UHASSELT

KNOWLEDGE IN ACTION

Faculteit Revalidatiewetenschappen

master in de revalidatiewetenschappen en de kinesitherapie

Masterthesis

Differences in sensory integration on balance between older fallers and non-fallers, a systematic review and meta-analysis

Quinten De Marie
Dries Van Thielen

Scriptie ingediend tot het behalen van de graad van master in de revalidatiewetenschappen en de kinesitherapie, afstudeerrichting revalidatiewetenschappen en kinesitherapie bij musculoskeletale aandoeningen

PROMOTOR :

Prof. dr. Pieter MEYNS

BEGELEIDER :

Mevrouw Esma Nur KOLBASI DOGAN

UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

2024
2025

UHASSELT

KNOWLEDGE IN ACTION

Faculteit Revalidatiewetenschappen

master in de revalidatiewetenschappen en de kinesitherapie

Masterthesis

Differences in sensory integration on balance between older fallers and non-fallers, a systematic review and meta-analysis

Quinten De Marie
Dries Van Thielen

Scriptie ingediend tot het behalen van de graad van master in de revalidatiewetenschappen en de kinesitherapie, afstudeerrichting revalidatiewetenschappen en kinesitherapie bij musculoskeletale aandoeningen

PROMOTOR :

Prof. dr. Pieter MEYNS

BEGELEIDER :

Mevrouw Esma Nur KOLBASI DOGAN

Differences in sensory integration on balance between older fallers and non-fallers, a systematic review and meta-analysis

Table of contents

Research context

This systematic review and meta-analysis is a part of the research domain: gait and balance within the Faculty of Rehabilitation Sciences of Hasselt University [1]. This review was written by two authors (D.V. and Q.D.) under the supervision of Prof. Dr. Pieter Meyns and Dr. Esma Nur Kolbaşı Doğan.

Falls are complex and involve an interaction of different factors, such as biomechanics, environmental factors, medical conditions, psychological influences, etc. [2,3]. Falls can often result from a combination of individual, psychological, physical [2], and environmental factors [3]. Despite the significant prevalence of falls in older populations and the consequences related to falls, more research is needed to understand which influences sensory integration has on balance. Despite the limited number of studies, the goal of this review is to comprehensively review what is currently reported in the available literature on sensory integration in older fallers and non-fallers. By doing so the researchers hope to inspire other researchers to perform more research on this topic in a more standardized manner.

In order to measure sensory integration and balance, the sensory organization test, the clinical test of sensory integration on balance, and center of pressure measurements under sensory changing conditions on a force plate are common to administer. This master thesis is situated in the field of research related to the PhD of Dr. Esma Nur Kolbaşı Doğan, which focuses on the sensory contributions to static and dynamic balance in young and older adults.

The goals of this review and the research question were established after consultation with promotor Prof. Dr. Pieter Meyns and co-promotor Dr. Esma Nur Kolbaşı Doğan. The literature search, data collection, and article screening were done independently (D.V. and Q.D.) under the supervision of our promotor and co-promotor. The academic writing in this study was performed by D.V. and Q.D. and corrected or adjusted after feedback from Prof. Dr. Pieter Meyns and Dr. Esma Nur Kolbaşı Doğan.

References

1. UHasselt. Gait & balance. REVAL 2022 [cited 2025 12 january 2025]; Available from: https://www.uhasselt.be/nl/onderzoeksgroepen/reval/gang-evenwicht.

2. Byun, M., J. Kim, and J.E. Kim, Physical and Psychological Factors Contributing to Incidental Falls in Older Adults Who Perceive Themselves as Unhealthy: A Cross-Sectional Study. International Journal of Environmental Research and Public Health, 2021. 18(7): p. 3738.

3. Lee, S., Falls associated with indoor and outdoor environmental hazards among community-dwelling older adults between men and women. BMC Geriatrics, 2021. 21(1).

1 Abstract:

Background: A key aspect of balance is sensory integration (SI), a process combining neural representations of physical variables derived from multiple sensory systems. Disturbed SI can affect balance, possibly resulting in falls. A comprehensive systematic review (SR) and meta-analysis (MA) on this topic is currently lacking, this thesis aims to address this.

Methods: The search strategy was performed on PEDro, Pubmed, and Scopus WoS. Subjects were divided into groups (fallers vs non-fallers) based on fall history. The CASP- and Downs&Black checklists determined the study quality. Three MAs compared older fallers to non-fallers during comparable conditions with similar outcome parameters, to study differences in; I)Composite equilibrium scores (CS) II)Positional Center of pressure (CoP) parameters while standing still on a firm surface with the eyes closed (EC-FS), and III)Dynamic CoP parameters during EC-FS.

Results: The search strategy resulted in 7,122 articles of which 13 articles (996 subjects) were included in the SR, and 7 studies (395 subjects) in the MA. 9/13 studies found significant differences in SI between older fallers and non-fallers. MA1 revealed a significant difference in CS between the groups (Hedges'g:-1.01 [-1.42,-0.60]; P:<0.005). Positional (Hedges'g:0.06 [-0.14,0.26]; P:0.58) and dynamic (Hedges'g:0.07 [-0.17,0.31]; P:0.57) CoP parameters revealed no differences between the groups.

Discussion and Conclusion: A significant difference with a large effect size was found between groups for the CS. No significant differences were found in the positional or dynamic CoP measurements during EC-FS. This indicates fallers exhibit worse SI when all conditions of the SOT are tested.
Important keywords: falls, sensory integration, balance, older adults

2 Introduction

2.1 Rationale

A fall has previously been defined as “an event causing a person to rest unintentionally on the ground or other lower level, not due to any intentional movement, a major intrinsic event, or extrinsic force.” [1]. Falls are a major public health problem, especially among older adults. Falls are one of the leading causes of injury and mortality among this population [2]. One-third of community-dwelling older adults aged 65 years or over have reported falling at least twice a year [3, 4]. Furthermore, the rate of falling rises to 50% in adults over the age of 80 [5]. It was found that experiencing a fall could have significant effects on the quality of life [6] and several psychological consequences, such as fear of falling and loss of confidence. Altogether the physical and psychological consequences of a fall could result in restricted activity levels, physical function, and a decrease in social interactions [7], which in turn increase the risk of falling, creating a vicious cycle [8]. The negative impact of a fall is significant for the individual and is a substantial economic burden on society [9].

In order to prevent falls, in-depth investigations are needed to understand how fall incidents are caused and which differences can be found between people prone to falling and those who are not.

One of the factors that influence one's risk of falling is the person's “ability to keep or regain the center of mass within the base of support during activities”, termed “balance”[10]. Balance is essential for the majority of daily motor activities. Balance impairments have previously been associated with increased mortality, loss of independence, hospitalization, lower quality of life, fractures, and fear of falling among older individuals [11-13].

Maintaining balance involves numerous different systems working in harmony to respond to disturbances. The visual, vestibular, and somatosensory systems provide information for the process of maintaining balance [14]. There is a constant series of small postural adjustments based on the received sensory information, that allows the body to generate appropriate

motor responses to counteract perturbations. Given the importance of balance, a thorough understanding of this ability is required [15]. Processing these small perturbations and reacting to them, requires a specific process involved in balance; sensory integration (SI), on which this thesis will focus.

SI is a process combining the neural representations of physical variables derived from multiple sensory systems [16]. Each individual sensory system provides us with an impression of our surroundings or the location and position of our body in these surroundings, which are essential for cognitive processing, control of action, and perception [17]. Furthermore, the brain is able to amplify important sensory information and filter out accompanying noise thereby detecting events in our surroundings faster and identifying these events more correctly [17].

Balance and SI are often assessed using force plates during sensory challenging tasks. Force plates can track many different outcomes. CoP is often used and is defined by Benda et al. (1994) [18] as "the projection on the ground plane of the centroid of the vertical force distribution". It is commonly interpreted more easily as the point where the resultant force vector would attach if imagined as a single point [18]. Analysis of the CoP provides an overview of how the CoP moved through space over a certain period giving researchers data about the CoP trajectory, which can then be interpreted to analyze stability. Greater postural sway, for instance, has been associated with an increased risk of falls [19]. Multiple different outcomes derived from these measurements can be analyzed, such as CoP velocity, the sway area, etc. These parameters can further be categorized into two groups; positional and dynamic CoP parameters. To assess SI specifically, several standardized tests have been developed which often include the use of force plates.

The sensory organization test (SOT) and the clinical test of sensory interaction on balance (CTSIB) with or without modifications are commonly used. The SOT is considered a gold standard for assessing the contributions of sensory systems to balance and the modified (m) CTSIB is the most widely utilized clinical version [20, 21]. Many testing procedures use CoP parameters to report a performance objectively. The SOT test is standardly performed on a force plate, but many studies using the CTSIB or mCTSIB, also report CoP-parameters.

Besides the standardized protocols, other conditions can also be performed on force plates to give an impression of the subject's sensory integration.

The SOT consists of six conditions, as described by Nashner et al. (1982) [22]. The first three conditions are performed on a normal steady surface and consist of standing still while; keeping the eyes open, keeping the eyes closed, and having sway-referenced vision (SRV). The last three conditions are the same tasks but performed on a sway-referenced platform (SRP). The CTSIB consists of six similar conditions, as described by Shumway-Cook and Horak (1986)[23]. In general, the stabilized support surface is replaced with a foam surface instead of the SRP used in the SOT. Conditions three and six differ between the SOT and CTSIB, as the SRV is replaced with a visual conflict dome instead.

The mCTSIB was first suggested by Cohen et al. (1993) [24] and is a shortened version of the CTSIB, which only studies the first four tasks from the original CTSIB. The mCTSIB is easier to conduct by leaving out the conditions with the conflict dome.

Several publications focus on differences in balance performance between older fallers and non-fallers due to their significant implications for understanding fall risk. Kozinc et al. (2020) [25] showed multiple balance tests and outcome measures were able to distinguish older fallers from non-fallers. This finding raised the question of whether these results are consistent across the different processes involved in maintaining balance. Among these processes, SI is a process that has not been studied as thoroughly compared to others. Since differences in balance performance have been shown between older fallers and non-fallers [25, 26], a difference in SI between both groups was hypothesized. There appears to be a lack of a comprehensive SR and MA summarizing the current literature on this topic.

2.2 Objectives

The objective of this thesis is to conduct an MA to systematically review and synthesize existing literature on the differences in SI on balance between older fallers and non-fallers.

3 Method:

This review was conducted following the Preferred Reporting Items for Systematic Reviews and MAs (PRISMA) guidelines [27].

3.1 Research question

This study aims to investigate the differences in SI on balance between older fallers and non-fallers. This SR and MA hypothesize that older fallers score significantly worse on SI tests compared to older non-fallers. This hypothesis was based on the large number of studies proving older fallers to have worse balance performance compared to older non-fallers, in which SI might be influential.

3.2 Eligibility criteria

For this review, specific inclusion and exclusion criteria were established before conducting the literature search to ensure that only relevant studies meeting predetermined criteria were considered for analysis.

The data used were exclusively extracted from studies explicitly analyzing the differences in SI between older fallers and non-fallers, or observational studies on this population looking into the effect on SI of a certain intervention or exposure, reporting their baseline measurements. The types of studies included in this review were; case-control studies, non-randomized control studies, randomized controlled studies, cohort studies, observational studies, and longitudinal studies. For the studies to be included, they were required to have a defined faller and non-faller group, based on the number of fall incidents prior to the SI testing.

The eligibility criteria are presented in Table 1.

Table 1

Overview of eligibility criteria

Inclusion criteria	Exclusion criteria
Studies published in English	Studies examining non-human subjects
Studies in which the interpretation of a fall is provided and include the following criteria: - It is unintentional - It results in the individual ending up on the ground or other lower level - It is not caused by any condition or factor due to which a normal healthy person would fall as well (being pushed, having a stroke, etc.)	Intervention studies in which the baseline data is not provided
Populations' mean age ≥ 65 years old	Non-applicable types of studies (case reports, narrative reviews, etc.)
Including a group of interest consisting of older fallers without any relevant medical conditions (an overview of this list is provided in the exclusion criteria)	Studies only analyzing tasks in which a perturbation is applied that forces the participant out of their limits of stability
Including a control group consisting of older non-fallers without any relevant medical conditions (an overview of this list is provided in the exclusion criteria)	Studies only analyzing tasks in which the base of support is altered
Studies reporting the results from the SOT, CTSIB, mCTSIB or other procedures analyzing SI that report CoP data.	Studies only analyzing populations with specific conditions that could influence balance:

	- Neurological conditions: Parkinson’s disease or forms of parkinsonism, multiple sclerosis, Huntington’s disease, stroke, spinal cord injury, etc. - Cognitive conditions: Alzheimer's disease or other forms of dementia - Systemic conditions: diabetes, cardiovascular or pulmonary conditions, etc. - Acute musculoskeletal conditions: acute knee, hip or shoulder replacement, muscle tears, muscle atrophy, osteoporosis, sarcopenia, etc. - Vestibular disorders: BPPV, Ménière’s disease, vertigo, dizziness, etc. - Visual disorders, which are currently impairing the patient (disorders not adjusted for with glasses or lenses)

Note: The in- and exclusion criteria established for both the systematic review and meta-analysis.

