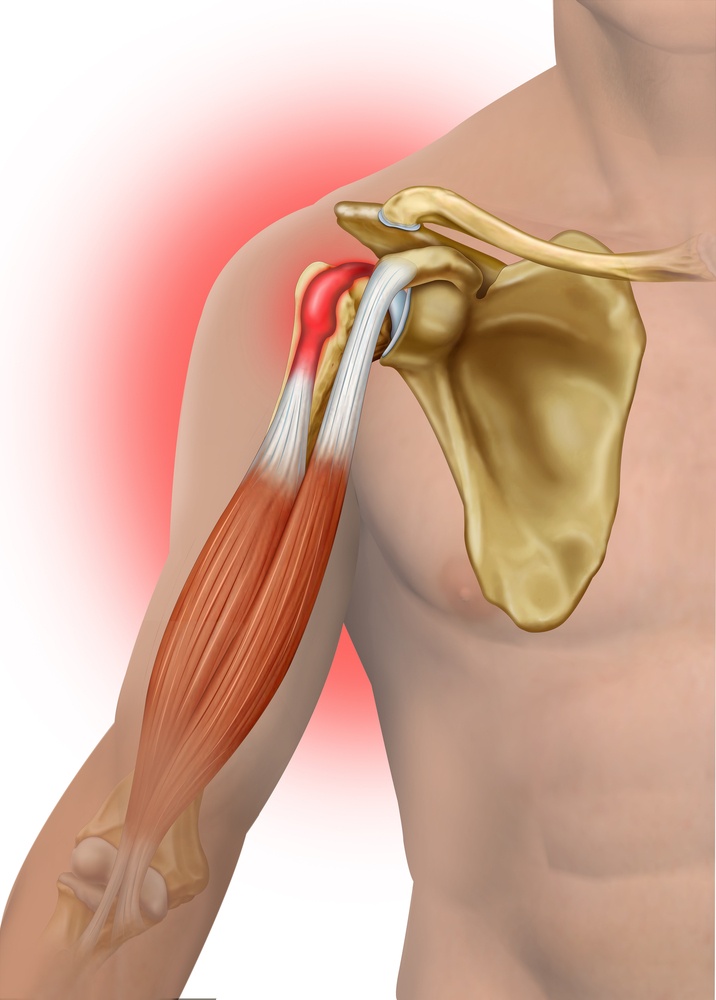
**‘Wat zijn de meest valide diagnostische middelen om caput longum bicepspees tendinopathie zo accuraat mogelijk te kunnen stellen in de algehele patiëntenpopulatie?’**



**Student:** Jesper Klijnstra



|

**Studentnummer**: 35453

**Scriptiebegeleider/ supervisor:** Marcel van Burgsteden

**Datum/Date:** 02-07-2021

# Voorwoord

Beste lezer,

In het kader van de afstudeerfase aan de Hanze Hogeschool Groningen voor de bacheloropleiding Fysiotherapie, presenteer ik u hier mijn literatuurstudie. In deze studie wordt de validiteit van de diagnostische middelen omtrent caput longum bicepspees tendinopathie onderzocht. Er is vanaf april 2021 tot juli 2021 gewerkt aan deze literatuurstudie.

Door het uitvoeren van een zelfstandig onderzoek, heb ik met name veel geleerd binnen de competentie wetenschap. Het academisch schrijven en het benaderen van de studies is in de afgelopen periode erg verbeterd. Tijdens de afstudeerfase ben ik begeleid door Marcel van Burgsteden. Marcel, bedankt voor de ondersteuning en hulp tijdens de wekelijkse Teams meetings.

Ten slotte wil ik mijn vrienden en familie bedanken voor hun ondersteuning in de afgelopen periode.

