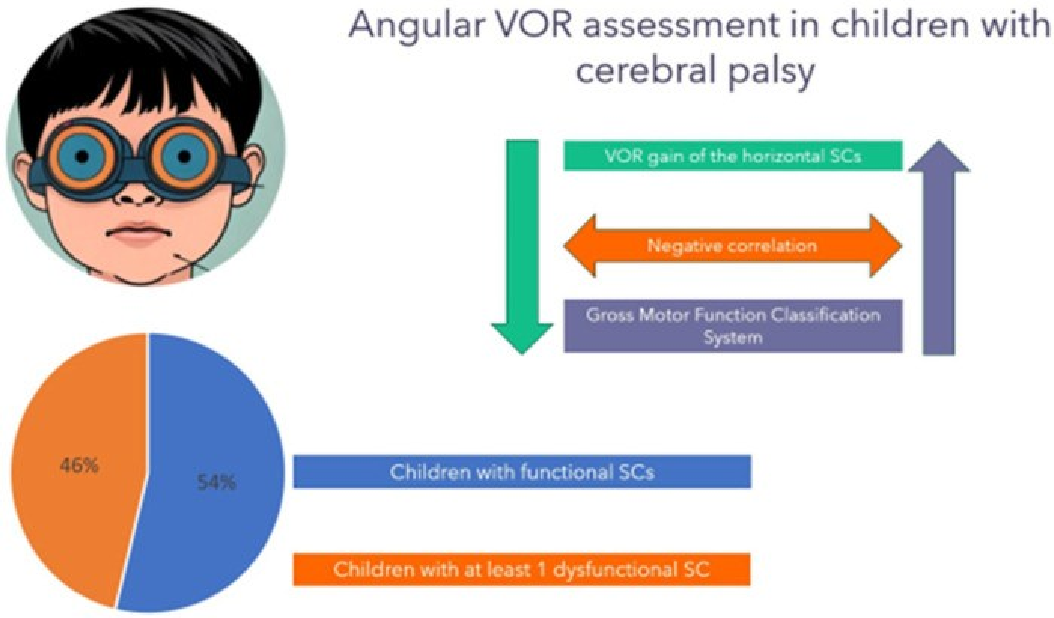

Figure 7. Summary findings (source: Casagrande Conti L et al., 2025)

Discussion

This study found that children with CP, especially those with greater motor impairment, often exhibit vestibular dysfunction in the horizontal semicircular canals. The observed reduction in aVOR gain underscores the value of vestibular assessment in rehabilitation, aligning with existing evidence that links vestibular deficits to the balance and postural control challenges in this population.[166]

The correlation between aVOR gain and motor function (GMFCS levels) provides crucial insight into the multisensory deficits in CP. As motor impairment increases, the vestibular system's ability to compensate for head movements diminishes, leading to problems with dynamic stability. Our findings are consistent with other studies that identify impaired vestibular function and sensory integration as key contributors to instability and gait difficulties in these children.[159,166,167] Notably, children with ataxia showed the most pronounced aVOR gain abnormalities, with half of the tested canals being impaired, suggesting significant challenges with visual stability and motor coordination.

Using the vHIT proved effective for identifying these horizontal canal deficits, confirming its clinical utility in a pediatric neurorehabilitation setting.[168,169] However, assessing the vertical canals was largely unfeasible due to difficulties in testing children with cognitive-motor disabilities, highlighting a need for child-friendly equipment and methods.

Clinically, these results argue for integrating vestibular assessments like vHIT into routine evaluations. Since vestibular rehabilitation can improve balance in other pediatric groups,[81,170–172] targeted therapies could likely benefit children with CP, improving daily function and motor outcomes.

This study has some limitations, including its observational design, small sample size, and inability to fully assess vertical canal function. Future research should involve larger, longitudinal studies that use comprehensive vestibular and sensory assessments to develop more effective, personalized rehabilitation strategies.

Conclusion

This study adds to the growing evidence of vestibular dysfunction in individuals with neurological disorders, including children with cerebral palsy. Our results confirm that instrumental vestibular testing is feasible in children with cognitive-motor disabilities, though currently limited to the horizontal canals using the HIMP paradigm. Future research should explore the function of the vertical canals and central vestibular pathways. Incorporating these assessments into standard clinical practice could enhance the understanding and management of balance and motor deficits in this population.

Study #9. Video Head Impulse Test in Survivors From Severe Traumatic Brain Injury: New Perspectives for Implementation of Assessment in Rehabilitation.

Published as: *Ferri N, Whitney SL, Verrecchia L, Casagrande Conti L, Turolla A, Lelli T, Formisano R, Buzzi MG, Pillastrini P, Manzari L, Tramontano M. Video Head Impulse Test in Survivors From Severe Traumatic Brain Injury: New Perspectives for Implementation of Assessment in Rehabilitation. J Head Trauma Rehabil. 2025;40(5):337-345.*[173]

Introduction

Traumatic brain injury (TBI) is a significant global health issue, with a prevalence of 599 per 100,000 people, primarily caused by falls and road accidents. It imposes high costs on healthcare systems and disrupts patients' families and work lives.[174,175] Managing TBI is challenging due to the cognitive, neuromotor, and behavioral changes that follow, underscoring the need for a person-centered, interdisciplinary approach.[176,177]

Imbalance is a prevalent symptom, affecting up to 80% of individuals with TBI[178] and is a major risk factor for falls.[179] This can stem from various vestibular issues, including central or peripheral dysfunctions,[180] benign paroxysmal positional vertigo,[181] or impaired sensory-motor integration.[182,183] Despite evidence of vestibular system involvement, specific vestibular assessments are not routine in trauma wards, complicating diagnosis and management.[184]

The vHIT is an efficient tool for assessing the aVOR. It uses two procedures, the HIMP and the SHIMP paradigms, to measure aVOR gain and its suppression, respectively.[105,164]

While vHIT has been used in various neurological conditions,[107,108,128–130] including TBI,[160] its application in survivors of severe TBI (sTBI) to assess semicircular canal function remains unexplored. Evaluating the aVOR slow phase in sTBI survivors could uncover vestibular dysfunction and help personalize neurorehabilitation. This study, therefore, aims to describe the aVOR function in people with sTBI and explore its relationship with balance measures.

Methods

This cross-sectional study, approved by the Local Independent Ethics Committee (Prot. CE/2022_011), was conducted at the Santa Lucia Foundation (Italy, Rome), a neurorehabilitation research hospital, following STROBE guidelines.[48]

Adults (18-65 years) with sTBI, defined by a Glasgow Coma Scale[185] score ≤ 8, were recruited from inpatient neurorehabilitation services. Eligible participants had a Level of Cognitive Functioning[186] ≥ 7, could understand commands, walk independently (Functional Ambulation Classification[135] >3), and were undergoing balance rehabilitation. All participants provided written informed consent.

Balance and gait were assessed using the BBS[187,188] and the MBT[112,189]. Vestibular function was evaluated with the vHIT using both the HIMP and SHIMP paradigms to measure the aVOR gain across all six semicircular canals. The assessment followed a strict protocol[113] administered by trained physiotherapists.

During HIMP, patients fixated on a stationary target while the examiner delivered brief, unpredictable head rotations. For SHIMP, patients tracked a head-fixed laser dot. The aVOR gains were classified as functional or dysfunctional (hypo-gain/hyper-gain) based on established normative ranges.[125,137,164]

Analyses were performed using STATA 18 (StataCorp, 2023, College Station, TX, USA). Data are presented as means/medians with standard deviations/interquartile ranges or as frequencies. The Shapiro-Wilk test assessed normality, while Spearman's rank correlation was used to examine relationships between aVOR gains and clinical balance scores.

Results

The study enrolled 22 patients with sTBI. Two participants were unable to complete the full vHIT protocol. Thus, a total of 124 semicircular canals (SCCs) were assessed using the HIMP paradigm, and 40 using the SHIMP paradigm (Table 9).

Table 9. Sample characteristics

Variables	Mean (SD)
Age (years)	48[a] (27.00)
Body mass (kg)	72.95 (10.72)
Height (m)	1.76[a] (0.08)
Time from diagnosis (months)	8.18[a] (7.38)
Education (years)	14.45 (2.39)
Fallers (n)	4[b] (20)
MBT	24[a] (7.5)
BBS	52.5[a] (7.5)

Data are reported as mean and standard deviation, or median and IQR for continuous variables (according to distribution properties), and frequency and percentage for categorical variables. [a]median and interquartile range, [b]counts and percentage. MBT: Mini-BESTest Scale, BBS: Berg Balance Scale.

VOR gain was abnormal in at least one canal for the vast majority of patients (19 out of 22, or 90.48%). Overall, 30% of the assessed canals showed abnormal gain. The right posterior canal was most frequently affected, with 71% of participants showing an abnormal gain, followed by the left anterior canal (33%). The SHIMP assessment identified 5 horizontal canals with functional gain and no anti-compensatory saccades. (Table 10).

Table 10. Vestibulo-ocular reflex gains

HIMP	n	Mean	Standard Deviation	Confidence Intervals	
Right Horizontal	21	0.97	0.21	0.87	1.06
Left Horizontal	21	0.87	0.14	0.80	0.94
Left Anterior	21	0.80	0.15	0.73	0.87
Right Posterior	21	0.70*	0.17	0.62	0.78
Right Anterior	20	0.91	0.25	0.79	1.03
Left Posterior	20	0.86	0.15	0.79	0.93
SHIMP					
Right Horizontal	20	0.80	0.26	0.68	0.92
Left Horizontal	20	0.73	0.22	0.63	0.84

**aVOR gain outside the functional range*

No correlation was found between balance scale scores and aVOR gain. Notably, only two patients had completely normal gains across all canals, and almost all (94%) of the identified abnormalities were reductions in gain (hypo-gain).

Discussion

This study used the vHIT to assess SCC function in survivors of sTBI. Results indicated that over 90% of participants had aVOR dysfunction in at least one canal, suggesting vestibular symptoms are a common, though variable, part of the complex sTBI clinical picture. A notably high prevalence of right posterior canal dysfunction was observed, aligning with previous findings in stroke survivors,[124] which may point to a need for comprehensive central and peripheral vestibular assessment in rehabilitation.

The dysfunction often presented in a seemingly random pattern across canals, underscoring the complexity of vestibular damage in sTBI and challenging straightforward diagnostic approaches. Notably, aVOR gains did not correlate with balance and gait scores using the BBS and MBT, indicating these tools may measure different constructs. The high balance scores observed may be partly due to the median post-injury time of 8 months, which allows for partial neuromotor recovery through conventional rehabilitation that may not specifically address vestibular deficits. This disconnect could also be influenced by vestibular agnosia, where patients are unaware of their vestibular impairment.[190]

Analysis suggested a generalized aVOR deficit, which can cause gaze instability during head movements, leading to dizziness and difficulty in dynamic environments. Moreover, whenever aVOR gain is normal, but with a complete absence of anti-compensatory saccades in the SHIMP test, it strongly suggests a CNS lesion affecting the aVOR suppression pathway. Thus, given the brainstem's role in processing vestibular signals, the observed impairments are likely due to central, not just peripheral, damage from the trauma.[191]

This research addresses a significant gap in the literature,[192] highlighting the need for early identification of vestibular dysfunction in sTBI. Future studies should investigate the underlying mechanisms and develop targeted rehabilitation strategies, such as customized vestibular physical therapy or virtual reality, to improve gaze stability, balance, and quality of life for sTBI survivors.

However, this study has some limitations. The observational design that prevents us to investigate a causal link, the small sample size limiting generalizability, and the use of balance scales that may not be sensitive enough to detect vestibular-specific deficits in this population.

Conclusion

Although individuals with sTBI can show impairments in the aVOR, these deficits were not correlated with motor or balance outcomes. The clinical significance of this finding remains uncertain and warrants further research. Nonetheless, it supports the growing consensus that vestibular assessment and rehabilitation should be considered a valuable component of personalized management for sTBI.

Study #10. Assessment of Vestibulo-Ocular Reflex Function in People with Parkinson's Disease: a Cross-Sectional Study in a Rehabilitation Setting Using the Video Head Impulse Test

Published as: *Ferri N, Casagrande Conti L, Manzari L, Piatti D, De Angelis D, Nocentini U, Tramontano M. Assessment of vestibulo-ocular reflex function in people with Parkinson's disease: a cross-sectional study in a rehabilitation setting using the video head impulse test. BMJ Open. 2025;15(8):e099765*.[193]

Introduction

PD is a progressive neurodegenerative disorder defined by motor symptoms such as bradykinesia, rigidity, tremor, and postural instability.[194] Recently, growing evidence underscores the role of vestibular dysfunction in worsening functional impairments, particularly postural instability and gait disturbances, which elevate fall risk and reduce quality of life.[130,184,195]

The vestibular system, via the aVOR, is essential for gaze stability and daily function.[196] However, while aVOR impairment is recognized in neurological disorders, its prevalence and impact in PD remain unclear. Dopaminergic deficits in the basal ganglia and vestibular nuclei may contribute to vestibular dysfunction in PD,[197,198] yet studies report inconsistent aVOR abnormalities and their association with functional outcomes like balance and gait.[198] The vHIT allows objective, quantitative aVOR assessment across all six semicircular canals.[199] It measures aVOR gains and detects corrective saccades in case of vestibular canal dysfunction. In literature, research usually reported group-level mean gains, thus potentially missing individual canal-level dysfunctions.[197,198,200] In fact, individual analysis can reveal subtle impairments that could be clinically relevant but overlooked in averaged data.[95,124]

In this context, conflicting findings exist. Some studies show no aVOR gain changes in PD,[200] while others note increased gains, possibly indicating compensatory mechanisms.[197] However, many studies reported and focused only on aVOR hypofunction, ignoring hyperfunction or compensatory increases.[197] Regardless, the ability of vHIT to detect subtle aVOR changes could definitely inform disease management and rehabilitation.[95,124]

This study used vHIT with both HIMP and SHIMP protocols to evaluate semicircular canal function in PD. Unlike previous work, we emphasized individualized aVOR gain analysis to determine the prevalence of canal-specific dysfunction at the patient level.

Methods

This cross-sectional study was conducted at the Santa Lucia Foundation Institute for Research and Healthcare (Rome, Italy) after local ethics committee approval (Prot. CE/2022_011). All procedures followed national and international ethical guidelines, including the Declaration of Helsinki, and it was reported following the STROBE reporting guidelines.[48] All participants provided written consent for the use of their clinical data.

Consecutive outpatients with PD were recruited from March to November 2024. Inclusion criteria were: no cognitive impairment (MMSE[136] > 25), Hoehn and Yahr[201] stage 2-3, and independent ambulation (Functional Ambulation Category[202] > 3). Exclusion criteria included significant visual or vestibular impairments, limited neck mobility, or other significant comorbidities, as determined by neurological screening and medical records.

The aVOR function was evaluated using vHIT (ICS Impulse, Otometrics/Natus, Denmark) with both HIMP and SHIMP paradigms. Functional aVOR gain was defined as 0.8-1.29 for HIMP and 0.66-1.29 for SHIMP.[126,137] Gains outside these ranges indicated dysfunction.

Data were analyzed using STATA 18.5 (StataCorp. 2023, College Station, TX, USA). Continuous data are presented as mean ± SD and categorical data as frequencies. A bootstrap analysis with 2000 replications calculated confidence intervals when there was an asymmetric data distribution. A post-hoc analysis determined prevalence confidence intervals using the Wilson method, as this exploratory study used a convenience sample due to lacking prior prevalence data.

Results

This study included 35 eligible participants with PD (mean age 69.9±8.4; 11 female). Their demographics are shown in Table 11.

Table 11. Sample characteristics

Variables	Mean
Age (years)	69.9 ± 8.4
Time from diagnosis (years)	6.5 ± 4.7
Education (years)	14.6 ± 3.6
m-UPDRS	27.7 ± 6.8

Data are reported as mean ± standard deviation.

The HIMP assessment was completed for 34 participants; one was excluded due to persistent blinking, which also prevented anterior canal assessment in another case. Finally, VOR gain for the SHIMP paradigm was successfully collected from all 35 participants (Table 12).