3.3 Literature search strategy

3.3.1 Information sources

The search for eligible studies was conducted in the following electronic databases;

- PEDro Physiotherapy Evidence Database (October 2004 - November 11th 2024)
- PubMed National Library of Medicine (March 13, 1971 - November 11th 2024)
- Clarivate Web of Science (January 01, 1967 - November 11th 2024)

3.3.2 Search strategy

The search strategy for this paper was conducted using a PECO which can be found in Table 2.

Table 2.

PECO used to conduct search strategy

Peco	Search strategy
P	Elder*[Title/Abstract],Older*[Title/Abstract], Aged[MeSH Terms], Aged[Title/Abstract], Aging[Title/Abstract], Geriatric*[Title/Abstract], Ageing[Title/Abstract]
E	Fall*[Title/Abstract], Fell[Title/Abstract], Accidental falls[MeSH Terms], Stumbl*[Title/Abstract], Slip*[Title/Abstract], Risk of fall*[Title/Abstract]
C	older non-fallers
O	sensory reweighting[Title/Abstract], sensory reweighing[Title/Abstract], sensory re-weighting[Title/Abstract], feedback, physiological[MeSH Terms], Sensory integration[Title/Abstract], sens*[Title/Abstract], CTSIB[Title/Abstract], SOT[Title/Abstract], posturograph*[Title/Abstract], sense organs[MeSH Terms], visu*[Title/Abstract], Balanc*[Title/Abstract], Stability[Title/Abstract], Stead*[Title/Abstract], Postur*[Title/Abstract], Postural balance[MeSH Terms],

	Unstab*[Title/Abstract], eyes open[Title/Abstract], eyes closed[Title/Abstract], vestibul*[Title/Abstract], propriocep*[Title/Abstract], somatosensory[Title/Abstract]

Note. P: population, E: exposure, C: comparison, O: outcome. The PECO framework was used to create a comprehensive search strategy. The Boolean operators used for the search string were "OR" and "AND". Within the categories of the PECO "OR" was used to include one or both terms, between different categories of the PECO "AND" was used to include both terms.

Appendix 1 includes a detailed description of every search string for the three different databases.

3.4 Selection process

The selection process of this review was performed by two individual researchers (Q.D. and D.V.) to ensure the inclusion of relevant studies. First, both researchers independently screened all articles based on their titles and abstracts and in- or excluded them in temporary overviews which were then combined to one complete overview of studies eligible for full-text screening. Secondly, a full-text screening was performed on those studies. For the articles that were found to be eligible after full-text screening, a more thorough screening of the results was performed by the two researchers independently. For every article separately, each of the researchers could mark whether it was eligible or not. In case of disagreement on an article's eligibility between the two researchers, a third independent researcher (F.V.) reviewed the article's title and abstract to make the final decision.

4 Data extraction

A data extraction sheet was developed and implemented by both authors. One author (Q.D.) extracted the data for the anthropometric data and the number of older fallers, verified by a second author (D.V.). Both researchers read the included studies and provided a summary of the results relevant to this thesis, provided in the column 'Results for the testing conditions'. The outcome measurements and testing conditions were extracted by one author (D.V.) and verified by a second author (Q.D.).

Outcomes are categorized as follows:

Table 3:

Overview of the extracted data

Author (year)	
Anthropometric measurements of the fallers and non-fallers groups	Age Height Weight Number of males and females
Testing conditions	
Outcome measurements of the SI tests	SOT outcomes (m)CTSIB outcomes CoP outcomes during sensory changing conditions
Results for the testing conditions	A summary of the relevant results

Note: An overview of the layout applied in the results table of the SR, provided in Appendix 4.

In studies that included a control group of young adults or a comparative group with a pathology, data from these groups were not included in this review.
To create more homogeneity in the analysis of CoP balance measurements, variables were categorized into positional CoP outcomes and dynamic CoP outcomes, as proposed by Quijoux et al. (2021) [28]. Data was categorized as a *positional CoP outcome* if it measured static balance characteristics, such as the area, mean distance, or range of the center of pressure displacement. *Dynamic CoP outcomes* were all data related to how the CoP moves over time, such as information about the CoP velocity, sway area per second, etc.
These subcategories were further subdivided based on the different test conditions reported in the studies. Different MAs focused on specific types of outcome measurements during specific test conditions, such as the different test conditions from the SOT or (m)CTSIB.
For this paper, no distinction was made between older single fallers and multiple fallers. If studies reported separate results for single fallers and multiple fallers, their data were combined to create one single group of "older fallers". The formulas used for these calculations, as proposed by Sen and Yildirim (2022) [29], are explained in detail in the statistical analysis section of this paper.
Similarly, studies in which CoP parameters were reported for different directions (e.g., antero-posterior and medio-lateral) also needed further calculations to create one combined value. To create a combined value of these results the same formulas as described by Sen and Yildirim 2022 [29] were used.
In case a study reported both a distinction in single fallers and multiple fallers, together with separate directional results, a combined mean was calculated across all subgroups.

To indicate the studies in which calculations had to be performed to create a combined group, a system of asterisks was applied in the data extraction and MA results section and in Appendix 4 and 5.

- One asterisk (Author, year*) indicates a combined group was calculated based on fall history (e.g., single fallers vs multiple fallers).
- Two asterisks (Author, year**) mean calculations were performed to create combined means for directional parameters (e.g., antero-posterior vs medio-lateral)

- Three asterisks (Author, year***) are presented in cases where calculations had to be done for both subgroups (fall history and directional results)

4.1 Quality and risk of bias assessment

For the quality assessment the Critical Appraisal Skills Programme (CASP) tool [30] was used. There are tools available with better validity, but for inexperienced researchers, the CASP checklist is easy to use [30]. The tool is divided into different versions for specific types of studies, the version consistently used in this thesis is the CASP checklist for case-control studies. No other versions were used because no other types of studies remained. This version consists of 11 questions divided into three major sections;

- "Are the results of the trial valid?"
    - Did the study address a clearly focused issue?
    - Did the author use an appropriate method to answer their question?
    - Were the cases recruited in an acceptable way?
    - Were the controls selected in an acceptable way?
    - Was the exposure accurately measured to minimize bias?
    - Aside from the experimental intervention, were the groups treated equally?
    - Have the authors taken account of the potential confounding factors in the design and/or in their analysis?
- "What are the results?"
    - How large was the treatment effect?
    - How precise was the estimate of the treatment effect?
    - Do you believe the results?
- "Will the results help locally?".
    - Can the results be applied to a local population?
    - Do the results of this study fit with other available evidence?

Most questions have three possible answers to check off: "Yes", "can't tell" and "No". The tool is interpreted qualitatively, meaning it does not add up to a certain numeric score, but rather an overall impression [31].

In this review, the researchers rated each of the three sections ‘Good’, ‘Moderate’, or ‘Poor’. For the first section, a study's validity is considered ‘Good’ if the aims are clearly defined, the methodology is robust, and the biases are defined or minimized. It is considered ‘Moderate’ when the aims are generally clear, the methodology is mostly appropriate but with some weaknesses (e.g., small sample size, partially applicable, etc.), and if there is a moderate risk of bias or missing details. The validity is considered ‘Poor’ if the study doesn’t define aims, has a weak methodology, and/or has a high risk of bias or confounding factors.
In the second section, a study’s results are considered ‘Good’ if the results are clearly reported with sufficient data and the statistical analyses are correct. It is considered ‘Moderate’ if the data are less clear, less transparent, and harder to interpret. The results are considered ‘Poor’ if they are unclear, incorrect and if no statistical analysis can be found or is incorrect.
In the last section, the applicability of a study is considered ‘Good’ if the findings are relevant to the research question of this SR and MA, if the conclusions are supported by evidence, and if the characteristics of the populations are similar. It is considered ‘Moderate’ if the findings are somewhat relevant to the research question of this SR and MA, the conclusion is partly supported by evidence and the population's characteristics mostly resemble each other. The applicability is considered ‘Poor’ if the findings are not relevant to the research question of this SR and MA, if there is no evidence that supports the conclusion, or if the population characteristics do not match.

Two individual researchers (Q.D. and D.V.) completed this checklist separately for all included studies and a third researcher (F.V.) was consulted to reach a consensus in case of disagreement. After reaching a consensus, a final version of the checklist was completed for each study combining the opinions of the different researchers to ensure an objective quality assessment.

The risk of bias was assessed using a modified Downs and Black checklist. The checklist was modified on item 5, which refers to the description of the distributions of principal confounders in each group of subjects to be compared. Instead of rating two points for 'Yes,' one point for 'Partially,' and zero points for 'No,' the rating was given as one point if a

description of the distributions of principal confounders in each group of subjects to be compared was clearly described, and zero points if this was not clearly described. Furthermore, item 27, which refers to the power of the study, was also modified. A rating of one point or zero points was used depending on whether the study calculated power, instead of scoring according to the available ranges of power in the study [32, 33]. The maximal score of this modified Downs and Black checklist equals 27 with ranges of 'Excellent' (26-27), 'Good' (20-25), 'Fair' (15-19), and 'Poor' (≤14) [34].
Two researchers (D.V. and Q.D.) individually scored each study according to the modified Downs and black checklist. If there was disagreement, a third researcher (F.V.) scored the study to make a final decision.

4.2 Statistical analysis

4.2.1 Effect measures

Separate MAs were conducted for different tasks and outcome measurements. To conduct an MA, at least three studies had to address the same task and use comparable outcome measures. All relevant data from these eligible studies were listed in Microsoft Excel. This table is shown in Appendix 5. This dataset was then imported into SPSS statistics (version 30) for further statistical analysis.

Due to a lack of studies measuring the same outcome measurements during the same conditions, multiple outcome measurements had to be combined to form larger groups. The outcomes of studies reporting CoP parameters were split into positional and dynamic CoP parameters, as suggested by Quijoux et al. (2021) [28]. Studies in which no suitable data for the MAs were provided, were excluded from the MA but still included in the SR.

Different MAs were performed, using Hedges' g as the effect size to analyze the varying sample sizes, as it is suggested as the most appropriate method for this purpose [34]. All statistical analyses were based on the mean values and standard deviations reported in the studies.

Some studies included in the MA reported a subclassification in their older fallers group. To implement these results into the MA, the combined means and standard deviations of these subgroups had to be calculated using the formulas recommended by Sen and Yildirim (2022) [29] as mentioned earlier.
If the results of one parameter were reported in different directions (i.e. antero-posterior & medio-lateral), the pooled mean and standard deviation of these values were calculated with the same formulas recommended by Sen and Yildirim (2022) [29].

The following formulas were used to calculate the mean and standard deviation for the combined group, based on the recommendations made by Sen and Yildirim (2022) [29]:
$mean_{total} = \frac{n_1 * mean_1 + n_2 * mean_2}{n_1 + n_2}$For the combined standard deviation, the following formula was used:

$SD_{total} = \sqrt{\frac{(n_1 - 1) * SD_1^2 + (n_2 - 1) * SD_2{}^2}{(n_1 - 1) + (n_2 - 1)}}$In both formulas, the abbreviations are as follows:

- Total = the combined value that will be used in the MA
- n_1 = number of subjects in the older single fallers group
- n_2 = number of subjects in the older multiple fallers group
- SD_1 = standard deviation for the older single fallers group
- SD_2 = standard deviation for the older multiple fallers group

4.2.2 Synthesis methods

Prior to pooling the effect sizes, between-study heterogeneity was accounted for by applying a random-effects model. Following this, the actual MAs were performed. Subsequently, heterogeneity was examined by calculating the I^2 statistic, to assess how much variation was present across studies. If high heterogeneity was observed, moderation and sensitivity analyses would be performed to check whether factors such as age, sex, weight, or length could possibly explain the heterogeneity.
The I^2 statistics were interpreted following the classification described by Higgins and Thompson, 2002, as cited in Heudo-Medina et al. (2006) [35], in which percentages of

around 25%, 50%, and 75% are interpreted as low, medium, and high heterogeneity respectively.

For academic writing, the two researchers worked together and received feedback from the promotor and co-promotor. To correct grammar and spelling, the text was checked by AI (chatGPT). For this purpose, full paragraphs written by the researchers were given as a prompt with the instruction “check for any spelling or grammatical mistakes”. This prompt was used in the same way several times for the different paragraphs and the outcome was then analyzed by the researchers and if necessary adjusted in the text.
No text was written by AI, all text given to AI was written by the researchers and the outcomes provided by AI were never copied literally.