Ik wens u veel leesplezier,

Jesper Klijnstra

Inhoud

[Voorwoord 2](#_Toc76137476)

[Samenvatting 4](#_Toc76137477)

[Abstract 5](#_Toc76137478)

[Inleiding 6](#_Toc76137479)

[Methode 7](#_Toc76137480)

[Databases en zoekstrategieën 7](#_Toc76137481)

[Selectiecriteria 8](#_Toc76137482)

[Methodologische kwaliteitsbepaling 8](#_Toc76137483)

[Data-extractie 9](#_Toc76137484)

[Data-analyse 9](#_Toc76137485)

[Resultaten 10](#_Toc76137486)

[Selectie van de studies 10](#_Toc76137487)

[Methodologische kwaliteit van artikelen 11](#_Toc76137488)

[Beschrijving van de studies 12](#_Toc76137489)

[Onderzoekspopulatie 13](#_Toc76137490)

[Interventie 13](#_Toc76137491)

[Uitkomstmaten 14](#_Toc76137492)

[Resultaten 14](#_Toc76137493)

[Discussie 18](#_Toc76137494)

[Belangrijkste resultaten studies 18](#_Toc76137495)

[Kritische beschouwing studies 19](#_Toc76137496)

[Vergelijking resultaten met bestaande literatuur 20](#_Toc76137497)

[Sterke punten literatuurstudie 20](#_Toc76137498)

[Zwakke punten literatuurstudie 21](#_Toc76137499)

[Aanbeveling voor vervolgonderzoek 21](#_Toc76137500)

[Aanbeveling voor de praktijk 22](#_Toc76137501)

[Conclusie 22](#_Toc76137502)

[Referentielijst 23](#_Toc76137503)

[Bijlagen 25](#_Toc76137504)

[Bijlage 1 Zoekstring 25](#_Toc76137505)

[Bijlage 2 Vrije zoektermen en MeSH term 25](#_Toc76137506)

[Bijlage 3 scorelijst STROBE 26](#_Toc76137507)

[Bijlage 4 scorelijst STROBE 28](#_Toc76137508)

[Bijlage 5 scorelijst STROBE 31](#_Toc76137509)

[Bijlage 6 scorelijst STROBE 34](#_Toc76137510)

[Bijlage 7 scorelijst STROBE 36](#_Toc76137511)

[Bijlage 8 scorelijst AMSTAR 39](#_Toc76137512)

# Samenvatting

**Aanleiding:** Jaarlijks gezien komen 31% van de patiënten bij een fysiotherapeut of huisarts om aan hun schouderklacht te worden geholpen. Subacromiale klachten veroorzaakt door rotator cuff pezen dan wel caput longum bicepspees zijn het meest voorkomend. Naast tendinopathie spelen instabiliteit en traumatische oorzaken een rol in het pathologische beeld van de caput longum bicepspees. Binnen de diagnostiek kunnen er meerdere middelen worden ingezet om tussen desbetreffende pathologieën onderscheid te maken. Gezien het feit dat tendinopathieën ofwel geïsoleerd ofwel ten gevolge van onderliggende pathologie voorkomen dient de diagnostische standaard zo valide mogelijk te zijn. De moeilijkheidsgraad van het diagnosticeren ligt hoog, waardoor accuraatheid op zijn plaats is in het proces van diagnostiek. De onderzoeksvraag die hieruit volgt luidt: ‘wat zijn de meest valide diagnostische middelen om caput longum bicepspees tendinopathie zo accuraat mogelijk te kunnen stellen in de algehele patiëntenpopulatie?’

**Doelstelling:** Het doel van deze literatuurstudie was het definiëren van de diagnostische validiteit, om zo antwoord te geven op de onderzoeksvraag en daarmee de benadering van de klacht te vereenvoudigen.

**Methode:** De literatuur die is gebruikt voor deze literatuurstudie is gevonden in de databanken PubMed, Google Scholar en ScienceDirect. De mogelijk relevante artikelen zijn geselecteerd en gefilterd aan de hand van de vooraf opgestelde in- en exclusiecriteria: de studies moesten gericht zijn op de diagnostische middelen fysieke tests en beeldvormend onderzoek als interventie. De methodologische kwaliteit van de geïncludeerde studies werd beoordeeld middels de AMSTAR en de STROBE. De sensitiviteit, specificiteit, PVW, NVW, accuraatheid, LR+ en LR- werden binnen de gebruikte interventies geëxtraheerd.