Table 12. Vestibulo-ocular reflex gains

Vestibular canal	n	VOR gain	95% Confidence Intervals	
Right Lateral	34	0.99 ± 0.20	0.92	1.06
Left Lateral	34	0.94 ± 0.16	0.88	0.99
Right Anterior	33	0.87 (0.25)	0.76	0.94
Left Posterior	34	0.87 (0.19)	0.77	0.92
Left Anterior	33	0.83 (0.15)	0.72	0.85
Right Posterior	34	0.82 (0.13)	0.71	0.84
Right Lateral (SHIMP)	35	0.83 ± 0.19	0.77	0.90
Left Lateral (SHIMP)	35	0.85 ± 0.20	0.78	0.92

Data are reported as mean ± SD, or median (IQR) according to their distribution.

The vHIT revealed abnormal aVOR gains across all canals and paradigms (prevalence: 14–38%), with the right posterior canal being most frequently affected (Figure 8).

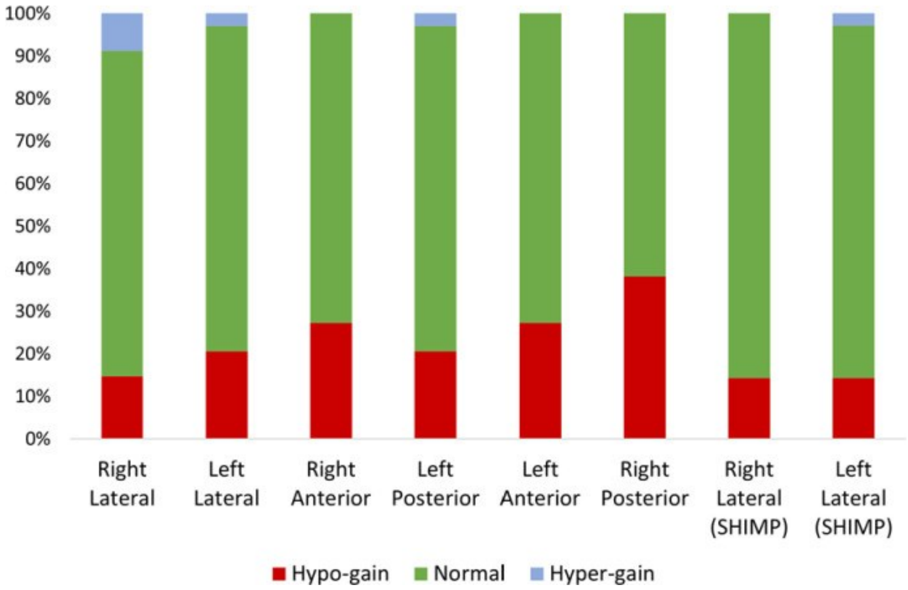

Figure 8. Prevalence of vestibulo-ocular reflex dysfunctions (Source: Ferri N et al., 2025).

Individually, 60% (95% CI: 43.6 to 74.5%) of participants had at least one canal with hypo-gain. This prevalence rose to 69% (95% CI: 52.5 to 81.8%) when including hyper-gain dysfunctions. When incorporating SHIMP results, 83% (95% CI: 67.5 to 92.0%) of participants exhibited an abnormal aVOR gain in at least one canal.

Discussion

This study investigated semicircular canal function in people with PD using the vHIT test with both HIMP and SHIMP paradigms. We demonstrated the feasibility of vestibular assessment in a neurorehabilitation setting and found that 69% of participants had at least one abnormal result, including both hypo- and hyper-gain dysfunctions.

These findings challenge a recent study reporting unchanged aVOR gain in PD.[200] Indeed, our individual-level analysis revealed isolated canal dysfunctions, consistent with our group's previous work in people with stroke,[124] and multiple sclerosis.[95] While the mean aVOR gain for the entire sample appeared normal, individual prevalence analysis uncovered underlying clinical conditions. This highlights the need for researchers to report the prevalence of dysfunction based on normative cut-offs, not just group means.

Our results align with Berkiten et al.[130] showing vestibular involvement in early-stage PD and its link to motor and non-motor symptoms. Moreover, vestibular rehabilitation may improve postural balance in PD,[77] completing the range of neurological conditions managed in rehabilitation.[53] Thus, systematic vestibular assessment should become routine in neurorehabilitation, especially given the increasing commercial availability of necessary equipment.

Our work further extends the literature on vestibular alterations in PD. While some studies reported preserved aVOR gain using HIMP,[200] subsequent research using SHIMP revealed impaired aVOR suppression,[203] and other works have shown abnormal otolith function and vestibulo-visual interactions.[200,204] This reinforces the need for multimodal assessment, including both HIMP/SHIMP paradigms, to fully characterize vestibular phenotype in PD.

However, some limitations have to be reported: the cross-sectional design preventing causal inferences, the focus solely on semicircular canal function via vHIT, and the inclusion of physically self-sufficient participants with mild to moderate disability, suggesting that future studies should assess a broader population.

Conclusion

An aVOR gain impairment is common in people with PD. Incorporating vestibular function assessment into routine balance evaluations could enhance rehabilitation care for this population. Further research is needed to confirm these results and to determine if aVOR-profile specific vestibular rehabilitation can yield functional benefits.

Rethinking Established Diagnostic Criteria: The Need to Consider Multiple Contributing Factors

Study #11. Revisiting Diagnostic Criteria for Bilateral Vestibulopathy: A New Comprehensive Instrumental Model.

Published as: *Manzari L, Ferri N, Tramontano M. Revisiting Diagnostic Criteria for Bilateral Vestibulopathy: A New Comprehensive Instrumental Model. Audiol Res. 2024;14(6):991-999.*[205]

Introduction

Bilateral vestibulopathy (BVP), or bilateral vestibular hypofunction (BVH), is a chronic condition involving significant or complete loss of vestibular function on both sides.[206] This impairment disrupts vestibular signals to the CNS, leading to symptoms such as persistent dizziness, imbalance, oscillopsia (i.e., visual instability during head movement), and often a subjective feeling of mental cloudiness or "brain fog".[207–211]

Prior to updated diagnostic criteria, BVP prevalence in U.S. adults was estimated at around 28 per 100,000, with incidence between 4% and 7%.[212] The condition affects a wide age range, though acquired forms are typically diagnosed between ages 50 and 60.[213,214] Etiologies are diverse, including toxic, infectious, traumatic, otologic, and congenital causes, making early diagnosis difficult.[206] Nonetheless, BVP significantly impairs daily function and quality of life, with common symptoms like imbalance (worsening in darkness or on uneven ground) and oscillopsia due to aVOR dysfunction.[215]

Currently diagnosis relies on clinical evaluation and vestibular testing. Indeed, according to the 2017 Bárány Society criteria, BVP is confirmed by chronic symptoms and bilaterally reduced or absent aVOR function, measurable via vHIT, caloric testing, or rotatory chair.[206] However, these criteria focus only on horizontal semicircular canal function and do not incorporate tests for vertical canals or otolithic organs (e.g., vestibular-evoked myogenic potentials, VEMPs).

Thus, we hypothesize that current diagnostic criteria may miss specific BVP presentations, such as patients with normal aVOR gain but abnormal VEMP results or vertical canal dysfunction.[216,217] This study aimed to assess both aVOR and otolithic function using a comprehensive battery of instrumental tests in individuals with chronic vestibular symptoms and significant dizziness-related handicap.

Methods

This diagnostic accuracy study analyzed the medical records of patients diagnosed with probable BVP at the MSA ENT Academy Center, a tertiary vestibular referral center, between July 2022 and December 2023. The study, approved by the Institutional Review Board, followed the STARD guidelines.[218]

Consecutive patients from the center's records were included if they reported symptoms like gait unsteadiness, movement-induced blurred vision, and had a Dizziness Handicap Inventory (DHI)[219] score >54, with an absence of symptoms while sitting or lying still.

The index test was a comprehensive vestibular assessment, including the vHIT to measure the aVOR gain across all six semicircular canals, and cervical/ocular VEMPs to evaluate otolith function. The reference standard was the Barany Society diagnostic criteria for BVP[206]:

- A. Chronic vestibular syndrome with the following symptoms:
    1. Unsteadiness when walking or standing plus at least one of 2 or 3;
    2. Movement-induced blurred vision or oscillopsia during walking or quick head/body movements and/or;
    3. Worsening of unsteadiness in darkness and/or on uneven ground.
- B. No symptoms while sitting or lying down under static conditions.
- C. Bilaterally reduced or absent aVOR function documented by:
    1. Bilaterally pathological horizontal aVOR gain <0.6, measured by the vHIT or scleral-coil technique and/or;
    2. Reduced caloric response (sum of bithermal max. peak SPV on each side <6°/s) and/or;

3. Reduced horizontal aVOR gain <0.1 upon sinusoidal stimulation on a rotatory chair (0.1 Hz, Vmax = 50°/s) and a phase lead >68 degrees (time constant <5 s).

- D. Not better accounted for by another disease.

For the vHIT, an expert clinician performed the test, applying brief, unpredictable head impulses. The aVOR gain (ratio of eye to head velocity) was the primary measure, with a gain of ≥0.60 considered functional.

For the VEMPs, ocular and cervical VEMPs were used to assess utricular and saccular function, respectively. Asymmetry ratios were calculated from the response amplitudes, and bilateral absence of responses also indicated BVP.

The diagnostic accuracy (sensitivity, specificity, ROC area) of the comprehensive index test model (vHIT + VEMPs) was compared to the Barany criteria model (horizontal vHIT only). A reverse analysis was also conducted, treating the comprehensive model as the new reference standard. Incomplete records were excluded from the final analysis.

Results

After screening 1,169 medical records, 78 patients (41 male, mean age 61.4 ± 12.99 years) meeting the criteria for BVP were included. The median (IQR) DHI score was 67 (6).

The new diagnostic criteria for BVP were compared against the established Barany criteria (Table 13).

Table 13. BPV diagnosis, according to different criteria

	BPV (n, %)
Barany criteria	36 (46)
Comprehensive model	78 (100)
6-canals vHIT	59 (75)
oVEMPs dysfunction	52 (66)
cVEMPs dysfunction	41 (52)

BPV: bilateral vestibulopathy. The comprehensive model includes cVEMPs, oVEMPs, and 6-canals vHIT.

Furthermore, the diagnostic accuracies of the three proposed new criteria are shown in Table 14. A comprehensive model combining the three tests could not be created due to structural zeros in the data.

Table 14. Multimodel diagnostic accuracy

Index test	Sensitivity	Specificity	LR+	LR-	ROC area	95% CI
6-canals vHIT	100%	45.24%	1.82	0.00	0.72	[0.65, 0.80]
oVEMPs	63.89%	30.95%	0.92	1.16	0.47	[0.36, 0.58]
cVEMPs	52.78%	47.62%	1.00	0.99	0.50	[0.38, 0.61]

When the combined new criteria were treated as the reference standard, the Barany criteria demonstrated low sensitivity (46%) and a high false-negative rate, indicating a poor ability to rule out BVP.

The DHI scores were significantly different between patients classified as true positives and false negatives by the Barany criteria (Wilcoxon test, p=0.0117, Figure 9).

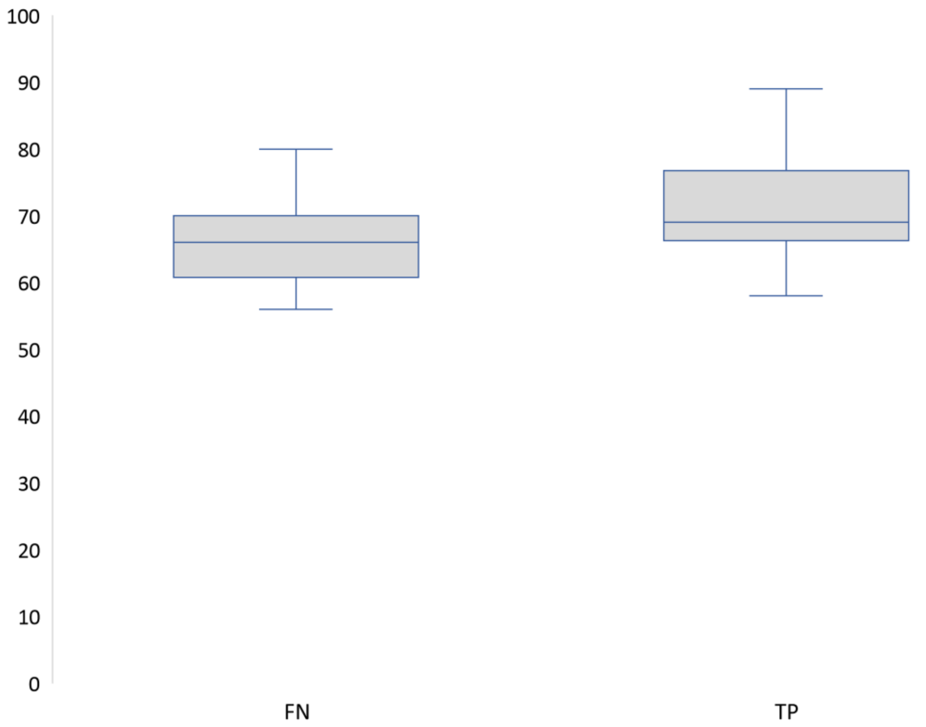

Figure 9. DHI median scores between false negatives and true positives (Barany Criteria). FN: false negatives, TP: true positives (source: Manzari L et al., 2024)

Median DHI scores across all models ranged from 66 to 69. A significant difference in DHI scores between test-positive and test-negative patients was found only for the Barany criteria and the six-canal vHIT model (Figure 10).

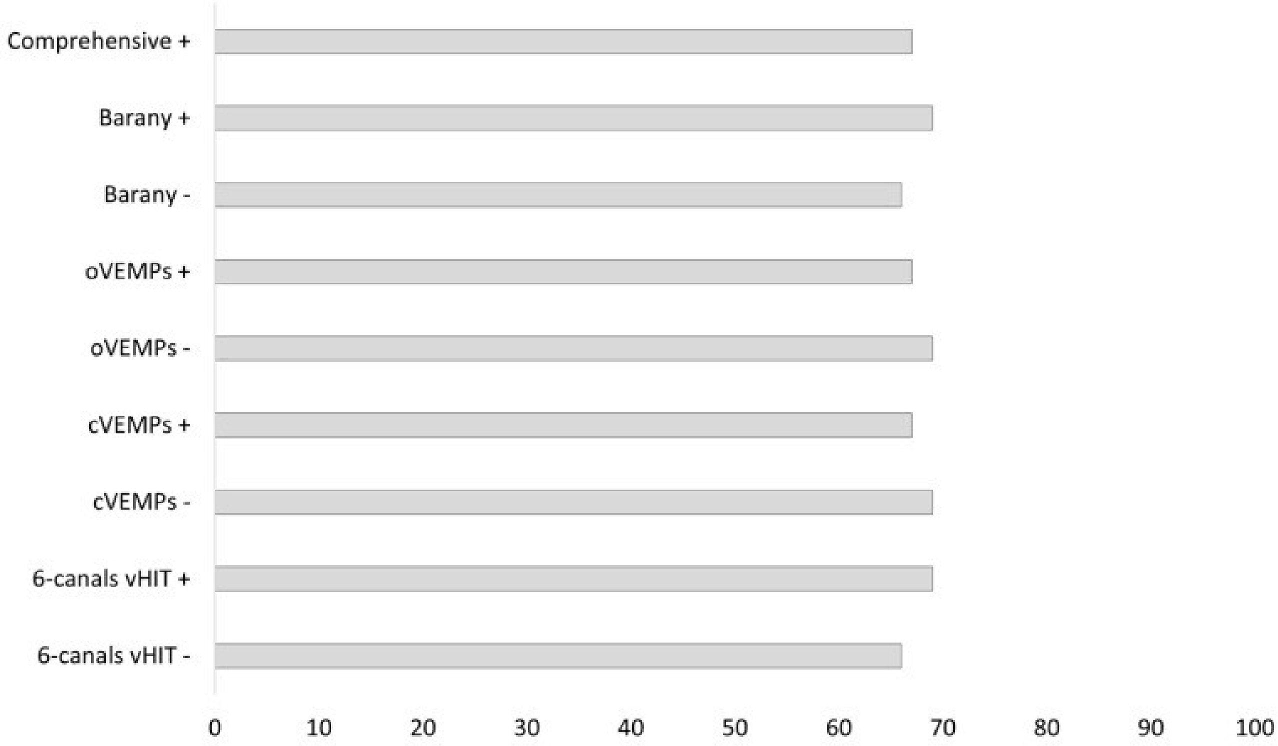

Figure 10. DHI scores according to different criteria (source: Manzari L et al., 2024)

Figure 11 illustrates different combinations of test results from the comprehensive model assessment.