5 Results:

5.1 Results study selection

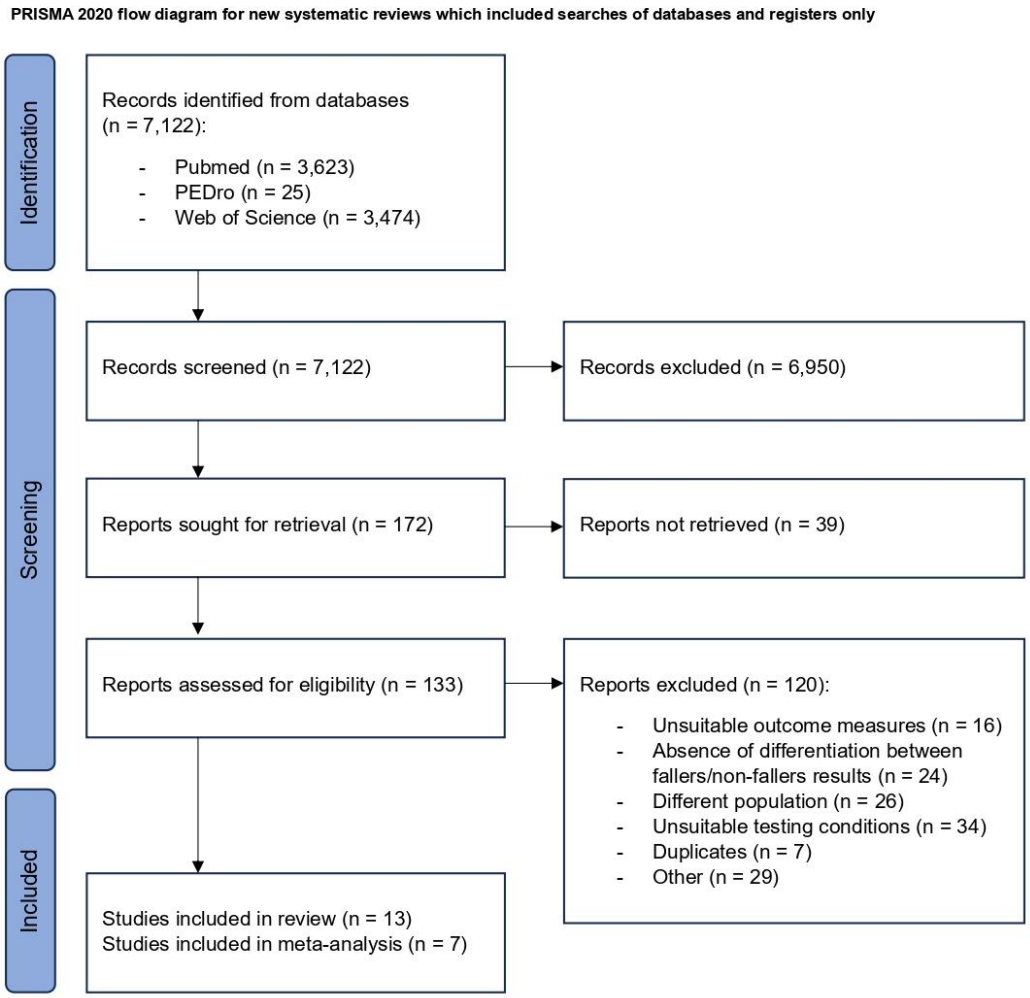


Fig. 1. Prisma flowchart of search results. Fig 1. displays the results of the search strategy and subsequent selection process

On November 11th, 2024 a total of 7,122 results were found from three different scientific databases, which were subsequently included for screening. At the end of the selection process, a large number of articles were excluded, leaving only 13 articles eligible. This resulted in a total of 996 subjects, consisting of 462 older fallers and 534 older non-fallers, across all studies included in the SR. The strict screening process ensured that the remaining

articles were most relevant to answer the research question. An overview of the selection process can be found in the PRISMA flowchart in the results section (Fig 1). An overview of all included studies can be found in Appendix 4.

5.2 Results quality assessment

The quality of the included articles was assessed using the CASP checklist for case-control studies. In the first section “Are the results of the study valid” nine out of the thirteen studies scored ‘Good’, while the other four studies: Liston et al. (2013) [36], Maranesi et al. (2016) [37], Müjdeci et al. (2012) [26] and Van den Hoorn et al. (2018) [38] were rated ‘Moderate’ because the selection process of the subjects was not clear or very little information was given about the subjects themselves. For example, some populations came from a database used in previous research without giving further information. For the second section “What are the results” all studies were rated ‘Good’. The third section “Will the results help locally” yielded more diverse outcomes with three studies scoring ‘Good’, eight studies scoring ‘Moderate’ and two studies scoring ‘Poor’. the separate CASP checklists can be found in Appendix 6, and Appendix 2 provides a brief overview of all the studies' performance on the CASP checklist. Kim et al. (2011) [3], received a ‘Poor’ score because of the specific population used. Supporting evidence for their results is indecisive as the evidence regarding their results is currently conflicted. Van Den Hoorn et al. (2018) [38] scored ‘Poor’ because no conclusion could be made regarding the applicability to a local population due to a lack of information about the subjects and similar to Kim et al. (2011) [3], the supporting evidence is currently conflicted.

The risk of bias was assessed using the Downs and Black checklist. An overview of the total score per article can be found in Appendix 3 along with the subscores. The mean score of the included studies was 14.69/27 (ranges from 12-18), which indicates an overall 'Fair' score on risk of bias. Among the thirteen studies evaluated, two studies scored 12/27 and another two scored 13/27. Only one scored 14/27, and the most prevalent score was 15/27 with four studies reaching this score. Two studies scored 16/27, and for the highest scores of 17/27 and 18/27, both only had one result. All the Downs and Black checklists can be

separately found in Appendix 7. For the reporting part, the average score was 7.4/10. For 'external validity' the average score was 2/3. For internal validity, two separate parts were made; 'internal validity - bias', where the average score was 2.7/7, and 'internal validity - confounding', for which the average score was 2.1/6. The last part of the Downs and Black checklist investigates the power, in this case, the average score was 0.5/1.

5.3 Results systematic review

For detailed data for every study, together with a summary of their results, Appendix 4 provides a clear overview. The average age of the participants included over all studies was 74.51 years and ranged from 61.49 [39] to 88 years [36]. However, Fino et al. (2016) [40], did not differentiate between the age of fallers and non-fallers. For which the mean age of all participants together is provided in both columns in the overview in Appendix 4, instead of reporting this for both groups separately, to complete the overview. Park et al. (2014)** [41] did not differentiate between any of the anthropometric data except for the number of fallers and non-fallers. The number of males and females is provided to assess potential gender differences. Not all articles reported age and height, in some cases, BMI was used instead, which is not presented in this table.
Most studies reported a significant difference in performance on SI tests between older fallers and non-fallers. Nine out of thirteen studies found such differences in one or more conditions.

Fino et al., 2016 [40], Howcroft et al., 2017 [42] and Park et al., 2014 [41] all compared older fallers and non-fallers while standing still with the eyes open on a firm surface (EO-FS) and while standing still with the eyes closed on a firm surface (EC-FS). None of the three studies reported any significant results during these conditions when looking at a wide range of CoP parameters and entropy measures. Yamagata et al., 2024 [43] looked at these two conditions as well but made the subjects stand with their feet together. When analyzing different CoP parameters for both the antero-posterior (AP) and medio-lateral (ML) direction, only the AP direction showed significant differences. These differences were

observable in the rambling mean velocity, CoP root mean square, and rambling root mean square.

Gregg et al., 2023 [44], Lázaro et al., 2011 [45], Maranesi et al., 2016 [37] and Petrella et al., 2012 [39] conducted the mCTSIB in the two groups. Gregg et al. 2023 [44] found older fallers to have significantly more right-directional control and less anterior maximum excursion overall. Lázaro et al., 2011 [45] only found significant differences in the two foam conditions. Both with the eyes open and with the eyes closed, older fallers showed a larger displacement when standing on a foam surface. Maranesi et al., 2016 [37] made different conclusions for frequent fallers and for infrequent fallers, which are presented in more detail in Appendix 4. Overall frequent fallers differed from infrequent fallers during standing still with the eyes open on an SRP (EO-SRP) and during both EO-FS and EC-FS. Frequent fallers differed from non-fallers during EO-FS, EC-FS, and EO-SRP for different CoP parameters in different directions. Petrella et al. [39] changed the test duration from 30 seconds to 60 seconds per condition and overall observed significantly more ML CoP displacement in fallers, compared to non-fallers.

Kim et al., 2011 [3], Liston et al., 2014 [36] and Müjdeci et al., 2012 [26] used the SOT test and both Kim et al., 2011 [3] Liston et al., 2014 [36] and Müjdeci et al., 2012 [26] found non-fallers to have a significantly lower composite equilibrium score (CS), indicating a worse performance on the test in general. Müjdeci et al., 2012 [26] also provided equilibrium scores (ES) for the different conditions. Only for the conditions with the SRV, a significantly lower ES was found, both while standing on a firm surface, and while standing on the SRP surface.

Ricci et al., 2009 [46] let the subjects perform the CTSIB and found significantly more fallers who were not able to maintain their balance during the first two conditions on the foam surface, both during eyes open and eyes closed. Van den Hoorn et al., 2018 [38] only looked at the performance of EC-FS and were not able to find significant differences for the different CoP parameters analyzed.

5.4 Results meta-analyses

The studies reported numerous different outcome variables related to SI.
The main testing procedures in which similar conditions and outcome parameters were reported, consisted of the SOT test and specific conditions of the mCTSIB or CTSIB test. The data was divided based on the performed test, for all studies that provided their results for both the older fallers and non-fallers group and were subdivided into comparable outcome measurements.
Seven studies used similar testing procedures and explicitly reported their outcome measurements. From the seven studies, three different MAs could be conducted. Other studies did report similar testing procedures and outcome measures, but did not provide the data for the two groups and could thus not be included in the MA. Seven studies [3, 26, 36, 37, 40, 41, 44] were included in the MA, of which three sub-meta-analyses could be conducted. The MAs focused on comparing older fallers and non-fallers to investigate differences in; I) the CS of the SOT, II) positional CoP parameters during EC-FS, and III) dynamic CoP parameters during EC-FS.

Meta-analysis I: composite equilibrium score of the SOT

The first MA, consisting of three studies [3, 26, 36], examined the CS of the SOT, which gives an interpretation of the performance across all six conditions of the test. This MA included a total number of 55 older fallers and 46 non-fallers. The effect size, displayed by Hedges' g, was -1.01 ($p < 0.001$), with a confidence interval ranging from -1.42 to -0.60. This suggests a significant difference between older fallers and non-fallers. The I^2 statistic revealed no heterogeneity ($I^2 = 0.00$). This result is visualized in Fig. 2a.

Meta-analysis II: positional CoP parameters during quiet standing with the eyes closed on a firm surface

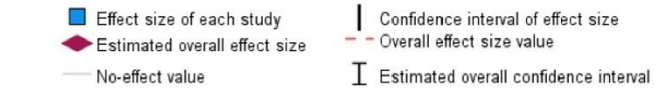
The second MA consisted of three studies [37, 40, 41] examining positional CoP parameters during one specific task part of the SOT test; EC-FS. The parameters grouped in this category were: CoP measures of 95% ellipsoidal area, CoP mean distance, mean CoP range, and the CoP distance.

Three studies [37, 40, 41] were included, of which two [37, 41] reported multiple useful outcomes. Both Maranesi et al. (2016)*** [37] and Park et al. (2014)** [41] reported on CoP mean distance and mean CoP range and were thus both included twice in the MA. Altogether this MA counts a total number of 172 older fallers and 221 older non-fallers. The effect size was 0.06 ($p = 0.58$) with a confidence interval from -0.14 to 0.26, indicating no significant difference was observed. Again, heterogeneity was absent ($I^2 = 0.00$). For a better understanding of these results, this data is further depicted in Fig. 2b.

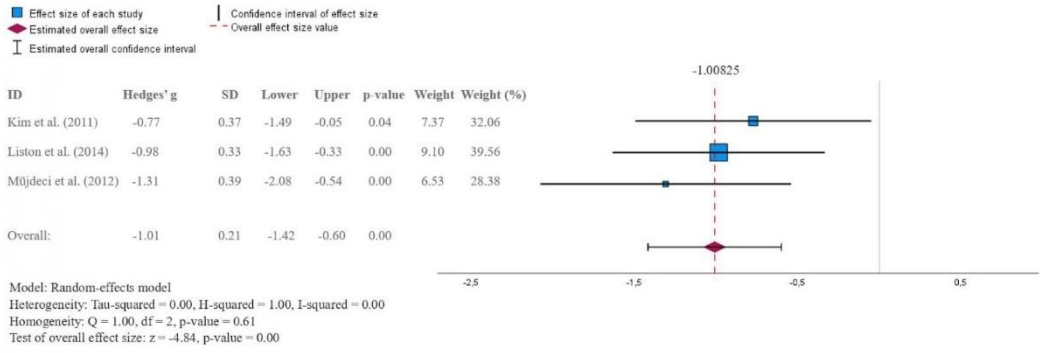
Meta-analysis III: dynamic CoP parameters during quiet standing with the eyes closed on a firm surface

The last MA consisted of three studies reporting dynamic CoP measures (CoP velocity) during EC-FS. The first two studies included were Fino et al. (2016) [40] and Maranesi et al. (2016)*** [37], which were also included in the previous MA on positional CoP outcomes. In total, this MA analyzed 114 older fallers and 151 older non-fallers. The effect size was 0.07 ($p = 0.57$), with a confidence interval from -0.17 to 0.31, indicating no significant difference. The I^2 statistic again showed no heterogeneity ($I^2 = 0.00$). This can be found in Fig. 2c.