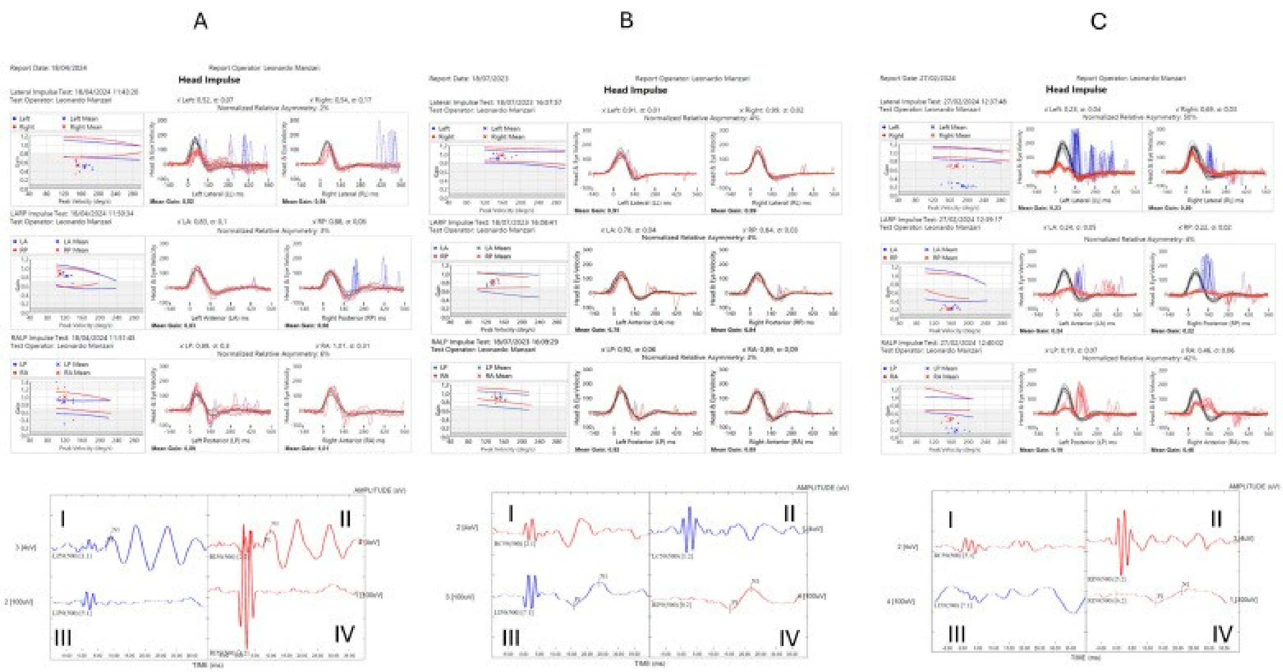

Figure 11. (A) Abnormal horizontal vHIT on the left with normal vertical canal function, combined with absent cVEMPs (saccular dysfunction) but normal oVEMPs (utricular function). (B) Normal vHIT function in all six canals, but absent oVEMPs (utricular dysfunction) with normal cVEMPs. (C) Abnormal vHIT in all six canals, combined with absent oVEMPs and absent cVEMPs, indicating complete utricular and saccular dysfunction (source Manzari et al., 2024)

Discussion

This study evaluated both semicircular canal and otolithic function in patients with BVP. While aVOR gain is a key measure of canal function, it often fails to capture the full scope of vestibular impairment. Moreover, the finding of significant otolithic dysfunction in patients with normal aVOR gain underscores the necessity of including otolith testing in the diagnostic workup for BVP.

Given the observed variability in vestibular function, current diagnostic criteria should be revised to incorporate a broader test battery, such as VEMPs. This would improve diagnostic accuracy, identify distinct BVP subtypes, and enable earlier, more targeted rehabilitation. Indeed, tailoring treatment to a patient's specific deficits, whether canal or otolith-based, is crucial for improving outcomes.

Future research should include longitudinal studies to determine if these BVP subtypes have different prognoses and to explore their underlying mechanisms. Furthermore, based on our normative data, adjusting the pathological aVOR gain cutoff from <0.6 to <0.76 should be considered, as the latter demonstrates superior sensitivity and specificity.

Some study limitations have to be acknowledged. In particular, the lack of an international consensus for ocular VEMPs protocols and the absence of true negative controls limited our analytical scope.

Conclusion

Our findings indicate that several forms of BVH exist that cannot be classified using the current Bárány criteria. It is essential to recognize these new clinical variants, as they can provide new insights into the etiology and potential progression of the condition. We therefore recommend that diagnostic criteria be updated promptly, taking into account modern instrumental tools such as the full vHIT assessment and VEMPs, now available to clinicians.

From Methods to Meaning: Evidence Synthesis with Translational Impact on Emergency Medicine

Study #12. Accuracy of Video Head Impulse Test in Stroke Diagnosis During Acute Vestibular Syndrome: A Systematic Review With Meta-Analysis.

Published as: *Ferri N, Bracci A, Metta Franceschelli G, Gnoli G, Turolla A, Pillastrini P, Manzari L, Tramontano M. Accuracy of video head impulse test in stroke diagnosis during acute vestibular syndrome: a systematic review with meta-analysis. J Neurol. 2025;272(6):436.*[139]

Introduction

Posterior circulation stroke (PCS) accounts for about 20% of ischemic strokes, impacting the brainstem, occipital lobes, and cerebellum areas.[220] Its diagnosis is often missed acutely because symptoms can be non-specific and may not be flagged by standard stroke scales, leading to it being overlooked three times more often than an anterior circulation stroke.[221–223]

Dizziness is the most frequent symptom of PCS, occurring in 47% of patients.[224] Consequently, many present with an acute vestibular syndrome (AVS), i.e., a persistent vertigo with nystagmus and nausea, or an acute imbalance syndrome.[225–229] However, only 3-4% of emergency department dizziness cases are actually due to a stroke.[230,231]

Current guidelines recommend the HINTS+ exam (Head-Impulse, Nystagmus, Test-of-Skew, plus hearing test) for evaluating acute dizziness, as it shows high sensitivity and specificity for detecting central causes.[232,233]

The vHIT is an instrumental version of the bedside[234] test, providing an objective measure of peripheral vestibular function.[125,217] It is valuable in the ED for differentiating stroke from peripheral causes in patients with AVS.[132,133] While a normal vHIT result often suggests a central problem, an abnormal result does not always rule one out.[121]

Although interest in vHIT for emergency triage is growing, inconsistent protocols have led to variable outcomes, and no consensus exists on its diagnostic accuracy for stroke in AVS patients.[235] This study aims to systematically review the literature to determine the sensitivity and specificity of vHIT for diagnosing stroke in individuals presenting with AVS, thereby clarifying its clinical utility.

Methods

This diagnostic test accuracy (DTA) systematic review and meta-analysis was conducted following Cochrane guidelines[236] and reported according to the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) statement.[237] The protocol was registered prospectively in PROSPERO (CRD42024532776).

We included primary studies evaluating the accuracy of the vHIT for diagnosing stroke in adults presenting with AVS in the Emergency Department. The reference standard was neuroimaging with clinical assessment. We searched four major databases (MEDLINE, CENTRAL, CINAHL, SCOPUS) from 2009 to July 2024, alongside grey literature sources. The search strategy was peer-reviewed by a librarian following the Peer Review of Electronic Search Strategies (PRESS) guidelines.[238]

Two reviewers independently screened records, assessed full texts, and extracted data using a predefined form. Disagreements were resolved by consensus.

The risk of bias was evaluated using the Quality Assessment of Diagnostic Test Accuracy Studies 2 (QUADAS-2)[239] tool. Publication bias was assessed using Egger's regression test.

A bivariate meta-analysis was performed to pool sensitivity and specificity, using a Hierarchical Summary Receiver Operating Characteristic (HSROC) model. In addition to pooled sensitivity and specificity, we also calculated diagnostic odds ratios and likelihood ratios and assessed between-study heterogeneity. Sensitivity analyses were conducted to check robustness, and subgroup analyses were performed to compare vHIT performance across the different device manufacturers.

Results

The initial search yielded 654 records. After removing duplicates and screening, 13 studies[129,132,133,240–249] involving 905 participants were included, with some potential sample overlap noted between a few of them (Figure 12).

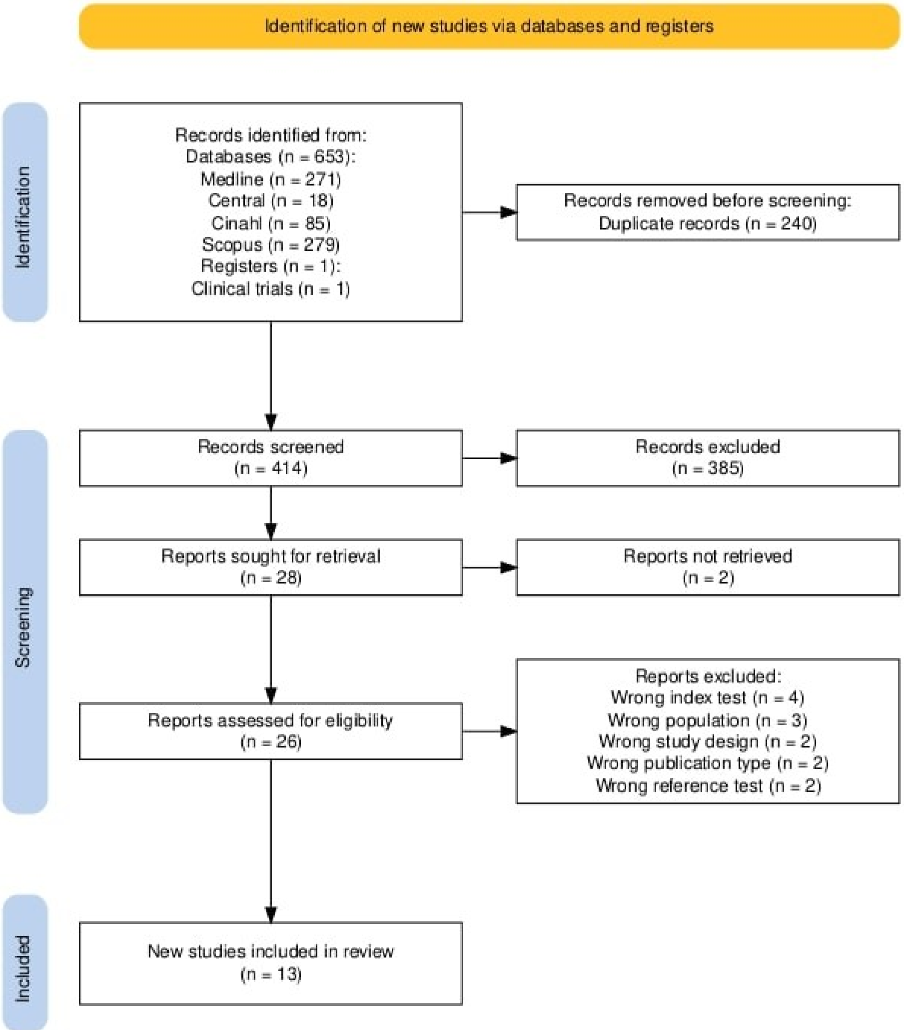

Figure 12. PRISMA-DTA flowchart (source: Ferri N et al., 2025)

Among the 13 studies, the most frequently used vHIT devices were EyeSeeCam (46%) and Otometrics (46%). Normative cutoffs for the horizontal canals were reported in nearly all studies, with EyeSeeCam devices generally having lower cutoff values (range: 0.57–0.70) than the others (range: 0.70–0.80). These cutoffs were derived from various sources, including ROC curves and previously published data. Table 15 reports the characteristics of the 13 studies.

Table 15. Sample characteristics

Author, year	Country	Sample	Age (y)	Sex (F, %)	vHIT	Normative cutoff (HC)
Guler, 2017	Turkey	n = 40 24 VN 16 strokes	49	35	Otometrics	>0.75
Korda, 2022a	Switzerland	n = 46 35 AUVP 11 strokes	55	46	EyeSeeCam	>0.69
Korda, 2022b	Switzerland	n = 57 39 AUVP 18 strokes	55	NR	EyeSeeCam	>0.57
Siepmann, 2021	Germany	n = 30 11 central (4 strokes, 7 TIA) 19 peripherals	55	47	EyeSeeCam	NR
Machner, 2021	Germany	n = 38 24 VN 14 strokes	69	34	EyeSeeCam	>0.70
Mantokoudis, 2015[35]	USA	n = 26 16 VN 10 strokes	60	27	Otometrics	≥0.70
Mantokoudis, 2016	USA	n = 23 16 VN 7 strokes	60	26	Otometrics	≥0.70
Michailidou, 2024	Switzerland	n = 76 52 VN 24 strokes	56	NR	EyeSeeCam	>0.67
Morrison, 2022	Switzerland	n = 152 58 VN 27 strokes 67 NR	56	NR	EyeSeeCam	>0.685
Nam, 2021	South Korea	n = 51 17 VN 17 strokes 17 healthy controls	61	NR	SLMED	>0.71
Nham, 2022	Australia	n = 148 61 VN 46 strokes 41 other	60	NR	Otometrics	≥0.74
Nham, 2023	Australia	n = 262 134 VN 128 strokes	62	36	Otometrics	≥0.74
Thomas, 2022	Australia	n = 133 53 VN 20 strokes	62	46	Otometrics	≥0.8

21 migraines
39 other

vHIT: video Head Impulse Test, HC: horizontal semicircular canals, NR: not reported, AUVP: acute unilateral vestibulopathy, TIA: transient ischemic attack, VN: vestibular neuritis

The risk of bias assessment using QUADAS-2 identified the 'flow and timing' domain as being at high risk in four studies, while a common concern was the inconsistent reporting on whether clinicians performing the vHIT were blinded to the patient's diagnosis. The timing of the vHIT test after symptom onset also varied widely (Figure 13).

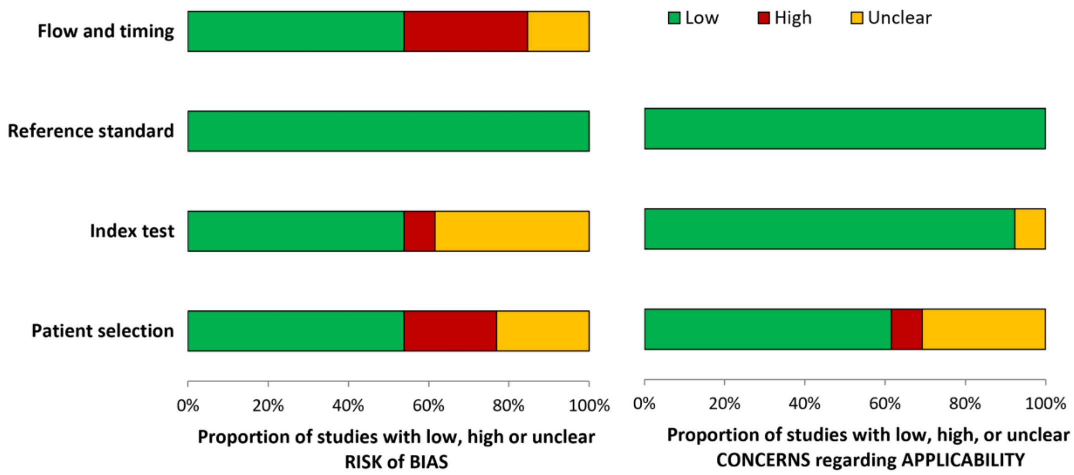

Figure 13. Risk of bias of included studies (source: Ferri N et al., 2025)

For the meta-analysis on diagnostic accuracy, eight studies[129,132,240,244–246,248,249] (624 participants) were included, excluding those that reported different results on the same sample. The vHIT demonstrated a pooled sensitivity of 84% and specificity of 85% for stroke diagnosis. Consequently, vHIT may be helpful in both ruling in and ruling out disease, showing excellent discriminative ability (Figure 14).

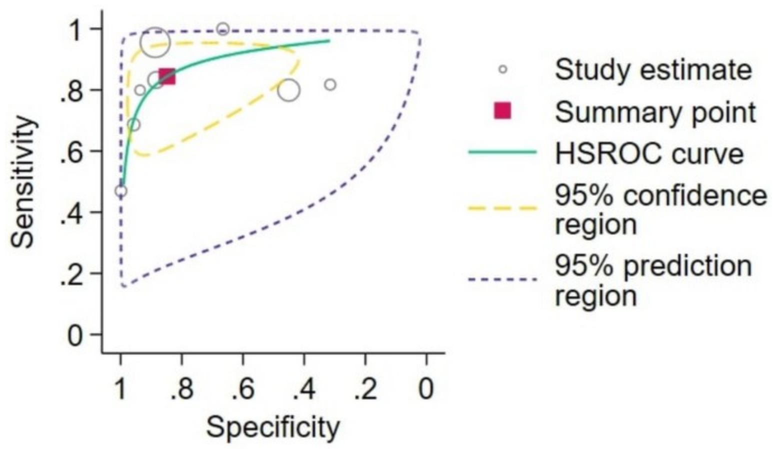

Figure 14. Bivariate model for overall accuracy (source: Ferri N et al., 2025)

However, some heterogeneity was observed among the studies. A subgroup analysis revealed that specificity might differ between device types (Otometrics, 4 studies: 81%; EyeSeeCam, 3 studies: 63%), but these results were uncertain due to small sample sizes. Statistical analyses like meta-regression were not feasible due to the limited sample size, and although the Egger test indicated no publication bias (p = 0.826), this result must be taken with caution, given the limited number of studies.

Discussion

This study systematically evaluated the accuracy of the vHIT for diagnosing stroke in patients with AVS. The results demonstrate that vHIT is a valuable clinical tool, with high sensitivity (84%) and specificity (85%) for differentiating central from peripheral causes of vertigo. Specifically, a negative vHIT result significantly lowers the probability of stroke, making it a practical rule-out test in emergency settings. This could expedite diagnosis and reduce reliance on more extensive neuroimaging.

However, significant variability in specificity was observed across studies, likely due to differences in vHIT devices, clinician expertise, testing procedures, and the criteria for defining abnormal results. While a comparison of the two primary devices, EyeSeeCam and Otometrics, suggested a slightly higher specificity for the latter, the difference was not statistically significant. This highlights a critical need for standardized testing protocols and normative values to ensure consistent performance.

When compared to the HINTS+ bedside exam, the gold standard for detecting stroke in AVS,[250] vHIT offers a key advantage as an objective, instrument-based measure that does not require specialized neuro-otological training. It thus shows promise as a complementary tool, particularly in settings where such expertise is unavailable.

Future research should also incorporate the assessment of compensatory saccades and employ data-driven methods to establish optimal, cross-device gain thresholds. This would significantly enhance the test's standardization and diagnostic power.

Conclusion

This systematic review and meta-analysis confirm the vHIT as a highly sensitive and specific tool for diagnosing stroke in emergency department patients with AVS. Its diagnostic value is highest when integrated into a comprehensive clinical evaluation, rather than used in isolation. However, significant variability in specificity was found, highlighting an urgent need for standardized protocols for test administration, interpretation, and training to ensure consistent results.

Future research should refine vHIT interpretation criteria, investigate the role of compensatory saccades, and evaluate its impact on clinical decisions, patient outcomes, and cost-effectiveness within stroke triage pathways.

General discussion and future perspectives

The evidence base from these studies, primarily focused on the objective quantification provided by the vHIT, clearly reveals a significant and often underappreciated prevalence of dynamic aVOR dysfunction across various neurological conditions in neurorehabilitation settings. The use of instrumental assessment tools, including the HIMP and the SHIMP paradigms, has demonstrated that vestibular impairment, characterized by both hypo- and hyper-gain dysfunctions, affects the vast majority of patients: over 83% of individuals with PD, more than 90% of sTBI survivors, and over 81% of people with MS. This high incidence highlights the critical need for individualized assessment, as relying solely on mean aVOR gain values often masks prevalent isolated canal dysfunctions, such as the specific impairment of vertical semicircular canals, which correlates significantly with compromised balance and gait in MS, particularly in those with higher EDSS scores (Class 2). Further supporting its clinical relevance, studies in children with CP demonstrate a strong negative correlation between aVOR gain in horizontal canals and functional motor ability (GMFCS level), suggesting that aVOR performance declines directly as motor impairment increases. The utility of vHIT is also confirmed in acute diagnostics, offering high sensitivity (84%) and specificity (85%) for differentiating stroke from peripheral causes in AVS, thereby serving as a valuable tool for ruling out central involvement due to its low negative likelihood ratio (0.18).

Despite mounting evidence, translating VPT knowledge into routine clinical practice remains significantly hampered by systemic barriers, primarily the perceived lack of specialized skills and the cost of equipment. Nearly half of the surveyed rehabilitation professionals report that they do not utilize instrumental assessments. Moving forward, the imperative for the field of neurorehabilitation involves not only the routine integration of objective vestibular assessment to complement standard balance evaluations but also a concerted effort toward educational enhancement at the basic academic level and the development of specific, standardized guidelines for VPT implementation in neurological disorders. Furthermore, the identified low sensitivity (46%) of the current Bárány Society criteria for BVP, which currently exclude the assessment of vertical canals and otolithic function (VEMPs), mandates an urgent revision of diagnostic standards to a comprehensive instrumental model that fully captures the heterogeneity and specific subtypes of vestibular loss. Ultimately, future interventions must be explicitly tailored to the patient's individual aVOR profile. This goal may necessitate innovative resource distribution strategies, such as establishing specialized Hub and Spoke centers for more complex instrumental testing.

From NICU to Developmental Milestones: The Role of the Vestibular System in Preterm Infants

Study #13. Effectiveness of Early Vestibular Habilitation in Preterm Infants: A Randomized Controlled Trial Protocol

Introduction

Prematurity is defined as birth occurring before 37 weeks of gestation. It has a global prevalence of approximately 10% and it is the leading cause of death in the first 5 years of life.[251] It is also a long-term risk factor for the development of cognitive and neuromotor disabilities at school age, with a risk 5 times greater compared to children born at term.[252] It is possible that the presence of stressful stimuli related to the Neonatal Intensive Care Unit (NICU) stay and the reduced presence of functional stimuli may influence neurocognitive development.[253] A recent systematic review with meta-analysis supported the efficacy of early multisensory care in increasing the weight of preterm infants at discharge; however, the available primary studies are still few and very heterogeneous in their treatment modalities.[254] Furthermore, the literature supports the efficacy of parental participation in neonatal care programs, facilitated by their early active involvement, which creates new perspectives for the parents, their child, and their mutual interaction and relationship (zero separation and family-integrated care).[255]

The objectives of the present study are:

- to assess the efficacy of early and ecological vestibular habilitation in improving short-term vital parameters, neurophysiological maturation, reducing the length of hospital stay, reducing the time to achieve full enteral nutrition (FEN), and improving neuromotor development at discharge and in the long term, compared to standard care.
- to assess the effect of movement during skin-to-skin contact on the level of perceived stress in parents.

Methods

The present research will be conducted as a randomized intervention study with a 1:1 allocation ratio, with blinded assessment for the primary outcome; the study design is identified as a superiority trial. This protocol is reported in accordance with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines.[256] The trial will be reported following the CONSORT guidelines.[257]

This clinical research will be conducted within the NICU of the “Infermi” Hospital in Rimini (Italy), which is also a NIDCAP (Newborn Individualized Developmental Care and Assessment Program) training center.

We will include:

- Preterm infants with a gestational age (GA) from 24 weeks (from 24+0 to 24+6 days) to 32 weeks (from 32+0 to 32+6 days) inclusive;
- Eligibility for skin-to-skin contact (SSC) according to national and international recommendations.

We will exclude:

- Cardiorespiratory instability
- Grade III-IV intraventricular haemorrhage according to the Volpe classification
- Hydrocephalus
- Major malformations
- Major surgery

The control group (CON) will receive standard ward care, according to the NIDCAP approach, as codified in a specific Unit procedure. Specifically, SSC will be offered using the traditional methodology, i.e., in a chair, in a vertical, horizontal, or semi-reclined position, with the infant placed securely and under constant clinical and instrumental monitoring on the mother's or father's chest. The NIDCAP approach and SSC have been the subject of numerous studies in the scientific literature, providing significant evidence of efficacy in improving the health status and psychomotor development of preterm infants.[258,259]

The experimental group (EXP) will be managed with the same NIDCAP approach as the control group. Furthermore, in this group, SSC will be supplemented by early vestibular integration: the parent will have the possibility to perform SSC dynamically, i.e., by moving around the infant's station or within the Unit while wearing the infant in a sling, securely and under constant clinical and multi-parametric monitoring. The duration, intensity, and frequency of this activity will be at the parent's discretion, with the recommendation to perform it at least once per day, for at least 15 minutes.

The CON and EXP interventions will begin after a minimum of 2 weeks of standard SSC for preterm infants below 30 weeks GA, and after a minimum of 1 week of standard SSC for preterm infants ≥ 30 weeks GA, to allow for proper assessment of exclusion criteria, clinical stabilization, and adequate parental training in SSC. In both groups, a certified and validated wearable device (HOWDY BABY©, Comftech, Italy) will monitor the infant's vital parameters, the total duration of the activity, and the accelerometric parameters related to movements performed at predefined intervals (Figure 13). Both parents will be allowed to perform SSC, even taking turns, but not during the same session.

Figure 15. Howdy Baby© (Comftech)

Primary Outcome

- Motor development at 24 months of post-conceptional age, measured with the Bayley Scales of Infant and Toddler Development (3rd Edition). A professional with at least 15 years of experience will perform the assessment, who will be blinded to group allocation.

Secondary Outcomes

- Vital parameters: heart rate (HR), respiratory rate (RR), transcutaneous oxygen saturation (SpO2), recorded during SSC
- Time to achieve FEF
- Assessment with the Burdjalov score[260] on amplitude-integrated electroencephalography (aEEG) at 36 weeks post-conceptional age
- Hammersmith Neonatal Neurological Examination (HNNE) at term-equivalent age (40 weeks)
- Number of desaturation episodes (SpO2<89%) lasting at least 10 seconds
- Heart Rate Variability (HRV)
- Number of supra-threshold heartbeats (HR>180 bpm) for a consecutive duration of at least 10 seconds
- Number of sub-threshold heartbeats (HR<90 bpm) for a consecutive duration of at least 10 seconds
- Parental Stressor Scale: Neonatal Intensive Care Unit (PSS: NICU), a scale to measure parental stress validated in Italian.[261]

Specifically:

- During SSC, the following physiological parameters will be monitored in both groups: HR, RR, SpO2, HRV, Desaturations, Supra-threshold beats, Sub-threshold beats.

- At 36 weeks PCA, the Burdjalov score on a 3-hour aEEG recording will be assessed.
- At discharge, data will be collected on: length of stay, time to achieve FEF, and the PSS: NICU scale administered to parents.
- At 40 weeks PCA, the HNNE will be administered.
- At 24 months corrected age, the Bayley Scales of Infant and Toddler Development (BSID) – Motor Domain will be administered.
- At school age (6 years), cognitive and neuromotor outcomes will be collected via the Griffiths Mental Development Scales – Extended Revised (GMDS-ER), which will be the subject of a secondary analysis.

The recruitment period is estimated to be approximately 30 months, from March 2026 to September 2028. The conclusion of the study is planned for September 2028 (endpoint: discharge), September 2030 (endpoint: 24 months), and August 2034 (endpoint: 6 years).

A cohort of 207 infants born at <32 weeks GA between September 2017 and June 2021, assessed at 24 months corrected age using the BSID scale, showed a mean (SD) motor subscale score of 89.7 (11.6). Assuming a pooled standard deviation of 11.6 in both the CON and EXP groups, to detect a statistically significant increase of 6 points (motor score 95.7) in the EXP group with 80% statistical power and a 5% alpha error, 59 infants per group are required (total 118 infants). Considering that the primary endpoint at 24 months is already a standard follow-up visit for preterm infant care and that the treatment is administered during hospitalization, a dropout rate of less than 5% is estimated. The final sample size, corrected for the dropout rate, is therefore 124 preterm infants, 62 in the EXP group, and 62 in the CON group. This population is recruitable within 36 months.
All infants admitted to the NICU of the "Infermi" Hospital will be assessed at 1 week of life according to the inclusion and exclusion criteria. If the infant is eligible, an information letter will be provided to the parents, and written informed consent for participation in the study will be obtained. The recruitment phase will continue until the predefined sample size is reached.

Infant allocation will occur after obtaining informed consent. A permuted block randomization list (blocks of 2 and 4) will be used for patient assignment, with stratification based on three gestational age levels: 24-26 weeks, 27-29 weeks, 30-32 weeks. An independent off-site researcher will perform this process, and the list will be kept concealed until assignment to the intervention or control group.

The quantity and direction of movement will be measured using a wearable monitoring device. This device will also record outcomes related to vital parameters. Movement and vital parameters will be collected using the HOWDY BABY© wearable textile sensor (Comftech - Monza, Italy), composed of an electronic unit (certified Class IIb medical device), a sensorized band (certified Class I medical device), and dedicated software for data reading, enabling movement during skin-to-skin contact thanks to wireless technology (Fig. 1). A recent study highlighted the reliability of this instrument compared to the gold standard represented by the electrocardiogram, with an average discrepancy of 1.3 bpm between the two instruments.[262]

The sensor will be positioned on the infant by clinically trained staff before the start of SSC. It will remain in situ and operational during SSC in both the EXP and CON groups to monitor the effect of these activities and ensure constant clinical monitoring.

The aEEG at 36 weeks will be performed by two independent assessors using the EB Neuro or Brainz device, with a total recording duration of 3 hours during a period of infant rest. A maturational score will be calculated according to the Burdjalov method.[260]

At discharge, the total length of stay (in days), the time to achieve FEF (in days) via medical record review, the results of the PSS:NICU scale administered to parents, and the HNNE administered at 40 weeks PCA will be collected. An assessor with at least 10 years of experience in neonatal neurological assessment will apply the latter. Long-term cognitive and neuromotor outcomes will be collected at 2 years via the BSID-Motor Domain and at 6 years via the GMDS-ER; both scales will be administered by a professional blinded to group allocation.

If, for primary clinical reasons, an infant cannot perform the activity as per allocation, they will not be excluded from the trial. Still, they will be temporarily suspended from the study until resumption is possible, if feasible. Adherence to the treatment will be monitored and subject to subsequent analysis between the two groups.

Data from the HOWDY BABY© sensor will be automatically transferred to the local server.
The aEEG data, length of stay, FEF, HNNE, BSID, and GMDS-ER data will be entered by the assessor into a predefined Excel sheet on the day of collection and stored in the local server.
Each infant will be assigned a unique identification code that will be used in data management and analysis through a pseudonymization procedure. The code attribution list, which links codes to identities, will be stored separately in a dedicated, locked archive located within the Operating Unit.

Descriptive and inferential analyses will be conducted. Specifically, the sample will be described using the mean and standard deviation (or median and interquartile range in case of asymmetric distribution) for continuous variables, and frequencies with percentages for categorical variables.
Inferential statistics will include within-group and between-group comparisons at baseline, weekly, discharge, 24 months, and 6 years. Analyses will be intention-to-treat, using a multiple imputation method for missing data. A regression analysis will be performed to evaluate the effect of treatment dosage (intensity, duration, frequency) on treatment efficacy.

Considering the functional characteristics of the intervention, no serious adverse events are expected. As with any external stimulus, a temporary increase in the variability of body temperature, HR, RR, and/or SpO2 may be expected; this will be constantly monitored and recorded daily. Any unexpected adverse event and the determination of a possible causal relationship will be promptly communicated to the Ethics Committee. The HOWDY BABY© monitoring sensor is a certified device, and its use in the trial is consistent with the technical data sheet and user manual.

This research has been positively evaluated by the Scientific Medical Board (CMS-Board) at the Romagnolo Institute for the Study of Tumors (IRST) in Meldola (Italy) with report No. 7 on June 11, 2024. Following approval by the National Pediatric Ethics Committee (AIFA), the protocol will be registered on ClinicalTrials.gov.