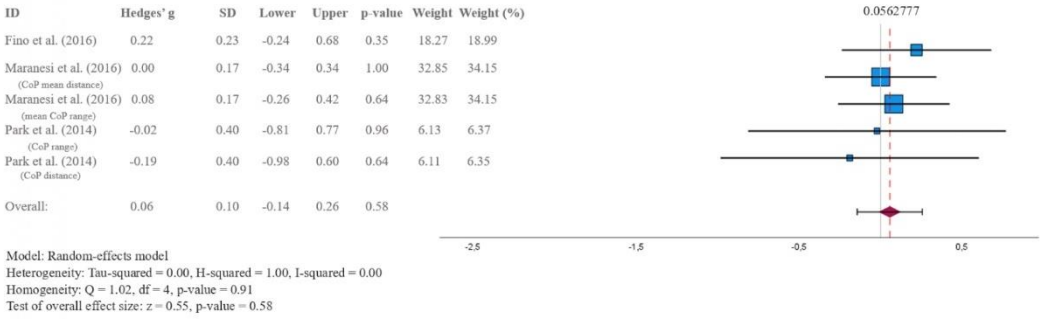
A comprehensive summary of the results from these articles is also included in Appendix 4.



A. MA 1: composite equilibrium score of the SOT



B. MA2: Positional CoP parameters during quiet standing with the eyes closed on a firm surface



C. MA3: Dynamic CoP parameters during quiet standing with the eyes closed on a firm surface

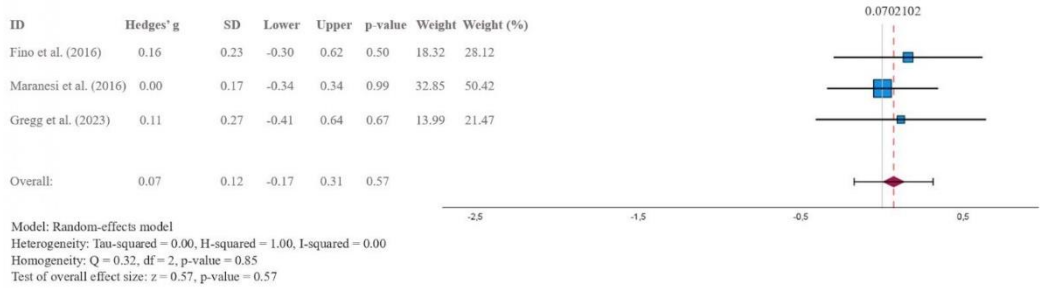


Fig. 2. Forest plots of meta-analyses. Fig 2. displays the forest plots of the three meta-analyses conducted.

6 Discussion:

This thesis aims to systematically review the current literature regarding the difference in SI on balance between older fallers and non-fallers. The findings of this thesis could only partially support our hypothesis concerning the different outcomes of older fallers and non-fallers on different SI tests. Our hypothesis is supported by most of the studies included in the SR and the significant differences found in the MA on the CS of the SOT. However, no significant difference could be found for the EC-FS task of the (m)CTSIB. These findings suggest that fallers might be able to use compensation strategies to complete individual conditions successfully [47]. The CS gives an overall impression of the SOT test which may be better to differentiate between fallers and non-fallers. None of the MAs showed high heterogeneity, indicating that differences in testing procedures were unlikely to influence the findings.

6.1 Reflection on quality

The quality of the studies included in this SR was assessed using the CASP checklist. The results of the checklist indicated 'Good' quality for the first section. This indicates the results of the studies selected for this SR and MA are valid, with clearly focused aims and appropriate methodology, contributing to the quality of this review. The second section: "what are the results" also scored 'Good', suggesting that the results of the included studies were clearly reported and that appropriate statistical analysis was used. This improves the quality of the studies further. However in the third section "Will the results help locally" most studies scored 'Moderate'. The moderate scores in this section could be attributed to the inclusion of studies conducted in diverse cultural and geographical contexts which affects the local applicability of the results. Additionally, some studies did not provide comprehensive data on their subjects or used small sample sizes, which negatively impacted the outcome.

The risk of bias was assessed using a modified Downs and Black checklist. The results indicate 'Fair' to 'Poor' outcomes on this checklist. These outcomes suggest an increased risk of bias which could lead to an over- or underestimation of the results in this thesis,

representing the population incorrectly. The lower score can be attributed to poor study design with a rather high risk of bias in the included studies.

6.2 Reflection on findings

Our findings reveal mixed results with a slight inclination in the literature supporting a difference in SI on balance between older fallers and non-fallers. Nine out of the thirteen studies included in the SR reported a difference with older fallers scoring worse compared to non-fallers. This is in accordance with existing literature on balance where differences between fallers and non-fallers have been described [48]. Fallers tend to perform worse on SI tests during more challenging conditions, which is according to the literature as well[49]. This might suggest that fallers can compensate for the altered sensory input during the less challenging conditions. This compensation is likely not possible when multiple different systems receive unreliable sensory information. Compensatory strategies have already been proven to play an important role in older adult's ability to maintain balance [47].
When people age, sensory input is diminished [2], therefore older adults have to rely on multiple sensory systems at once to maintain balance. Older fallers seem to be less capable of integrating sensory information, which could be due to receiving less input from sensory systems compared to older non-fallers [47]. This might suggest that the sensory systems of fallers decline faster than non-fallers.
The results of the studies included in this SR often only describe a significant difference during a specific task or only in a specific direction. This could be attributed to a difference in the complexity between the tasks. Implying that some tasks challenge SI more than others, revealing difficulties that are not present during simpler tasks [50]. Challenges in different directions such as ML displacement could be the result of reduced limits of stability [51].

The MA on the CS of the SOT revealed a significant difference, with fallers scoring a consistently lower CS compared to non-fallers in the MA. These results indicate that fallers struggle to reweigh sensory inputs, affecting their balance during challenging conditions and increasing the risk of a fall. Other studies report similar outcomes [52]. This seems to imply that fallers tend to rely more on a combination of their sensory systems. If one of these

sensory systems becomes unreliable, the ability to maintain their balance decreases significantly. Furthermore, the lower CS of fallers could be attributed to slower postural adjustments or other consequences of aging, which leads to greater sway.
However, the CS does not differentiate between the different tasks of the SOT, making it more difficult to identify the sensory input system most responsible for the worse performance.
The current SR shows no significant differences in equilibrium scores for the separate conditions between fallers and non-fallers for every condition. Only conditions in which the SRV was applied, revealed significant differences. The CS is a weighted score encompassing all the different testing conditions, in which the conditions with more trials contribute more data points [53]. This emphasizes the greater difference in performance for the more difficult tasks. When only one condition is analyzed, compensatory strategies may be able to mask the differences in SI. This is supported by better test-retest reliability of the CS compared to the separate ESs [50]. When all conditions are combined, all sensory systems are challenged which fully encompasses the ability to perform SI.

The MA on (m)CTSIB did not show any significant difference during the EC-FS condition. This indicates that the difference in SI is not significant between fallers and non-fallers when only visual information is altered. However this condition did reveal differences in performance, but these were not found to be significant. Although the results were not significant, the small P-values of 0.06 and 0.07 do suggest a general trend. A potential explanation for this marginally non-significant result could be the relatively low difficulty of this condition, compared to the other testing conditions. Only one of the sensory input systems, vision, is altered, which could be compensated for. Fallers may use compensatory strategies to maintain their balance when only visual input is removed [47].
As evident from the findings of the SR on (m)CTSIB most studies investigating this test resulted in significantly worse performance in older fallers than in non-fallers. One study however reported more right-directional control and less anterior-displacement in fallers. The right-directional control may be a compensation mechanism by trying to maintain their CoP within a range where they have better control over it [54]. Whereas the reduced anterior displacement may indicate a restricted ability to shift weight forward [55].

In general, the results of the (m)CTSIB tend to show worse performance in older fallers than in non-fallers for at least one condition.

6.3 Reflection on strengths and weaknesses

Weaknesses:

During the literature search many different outcomes, for many different tests, were found making it difficult to consider which outcomes could be relevant for this thesis to assess SI and falls. Even within the chosen tests, the SOT and the (m)CTSIB, many varying testing procedures and outcomes were used. Two of the MAs in the current study investigated the EC-FS condition due to a lack of studies reporting other, more challenging, conditions. This resulted in less interpretable findings since these conditions are less challenging and could be compensated for.

There are numerous varying definitions of falling or other important terms related to this topic, leaving much room for interpretation and challenging researchers to find studies with similar interpretations. Furthermore, because of the complexity of the human sensory system, it is hard for one test to comprise all of the different input and output processes.

The inclusion and exclusion criteria were applied strictly but were hard to apply for the age range. Another weakness of this study was the inclusion of studies with fair to moderate risk of bias, making the result susceptible to bias and influencing the quality of this SR and MA.

Many studies with promising titles and abstracts could not be included because access to their full text was restricted. Furthermore, a few studies eligible for the MAs did not provide their data. The authors of the studies could have been contacted to obtain this information, however, this was not done, being an important limitation of this study.

Finally, although this review tried to minimize the underlying factors and conditions of the participants, due to the multifaceted and broad nature of falls it is almost impossible to research only one aspect of falling with a single test.

Strengths:

The primary strength of this thesis is that this is the first SR and MA conducted on this topic. No other study ever compared data from multiple studies investigating the differences in SI on balance between older fallers and non-fallers.

An extensive literature search was performed, ultimately leading to over 7.000 studies being screened for inclusion. Due to the rigorous in- and exclusion criteria, 13 relevant studies remained for the SR, and 7 were eligible for the MA.

Through the use of validated tools such as the Downs & Black checklist and CASP, the quality of included studies was assessed objectively and systematically.

6.4 Recommendations for future research

This thesis highlights a limited body of evidence in the current literature concerning the differences in SI between older fallers and non-fallers. More high-quality research is needed to improve and expand upon what is currently known about SI and falls. Only thirteen studies could be included in this SR, resulting in 996 subjects, and only seven studies were eligible for MA with a total of 395 different subjects. The MA was limited to the CS of the SOT and the EC-FS conditions of the (m)CTSIB. For future research, we suggest that all SI conditions are included. Testing only one condition seems insufficient to differentiate between older fallers and non-fallers. Furthermore, the population studied should be more extensive. Larger populations and more rigorous subject selection could contribute to larger variability of the older population. This would result in more representative effect sizes applicable to the general older population.

The interactions between the different sensory systems require further investigation to draw conclusions on whether or not they can be a predictor for falls. Overall more research is needed to better understand falls

7 Conclusion

This SR and MA have focused on the SOT, the (m)CTSIB, and CoP measures during quiet standing. The results found in the SR suggest that older fallers score significantly worse compared to older non-fallers on the SOT. This was confirmed in the MA with the CS of the SOT being significantly lower in older fallers. However, the EC-FS task of the (m)CTSIB did not find a significant difference. These results suggest that there is a difference in SI, however not for every condition. It seems that multiple sensory systems should be challenged. This could imply that fallers can compensate and maintain their balance when a single sensory system is compromised. Future research could focus on including all SI conditions from the SOT or (m)CTSIB, as focusing on a single condition does not seem to be able to differentiate between older fallers and non-fallers.

8 Acknowledgements

We would like to thank the Faculty of Rehabilitation Sciences and the University of Hasselt for providing the educational foundation and resources needed to complete this master's thesis.

We would like to genuinely thank Prof. Dr. Pieter Meyns and Dr. Esma Nur Kolbaşı Doğan. This thesis was written under their supervision and their expertise has been invaluable and greatly contributed to this thesis. We are deeply grateful for their support and feedback.

Lastly, we would like to express our gratitude to our friends and family for their support during this project.

9 References

1. Chu, L.W., A.Y. Chiu, and I. Chi, Falls and subsequent health service utilization in community-dwelling Chinese older adults. Arch Gerontol Geriatr, 2008. 46(2): p. 125-35.

2. Patti, A.P., ∗; Zangla, Daniele MSca; Sahin, Fatma Nese PhDc; Cataldi, Stefania MScb; Lavanco, Gioacchino MSca; Palma, Antonio MDa; Fischietti, Francesco MScb, Physical exercise and prevention of falls. Effects of a Pilates training method compared with a general physical activity program

A randomized controlled trial. 2021.

3. Kim, Y.-W., et al., Comparison of Toe Plantar Flexors Strength and Balancing Ability between Elderly Fallers and Non-fallers. Journal of Physical Therapy Science, 2011. 23(1): p. 127-132.

4. Hausdorff, J.M., D.A. Rios, and H.K. Edelberg, Gait variability and fall risk in community-living older adults: A 1-year prospective study. Archives of Physical Medicine and Rehabilitation, 2001. 82(8): p. 1050-1056.