Any deviation from the protocol will be reported to the Ethics Committee, all researchers, the parents of the trial participants, the ClinicalTrials.gov registry, corporate regulators, and reported in the final manuscript. Informed consent will be obtained in written form by a researcher during an informational meeting with the parents, where sufficient time will be granted to clarify any requests for additional information. The authors declare no conflicts of interest; the HOWDY BABY© sensors will be provided free of charge by the manufacturing company, which will have no role in the conduct of the study, analyses, or reporting.

Disclaimer

I have no conflicts of interest to declare and have received no grants outside of the PhD scholarship. Artificial intelligence tools have been used to edit reports and enforce rules against self-plagiarism, primarily because these involve studies we published in scientific journals.

References

1. Lunny C, Thirugnanasampanthar SS, Kanji S, et al. How can clinicians choose between conflicting and discordant systematic reviews? A replication study of the Jadad algorithm. *BMC Med Res Methodol*. 2022;22(1):276. doi:10.1186/s12874-022-01750-2

2. Ioannidis JPA. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *Milbank Q*. 2016;94(3):485-514. doi:10.1111/1468-0009.12210

3. Hoffmann F, Allers K, Rombey T, et al. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. *J Clin Epidemiol*. 2021;138:1-11. doi:10.1016/j.jclinepi.2021.05.022

4. Niforatos JD, Weaver M, Johansen ME. Assessment of Publication Trends of Systematic Reviews and Randomized Clinical Trials, 1995 to 2017. *JAMA Intern Med*. 2019;179(11):1593-1594. doi:10.1001/jamainternmed.2019.3013

5. Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ Can Med Assoc J J Assoc Medicale Can*. 1997;156(10):1411-1416.

6. Lunny C, Neelakant T, Chen A, et al. Bibliometric study of “overviews of systematic reviews” of health interventions: Evaluation of prevalence, citation and journal impact factor. *Res Synth Methods*. 2022;13(1):109-120. doi:10.1002/jrsm.1530

7. Lunny C, Thirugnanasampanthar SS, Kanji S, et al. Identifying and addressing conflicting results across multiple discordant systematic reviews on the same question: protocol for a replication study of the Jadad algorithm. *BMJ Open*. 2022;12(4):e054223. doi:10.1136/bmjopen-2021-054223

8. Bakdach WMM, Hadad R. Effectiveness of different adjunctive interventions in the management of orthodontically induced white spot lesions: A systematic review of systematic reviews and meta-analyses. *Dent Med Probl*. 2020;57(3):305-325. doi:10.17219/dmp/118330

9. Bolland MJ, Grey A. A case study of discordant overlapping meta-analyses: vitamin d supplements and fracture. *PloS One*. 2014;9(12):e115934. doi:10.1371/journal.pone.0115934

10. Chalmers PN, Mascarenhas R, Leroux T, et al. Do arthroscopic and open stabilization techniques restore equivalent stability to the shoulder in the setting of anterior glenohumeral instability? a systematic review of overlapping meta-analyses. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2015;31(2):355-363. doi:10.1016/j.arthro.2014.07.008

11. Chen X, Guo W, Li Q, et al. Is Unilateral Percutaneous Kyphoplasty Superior to Bilateral Percutaneous Kyphoplasty for Osteoporotic Vertebral Compression Fractures? Evidence from a Systematic Review of Discordant Meta-Analyses. *Pain Physician*. 2018;21(4):327-336.

12. Erickson BJ, Mascarenhas R, Sayegh ET, et al. Does Operative Treatment of First-Time Patellar Dislocations Lead to Increased Patellofemoral Stability? A Systematic Review of Overlapping Meta-analyses. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2015;31(6):1207-1215. doi:10.1016/j.arthro.2014.11.040

13. Fu BS, Jia HL, Zhou DS, Liu FX. Surgical and Non-Surgical Treatment for 3-Part and 4-Part Fractures of the Proximal Humerus: A Systematic Review of Overlapping Meta-Analyses. *Orthop Surg*. 2019;11(3):356-365. doi:10.1111/os.12486

14. Grassi A, Compagnoni R, Ferrua P, et al. Patellar resurfacing versus patellar retention in primary total knee arthroplasty: a systematic review of overlapping meta-analyses. *Knee Surg Sports Traumatol Arthrosc Off J ESSKA*. 2018;26(11):3206-3218. doi:10.1007/s00167-018-4831-8

15. Guo XW, Hu ND, Sun GZ, Li M, Zhang PT. Shenyi Capsule () plus Chemotherapy versus Chemotherapy for Non-Small Cell Lung Cancer: A Systematic Review of Overlapping Meta-Analyses. *Chin J Integr Med*. 2018;24(3):227-231. doi:10.1007/s11655-017-2951-5

16. Houck DA, Kraeutler MJ, Schuette HB, McCarty EC, Bravman JT. Early Versus Delayed Motion After Rotator Cuff Repair: A Systematic Review of Overlapping Meta-analyses. *Am J Sports Med*. 2017;45(12):2911-2915. doi:10.1177/0363546517692543

17. Mascarenhas R, Cvetanovich GL, Sayegh ET, et al. Does Double-Bundle Anterior Cruciate Ligament Reconstruction Improve Postoperative Knee Stability Compared With Single-Bundle Techniques? A Systematic Review of Overlapping Meta-analyses. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2015;31(6):1185-1196. doi:10.1016/j.arthro.2014.11.014

18. Pekala PA, Henry BM, Taterra D, et al. FokI as a genetic factor of intervertebral disc degeneration: A PRISMA-compliant systematic review of overlapping meta-analyses. *J Clin Neurosci Off J Neurosurg Soc Australas*. 2019;60:36-43. doi:10.1016/j.jocn.2018.09.028

19. Poolman RW, Abouali JAK, Conter HJ, Bhandari M. Overlapping systematic reviews of anterior cruciate ligament reconstruction comparing hamstring autograft with bone-patellar tendon-bone autograft: why are they different? *J Bone Joint Surg Am*. 2007;89(7):1542-1552. doi:10.2106/JBJS.F.01292

20. Song GM, Bian W, Zeng XT, Zhou JG, Luo YQ, Tian X. Laparoscopic cholecystectomy for acute cholecystitis: early or delayed?: Evidence from a systematic review of discordant meta-analyses. *Medicine (Baltimore)*. 2016;95(23):e3835. doi:10.1097/MD.0000000000003835

21. Tan G, Li F, Zhou D, Cai X, Huang Y, Liu F. Unilateral versus bilateral percutaneous balloon kyphoplasty for osteoporotic vertebral compression fractures. *Medicine (Baltimore)*. 2018;97(33):e11968. doi:10.1097/MD.0000000000011968

22. Xing D, Wang B, Liu Q, et al. Intra-articular Hyaluronic Acid in Treating Knee Osteoarthritis: a PRISMA-Compliant Systematic Review of Overlapping Meta-analysis. *Sci Rep*. 2016;6:32790. doi:10.1038/srep32790

23. Xu B, Lin Y, Wang Z, et al. Is intramedullary fixation of displaced midshaft clavicle fracture superior to plate fixation? Evidence from a systematic review of discordant meta-analyses. *Int J Surg Lond Engl*. 2017;43:155-162. doi:10.1016/j.ijsu.2017.05.069

24. Zhang X, Huang J, Zhao L, et al. Inferomedial cortical bone contact and fixation with calcar screws on the dynamic and static mechanical stability of proximal humerus fractures. *J Orthop Surg*. 2019;14(1):1. doi:10.1186/s13018-018-1031-7

25. Zhao JG, Wang J, Long L. Surgical Versus Conservative Treatments for Displaced Midshaft Clavicular Fractures. *Medicine (Baltimore)*. 2015;94(26):e1057. doi:10.1097/MD.0000000000001057

26. Zhao JG, Wang J, Wang C, Kan SL. Intramedullary nail versus plate fixation for humeral shaft fractures: a systematic review of overlapping meta-analyses. *Medicine (Baltimore)*. 2015;94(11):e599. doi:10.1097/MD.0000000000000599

27. Zhiyong C, Yun T, Hui F, Zhongwei Y, Zhaorui L. Unilateral Versus Bilateral Balloon Kyphoplasty for Osteoporotic Vertebral Compression Fractures: A Systematic Review of Overlapping Meta-analyses. *Pain Physician*. 2019;22(1):15-28.

28. Mascarenhas R, Chalmers PN, Sayegh ET, et al. Is double-row rotator cuff repair clinically superior to single-row rotator cuff repair: a systematic review of overlapping meta-analyses. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2014;30(9):1156-1165. doi:10.1016/j.arthro.2014.03.015

29. Lunny C, Whitelaw S, Reid EK, et al. Exploring decision-makers' challenges and strategies when selecting multiple systematic reviews: insights for AI decision support tools in healthcare. *BMJ Open*. 2024;14(7):e084124. doi:10.1136/bmjopen-2024-084124

30. Sackett DL. Evidence-based medicine. *Semin Perinatol*. 1997;21(1):3-5. doi:10.1016/s0146-0005(97)80013-4

31. Moher D. The problem of duplicate systematic reviews. *BMJ*. 2013;347:f5040. doi:10.1136/bmj.f5040

32. Page MJ, Moher D. Mass Production of Systematic Reviews and Meta-analyses: An Exercise in Mega-silliness? *Milbank Q*. 2016;94(3):515-519. doi:10.1111/1468-0009.12211

33. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res*. 2004;6(3):e34. doi:10.2196/jmir.6.3.e34

34. Dillman DA. *Mail and Internet Surveys: The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. John Wiley & Sons; 2011.

35. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71

36. Chiu K, Grundy Q, Bero L. "Spin" in published biomedical literature: A methodological systematic review. *PLoS Biol*. 2017;15(9):e2002173. doi:10.1371/journal.pbio.2002173

37. Demla S, Shinn E, Ottwell R, et al. Evaluation of spin in the abstracts of systematic reviews and meta-analyses focused on cataract therapies. *Am J Ophthalmol*. 2021;228:47-57. doi:10.1016/j.ajo.2021.03.032

38. McGrath TA, McInnes MDF, van Es N, Leeflang MMG, Korevaar DA, Bossuyt PMM. Overinterpretation of Research Findings: Evidence of "Spin" in Systematic Reviews of Diagnostic Accuracy Studies. *Clin Chem*. 2017;63(8):1353-1362. doi:10.1373/clinchem.2017.271544

39. Nascimento DP, Gonzalez GZ, Araujo AC, Moseley AM, Maher CG, Costa LOP. Eight in Every 10 Abstracts of Low Back Pain Systematic Reviews Presented Spin and Inconsistencies With the Full Text: An Analysis of 66 Systematic Reviews. *J Orthop Sports Phys Ther*. 2020;50(1):17-23. doi:10.2519/jospt.2020.8962

40. Whiting P, Savović J, Higgins JPT, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-234. doi:10.1016/j.jclinepi.2015.06.005

41. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. doi:10.1136/bmj.j4008

42. Bojcic R, Todoric M, Puljak L. Adopting AMSTAR 2 critical appraisal tool for systematic reviews: speed of the tool uptake and barriers for its adoption. *BMC Med Res Methodol*. 2022;22(1):104. doi:10.1186/s12874-022-01592-y

43. Burda BU, Holmer HK, Norris SL. Limitations of A Measurement Tool to Assess Systematic Reviews (AMSTAR) and suggestions for improvement. *Syst Rev*. 2016;5:58. doi:10.1186/s13643-016-0237-1

44. Wegewitz U, Weikert B, Fishta A, Jacobs A, Pieper D. Resuming the discussion of AMSTAR: What can (should) be made better? *BMC Med Res Methodol*. 2016;16(1):111. doi:10.1186/s12874-016-0183-6

45. Banzi R, Cinquini M, Gonzalez-Lorenzo M, Pecoraro V, Capobussi M, Minozzi S. Quality assessment versus risk of bias in systematic reviews: AMSTAR and ROBIS had similar reliability but differed in their construct and applicability. *J Clin Epidemiol*. 2018;99:24-32. doi:10.1016/j.jclinepi.2018.02.024

46. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. *J Clin Epidemiol*. 2019;108:26-33. doi:10.1016/j.jclinepi.2018.12.004

47. Higgins J, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Healthcare Interventions*. 2nd Edition. John Wiley & Sons; 2019.

48. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-349. doi:10.1016/j.jclinepi.2007.11.008

49. Smires S, Afach S, Mazaud C, et al. Quality and Reporting Completeness of Systematic Reviews and Meta-Analyses in Dermatology. *J Invest Dermatol*. 2021;141(1):64-71. doi:10.1016/j.jid.2020.05.109

50. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol*. 2020;122:129-141. doi:10.1016/j.jclinepi.2019.12.020

51. Ferri N, Ravizzotti E, Bracci A, Carreras G, Pillastrini P, Di Bari M. The confidence in the results of physiotherapy systematic reviews in the musculoskeletal field is not increasing over time: a meta-epidemiological study using AMSTAR 2 tool. *J Clin Epidemiol*. 2024;169:111303. doi:10.1016/j.jclinepi.2024.111303

52. Clarke J. What is a systematic review? *Evid Based Nurs*. 2011;14(3):64. doi:10.1136/ebn.2011.0049

53. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Lond Engl*. 2021;396(10267):2006-2017. doi:10.1016/S0140-6736(20)32340-0

54. Zhong CCW, Zhao J, Wong CHL, et al. Methodological quality of systematic reviews on treatments for Alzheimer's disease: a cross-sectional study. *Alzheimers Res Ther*. 2022;14(1):159. doi:10.1186/s13195-022-01100-w

55. Matthias K, Rissling O, Pieper D, et al. The methodological quality of systematic reviews on the treatment of adult major depression needs improvement according to AMSTAR 2: A cross-sectional study. *Heliyon*. 2020;6(9):e04776. doi:10.1016/j.heliyon.2020.e04776

56. De Santis KK, Lorenz RC, Lakeberg M, Matthias K. The application of AMSTAR2 in 32 overviews of systematic reviews of interventions for mental and behavioural disorders: A cross-sectional study. *Res Synth Methods*. 2022;13(4):424-433. doi:10.1002/jrsm.1532

57. De Santis KK, Matthias K. Different Approaches to Appraising Systematic Reviews of Digital Interventions for Physical Activity Promotion Using AMSTAR 2 Tool: Cross-Sectional Study. *Int J Environ Res Public Health*. 2023;20(6):4689. doi:10.3390/ijerph20064689

58. Hammel C, Pandis N, Pieper D, Faggion CM. Methodological assessment of systematic reviews of in-vitro dental studies. *BMC Med Res Methodol*. 2022;22(1):110. doi:10.1186/s12874-022-01575-z

59. Matterne U, Theurich MA, Pröbstl S, et al. Quality of systematic reviews on timing of complementary feeding for early childhood allergy prevention. *BMC Med Res Methodol*. 2023;23(1):80. doi:10.1186/s12874-023-01899-4

60. Zhou X, Li L, Lin L, Ju K, Kwong JSW, Xu C. Methodological quality for systematic reviews of adverse events with surgical interventions: a cross-sectional survey. *BMC Med Res Methodol*. 2021;21(1):223. doi:10.1186/s12874-021-01423-6

61. Siegel MG, Brand JC, Rossi MJ, Lubowitz JH. “Publish or Perish” Promotes Medical Literature Quantity Over Quality. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2018;34(11):2941-2942. doi:10.1016/j.arthro.2018.08.029

62. Bojcic R, Todoric M, Puljak L. Most systematic reviews reporting adherence to AMSTAR 2 had critically low methodological quality: a cross-sectional meta-research study. *J Clin Epidemiol*. 2024;165:111210. doi:10.1016/j.jclinepi.2023.10.026

63. Gates M, Gates A, Pieper D, et al. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. *BMJ*. 2022;378:e070849. doi:10.1136/bmj-2022-070849

64. Gattrell WT, Logullo P, van Zuuren EJ, et al. ACCORD (ACcurate COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine developed via a modified Delphi. *PLoS Med*. 2024;21(1):e1004326. doi:10.1371/journal.pmed.1004326