5. Tinetti, M.E., M. Speechley, and S.F. Ginter, Risk factors for falls among elderly persons living in the community. N Engl J Med, 1988. 319(26): p. 1701-7.

6. Sherrington, C., et al., Exercise for preventing falls in older people living in the community. Cochrane Database of Systematic Reviews, 2019. 2019(1).

7. Yardley, L. and H. Smith, A Prospective Study of the Relationship Between Feared Consequences of Falling and Avoidance of Activity in Community-Living Older People. The Gerontologist, 2002. 42(1): p. 17-23.

8. Voermans, N.C., et al., Why old people fall (and how to stop them). Practical Neurology, 2007. 7(3): p. 158.

9. Davis, J.C., et al., International comparison of cost of falls in older adults living in the community: a systematic review. Osteoporosis International, 2010. 21(8): p. 1295-1306.

10. Verbecque, E., et al., Balance control in individuals with developmental coordination disorder: A systematic review and meta-analysis. Gait Posture, 2021. 83: p. 268-279.

11. Poncumhak, P., et al., Comparison of the Ability of Static and Dynamic Balance Tests to Determine the Risk of Falls among Older Community-Dwelling Individuals. Journal of Functional Morphology and Kinesiology, 2023. 8(2): p. 43.

12. Salzman, B., Gait and balance disorders in older adults. Am Fam Physician, 2010. 82(1): p. 61-8.

13. Lin, M.R., et al., Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. J Am Geriatr Soc, 2004. 52(8): p. 1343-8.

14. Berg RL, C.J., The Second Fifty Years: Promoting Health and Preventing Disability. 1992, Washington (DC): National Academies Press (US): Institute of Medicine (US) Division of Health Promotion and Disease Prevention.

15. Shokouhyan, S.M., et al., Time-delay estimation in biomechanical stability: a scoping review. Front Hum Neurosci, 2024. 18: p. 1329269.

16. Peterka, R.J., Chapter 2 - Sensory integration for human balance control, in Handbook of Clinical Neurology, B.L. Day and S.R. Lord, Editors. 2018, Elsevier. p. 27-42.

17. Freiherr, J., et al., Multisensory integration mechanisms during aging. Frontiers in Human Neuroscience, 2013.

18. Benda, B., P. Riley, and D. Krebs, Biomechanical relationship between center of gravity and center of pressure during standing. Rehabilitation Engineering, IEEE Transactions on, 1994. 2: p. 3-10.

19. Johansson, J., et al., Predicting incident falls: Relationship between postural sway and limits of stability in older adults. Human Movement Science, 2019. 66: p. 117-123.

20. Grove, C.R., et al., Validation of a next-generation sensory organization test in adults with and without vestibular dysfunction. J Vestib Res, 2021. 31(1): p. 33-45.

21. Goble, D., et al., Test–Retest Reliability of the Balance Tracking System Modified Clinical Test of Sensory Integration and Balance Protocol Across Multiple Time Durations. Medical Devices: Evidence and Research, 2021. 14: p. 355-361.

22. Nashner, L., F. Black, and C. Wall, Adaptation to altered support and visual conditions during stance: patients with vestibular deficits. The Journal of Neuroscience, 1982. 2(5): p. 536-544.

23. Shumway-Cook, A. and F.B. Horak, Assessing the influence of sensory interaction of balance. Suggestion from the field. Phys Ther, 1986. 66(10): p. 1548-50.

24. Cohen, H., C.A. Blatchly, and L.L. Gombash, A Study of the Clinical Test of Sensory Interaction and Balance. Physical Therapy, 1993. 73(6): p. 346-351.

25. Kozinc, Ž., et al., Diagnostic Balance Tests for Assessing Risk of Falls and Distinguishing Older Adult Fallers and Non-Fallers: A Systematic Review with Meta-Analysis. Diagnostics (Basel), 2020. 10(9).

26. Müjdeci, B., S. Aksoy, and A. Atas, Evaluation of balance in fallers and nonfallers elderly. Brazilian Journal of Otorhinolaryngology, 2012. 78(5): p. 104-109.

27. Liberati, A., et al., The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Journal of Clinical Epidemiology, 2009. 62(10): p. e1-e34.

28. Quijoux, F., et al., A review of center of pressure (COP) variables to quantify standing balance in elderly people: Algorithms and open-access code*. Physiological Reports, 2021. 9(22).

29. Sen, S. and I. Yildirim, A Tutorial on How to Conduct Meta-Analysis with IBM SPSS Statistics. Psych, 2022. 4(4): p. 640-667.

30. Hannes, K., C. Lockwood, and A. Pearson, A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. Qual Health Res, 2010. 20(12): p. 1736-43.

31. Long, H., D. French, and J. Brooks, Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. Research Methods in Medicine & Health Sciences, 2020. 1: p. 263208432094755.

32. Hindiyeh, N.A., et al., The Role of Diet and Nutrition in Migraine Triggers and Treatment: A Systematic Literature Review. Headache, 2020. 60(7): p. 1300-1316.

33. Downs, S.H. and N. Black, The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health, 1998. 52(6): p. 377-84.

34. Dhakal, C., Vitality and Application of Effect Size for Quality Research. Journal of the Institute of Agriculture and Animal Science, 2023: p. 49-57.

35. Huedo-Medina, T.B., et al., Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods, 2006. 11(2): p. 193-206.

36. Liston, M.B., et al., Peripheral vestibular dysfunction is prevalent in older adults experiencing multiple non-syncopal falls versus age-matched non-fallers: a pilot study. Age and Ageing, 2014. 43(1): p. 38-43.

37. Maranesi, E., et al., A statistical approach to discriminate between non-fallers, rare fallers and frequent fallers in older adults based on posturographic data. Clinical Biomechanics, 2016. 32: p. 8-13.

38. Van Den Hoorn, W., et al., Center of Pressure Motion After Calf Vibration Is More Random in Fallers Than Non-fallers: Prospective Study of Older Individuals. Frontiers in Physiology, 2018. 9.

39. Petrella, M., et al., Parâmetros do controle postural em mulheres idosas com ou sem histórico de quedas associadas ou não à osteoartrite de joelhos. Revista Brasileira de Reumatologia, 2012. 52(4): p. 512-517.

40. Fino, P.C., et al., Comparing Postural Stability Entropy Analyses to Differentiate Fallers and Non-fallers. Annals of Biomedical Engineering, 2016. 44(5): p. 1636-1645.

41. Park, J.W., M. Jung, and M. Kweon, The Mediolateral CoP Parameters can Differentiate the Fallers among the Community-dwelling Elderly Population. Journal of Physical Therapy Science, 2014. 26(3): p. 381-384.

42. Howcroft, J., et al., Elderly fall risk prediction using static posturography. PLOS ONE, 2017. 12(2): p. e0172398.

43. Yamagata, M., et al., Fall assessment in healthy older adults: Approach using rambling-trembling decomposition method. Clin Biomech (Bristol, Avon), 2024. 120: p. 106355.

44. Gregg, E., et al., A machine learning approach to identify important variables for distinguishing between fallers and non-fallers in older women. PLOS ONE, 2023. 18(10): p. e0293729.

45. Lázaro, M., et al., Postural stability in the elderly: Fallers versus non-fallers. European Geriatric Medicine, 2011. 2(1): p. 1-5.

46. Ricci, N.A., et al., Sensory interaction on static balance: A comparison concerning the history of falls of community-dwelling elderly. Geriatrics & Gerontology International, 2009. 9(2): p. 165-171.

47. Borel, L. and B. Alescio-Lautier, Posture and cognition in the elderly: Interaction and contribution to the rehabilitation strategies. Neurophysiologie Clinique/Clinical Neurophysiology, 2014. 44(1): p. 95-107.

48. Melzer, I., I. Kurz, and L.I.E. Oddsson, A retrospective analysis of balance control parameters in elderly fallers and non-fallers. Clinical Biomechanics, 2010. 25(10): p. 984-988.

49. Mahoney, J.R., K. Cotton, and J. Verghese, Multisensory Integration Predicts Balance and Falls in Older Adults. J Gerontol A Biol Sci Med Sci, 2019. 74(9): p. 1429-1435.

50. Wrisley, D.M., et al., Learning effects of repetitive administrations of the sensory organization test in healthy young adults. Arch Phys Med Rehabil, 2007. 88(8): p. 1049-54.

51. Bent, L.R., et al., Medio-lateral balance adjustments preceding reflexive limb withdrawal are modified by postural demands. Brain Research, 2001. 914(1): p. 100-105.

52. Pierchała, K., et al., Evaluation of the Sensory Organization Test to differentiate non-fallers from single- and multi-fallers. Adv Clin Exp Med, 2019. 28(1): p. 35-43.

53. Gera, G., et al., Identification of Balance Deficits in People with Parkinson Disease; is the Sensory Organization Test Enough? Int J Phys Med Rehabil, 2016. 4(1).

54. Zhu, R.T.-L., et al., Older Fallers and Non-fallers' Neuromuscular and Kinematic Alterations in Reactive Balance Control: Indicators of Balance Decline or Compensation? 2024.

55. Carty, C.P., P. Mills, and R. Barrett, Recovery from forward loss of balance in young and older adults using the stepping strategy. Gait & Posture, 2011. 33(2): p. 261-267.

10 Appendix

Appendix 1: search strategy:

Database	search strategy
Pedro (25 results)	*fall* balance elder* sensor**
Pubmed (3,623 results)	*((elder*[Title/Abstract]) OR (older*[Title/Abstract]) OR (aged[MeSH Terms]) OR (aged[Title/Abstract]) OR (Aging[Title/Abstract]) OR (Geriatric*[Title/Abstract]) OR (ageing[Title/Abstract])) AND ((Balanc*[Title/Abstract]) OR (stability[Title/Abstract]) OR (Stead*[Title/Abstract]) OR (Postur*[Title/Abstract]) OR (postural balance[MeSH Terms]) OR (Unstab*[Title/Abstract])) AND ((Fall*[Title/abstract]) OR (Fell[Title/abstract]) OR (Accidental falls[Mesh]) OR (Stumbl*[Title/abstract]) OR (Slip*[Title/abstract]) OR (Risk of fall*[Title/Abstract])) AND ((sensory reweighting[Title/Abstract]) OR (sensory reweighing[Title/Abstract) OR (sensory re-weighting[Title/abstract]) OR (Feedback, Physiological[Mesh]) OR (sensory integration[Title/abstract]) OR (sens*[Title/abstract]) OR (CTSIB[Title/abstract]) OR (SOT[Title/abstract]) OR (posturograph*[Title/abstract]) OR (sense organs[MeSH Terms]) OR (visu*[Title/Abstract]) OR (eyes open[Title/abstract]) OR (eyes closed[Title/abstract]) OR (vestibul*[Title/Abstract]) OR (propriocep*[Title/Abstract]) OR (somatosensory[Title/abstract]))*

Web of science (3,474 results)	*((AB=elder*) OR (AB=older*) OR (AB=aged]) OR (AB=Aging) OR (AB=Geriatric*) OR AB=(ageing)) AND ((AB=Balanc*) OR (AB=stability) OR (AB=Stead*]) OR (AB=Postur*) OR (AB=Unstab*)) AND ((AB=Fall*) OR (AB=Fell) OR (AB=Accidental falls) OR (AB=Stumbl*) OR (AB=Slip*) OR (AB=Risk of fall*)) AND ((AB=sensory reweighting) OR (AB=sensory reweighing) OR (AB=sensory re-weighting) OR (AB=sensory integration) OR (AB=sens*) OR (AB=CTSIB) OR (AB=SOT) OR (AB=posturograph*) OR (AB=sense organs) OR (AB=visu*) OR (AB=eyes open) OR (AB=eyes closed) OR (AB=vestibul*) OR (AB=propriocep*) OR (AB=somatosensory))*

Appendix 2: Table results quality assessment CASP:

Article	CASP		
	Section 1 Are the results of the study valid?	Section 2 What are the results?	Section 3 Will the results help locally?
Fino et al. (2016)[40]	Good	Good	Good
Gregg et al. (2023) [44]	Good	Good	Moderate
Howcroft et al. (2017) [42]	Good	Good	Moderate
Kim et al. (2011) [3]	Good	Good	Poor
Lázaro et al. (2011) [45]	Good	Good	Good
Liston et al. (2014) [36]	Moderate	Good	Good
Maranesi et al. (2016) [37]	Moderate	Good	Moderate
Müjdeci et al. (2012) [26]	Moderate	Good	Moderate
Park et al. (2014) [41]	Good	Good	Moderate
Petrella et al. (2012) [39]	Good	Good	Moderate
Ricci et al. (2009) [46]	Good	Good	Moderate
Van den Hoorn et al. (2018) [38]	Moderate	Good	Poor
Yamagata et al. (2024) [43]	Good	Good	Moderate

Appendix 3: Table results risk of bias assessment Downs and Black:

Article	Subscores					Total score /27	Interpretation: + : good = : fair - : poor
	Reporting	External Validity	Internal Validity Bias	Internal Validity Confounding	Power		
Fino et al. (2016) [40]	6	2	3	1	1	13	-
Gregg et al. (2023) [44]	8	2	3	1	1	15	=
Howcroft et al. (2017) [42]	8	2	3	1	0	14	-
Kim et al. (2011) [3]	6	2	2	2	0	12	-
Lázaro et al. (2011) [45]	8	3	3	3	1	18	=
Liston et al. (2014) [36]	7	3	2	3	1	16	=
Maranesi et al. (2016) [37]	8	3	4	3	0	18	=
Müjdeci et al. (2012) [26]	7	3	3	2	0	15	=
Park et al. (2014) [41]	7	2	4	3	0	16	=
Petrella et al. (2012) [39]	8	1	2	3	1	15	=

Ricci et al. (2009) [46]	8	3	2	3	1	17	=
Van den Hoorn et al. (2018) [38]	9	0	2	4	0	15	=
Yamagata et al. (2024) [43]	7	1	3	0	1	12	-

Appendix 4: Table results Systematic Review:

Author (year)	Characteristics of falling group N Age Height Weight # males # females	Characteristics of non-falling group N Age Height Weight # males # females	Test conditions	Outcome measurements	Results for test conditions	Included in meta-analysis #

Fino et al., 2016 [40]	30 74.4 (9) 167 (9) 76.8 (18) 23 7	45 74.4 (9) 167 (10) 73.9 (15.2) / /	Standing EO-FS Standing EC-FS	CoP parameters: - 95% ellipsoidal CoP area (cm^2) - CoP velocity (cm/s) - ML & AP standard deviation (x & y SD) (mm) Entropy measures: - RenyEn (mm) - ShanEn (mm) - ApEn (mm) - SaEn (mm) - MSE(mm) - CompMSE (mm) - RQAEn (mm)	No significant results were found	2 & 3

Gregg et al., 2023** [44]	21 71.52 (4.33) 162 (5) 65.43 (10.67) / 21	39 68.87 (3.41) 160 (6) 62.31 (10.55) / 39	mCTSIB: - EO-FS - EC-FS - EO-foam - EC-foam	CoP sway velocity (°/s) Symmetry angle EO/EC (%)	Right directional control: F > NF Anterior maximum excursion: F < NF	3
Howcroft et al., 2017 [42]	24 76.3 (7.0) 165.2 (10.3) 71.9 (14.3) 13 11	76 75.2 (6.6) 165.1 (10.0) 73.1 (13.4) 31 45	Standing in: - EO-FS - EC-FS	CoP parameters: - CoP Range, AP (mm) - CoP Range, ML (mm) - CoP RMS, AP (mm) - CoP RMS, ML (mm) - CoP Velocity, AP (mm/s) - CoP Velocity, ML (mm/s) - CoP Velocity, VSM (mm/s)	No significant results were found	/

Kim et al., 2011 [3]	15 71.4 (4.3) 156.3 (0.54) 61.3 (7.5) 8 7	15 72.1 (5.0) 156.5 (0.45) 60.9 (6.9) 8 7	SOT: - EO-FS - EC-FS - SRV-FS - EO-SRP - EC-SRP - SRV-SRP	CS (x/100)	CS: NF > F	1
Lázaro et al., 2011 [45]	99 78 (5) / / / /	113 78 (5) / / / /	mCTSIB - EO-FS - EC-FS - EO-foam - EC-foam	CoG sway velocity (°/s) CoG displacement (no unit reported)	Displacement: EO-foam: F> NF EC-foam: F> NF	/
Liston et al., 2014 [36]	25 76.6 (68-88) / / 21 4	16 74.5 (65-84) / / 13 3	SOT: - EO-FS - EC-FS - SRV-FS - EO-SRP - EC-SRP - SRV-SRP	CS (x/100)	CS: F< NF	1

Maranesi et al., 2016*** [37]	63 79.6 (6) 1.62 (0.07) 65.7 (14.6) 23 42	67 79 (5) 1.67 (0.10) 69 (14) 29 38	mCTSIB: - EO-FS - EC-FS - EO-foam - EC-foam	CoP parameters (no units reported): - CoP mean distance - CoP rms - CoP range - CoP mean velocity - CoP sway area	FF ≠ IF for: - CoP mean distance: EO-SRP AP - CoP rms distance: EC-FS AP - CoP range: EO-FS AP & ML EO-SRP AP & ML - CoP sway area: EO-SRP FF ≠ NF for: - CoP mean distance: EO-SRP AP - CoP mean distance: EO-FS AP EC-FS AP - CoP range: EO-FS AP & ML EC-FS ML EO-SRP AP & ML	2 & 3

					- CoP sway area: EO-SRP EC-FS	
Müjdeci et al., 2012 [26]	15 70.2 (4.39) / / / /	15 71.93 (6.11) / / / /	SOT: - EO-FS - EC-FS - SRV-FS - EO-SRP - EC-SRP - SRV-SRP	ES (x/100) CS (x/100)	ES SRV-FS: F< NF ES SRV-SRP: F< NF CS: F< NF	1
Park et al., 2014** [41]	8 78.9 / 56.16 3 26	21 78.9 / 56.16 3 26	2 conditions of the Dynamic Balance Measures: - EO-FS - EC-FS	CoP parameters (no units reported): - CoP range x/y axis - CoP distance x/y axis	No significant results were found	2

Petrella et al., 2012 [39]	11 72.72 (4.90) 153 (6) 65.00 (8.25) 21 0	21 66.62 (5.13) 154 (6) 62.47 (8.08) 35 0	mCTSIB for 60s: - EO-FS - EC-FS - EO-foam - EC-foam	CoP parameters: - CoP displacement AP & ML (cm)	ML CoP displacement: F>NF	/
Ricci et al., 2009* [46]	64 74.83 (6.87) / / 32 32	32 74.81 (7.25) / / 16 16	CTSIB: - EO-FS - EC-FS - SRV-FS - EO-foam - EC-foam - SRV-foam	Frequency of normal/abnormal cases in each group (n/%); - normal: able to maintain balance for 30s - abnormal: unable to maintain balance for 30s	Significant difference: - EO-foam: NF<FF & IF<FF - EC-foam: NF<FF	/

van den Hoorn et al., 2018 [38]	41 76 (5) 170 (8) 78 (17) 13 28	58 75 (6) 168 (10) 79 (15) 27 31	EC-FS	AP CoP - SP (mm/s) - %DET - Lmean - %LAM - TT - DFA1 - DFA2 - DFAtau	No significant results were found	/
Yamagata et al., 2024 [43]	46 79 (6) 156.5 (9.2) 52.9 (8.7) 19 27	16 77 (6) 31.9 (5.9) 55.9 (6.7) 10 6	EO-FS-FT EC-FS-FT	CoP parameters for both AP and ML: - CoP_mv (cm/s) - Rm_mv (cm/s) - Tr_mv (cm/s) - CoP_rms (cm) - Rm_rms (cm) - Tr_rms (cm)	Group effects (falling older adults VS non-falling older adults): - AP: Rm_mv Interaction effects: - AP: Rm_mv, CoP_rms, Rm_rms	/

Appendix 5: Data for meta-analysis:

MA #	Name of MA	Included studies	Outcome measurement per study	Fallers			Non-fallers		
				Mean	SD	n	Mean	SD	n
1	SOT composite score (0-100)	Kim et al. (2011) [3]	SOT composite equilibrium score	70.6	7.34	15	75.8	5.65	15
		Liston et al. (2014) [36]	SOT composite equilibrium score	48.71	16.58	25	65.19	16.17	16
		Müjdeci et al. (2012) [26]	SOT composite equilibrium score	77.06	3.47	15	81.66	3.37	15
2	Positional CoP measurements during ECFS	Fino et al. (2016) [40]	CoP 95% ellipsoidal area	5.77	6.51	30	4.51	5.02	45
		Maranesi et al. (2016) [37]	CoP mean distance	2.957***	0.264***	63*	2.85**	1.170**	67
		Maranesi et al. (2016) [37]	CoP mean range	18.921***	4.415***	63	18.35	8.758**	67
		Park et al. (2014) [41]	CoP mean range	2.37**	0.600**	8	2.385**	0.680**	21
		Park et al. (2014) [41]	CoP mean distance	26.95**	11.756**	8	28.75**	8.163**	21
3	Dynamic CoP Measurements during ECFS	Fino et al. (2016) [40]	Mean CoP velocity	2.91	1.34	30	2.7	1.29	45
		Maranesi et al. (2016) [37]	Mean CoP velocity	10.707***	0.741***	63*	9.7**	6.775**	67

	Gregg et al. (2023) [44]	CoP sway velocity	6.345**	3.375**	21	6.0**	2.730**	39

Notes: some values are added with one or more asterisks based on which calculations were performed to calculate this pooled mean or standard deviation.

one asterisk (*) indicates subgroups based on fall history were combined (e.g., single-fallers + multiple-fallers)

two asterisks (**) indicate results from different directions were combined (e.g., antero-posterior + medio-lateral)

three asterisks (***) indicate both of the above calculations were combined (all subgroups together)

Appendix 6: Checklist CASP

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Fino et al., 2016

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: This study compared the discriminatory ability of several entropy methods at differentiating two paradigms in the center-of-pressure (COP) of elderly individuals

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: They used a case controle study to find the difference between fallers and non fallers

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:
The data was collected as part of a free Community Fall Risk Screening Event that involved meetings with medical personnel and pharmacists, and a series of mobility tasks for the elderly at local community centers

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:
The data was collected as part of a free Community Fall Risk Screening Event that involved meetings with medical personnel and pharmacists, and a series of mobility tasks for the elderly at local community centers

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ☑
Can't Tell ☐
No ☐

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
Participants were asked to stand quietly with the arms at their sides for one minute while standing on a force plate (AMTI, Model # BP400600-2K, Advanced Mechanical Technology, Inc., Watertown, MA, USA) with their eyes open (EO). The participants were then given a short break to sit down before repeating the procedure with their eyes closed (EC).

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: Yes

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ☑
Can't Tell ☐
No ☐

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: the testing environment of the participants was not identical, nor completely free of distractions, The second limitation was that only 60 seconds of postural stability data were collected for each condition due to the difficulties with testing elderly subjects and the possibility of early fatigue

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

The cumulative classifiers had AUC ' s of 0.796 for the EO/EC paradigm and 0.908 for the
F/NF groups. The optimal EO/EC classifier had an AUC of 0.802 and consisted of all
entropy measures except MSESlope R, and MSESlope Ψ, ShanEn. The optimal F/NF classifier
had an AUC of 0.911 and consisted of all measures except ApEn R, ApEn Ψ, CompMSEArea
Ψ.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: All important variables were considerd

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The results were in accordacnce with our expectations

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The setting of this study is similar to other settings in western countries

11. Do the results of this study fit with other available evidence?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: statistical differences between fallers and non-fallers have been reported using ApEn 11, MSE 9 , and RQAEn 13, the present comparison between methods shows MSE can be a slightly better classifier than ApEn

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Gregg et al., 2023

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: this cross-sectional study aimed to identify those functional variables (i.e. balance, gait and clinical measures) and physical characteristics (i.e. strength and body composition) that could best distinguish between older female fallers and non-fallers, using a machine learning approach

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: A case control study is ideal because every subject undergoes the same test and is retrospectively divided in different groups, in this case fallers and non fallers

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:
randomly recruited from within the local community through a range of avenues, including liaising with Neighbourhood Network Schemes and the University of the Third Age. To advertise this study, several recruitment presentations were conducted, and flyers were distributed physically and online

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:
randomly recruited from within the local community through a range of avenues, including liaising with Neighbourhood Network Schemes and the University of the Third Age. To advertise this study, several recruitment presentations were conducted, and flyers were distributed physically and online

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias
- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
Participants attended three data collection sessions participants underwent screening of baseline characteristics, clinical measures, and balance assessment. During session two, gait analysis was performed, and during the final session, body composition and strength measurements were conducted

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed
- genetic
- environmental
- socio-economic

List: All subjects were treated identically

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ✔
Can't Tell
No

HINT: Look for
- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: general applicability of these results is perhaps limited to community-dwelling older women who were healthy and relatively active
medication usage was not considered in the exclusion criteria
sample size of the study was relatively smal

CASP
Critical Appraisal
Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

it is possible to discriminate between older female fallers and
non-fallers using a refined combination of variables and a multivariate machine learning approach.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments:. This model demonstrated an outstanding ability to distinguish between fallers and non-fallers (AUC0.90), with
19/21 fallers (sensitivity = 90%) and 34/39 non-fallers (specificity = 87%) classified correctly

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The results are more or less in accordance with our expectations

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: but only for community-dwelling older women who were healthy and relatively active

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments:

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Howcroft et al., 2017

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP
Critical Appraisal
Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes	✔
Can't Tell	
No	

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: This study investigated eyes open and eyes closed standing posturography with elderly adults to identify differences and determine appropriate outcome measure cut-off scores for prospective faller, single-faller, multi-faller, and non-faller classifications

2. Did the authors use an appropriate method to answer their question?

Yes	✔
Can't Tell	
No	

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: They used a case controle study to find differences for prospective faller, single-faller, multi-faller, and non-faller classifications

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:
local churches, retiree associations, and independent-living retirement homes

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:
local churches, retiree associations, and independent-living retirement homes

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes	✔
Can't Tell	
No	

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:

Two Wii Balance Boards (WBB) were placed such that their long axes were oriented parallel
to the AP axis, as described in [22]. Recent WBB studies reported good correspondence with
force platform measures [26,27], excellent test-retest reliability for one-board [26,28] and twoboard [29] configurations, and good to excellent concurrent validity [26,28] for CoP displacement measures. Participants stood in comfortable stance on two WBBs, with one foot on each
board. Participants stood quietly with eyes open and then with eyes closed, while WBB data
[30] were collected at 100 Hz for 30 seconds for each condition. Each condition was assessed
once for each participant

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: There was no experimentel intervention, every participant went through the same tests

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes	✔
Can't Tell	
No	

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: only six multi-fallers, the multi-faller sensitivity results for the determined cut-off scores were limited to one of seven levels, The
PMF population also weighed more than the other study participants, which could have influenced the posturography results,Medical information such as disease status, alcoholism,
etc. was not collected, and therefore no evaluation of the impact of these factors on posturography results could be performed, While longer or repeated posturography trials may be preferred for postural stability assessments, a single, short trial was used

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

Eyes-closed results were significantly greater than for eyes-open, for PNF, PF, and PSF group,
RQ for AP range and AP RMS were significantly greater for PNF than PA

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: All important variables were considerd

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The results were in accordacnce with our expectations

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The setting of this study is similar to other settings in western countries

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments:

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

CASP Critical Appraisal Skills Programme

Kim et al., 2011

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

CASP Checklist: 11 questions to help you make sense of a **Case Control Study**

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP
Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: They clearly stated three hypotheses they wanted to test and chose the apropriate population for this, namely elderly with and without fall-incidents in the past 12 months.