65. Riley S, Swanson BT, Sawyer SF, Brismée JM, Staysniak G. Should low-quality evidence dominate high-level evidence? A systematic review and meta-analysis of systematic reviews of musculoskeletal physical therapy interventions. *J Man Manip Ther*. 2021;29(4):203-215. doi:10.1080/10669817.2020.1839728

66. Lorenz RC, Matthias K, Pieper D, et al. AMSTAR 2 overall confidence rating: lacking discriminating capacity or requirement of high methodological quality? *J Clin Epidemiol*. 2020;119:142-144. doi:10.1016/j.jclinepi.2019.10.006

67. Curry S. Let’s move beyond the rhetoric: it’s time to change how we judge research. *Nature*. 2018;554(7691):147. doi:10.1038/d41586-018-01642-w

68. Hunter PR. Bibliometrics, research quality, and neglected tropical diseases. *Lancet Lond Engl*. 2009;373(9664):630-631. doi:10.1016/S0140-6736(09)60397-4

69. Singh HP. Alternative research bibliometrics: It’s about quality and not quantity. *Shoulder Elb*. 2022;14(2):121-122. doi:10.1177/17585732211058453

70. Weightman AL, Butler CC. Using bibliometrics to define the quality of primary care research. *BMJ*. 2011;342:d1083. doi:10.1136/bmj.d1083

71. Ferri N, Morone G, Piermaria J, et al. Knowledge, Barriers, and Future Directions of Vestibular Rehabilitation Practice in Neurorehabilitation: An Italian Survey. *Healthc Basel Switz*. 2024;13(1):22. doi:10.3390/healthcare13010022

72. Sulway S, Whitney SL. Advances in Vestibular Rehabilitation. *Adv Otorhinolaryngol*. 2019;82:164-169. doi:10.1159/000490285

73. McDonnell MN, Hillier SL. Vestibular rehabilitation for unilateral peripheral vestibular dysfunction. *Cochrane Database Syst Rev*. 2015;1(1):CD005397. doi:10.1002/14651858.CD005397.pub4

74. Porciuncula F, Johnson CC, Glickman LB. The effect of vestibular rehabilitation on adults with bilateral vestibular hypofunction: a systematic review. *J Vestib Res Equilib Orientat*. 2012;22(5-6):283-298. doi:10.3233/VES-120464

75. Tramontano M, Russo V, Spitoni GF, et al. Efficacy of Vestibular Rehabilitation in Patients With Neurologic Disorders: A Systematic Review. *Arch Phys Med Rehabil*. 2021;102(7):1379-1389. doi:10.1016/j.apmr.2020.11.017

76. García-Muñoz C, Cortés-Vega MD, Heredia-Rizo AM, Martín-Valero R, García-Bernal MI, Casuso-Holgado MJ. Effectiveness of Vestibular Training for Balance and Dizziness Rehabilitation in People with Multiple Sclerosis: A Systematic Review and Meta-Analysis. *J Clin Med*. 2020;9(2):590. doi:10.3390/jcm9020590

77. do Amaral CMS, de Almeida SB, de Almeida RP, do Nascimento SL, Ribeiro RM, Braga-Neto P. Effectiveness of vestibular rehabilitation on postural balance in Parkinson's disease: a systematic review and meta-analysis of randomized controlled trials. *BMC Neurol*. 2024;24(1):161. doi:10.1186/s12883-024-03649-5

78. Babula G, Warunek E, Cure K, Nikolski G, Fritz H, Barker S. Vestibular Rehabilitation as an Early Intervention in Athletes Who are Post-concussion: A Systematic Review. *Int J Sports Phys Ther*. 2023;V18(3):577-586. doi:10.26603/001c.75369

79. Tramontano M, Belluscio V, Bergamini E, et al. Vestibular Rehabilitation Improves Gait Quality and Activities of Daily Living in People with Severe Traumatic Brain Injury: A Randomized Clinical Trial. *Sensors*. 2022;22(21):8553. doi:10.3390/s22218553

80. Aljabri A, Halawani A, Ashqar A, Alageely O, Alhazzani A. The Efficacy of Vestibular Rehabilitation Therapy for Mild Traumatic Brain Injury: A Systematic Review and Meta-analysis. *J Head Trauma Rehabil*. 2024;39(2):E59-E69. doi:10.1097/HTR.0000000000000882

81. Tramontano M, Medici A, Iosa M, et al. The Effect of Vestibular Stimulation on Motor Functions of Children With Cerebral Palsy. *Motor Control*. 2017;21(3):299-311. doi:10.1123/mc.2015-0089

82. Meng L, Liang Q, Yuan J, et al. Vestibular rehabilitation therapy on balance and gait in patients after stroke: a systematic review and meta-analysis. *BMC Med*. 2023;21(1):322. doi:10.1186/s12916-023-03029-9

83. Feigin VL, Vos T, Nichols E, et al. The global burden of neurological disorders: translating evidence into policy. *Lancet Neurol*. 2020;19(3):255-265. doi:10.1016/S1474-4422(19)30411-9

84. Šarkić B, Douglas JM, Simpson A. A cross-sectional survey of non-specialist Australian audio-vestibular clinical practice for traumatic brain injury and rehabilitation. *Brain Impair Multidiscip J Aust Soc Study Brain Impair*. 2023;24(3):611-628. doi:10.1017/BrImp.2022.35

85. Alyahya D, Kashoo FZ. Perception, knowledge, and attitude of medical doctors in Saudi Arabia about the role of physiotherapists in vestibular rehabilitation: a cross-sectional survey. *PeerJ*. 2022;10:e13035. doi:10.7717/peerj.13035

86. Kumar K, Shaju AE, Premkumar PK, Ebenezer A, Roushan R. Practice Patterns of Audiologists Involved in Assessment and Management of Vestibular Disorders: An Indian Cross-Sectional Survey. *Int Tinnitus J*. 2022;25(2):137-142. doi:10.5935/0946-5448.20210025

87. Male AJ, Ramdharry GM, Grant R, Davies RA, Beith ID. A survey of current management of Benign Paroxysmal Positional Vertigo (BPPV) by physiotherapists' interested in vestibular rehabilitation in the UK. *Physiotherapy*. 2019;105(3):307-314. doi:10.1016/j.physio.2018.08.007

88. van Vugt VA, Diaz Nerio PM, van der Wouden JC, van der Horst HE, Maarsingh OR. Use of canalith repositioning manoeuvres and vestibular rehabilitation: a GP survey. *Scand J Prim Health Care*. 2017;35(1):19-26. doi:10.1080/02813432.2017.1288683

89. Meldrum D, Burrows L, Cakrt O, et al. Vestibular rehabilitation in Europe: a survey of clinical and research practice. *J Neurol*. 2020;267(Suppl 1):24-35. doi:10.1007/s00415-020-10228-4

90. Sharma A, Minh Duc NT, Luu Lam Thang T, et al. A Consensus-Based Checklist for Reporting of Survey Studies (CROSS). *J Gen Intern Med*. 2021;36(10):3179-3187. doi:10.1007/s11606-021-06737-1

91. Olesen J, Gustavsson A, Svensson M, et al. The economic cost of brain disorders in Europe. *Eur J Neurol*. 2012;19(1):155-162. doi:10.1111/j.1468-1331.2011.03590.x

92. Cawthorne T. Vestibular Injuries. *Proc R Soc Med*. 1946;39(5):270-273. doi:10.1177/003591574603900522

93. Cooksey FS. Rehabilitation in Vestibular Injuries. *Proc R Soc Med*. 1946;39(5):273-278. doi:10.1177/003591574603900523

94. Dunlap PM, Holmberg JM, Whitney SL. Vestibular rehabilitation: advances in peripheral and central vestibular disorders. *Curr Opin Neurol*. 2019;32(1):137-144. doi:10.1097/WCO.0000000000000632

95. Tramontano M, Casagrande Conti L, Orejel Bustos AS, et al. Abnormal Vestibulo-Ocular Reflex Function Correlates with Balance and Gait Impairment in People with Multiple Sclerosis. *Audiol Res*. 2024;14(5):799-808. doi:10.3390/audiolres14050067

96. Browne P, Chandraratna D, Angood C, et al. Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology*. 2014;83(11):1022-1024. doi:10.1212/WNL.0000000000000768

97. Walton C, King R, Rechtman L, et al. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Mult Scler Houndmills Basingstoke Engl*. 2020;26(14):1816-1821. doi:10.1177/1352458520970841

98. Young CA, Mills R, Rog D, et al. Quality of life in multiple sclerosis is dominated by fatigue, disability and self-efficacy. *J Neurol Sci*. 2021;426:117437. doi:10.1016/j.jns.2021.117437

99. Amatya B, Khan F, Galea M. Rehabilitation for people with multiple sclerosis: an overview of Cochrane Reviews. *Cochrane Database Syst Rev*. 2019;1(1):CD012732. doi:10.1002/14651858.CD012732.pub2

100. Cochrane GD, Christy JB, Motl RW. Central Vestibular Functions Correlate With Fatigue and Walking Capacity in People With Multiple Sclerosis. *Phys Ther*. 2021;101(9):pzab168. doi:10.1093/ptj/pzab168

101. Guo J, Wang J, Liang P, et al. Vestibular dysfunction leads to cognitive impairments: State of knowledge in the field and clinical perspectives (Review). *Int J Mol Med*. 2024;53(4):36. doi:10.3892/ijmm.2024.5360

102. Tramontano M, Martino Cinnera A, Manzari L, et al. Vestibular rehabilitation has positive effects on balance, fatigue and activities of daily living in highly disabled multiple sclerosis people: A preliminary randomized controlled trial. *Restor Neurol Neurosci*. 2018;36(6):709-718. doi:10.3233/RNN-180850

103. Tramontano M, Grasso MG, Soldi S, et al. Cerebellar Intermittent Theta-Burst Stimulation Combined with Vestibular Rehabilitation Improves Gait and Balance in Patients with Multiple Sclerosis: a Preliminary Double-Blind Randomized Controlled Trial. *Cerebellum Lond Engl*. 2020;19(6):897-901. doi:10.1007/s12311-020-01166-y

104. Weber KP, MacDougall HG, Halmagyi GM, Curthoys IS. Impulsive testing of semicircular-canal function using video-oculography. *Ann N Y Acad Sci*. 2009;1164:486-491. doi:10.1111/j.1749-6632.2008.03730.x

105. MacDougall HG, McGarvie LA, Halmagyi GM, et al. A new saccadic indicator of peripheral vestibular function based on the video head impulse test. *Neurology*. 2016;87(4):410-418. doi:10.1212/WNL.0000000000002827

106. Eğilmez OK, Tunç A, Yılmaz MS, et al. Cervical vestibular evoked myogenic potentials and video head impulse test studies: alternative methods for detecting brainstem involvement in multiple sclerosis. *Acta Otolaryngol (Stockh)*. 2022;142(2):168-174. doi:10.1080/00016489.2022.2039759

107. Pavlović I, Ruška B, Pavičić T, et al. Video head impulse test can detect brainstem dysfunction in multiple sclerosis. *Mult Scler Relat Disord*. 2017;14:68-71. doi:10.1016/j.msard.2017.04.001

108. Grove CR, Wagner A, Yang VB, Loyd BJ, Dibble LE, Schubert MC. Greater Disability Is Associated with Worse Vestibular and Compensatory Oculomotor Functions in People Living with Multiple Sclerosis. *Brain Sci*. 2022;12(11):1519. doi:10.3390/brainsci12111519

109. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol*. 2011;69(2):292-302. doi:10.1002/ana.22366

110. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983;33(11):1444-1452. doi:10.1212/wnl.33.11.1444

111. Cattaneo D, Regola A, Meotti M. Validity of six balance disorders scales in persons with multiple sclerosis. *Disabil Rehabil*. 2006;28(12):789-795. doi:10.1080/09638280500404289

112. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. *J Rehabil Med*. 2010;42(4):323-331. doi:10.2340/16501977-0537

113. Money-Nolan LE, Flagge AG. Factors affecting variability in vestibulo-ocular reflex gain in the Video Head Impulse Test in individuals without vestibulopathy: A systematic review of literature. *Front Neurol*. 2023;14:1125951. doi:10.3389/fneur.2023.1125951

114. Afrasiabifar A, Karami F, Najafi Doulatabad S. Comparing the effect of Cawthorne-Cooksey and Frenkel exercises on balance in patients with multiple sclerosis: a randomized controlled trial. *Clin Rehabil*. 2018;32(1):57-65. doi:10.1177/0269215517714592

115. Loyd BJ, Fangman A, Peterson DS, et al. Rehabilitation to Improve Gaze and Postural Stability in People With Multiple Sclerosis: A Randomized Clinical Trial. *Neurorehabil Neural Repair*. 2022;36(10-11):678-688. doi:10.1177/15459683221124126

116. Hebert JR, Corboy JR, Manago MM, Schenkman M. Effects of vestibular rehabilitation on multiple sclerosis-related fatigue and upright postural control: a randomized controlled trial. *Phys Ther*. 2011;91(8):1166-1183. doi:10.2522/ptj.20100399

117. Hebert JR, Corboy JR, Vollmer T, Forster JE, Schenkman M. Efficacy of Balance and Eye-Movement Exercises for Persons With Multiple Sclerosis (BEEMS). *Neurology*. 2018;90(9):e797-e807. doi:10.1212/WNL.0000000000005013

118. Ozgen G, Karapolat H, Akkoc Y, Yuceyar N. Is customized vestibular rehabilitation effective in patients with multiple sclerosis? A randomized controlled trial. *Eur J Phys Rehabil Med*. 2016;52(4):466-478.

119. Manzari L, Tramontano M. Identifying the Clinical Signs on the Healthy Side Using Video Head Impulse Test During Different Stages of Vestibular Neuritis. *J Int Adv Otol*. 2021;17(5):433-437. doi:10.5152/iao.2021.21155

120. Gür E, Binkhamis G, Kluk K. Effects of multiple sclerosis on the audio-vestibular system: a systematic review. *BMJ Open*. 2022;12(8):e060540. doi:10.1136/bmjopen-2021-060540

121. Choi JY, Kim HJ, Kim JS. Recent advances in head impulse test findings in central vestibular disorders. *Neurology*. 2018;90(13):602-612. doi:10.1212/WNL.0000000000005206

122. Jeong SH, Kim JS, Baek IC, et al. Perverted head impulse test in cerebellar ataxia. *Cerebellum Lond Engl*. 2013;12(5):773-775. doi:10.1007/s12311-013-0480-0

123. Walker MF, Zee DS. Cerebellar disease alters the axis of the high-acceleration vestibuloocular reflex. *J Neurophysiol*. 2005;94(5):3417-3429. doi:10.1152/jn.00375.2005

124. Tramontano M, Ferri N, Turolla A, et al. Video head impulse test in subacute and chronic stroke survivors: new perspectives for implementation of assessment in rehabilitation. *Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol - Head Neck Surg*. 2024;281(10):5129-5134. doi:10.1007/s00405-024-08721-x

125. MacDougall HG, McGarvie LA, Halmagyi GM, Curthoys IS, Weber KP. Application of the video head impulse test to detect vertical semicircular canal dysfunction. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol*. 2013;34(6):974-979. doi:10.1097/MAO.0b013e31828d676d

126. Manzari L, Orejel Bustos AS, Princi AA, Tramontano M. Video Suppression Head Impulses and Head Impulses Paradigms in Patients with Vestibular Neuritis: A Comparative Study. *Healthc Basel Switz*. 2022;10(10):1926. doi:10.3390/healthcare10101926

127. MacDougall HG, Weber KP, McGarvie LA, Halmagyi GM, Curthoys IS. The video head impulse test: diagnostic accuracy in peripheral vestibulopathy. *Neurology*. 2009;73(14):1134-1141. doi:10.1212/WNL.0b013e3181bacf85

128. Ha SH, Lee DK, Park G, et al. Prospective analysis of video head impulse tests in patients with acute posterior circulation stroke. *Front Neurol*. 2023;14:1256826. doi:10.3389/fneur.2023.1256826

129. Nham B, Wang C, Reid N, et al. Modern vestibular tests can accurately separate stroke and vestibular neuritis. *J Neurol*. 2023;270(4):2031-2041. doi:10.1007/s00415-022-11473-5

130. Berkiten G, Tutar B, Atar S, et al. Assessment of the Clinical Use of Vestibular Evoked Myogenic Potentials and the Video Head Impulse Test in the Diagnosis of Early-Stage Parkinson's Disease. *Ann Otol Rhinol Laryngol*. 2023;132(1):41-49. doi:10.1177/00034894211067838

131. Heravian Shandiz J, Jafarzadeh S, Fathi H, Foroughipour M, Karimpour M. Vestibulo ocular reflex in multiple sclerosis patients without any optic neuritis. *J Optom*. 2021;14(3):282-286. doi:10.1016/j.optom.2020.07.001

132. Machner B, Erber K, Choi JH, Trillenberg P, Sprenger A, Helmchen C. Usability of the head impulse test in routine clinical practice in the emergency department to differentiate vestibular neuritis from stroke. *Eur J Neurol*. 2021;28(5):1737-1744. doi:10.1111/ene.14707

133. Morrison M, Korda A, Zamaro E, et al. Paradigm shift in acute dizziness: is caloric testing obsolete? *J Neurol*. 2022;269(2):853-860. doi:10.1007/s00415-021-10667-7

134. Sayed SZ, Abdul Wahat NH, Raymond AA, Hussein N, Wan Asyraf WZ, Omar M. Quantitative vestibular function tests in posterior circulation stroke patients: A review. *Med J Malaysia*. 2021;76(6):898-905.

135. Viosca E, Martínez JL, Almagro PL, Gracia A, González C. Proposal and validation of a new functional ambulation classification scale for clinical use. *Arch Phys Med Rehabil*. 2005;86(6):1234-1238. doi:10.1016/j.apmr.2004.11.016

136. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6

137. Curthoys IS, Manzari L, Rey-Martinez J, Dlugaiczyk J, Burgess AM. Enhanced Eye Velocity in Head Impulse Testing-A Possible Indicator of Endolymphatic Hydrops. *Front Surg*. 2021;8:666390. doi:10.3389/fsurg.2021.666390

138. Zamaro E, Saber Tehrani AS, Kattah JC, et al. VOR gain calculation methods in video head impulse recordings. *J Vestib Res Equilib Orientat*. 2020;30(4):225-234. doi:10.3233/VES-200708

139. Ferri N, Bracci A, Metta Franceschelli G, et al. Accuracy of video head impulse test in stroke diagnosis during acute vestibular syndrome: a systematic review with meta-analysis. *J Neurol*. 2025;272(6):436. doi:10.1007/s00415-025-13174-1

140. Manzari L, De Angelis S, Princi AA, Galeoto G, Tramontano M. The Clinical Use of the Suppression Head Impulse Paradigm in Patients with Vestibulopathy: A Systematic Review. *Healthc Basel Switz*. 2022;10(7):1182. doi:10.3390/healthcare10071182

141. Bonnì S, Ponzo V, Tramontano M, et al. Neurophysiological and clinical effects of blindfolded balance training (BBT) in Parkinson's disease patients: a preliminary study. *Eur J Phys Rehabil Med*. 2019;55(2):176-182. doi:10.23736/S1973-9087.18.05126-2

142. Tramontano M, Piermaria J, Morone G, Reali A, Vergara M, Tamburella F. Postural Changes During Exteroceptive Thin Plantar Stimulation: The Effect of Prolonged Use and Different Plantar Localizations. *Front Syst Neurosci*. 2019;13:49. doi:10.3389/fnsys.2019.00049

143. Bent LR, McFadyen BJ, Inglis JT. Vestibular contributions during human locomotor tasks. *Exerc Sport Sci Rev*. 2005;33(3):107-113. doi:10.1097/00003677-200507000-00002

144. Casagrande Conti L, Ferri N, Manzari L, et al. Vestibulo-Oculomotor Reflex Dysfunction in Children with Cerebral Palsy Correlates with Gross Motor Function Classification System. *Audiol Res*. 2025;15(2):21. doi:10.3390/audiolres15020021

145. Rosenbaum P, Paneth N, Leviton A, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl*. 2007;109:8-14.

146. Bax MCO, Flodmark O, Tydeman C. Definition and classification of cerebral palsy. From syndrome toward disease. *Dev Med Child Neurol Suppl*. 2007;109:39-41.

147. Anaby D, Hand C, Bradley L, et al. The effect of the environment on participation of children and youth with disabilities: a scoping review. *Disabil Rehabil*. 2013;35(19):1589-1598. doi:10.3109/09638288.2012.748840

148. Bedell G, Coster W, Law M, et al. Community participation, supports, and barriers of school-age children with and without disabilities. *Arch Phys Med Rehabil*. 2013;94(2):315-323. doi:10.1016/j.apmr.2012.09.024

149. Cullen KE. Vestibular motor control. *Handb Clin Neurol*. 2023;195:31-54. doi:10.1016/B978-0-323-98818-6.00022-4

150. Biju K, Oh E, Rosenberg P, et al. Vestibular Function Predicts Balance and Fall Risk in Patients with Alzheimer's Disease. *J Alzheimers Dis JAD*. 2022;86(3):1159-1168. doi:10.3233/JAD-215366

151. Choi SY, Choi KD, Choi JH, Kim JS. Abnormal vestibular-evoked myogenic potentials as a risk factor for unpredicted falls in spinocerebellar ataxia: a preliminary study. *J Neurol*. 2024;271(5):2539-2546. doi:10.1007/s00415-024-12195-6

152. Politi L, Salerni L, Bubbico L, et al. Risk of falls, vestibular multimodal processing, and multisensory integration decline in the elderly-Predictive role of the functional head impulse test. *Front Neurol*. 2022;13:964017. doi:10.3389/fneur.2022.964017

153. Fazzi E, Signorini SG, LA Piana R, et al. Neuro-ophthalmological disorders in cerebral palsy: ophthalmological, oculomotor, and visual aspects. *Dev Med Child Neurol*. 2012;54(8):730-736. doi:10.1111/j.1469-8749.2012.04324.x

154. Galli J, Loi E, Calza S, et al. Natural history of cerebral visual impairment in children with cerebral palsy. *Dev Med Child Neurol*. 2025;67(4):486-495. doi:10.1111/dmcn.16096

155. Tinelli F, Guzzetta A, Purpura G, Pasquariello R, Cioni G, Fiori S. Structural brain damage and visual disorders in children with cerebral palsy due to periventricular leukomalacia. *NeuroImage Clin*. 2020;28:102430. doi:10.1016/j.nicl.2020.102430

156. Longridge NS, Mallinson AI. A New Perspective to Interpret How the Vestibular Efferent System Correlates the Complexity of Routine Balance Maintenance with Management of Emergency Fall Prevention Strategies. *Audiol Res*. 2024;14(3):518-544. doi:10.3390/audiolres14030044

157. Duke RE, Nwachukuw J, Torty C, et al. Visual impairment and perceptual visual disorders in children with cerebral palsy in Nigeria. *Br J Ophthalmol*. 2022;106(3):427-434. doi:10.1136/bjophthalmol-2020-317768

158. Dasgupta S, Mukherjee AL, Crunkhorn R, et al. Video Head Impulse Test in Children-A Systematic Review of Literature. *J Clin Med*. 2025;14(2):369. doi:10.3390/jcm14020369

159. Janky KL, Patterson J, Thomas M, Al-Salim S, Robinson S. The effects of vestibular dysfunction on balance and self-concept in children with cochlear implants. *Int J Pediatr Otorhinolaryngol*. 2023;171:111642. doi:10.1016/j.ijporl.2023.111642

160. Taylor RL, Wise KJ, Taylor D, Chaudhary S, Thorne PR. Patterns of vestibular dysfunction in chronic traumatic brain injury. *Front Neurol*. 2022;13:942349. doi:10.3389/fneur.2022.942349

161. Tofani M, Candeloro C, Sabbadini M, et al. A study validating the Italian version of the Level of Sitting Scale in children with cerebral palsy. *Clin Rehabil*. 2019;33(11):1810-1818. doi:10.1177/0269215519858387

162. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997;39(4):214-223. doi:10.1111/j.1469-8749.1997.tb07414.x

163. Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH. Content validity of the expanded and revised Gross Motor Function Classification System. *Dev Med Child Neurol*. 2008;50(10):744-750. doi:10.1111/j.1469-8749.2008.03089.x

164. Manzari L, Tramontano M. Suppression Head Impulse Paradigm (SHIMP) in evaluating the vestibulo-saccadic interaction in patients with vestibular neuritis. *Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol - Head Neck Surg*. 2020;277(11):3205-3212. doi:10.1007/s00405-020-06085-6

165. Shen Q, Magnani C, Sterkers O, et al. Saccadic Velocity in the New Suppression Head Impulse Test: A New Indicator of Horizontal Vestibular Canal Paresis and of Vestibular Compensation. *Front Neurol*. 2016;7:160. doi:10.3389/fneur.2016.00160

166. Ghai S, Hakim M, Dannenbaum E, Lamontagne A. Prevalence of Vestibular Dysfunction in Children With Neurological Disabilities: A Systematic Review. *Front Neurol*. 2019;10:1294. doi:10.3389/fneur.2019.01294

167. Almutairi A, Cochrane GD, Christy JB. Vestibular and oculomotor function in children with CP: Descriptive study. *Int J Pediatr Otorhinolaryngol*. 2019;119:15-21. doi:10.1016/j.ijporl.2018.12.038

168. Almutairi A, Christy JB, Vogtle L. Vestibular and Oculomotor Function in Children with Cerebral Palsy: A Scoping Review. *Semin Hear*. 2018;39(3):288-304. doi:10.1055/s-0038-1666819

169. Zhou G, Yun A, Wang A, Brodsky JR. Comparing Video Head Impulse Testing With Rotary Chair in Pediatric Patients: A Controlled Trial. *Otolaryngol--Head Neck Surg Off J Am Acad Otolaryngol-Head Neck Surg*. 2024;171(4):1190-1196. doi:10.1002/ohn.820

170. Lotfi Y, Rezazadeh N, Moossavi A, et al. Preliminary evidence of improved cognitive performance following vestibular rehabilitation in children with combined ADHD (cADHD) and concurrent vestibular impairment. *Auris Nasus Larynx*. 2017;44(6):700-707. doi:10.1016/j.anl.2017.01.011

171. Melo RS, Lemos A, Paiva GS, et al. Vestibular rehabilitation exercises programs to improve the postural control, balance and gait of children with sensorineural hearing loss: A systematic review. *Int J Pediatr Otorhinolaryngol*. 2019;127:109650. doi:10.1016/j.ijporl.2019.109650

172. Mohamed ST, Hazzaa N, Abdel Rahman T, Ezz Eldin DM, Elhusseiny AM. Efficacy of vestibular rehabilitation program in children with balance disorders and sensorineural hearing loss. *Int J Pediatr Otorhinolaryngol*. 2024;179:111931. doi:10.1016/j.ijporl.2024.111931

173. Ferri N, Whitney SL, Verrecchia L, et al. Video Head Impulse Test in Survivors From Severe Traumatic Brain Injury: New Perspectives for Implementation of Assessment in Rehabilitation. *J Head Trauma Rehabil*. 2025;40(5):337-345. doi:10.1097/HTR.0000000000001044

174. Curtiss G, Klemz S, Vanderploeg RD. Acute impact of severe traumatic brain injury on family structure and coping responses. *J Head Trauma Rehabil*. 2000;15(5):1113-1122. doi:10.1097/00001199-200010000-00005

175. Guan B, Anderson DB, Chen L, Feng S, Zhou H. Global, regional and national burden of traumatic brain injury and spinal cord injury, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *BMJ Open*. 2023;13(10):e075049. doi:10.1136/bmjopen-2023-075049

176. DeGraba TJ, Williams K, Koffman R, et al. Efficacy of an Interdisciplinary Intensive Outpatient Program in Treating Combat-Related Traumatic Brain Injury and Psychological Health Conditions. *Front Neurol*. 2020;11:580182. doi:10.3389/fneur.2020.580182

177. Wallace TD, McCauley KL, Hodge AT, et al. Use of person-centered goals to direct interdisciplinary care for military service members and Veterans with chronic mTBI and co-occurring psychological conditions. *Front Neurol*. 2022;13:1015591. doi:10.3389/fneur.2022.1015591

178. Maskell F, Chiarelli P, Isles R. Dizziness after traumatic brain injury: overview and measurement in the clinical setting. *Brain Inj*. 2006;20(3):293-305. doi:10.1080/02699050500488041

179. Murphy MP, Carmine H, Kolakowsky-Hayner S. Modifiable and nonmodifiable risk factors for falls after traumatic brain injury: an exploratory investigation with implications for medication use. *Rehabil Nurs Off J Assoc Rehabil Nurses*. 2014;39(3):113-122. doi:10.1002/rnj.89

180. Marcus HJ, Paine H, Sargeant M, et al. Vestibular dysfunction in acute traumatic brain injury. *J Neurol*. 2019;266(10):2430-2433. doi:10.1007/s00415-019-09403-z

181. Harrell R, Manetta C, Guthrie M, Enam N. The Prevalence of Symptom Reporting for Benign Paroxysmal Positional Vertigo in a Traumatic Brain Injury Population. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol*. 2023;44(2):172-176. doi:10.1097/MAO.0000000000003770

182. Campbell KR, King LA, Parrington L, Fino PC, Antonellis P, Peterka RJ. Central sensorimotor integration assessment reveals deficits in standing balance control in people with chronic mild traumatic brain injury. *Front Neurol*. 2022;13:897454. doi:10.3389/fneur.2022.897454

183. Hadi Z, Mahmud M, Seemungal BM. Brain Mechanisms Explaining Postural Imbalance in Traumatic Brain Injury: A Systematic Review. *Brain Connect*. 2024;14(3):144-177. doi:10.1089/brain.2023.0064

184. Smith RM, Burgess C, Tahtis V, Marsden J, Seemungal BM. Why are patients with acute traumatic brain injury not routinely assessed or treated for vestibular dysfunction in the UK? A qualitative study. *BMJ Open*. 2023;13(1):e067967. doi:10.1136/bmjopen-2022-067967

185. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet Lond Engl*. 1975;1(7905):480-484. doi:10.1016/s0140-6736(75)92830-5

186. Gouvier WD, Blanton PD, LaPorte KK, Nepomuceno C. Reliability and validity of the Disability Rating Scale and the Levels of Cognitive Functioning Scale in monitoring recovery from severe head injury. *Arch Phys Med Rehabil*. 1987;68(2):94-97.