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: They conducted a case-control study which is able to answer all three of the hypotheses because they all try to investigate characteristics between groups.

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:

The study doesn't mention where the subjects were recruited from, but does have strict criteria for inclusion and exclusion. to differentiate between case and control, a clear definition was provided on what is considered a fall.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:

The study doesn't mention where the subjects were recruited from, but does have strict criteria for inclusion and exclusion. to be included as a control subject the participants can't have experienced a fall incident in the past 12 months, which is clearly described what can be considered as a fall.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
The Sensory organization test is a valid tool to measure the sensory integration, the other tests used (BBS, FRT, dynamometer) are valid tools as well. All tests were conducted for both cases and controls in the same manner.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ✔
Can't Tell
No

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: cases and controls were age- and sex-matched. All data was analyzed with multiple statistical tests as well.

Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

The FRT and the SOT in the falling group were significantly lower than that of the non-falling group. confounding factors were taken into account.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: All important variables were considered.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study is conducted in a well-organized manner, making bias less obvious.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☐
No ☑

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The sample size of the study is too small to generalize this result to all falling/non-falling elderly. Also the study was conducted in Korea where the lifestyle is very different to the one in western Europe.

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: One study found similar results, but another found the oposite results regarding differences in SOT-scores between fallers and non-fallers.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Lázaro et al., 2011

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The population is clearly stated (2 or more falls in the past 6 months vs no falls in past 6 months)

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: The study aims to compare static and dynamic balance alterations in recurrent fallers and non-fallers elderly patients and also tries to assess the functional repercussions of balance disorders and falls in these subjects. They performed a retrospective observational case-control study to do this which is an appropriate way to study this.

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:

All participants were recruited from 8 different institution where they went because of a fall incident.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:

All controls were recruited from the same 8 institutions but didn't visit the institution because of a fall incident.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias
- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
The testing and analysis was done very extensively and thus doesn't seem to include bias.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed
- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes
Can't Tell ✔
No

HINT: Look for
- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: There is no information given on certain tests or actions to minimise bias.

Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

They found recurrent fallers to show greater postural instability than non-fallers when visual and proprioceptive conditions changed. The coclusion was that mCTSIB assessment appears to be a sensitive tool for identifying those at high risk of recurrent falls.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: A 95% confidence interval was used for assessing all parameters.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ✔
No

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study is conducted in a proper way and doesn't have personal advantages by influencing the results.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ✔
Can't Tell
No

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: It's a broad population that is tested and it is done in a population similar to the Belgian population.

11. Do the results of this study fit with other available evidence?

Yes ✔
Can't Tell
No

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: This study also suggests that the center of gravity will displace more in people with recurrent falls when compared to people without falls in the past 6 months.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Liston et al., 2014

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The study aims were to (i) assess whether a greater proportion of older adult fallers have a peripheral vestibular impairment identified on the basis of history and comprehensive testing compared with age-matched healthy controls and (ii) compare subjective symptoms, balance confidence, falls risk and functional gait assessment (FGA) between older adults referred to a falls versus a neuro-otology clinic.

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: They used a case controle study to investigate if older fallers have vestibular impairments compared to non fallers and compare fall risk between the groups

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	✔
Can't Tell	
No	

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

Comments:

from falls clinics within
the Southwark and Lambeth Integrated Care Pathway for
Fallers (London, UK

Neuro-otology Department,
National Hospital for Neurology and Neurosurgery

4. Were the controls selected in an acceptable way?

Yes	
Can't Tell	✔
No	

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

Comments:

from community exercise classes

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ☑
Can't Tell ☐
No ☐

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
All the participants completed a routine neuro-otological examination
The Sensory Organization Test was performed according to a published protocol

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: There was no experimentel intervention, every participant went through the same tests

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ☐
Can't Tell ☐
No ☑

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments:

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

80% of fallers have a vestibular impairment compared with non-fallers.
The majority of older adult fallers with vestibular impairment do not report dizziness or vertigo symptoms.
Older adults referred to a falls versus a neuro-otology clinic
are older, more impaired, and report more falls.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: All important variables were considerd

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ✔
No

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The results were in accordacnce with our expectations

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ✔
Can't Tell
No

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The subjects and conditions are similar to other subjects and conditions in western communities

11. Do the results of this study fit with other available evidence?

Yes ✔
Can't Tell
No

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: The British and American Geriatrics Society joint guideline for falls prevention [30] recommends a balance, gait and mobility assessment, but vestibular testing is not explicitly incorporated. Our findings demonstrate current guidance may benefit from including vestibular testing, or increasing awareness of vestibular pathology when assessing older adults presenting with multiple non-syncopal falls

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Maranesi et al., 2016

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes	✔
Can't Tell	
No	

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: They specify the population they want to study thouroughly based on the number of fall incidents in the last year

2. Did the authors use an appropriate method to answer their question?

Yes	✔
Can't Tell	
No	

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: The study aims to analyze the association between the fall history of elderly and the values of a PCA-posturography conducted on those elderly. They used a case-control study to do so which is an appropriate way. After the analysis they want to determine a set of parameters able to differentiate groups characterized by different fall histories.

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	
Can't Tell	✔
No	

Comments:

The demographic, clinical and functional characteristics of all participants were compared for both groups and seem to be very similar. The cases are all visitors from the memory hospital so have an increased chance of being frail.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	
Can't Tell	✔
No	

Comments:

The controls are all visitors from the memory hospital so have an increased chance of being frail.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes	
Can't Tell	✔
No	

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
All participants were recruited from the same institute but were checked for important conditions. Also participants were controlled for age and gender.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes	✔
Can't Tell	
No	

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: For each clinical, functional and posturographic parameter, the Kruskal – Wallis test was applied to detect statistically significant differences. Also the demographic data seemed to be very similar between all groups.

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

They found a relationship between the number of falls in the past and the current findings on posturography.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: They used a 95% confidence interval to determine the significance of the results.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study is conducted in a proper way and the results are believeable.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The study has a large sample size and is conducted in Western Europe so is appliable to the local population.

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☐
No ☑

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: The results concerning the postural sway comparison between frequent fallers and non-fallers were in accordance with other studies. The results from looking at single-CoP parameters did not show a difference while literature does suppose a difference.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Müjdeci et al., 2012

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The population and method is clearly stated and appropriate for the topic.

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: The purpose of the study was to evaluate balance between fallers (2 or more falls in the last year) and non-fallers amongst the elderly. They used a cross-sectional case control study which is an appropriate design to study this.

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	
Can't Tell	✔
No	

Comments:

There is very little information on the recruitment of the both the cases and the controls.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	
Can't Tell	✔
No	

Comments:

There is very little information on the recruitment of the both the cases and the controls.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
The definition of falls is stated very clearly and the tests used in the study have been shown to be reliable for this parameter.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes
Can't Tell ✔
No

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: Very little information is given about the statistical tests conducted for the study.

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

CDP and BBS scores in fallers were found to be significiantly lower as compared to the non-fallers elderly in this study.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: Results were seen as significant when they had a P-value < 0,05. The important variables were taken into account in the measurements.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study has been conducted properly but contains some language mistakes that make the believeability less obvious.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The population was only 30 people (15 fallers vs 15 non-fallers) and there is no information about the global region the participants were recruited from.

11. Do the results of this study fit with other available evidence?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: Previous studies found similar results about detecting a significant correlation between SOT and falling. This study however did not verify another study that states LOS can be used for forecasting the risk of falling.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Park et al., 2014

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The study clearly states which topic and population it tries to investigate.

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: The study aims to investigate the differences in ML CoP parameters between elderly fallers and non-fallers during dynamic postural control.

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:

They don't mention in which geografical area the participants were recruited. The inclusion and exclusion criteria were clearly stated.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:

The criteria for inclusion/exclusion were the same as those for the cases and were appropriate for this study.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
The participants had to report their falling history of the past year themselves which is a subjective measurement prone to recall bias. The tests were conducted on a very standardize manner with clear instructions (movements on a specific pace and position during tasks being determined by the protocol).

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ✔
Can't Tell
No

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: Fatigue was taken into account during the testing, for which an appropriate resting period was given. Also the learning effect was taken into account for and diminished by using a random order of tasks. No statistical differences were found in general characteristics between the two groups.

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

The results were adjusted to minimize influencing factors. To minimize noise, the data was filtered and the first 6 and last 3 seconds were cut off

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: A 95% confidence interval was used and showed a significant difference in postural sway between fallers and non-fallers.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study is done in a trustful manner which makes the results likely to be believed. They did a lot of effort to minimize bias and confounding.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☐
No ☑

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: It only includes a small number of subjects and the distribution of males vs females in both groups is not representative for the population (very large number of women compared to men)

11. Do the results of this study fit with other available evidence?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: The fact that fallers demonstrated increased postural sway in static postural control tasks compared with the non-faller group was consistent with the result of a previous study.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

CASP Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Petrella et al., 2012

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ✔
Can't Tell
No

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The study clearly defines the different populations studied (FOA, FNOA, NFOA, NFNOA) and the different parameters assessed (APD, MLD)

2. Did the authors use an appropriate method to answer their question?

Yes ✔
Can't Tell
No

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: They conducted a case control study which is an appropriate protocol to compare stabilometric parameters of elderly female fallers and non-fallers associated or not with knee osteoarthritis.

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:

The inclusion and exclusion criteria are strict which limits the influence of different factors. The way patients were recruited and the geographical area is not mentioned. There is a clear discrimination between the different groups.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☑
Can't Tell ☐
No ☐

Comments:

The way they recruited participants and the geographical area are not mentioned, but the inclusion and exclusion criteria make influencing factors less possible. There is a clear discrimination between the different groups.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes	✔
Can't Tell	
No	

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
The measurements used are done with standardized tests and are appropriate to measure the sensory integration.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes	✔
Can't Tell	
No	

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: the anthropometric characteristics were compared between the groups using a two-way ANOVA.

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

There was no interaction between the factors fall and disease (OA) for both APD and MLD. Elderly females with OA had a greater AP displacement compared to the group without OA. Fall has a greater influence on the ML balance, regardless of the presence of OA. Confounding was taken into account.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: Statistical tests were conducted with a 95% confidence interval.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study adressed everything they did properly and doesn't seem to be influenced by important factors and thus presents with believeable results.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☐
No ☑

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The study was conducted only on females and thus isn't representative for the whole population.

11. Do the results of this study fit with other available evidence?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: In general most of the findings support previously reported findings, some factors differ a little from what was shown in other studies but have conflicting evidence in other studies as well.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

Ricci et al., 2009

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes	✔
Can't Tell	
No	

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The population and test procedure were clearly stated and not too broad.

2. Did the authors use an appropriate method to answer their question?

Yes	✔
Can't Tell	
No	

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: They wanted to determine whether elderly subjects with distinct histories of falls presented differences concerning the influence of sensory interaction on balance, and did this using a cross-sectional research which is an appropriate way to do this.

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:

The participants were recruited in Brazil without a clearly stated system to recruit them. The cases and controls represented the same age and sex.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:

The participants were recruited in Brazil without a clearly stated system to recruit them. The cases and controls represented the same age and sex.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
They used an objective tool to measure the parameter they wanted to measure and tried to minimize bias by making both groups around the same age or sex.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List: /

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ✔
Can't Tell
No

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: They compared data from both groups to check that there were no big differences between the two groups.