187. Feld JA, Rabadi MH, Blau AD, Jordan BD. Berg balance scale and outcome measures in acquired brain injury. *Neurorehabil Neural Repair*. 2001;15(3):239-244. doi:10.1177/154596830101500312

188. Newstead AH, Hinman MR, Tomberlin JA. Reliability of the Berg Balance Scale and balance master limits of stability tests for individuals with brain injury. *J Neurol Phys Ther JNPT*. 2005;29(1):18-23. doi:10.1097/01.npt.0000282258.74325.cf

189. Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. *Phys Ther*. 2009;89(5):484-498. doi:10.2522/ptj.20080071

190. Calzolari E, Chepisheva M, Smith RM, et al. Vestibular agnosia in traumatic brain injury and its link to imbalance. *Brain J Neurol*. 2021;144(1):128-143. doi:10.1093/brain/awaa386

191. Wallace B, Lifshitz J. Traumatic brain injury and vestibulo-ocular function: current challenges and future prospects. *Eye Brain*. 2016;8:153-164. doi:10.2147/EB.S82670

192. Crampton A, Garat A, Shepherd HA, et al. Evaluating the Vestibulo-Ocular Reflex Following Traumatic Brain Injury: A Scoping Review. *Brain Inj*. 2021;35(12-13):1496-1509. doi:10.1080/02699052.2021.1972450

193. Ferri N, Casagrande Conti L, Manzari L, et al. Assessment of vestibulo-ocular reflex function in people with Parkinson's disease: a cross-sectional study in a rehabilitation setting using the video head impulse test. *BMJ Open*. 2025;15(8):e099765. doi:10.1136/bmjopen-2025-099765

194. Kalia LV, Lang AE. Parkinson's disease. *Lancet Lond Engl*. 2015;386(9996):896-912. doi:10.1016/S0140-6736(14)61393-3

195. Gui M, Lv L, Qin L, Wang C. Vestibular dysfunction in Parkinson's disease: a neglected topic. *Front Neurol*. 2024;15:1398764. doi:10.3389/fneur.2024.1398764

196. Wang L, Zobeiri OA, Millar JL, Schubert MC, Cullen KE. Head movement kinematics are altered during gaze stability exercises in vestibular schwannoma patients. *Sci Rep*. 2021;11(1):7139. doi:10.1038/s41598-021-86533-3

197. Lv W, Guan Q, Hu X, et al. Vestibulo-ocular reflex abnormality in Parkinson's disease detected by video head impulse test. *Neurosci Lett*. 2017;657:211-214. doi:10.1016/j.neulet.2017.08.021

198. Smith PF. Vestibular Functions and Parkinson's Disease. *Front Neurol*. 2018;9:1085. doi:10.3389/fneur.2018.01085

199. Macdougall HG, McGarvie LA, Halmagyi GM, Curthoys IS, Weber KP. The video Head Impulse Test (vHIT) detects vertical semicircular canal dysfunction. *PloS One*. 2013;8(4):e61488. doi:10.1371/journal.pone.0061488

200. Hawkins KE, Chiarovano E, Paul SS, Burgess AM, MacDougall HG, Curthoys IS. Vestibular semicircular canal function as detected by video Head Impulse Test (vHIT) is essentially unchanged in people with Parkinson's disease compared to healthy controls. *J Vestib Res Equilib Orientat*. 2022;32(3):261-269. doi:10.3233/VES-201626

201. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology*. 1967;17(5):427-442. doi:10.1212/wnl.17.5.427

202. Wade DT. Measurement in neurological rehabilitation. *Curr Opin Neurol Neurosurg*. 1992;5(5):682-686.

203. Hawkins KE, Rey-Martinez J, Chiarovano E, et al. Suppression head impulse test paradigm (SHIMP) characteristics in people with Parkinson's disease compared to healthy controls. *Exp Brain Res*. 2021;239(6):1853-1862. doi:10.1007/s00221-021-06107-7

204. Hawkins KE, Paul SS, Chiarovano E, Curthoys IS. Using virtual reality to assess vestibulo-visual interaction in people with Parkinson's disease compared to healthy controls. *Exp Brain Res*. 2021;239(12):3553-3564. doi:10.1007/s00221-021-06219-0

205. Manzari L, Ferri N, Tramontano M. Revisiting Diagnostic Criteria for Bilateral Vestibulopathy: A New Comprehensive Instrumental Model. *Audiol Res*. 2024;14(6):991-999. doi:10.3390/audiolres14060082

206. Strupp M, Kim JS, Murofushi T, et al. Bilateral vestibulopathy: Diagnostic criteria Consensus document of the Classification Committee of the Bárány Society. *J Vestib Res Equilib Orientat*. 2017;27(4):177-189. doi:10.3233/VES-170619

207. Hain TC, Cherchi M, Yacovino DA. Bilateral vestibular loss. *Semin Neurol*. 2013;33(3):195-203. doi:10.1055/s-0033-1354597

208. Herssens N, Verbecque E, McCrum C, et al. A Systematic Review on Balance Performance in Patients With Bilateral Vestibulopathy. *Phys Ther*. 2020;100(9):1582-1594. doi:10.1093/ptj/pzaa083

209. Lucieer F, Duijn S, Van Rompaey V, et al. Full Spectrum of Reported Symptoms of Bilateral Vestibulopathy Needs Further Investigation-A Systematic Review. *Front Neurol*. 2018;9:352. doi:10.3389/fneur.2018.00352

210. Kremmyda O, Hüfner K, Flanagin VL, et al. Beyond Dizziness: Virtual Navigation, Spatial Anxiety and Hippocampal Volume in Bilateral Vestibulopathy. *Front Hum Neurosci*. 2016;10:139. doi:10.3389/fnhum.2016.00139

211. Brandt T, Schautzer F, Hamilton DA, et al. Vestibular loss causes hippocampal atrophy and impaired spatial memory in humans. *Brain J Neurol*. 2005;128(Pt 11):2732-2741. doi:10.1093/brain/awh617

212. Ward BK, Agrawal Y, Hoffman HJ, Carey JP, Della Santina CC. Prevalence and impact of bilateral vestibular hypofunction: results from the 2008 US National Health Interview Survey. *JAMA Otolaryngol--Head Neck Surg*. 2013;139(8):803-810. doi:10.1001/jamaoto.2013.3913

213. Kim S, Oh YM, Koo JW, Kim JS. Bilateral vestibulopathy: clinical characteristics and diagnostic criteria. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol*. 2011;32(5):812-817. doi:10.1097/MAO.0b013e31821a3b7d

214. Rinne T, Bronstein AM, Rudge P, Gresty MA, Luxon LM. Bilateral loss of vestibular function: clinical findings in 53 patients. *J Neurol*. 1998;245(6-7):314-321. doi:10.1007/s004150050225

215. Guinand N, Boselie F, Guyot JP, Kingma H. Quality of life of patients with bilateral vestibulopathy. *Ann Otol Rhinol Laryngol*. 2012;121(7):471-477. doi:10.1177/000348941212100708

216. Fujimoto C, Murofushi T, Chihara Y, Suzuki M, Yamasoba T, Iwasaki S. Novel subtype of idiopathic bilateral vestibulopathy: bilateral absence of vestibular evoked myogenic potentials in the presence of normal caloric responses. *J Neurol*. 2009;256(9):1488-1492. doi:10.1007/s00415-009-5147-x

217. Manzari L, Princi AA, De Angelis S, Tramontano M. Clinical value of the video head impulse test in patients with vestibular neuritis: a systematic review. *Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol - Head Neck Surg*. 2021;278(11):4155-4167. doi:10.1007/s00405-021-06803-8

218. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi:10.1136/bmj.h5527

219. Jacobson GP, Newman CW. The development of the Dizziness Handicap Inventory. *Arch Otolaryngol Head Neck Surg*. 1990;116(4):424-427. doi:10.1001/archotol.1990.01870040046011

220. Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet Lond Engl*. 1991;337(8756):1521-1526. doi:10.1016/0140-6736(91)93206-o

221. Banerjee G, Stone SP, Werring DJ. Posterior circulation ischaemic stroke. *BMJ*. 2018;361:k1185. doi:10.1136/bmj.k1185

222. Schulz UG, Fischer U. Posterior circulation cerebrovascular syndromes: diagnosis and management. *J Neurol Neurosurg Psychiatry*. 2017;88(1):45-53. doi:10.1136/jnnp-2015-311299

223. Tarnutzer AA, Lee SH, Robinson KA, Wang Z, Edlow JA, Newman-Toker DE. ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: A meta-analysis. *Neurology*. 2017;88(15):1468-1477. doi:10.1212/WNL.0000000000003814

224. Searls DE, Pazdera L, Korbel E, Vysata O, Caplan LR. Symptoms and signs of posterior circulation ischemia in the new England medical center posterior circulation registry. *Arch Neurol*. 2012;69(3):346-351. doi:10.1001/archneurol.2011.2083

225. Kattah JC, Talkad AV, Wang DZ, Hsieh YH, Newman-Toker DE. HINTS to diagnose stroke in the acute vestibular syndrome: three-step bedside oculomotor examination more sensitive than early MRI diffusion-weighted imaging. *Stroke*. 2009;40(11):3504-3510. doi:10.1161/STROKEAHA.109.551234

226. Lee SH, Kim JS. Acute Diagnosis and Management of Stroke Presenting Dizziness or Vertigo. *Neurol Clin*. 2015;33(3):687-698, xi. doi:10.1016/j.ncl.2015.04.006

227. Machner B, Choi JH, Trillenberg P, Heide W, Helmchen C. Risk of acute brain lesions in dizzy patients presenting to the emergency room: who needs imaging and who does not? *J Neurol*. 2020;267(Suppl 1):126-135. doi:10.1007/s00415-020-09909-x

228. Tarnutzer AA, Berkowitz AL, Robinson KA, Hsieh YH, Newman-Toker DE. Does my dizzy patient have a stroke? A systematic review of bedside diagnosis in acute vestibular syndrome. *CMAJ Can Med Assoc J J Assoc Medicale Can*. 2011;183(9):E571-592. doi:10.1503/cmaj.100174

229. Zwergal A, Dieterich M. Vertigo and dizziness in the emergency room. *Curr Opin Neurol*. 2020;33(1):117-125. doi:10.1097/WCO.0000000000000769

230. Kerber KA, Brown DL, Lisabeth LD, Smith MA, Morgenstern LB. Stroke among patients with dizziness, vertigo, and imbalance in the emergency department: a population-based study. *Stroke*. 2006;37(10):2484-2487. doi:10.1161/01.STR.0000240329.48263.0d

231. Newman-Toker DE, Hsieh YH, Camargo CA, Pelletier AJ, Butchy GT, Edlow JA. Spectrum of dizziness visits to US emergency departments: cross-sectional analysis from a nationally representative sample. *Mayo Clin Proc*. 2008;83(7):765-775. doi:10.4065/83.7.765

232. Edlow JA, Carpenter C, Akhter M, et al. Guidelines for reasonable and appropriate care in the emergency department 3 (GRACE-3): Acute dizziness and vertigo in the emergency department. *Acad Emerg Med Off J Soc Acad Emerg Med*. 2023;30(5):442-486. doi:10.1111/acem.14728

233. Shah VP, Oliveira J E Silva L, Farah W, et al. Diagnostic accuracy of the physical examination in emergency department patients with acute vertigo or dizziness: A systematic review and meta-analysis for GRACE-3. *Acad Emerg Med Off J Soc Acad Emerg Med*. 2023;30(5):552-578. doi:10.1111/acem.14630

234. Halmagyi GM, Curthoys IS. A clinical sign of canal paresis. *Arch Neurol*. 1988;45(7):737-739. doi:10.1001/archneur.1988.00520310043015

235. Jaganathan N, Mohamed MH, Md Pauzi AL, et al. Video head impulse test in stroke: a review of published studies. *Front Neurol*. 2024;15:1339039. doi:10.3389/fneur.2024.1339039

236. Deeks J, Bossuyt P, Leeflang M, Takwoingi Y. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. 1st edition. John Wiley & Sons; 2023.

237. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA*. 2018;319(4):388-396. doi:10.1001/jama.2017.19163

238. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol*. 2016;75:40-46. doi:10.1016/j.jclinepi.2016.01.021

239. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. doi:10.7326/0003-4819-155-8-201110180-00009

240. Guler A, Karbek Akarca F, Eraslan C, et al. Clinical and video head impulse test in the diagnosis of posterior circulation stroke presenting as acute vestibular syndrome in the emergency department. *J Vestib Res Equilib Orientat*. 2017;27(4):233-242. doi:10.3233/VES-170620

241. Korda A, Wimmer W, Zamaro E, et al. Videooculography “HINTS” in Acute Vestibular Syndrome: A Prospective Study. *Front Neurol*. 2022;13:920357. doi:10.3389/fneur.2022.920357

242. Korda A, Wimmer W, Wyss T, et al. Artificial intelligence for early stroke diagnosis in acute vestibular syndrome. *Front Neurol*. 2022;13:919777. doi:10.3389/fneur.2022.919777

243. Mantokoudis G, Saber Tehrani AS, Wozniak A, et al. Impact of artifacts on VOR gain measures by video-oculography in the acute vestibular syndrome. *J Vestib Res Equilib Orientat*. 2016;26(4):375-385. doi:10.3233/VES-160587

244. Mantokoudis G, Tehrani ASS, Wozniak A, et al. VOR gain by head impulse video-oculography differentiates acute vestibular neuritis from stroke. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol*. 2015;36(3):457-465. doi:10.1097/MAO.0000000000000638

245. Siepmann T, Gruener C, Simon E, et al. Video-Oculography-Assisted Head Impulse Test and Caloric Testing for Detecting Stroke in Acute Vertigo Patients via Modified HINTS Plus. *J Clin Med*. 2021;10(19):4471. doi:10.3390/jcm10194471

246. Michailidou E, Korda A, Wyss T, et al. The value of saccade metrics and VOR gain in detecting a vestibular stroke. *J Vestib Res Equilib Orientat*. 2024;34(1):49-61. doi:10.3233/VES-230083

247. Nham B, Reid N, Bein K, et al. Capturing vertigo in the emergency room: three tools to double the rate of diagnosis. *J Neurol*. 2022;269(1):294-306. doi:10.1007/s00415-021-10627-1

248. Nam GS, Shin HJ, Kang JJ, Lee NR, Oh SY. Clinical Implication of Corrective Saccades in the Video Head Impulse Test for the Diagnosis of Posterior Inferior Cerebellar Artery Infarction. *Front Neurol*. 2021;12:605040. doi:10.3389/fneur.2021.605040

249. Thomas JO, Sharobeam A, Venkat A, et al. Video head impulse testing to differentiate vestibular neuritis from posterior circulation stroke in the emergency department: a prospective observational study. *BMJ Neurol Open*. 2022;4(1):e000284. doi:10.1136/bmjno-2022-000284

250. Gottlieb M, Peksa GD, Carlson JN. Head impulse, nystagmus, and test of skew examination for diagnosing central causes of acute vestibular syndrome. *Cochrane Database Syst Rev*. 2023;11(11):CD015089. doi:10.1002/14651858.CD015089.pub2

251. Lawn JE, Ohuma EO, Bradley E, et al. Small babies, big risks: global estimates of prevalence and mortality for vulnerable newborns to accelerate change and improve counting. *Lancet Lond Engl*. 2023;401(10389):1707-1719. doi:10.1016/S0140-6736(23)00522-6

252. Roze E, Reijneveld SA, Stewart RE, Bos AF. Multi-domain cognitive impairments at school age in very preterm-born children compared to term-born peers. *BMC Pediatr*. 2021;21(1):169. doi:10.1186/s12887-021-02641-z

253. Zhang X, Spear E, Hsu HHL, Gennings C, Stroustrup A. NICU-based stress response and preterm infant neurobehavior: exploring the critical windows for exposure. *Pediatr Res*. 2022;92(5):1470-1478. doi:10.1038/s41390-022-01983-3

254. Rodovanski GP, Réus BAB, Neves Dos Santos A. The effects of multisensory stimulation on the length of hospital stay and weight gain in hospitalized preterm infants: A systematic review with meta-analysis. *Braz J Phys Ther*. 2023;27(1):100468. doi:10.1016/j.bjpt.2022.100468

255. Øberg GK, Sørvoll M, Labori C, Girolami GL, Håkstad RB. A systematic synthesis of qualitative studies on parents' experiences of participating in early intervention programs with their infant born preterm. *Front Psychol*. 2023;14:1172578. doi:10.3389/fpsyg.2023.1172578

256. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: Updated guideline for protocols of randomised trials. *PLoS Med*. 2025;22(4):e1004589. doi:10.1371/journal.pmed.1004589

257. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *Lancet Lond Engl*. Published online April 14, 2025:S0140-6736(25)00672-5. doi:10.1016/S0140-6736(25)00672-5

258. Conde-Agudelo A, Díaz-Rossello JL. Kangaroo mother care to reduce morbidity and mortality in low birthweight infants. *Cochrane Database Syst Rev*. 2016;2016(8):CD002771. doi:10.1002/14651858.CD002771.pub4

259. Vanderveen JA, Bassler D, Robertson CMT, Kirpalani H. Early interventions involving parents to improve neurodevelopmental outcomes of premature infants: a meta-analysis. *J Perinatol Off J Calif Perinat Assoc*. 2009;29(5):343-351. doi:10.1038/jp.2008.229

260. Burdjalov VF, Baumgart S, Spitzer AR. Cerebral function monitoring: a new scoring system for the evaluation of brain maturation in neonates. *Pediatrics*. 2003;112(4):855-861. doi:10.1542/peds.112.4.855

261. Montirosso R, Provenzi L, Calciolari G, Borgatti R, NEO-ACQUA Study Group. Measuring maternal stress and perceived support in 25 Italian NICUs. *Acta Paediatr Oslo Nor 1992*. 2012;101(2):136-142. doi:10.1111/j.1651-2227.2011.02440.x

262. Lavizzari A, Falgari R, Pesenti N, et al. Heart-rate agreement between ECG and a new, wireless device during early skin-to-skin contact. *Acta Paediatr Oslo Nor 1992*. 2021;110(6):1803-1809. doi:10.1111/apa.15769