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

The study found a variation in sensorial interaction in the elderly according to their history of falls. The recurrent fallers were found to be more dependent on visual information than the non-falling group.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: The P-values for the measured differences are all close to P = 0,05. Important variables were considered and mentioned in the study.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The results show results that are objectively measured in a well-conducted study and are easily believeable.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: The study is performed in Brazil which makes it hard to determine if the differences in lifestyle and healthsystem could lead to different results than in a Western European population.

11. Do the results of this study fit with other available evidence?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: The results they found fit with other available evidence also showing differences in sensory integration between fallers and non-fallers.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

CASP
Critical Appraisal Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

van den Hoorn et al., 2018

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: The topic and population of the study is stated very clearly and extensively.

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: The study aimed to investigate the influence of calf vibration on balance control in falling and non-falling elderly and in young adults. They used a prospective case-control study to do this.

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	
Can't Tell	✔
No	

Comments:

The participants were a subset of another study but they didn't randomly take them, instead they tried to take PPA values under 0,5 or above 1 to have a wide variety in results. This might influence the results since the population in between this score is not included.

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	
Can't Tell	✔
No	

Comments:

The participants were a subset of another study but they didn't randomly take them, instead they tried to take PPA values under 0,5 or above 1 to have a wide variety in results. This might influence the results since the population in between this score is not included.

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias
- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
The falls were measured with a weekly fall diary to minimize recall bias.

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed
- genetic
- environmental
- socio-economic

List: Yes

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ✔
Can't Tell
No

HINT: Look for
- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: Differences between the demographics and PPA of the fallers and non-fallers were assessed and did not differ significantly. Differences between groups at baseline were assessed and bonferonni was used post-hoc.

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

non-linear measures of CoP motion reveal differences in the quality of balance control between young and old individuals and between older individuals who do and do not go on to fall.

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: The study was conducted with a 95% confidence interval. All subjects that dropped out of the study were not included in the statistics but were mentioned.

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ☑
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: The study uses a lot of specific parameters to analyse the postural control and does find clear differences in multiple parameters which makes the result easy to believe.

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: It is a large population studied but the geographic location of the recruitment is not mentioned an thus deciding if the results are appliable to the local population is not easily done.

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☑
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: Some studies resulted in the same findings but others didn't which makes the literature hard to interpret.

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Yamagata et al., 2014

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP Critical Appraisal Skills Programme

Paper for appraisal and reference:

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☑
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments: It tries to find indices of potential fall risk in posturography

2. Did the authors use an appropriate method to answer their question?

Yes ☑
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments: The study compares the outcomes of fallers and non-fallers to find a difference

CASP
Critical Appraisal Skills Programme

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:
an invitation was sent to participants of a previous study that adhered to the recuired in and exclusion criteria

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes	✔
Can't Tell	
No	

Comments:
Strict in and exclusion criteria were used

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

CASP
Critical Appraisal Skills Programme

5. Was the exposure accurately measured to minimise bias?

Yes ✔
Can't Tell
No

HINT: We are looking for measurement, recall or classification bias
- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:
: timed
up and go test, 10-m walk test, functional reach test, one-leg stand test,
isometric muscle strength tests, open-close stepping test, and sit-to-stand
test. All clinical measures were performed with shoes on and evaluated
by several physical therapists with adequate skills and knowledge

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed
- genetic
- environmental
- socio-economic

List: There was no difference in treatment of both groups as there was no experiment

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ✔
Can't Tell
No

HINT: Look for
- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments: The study has dedicated a section to the study limitations

CASP
Critical Appraisal Skills Programme

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

In the AP direction, a significant Vision effect was found (CoP_rms F (1, 60) = 29.3; p < 0.001; ηp2 = 0.33, Rm_rms F (1, 60) = 7.8; p < 0.01; ηp2 = 0.12, Tr_rms F (1, 60) = 34.9; p < 0.001; ηp2 = 0.36); however, no Group effects were found. We found a significant Group x Vision interaction in CoP_rms (F (1, 60) = 8.7, p < 0.01; ηp2 = 0.13) and Rm_rms (F (1, 60) = 7.8, p < 0.01; ηp2 = 0.12) in the AP direction; however, no significant differences between fallers and non-fallers were found in either the EO or EC conditions

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments: All important variables were considerd

CASP
Critical Appraisal Skills Programme

9. Do you believe the results?

Yes ✔
No

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments: No effects of clinical test can be found which is in accordance with other literature, however for the posturograph does this study use a method i am not familiar with and this seems to find a difference in a particular part

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ✔
Can't Tell
No

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments: These are all healthy older adults

11. Do the results of this study fit with other available evidence?

Yes ✔
Can't Tell ✔
No

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments: Parts do, other parts i can't tell

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Appendix 7: Checklist Downs and Black

Fino et al., 2016

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (1)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	x			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	x			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	x			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	x			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***		x		
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	x			
7. **Does the study provide estimates of the random variability in thedata forthe mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	x			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***		X		

9. **Have the characteristics of subjects lost to follow-up been described?** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			x	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**		x		
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	x			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***				x
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***	x			

Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes*.**			x	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		x		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***		x		
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no*.**			x	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	x			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***	x			

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				

20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	x			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***			x	
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***			x	
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			x	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, itshould be answered no.***			x	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	x			

26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow- up was too small to affect the main findings, the question should be answered yes.***			x	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detecta difference of x% and y.***	x			
Totale score:	13/27			

Gregg et al., 2023

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	x			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	x			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	x			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	x			
5. **Are the distributions of principal confounders in each group of subjects to be compared clearly described?** ***A list of principal confounders is provided***	x			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	x			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	x			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			x	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			x	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	x			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of***	x			

all members of the relevant population exists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	x			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***			x	
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes*.**			x	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?			x	
16. **If any of the results of the study were based on "data dredging", was this made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***	x			
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no*.**			x	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	x			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			x	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? *For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.*	x			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***			x	
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***			x	
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			x	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			x	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	x			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***			x	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***	x			
Totale score:	15/27			

Howcroft et al., 2017

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	x			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	x			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	x			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	x			
5. **Are the distributions of principal confounders in each group of subjects to be compared clearly described?** ***A list of principal confounders is provided***	x			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	x			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	x			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			x	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			x	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	x			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of***	x			

all members of the relevant population exists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	x			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***			x	
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes*.**			x	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?			x	
16. **If any of the results of the study were based on "data dredging", was this made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***	x			
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no*.**			x	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	x			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			x	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? *For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.*	x			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***			x	
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***			x	
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			x	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			x	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	x			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***			x	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***		X		
Totale score:				

Kim et al., 2011

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***			X	
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***		X		
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			X	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			X	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	X			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***			X	
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.***			X	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***		X		
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.***			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***		X		
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***			X	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***		X		
Totale score:	12			

Lázaro et al., 2011

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***	X			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			X	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			X	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	X			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** *The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.*	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** *For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.*	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** *For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.*			X	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?				X
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** *Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.*			X	
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** *Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.*	X			
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** *Where there was noncompliance with the allocated treatment or*			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***			X	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***	X			
Totale score:	18			

Liston et al., 2014

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***		X		
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***		X		
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***		X		
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	X			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.***		X		
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***		X		
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.***			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***		X		
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***		X		
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***		X		
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***	X			
Totale score:	16			

Maranesi et al., 2016

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***	X			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***		X		
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			X	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	X			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.***			X	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***	X			
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.***			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***	X			

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***			X	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***		X		
Totale score:	18			

Müjdeci et al., 2012

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***		X		
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***		X		
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***		X		
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	X			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.***		X		
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***		X		
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.***	X			
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***		X		
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***		X		
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***		X		
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***		X		
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***		X		
Totale score:	15			

Park et al., 2014

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***		X		
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			X	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***		X		
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***		X		

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** *The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.*	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** *For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.*	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** *For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.*		X		
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** *Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.*	X			
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** *Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.*			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** *Where there was noncompliance with the allocated treatment or*	X			

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***		X		
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***		X		
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***		X		
Totale score:	16			

Petrella et al., 2012

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***	X			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			X	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***			X	
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***				X

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***				X
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.***			X	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?				X
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***			X	
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.***			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***			X	
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***	X			
Totale score:	15			

Ricci et al., 2009

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***	X			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			X	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***		X		
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***	X			

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***	X			
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***	X			
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.***			X	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***			X	
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.***			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***		X		
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***	X			
Totale score:	17			

van den Hoorn et al., 2018

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to becompared clearly described?** ***A list of principal confounders is provided***	X			
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. Innormally distributed data the standard error, standard deviation or confidence intervalsshould be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the interventionbeen reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***		X		
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***	X			
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of all members***		X		

of the relevant populationexists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** *The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.*		X		
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** *For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.*				X
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** *For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.*				
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?				
16. **If any of the results of the study were based on "data dredging", wasthis made clear?** *Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.*			X	
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** *Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no.*				X
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** *Where there was noncompliance with the allocated treatment or*			X	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	X			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***	X			
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***	X			
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***	X			
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***	X			
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***				X
Totale score:	15			

Yamagata et al., 2024

Downs and Black: Risk of bias assessment

Reporting	Yes (1)	No (0)	N/A (0)	Unable to determine (0)
1. Is the hypothesis/aim/objective of the study clearly described?	X			
2. **Are the main outcomes to be measured clearly described in the Introduction or methods section?** ***If the main outcomes are first mentioned in the results section, the question should be answered as no.***	X			
3. **Are the characteristics of the subjects included in the study clearly described?** ***In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and source for controls should be given.***	X			
4. **Are the interventions of interest clearly described?** ***Treatments and placebo (where relevant) that are to be compared should be clearlydescribed.***	X			
5. **Are the distributions of principal confounders in each group of subjects to be compared clearly described?** ***A list of principal confounders is provided***		X		
6. **Are the main findings of the study clearly described?** ***Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical test which are considered below)***	X			
7. **Does the study provide estimates of the random variability in the data for the mainoutcomes?** ***In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If distribution data is not described, it must be assumed that the estimates used were appropriate and the question should be answered with yes.***	X			
8. **Have all important adverse events that may be a consequence of the intervention been reported?** ***This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).***			X	
9. **Have the characteristics of subjects lost to follow-up been described*?*** ***This should be answered yes where there were no losses to follow-up or where lossesto follow-up were so small that findings would be unaffected by their inclusion. This should be answered no, where a study does not report the number of patients lost to follow-up.***		X		
10. **Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than0.001?**	X			
External validity				
11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** ***The study must identify the source population for subjects and describe how the subjects were selected. Subjects would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample.Random sampling is only feasible where a list of***	x			

all members of the relevant population exists. Where a study does not report the proportion of the source population from which the subjects are derived, the question should be answered as unable to determine.				
12. **Were those subjects who were prepared to participate representative of theentire population from which they were recruited?** ***The proportion of those asked who agreed should be stated. Validation thatthe sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.***		X		
13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** ***For the question to be answered yes, the study should demonstrate that theintervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.***		x		
Internal Validity				
14. **Was an attempt made to blind study subjects to the intervention theyhave received?** ***For studies where the patients would have no way of knowing which intervention they received, this should be answered yes*.**			X	
15. Was an attempt made to blind those measuring the main outcomes ofthe intervention?		X		
16. **If any of the results of the study were based on "data dredging", was this made clear?** ***Any analyses that had not been planned at the outset of the study should beclearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.***	X			
17. **In trials and cohort studies, do the analyses adjust for different lengths offollow-up, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** ***Where follow-up was the same for all study subjects the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow up areignored should be answered no*.**			X	
18. Were the statistical tests used to assess the main outcomes appropriate?The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.	X			
19. **Was compliance with the intervention/s reliable?** ***Where there was noncompliance with the allocated treatment or***			x	

where there was contamination of one group, the question should be answered no.For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.				
20. Were the main outcome measures used accurate (valid and reliable)? ***For studies where the outcome measures are clearly described, the questionshould be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.***	x			
Internal Validity – confounding (selection bias)				
21. **Were the subjects in different intervention groups or were they recruited from the same population?** ***For example, patients for all comparison groups should be selected from thesame hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.***			X	
22. **Were study subjects in different intervention groups or were theyrecruited over the same period of time?** ***For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.***			X	
23. **Were study subjects randomised to intervention groups?** ***Studies which state that subjects were randomized should be answered yesexcept where method of randomisation would not ensure random allocation.For example, alternate allocation would score no because it is predictable.***			X	
24. **Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** ***All non-randomised studies should be answered no. If assignment wasconcealed from patients but not from staff, it should be answered no.***			X	
25. **Was there adequate adjustment for confounding in the analyses from which themain findings were drawn?** ***This question should be answered no for trials if: the main conclusions of thestudy were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.***		X		
26. **Were losses of subjects to follow-up taken into account?** ***If the numbers of subjects lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.***		x		
Power				

27. **Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than5%?** ***Sample sizes have been calculated to detect a difference of x% and y.***	X			
Totale score:	12/27			