**Resultaten:** van de in totaal zes geïncludeerde studies, is er onderzoek uitgevoerd naar vijf fysieke tests en 3 naar beeldvormend onderzoek. De Yergason’s test heeft een gemiddelde sensitiviteit van 45,4%, specificiteit van 73,4%, PVW van 57%, NVW van 59,3%, accuraatheid van 60,8%, LR+ van 2.10 en een LR- van 0.76. De Speed’s test heeft een gemiddelde sensitiviteit van 72,5%, specificiteit van 50,1%, PVW van 77,3%, NVW van 67,1%, accuraatheid van 57,9%, LR+ van 1.55 en een LR- van 0.54. De palpatie van de sulcus intertubercularis heeft een sensitiviteit van 57,1%, specificiteit van 72,4%, PVW van 57%, NVW van 72%, LR+ van 2.07 en een LR- 0.59. De interventie CT-artrografie heeft een gemiddelde sensitiviteit van 74,8%, specificiteit van 96%, PVW van 96,3%, NVW van 71,9%, accuraatheid van 84%, LR+ van 12.97 en een LR- van 0.59.

**Conclusie:** CT-artrografie dient als meest valide diagnostische middel voor het in- en uitsluiten van de diagnose caput longum bicepspees tendinopathie. Daarnaast zijn de Yergason’s test en de Speed’s test geschikt voor het in- of uitsluiten van de diagnose. Vervolgonderzoek wordt aanbevolen om de validiteit van een cluster fysieke tests te onderzoeken.

# Abstract

**Introduction**: Every year, 31% of patients visit a physiotherapist or general practitioner to be helped with their shoulder complaints. Subacromial complaints caused by rotator cuff tendons or caput longum biceps tendon are the most common. In addition to tendinopathy, instability and traumatic causes play a role in the pathological picture of the caput longum biceps tendon. Within diagnostics, several means can be used to distinguish between the relevant pathologies. Given that tendinopathies occur either in isolation or as a result of underlying pathology, the diagnostic standard should be as valid as possible. The difficulty of diagnosing is high, which means that accuracy is in order in the process of diagnosis. The research question that follows from this is: 'what are the most valid diagnostic tools to determine caput longum biceps tendinopathy as accurately as possible in the general patient population?'

**Methods**: The literature used for this literature study was found in the PubMed, Google Scholar and ScienceDirect databases. The possibly relevant articles were selected and filtered according to the predefined inclusion and exclusion criteria: the studies had to focus on the diagnostic means, physical tests and imaging research as intervention. The methodological quality of the included studies was assessed using the AMSTAR and the STROBE. The sensitivity, specificity, PVW, NVW, accuracy, LR+ and LR- were extracted within the used interventions.

**Results**: Out of a total of six studies included, five physical tests and three imaging tests were investigated. The Yergason's test has an average sensitivity of 45.4%, specificity of 73.4%, PVV of 57%, NPV of 59.3%, accuracy of 60.8%, LR+ of 2.10 and an LR- of 0.76. The Speed's test has an average sensitivity of 72.5%, specificity of 50.1%, PVV of 77.3%, NPV of 67.1%, accuracy of 57.9%, LR+ of 1.55 and an LR- of 0.54 . The palpation of the intertubercular sulcus has a sensitivity of 57.1%, specificity of 72.4%, PVV of 57%, NPV of 72%, LR+ of 2.07 and an LR-0.59. The interventional CT arthrography has a mean sensitivity of 74.8%, specificity of 96%, PV of 96.3%, NPV of 71.9%, accuracy of 84%, LR+ of 12.97 and an LR- of 0.59.

**Conclusion**: CT arthrography is the most valid diagnostic tool for including and excluding the diagnosis of caput longum biceps tendinopathy. In addition, the Yergason's test and the Speed's test are suitable for including or excluding the diagnosis. Follow-up research is recommended to investigate the validity of a cluster of physical tests.

# Inleiding

Schouderklachten zijn veelvoorkomende klachten van het bewegingsapparaat. Gemiddeld gezien komen jaarlijks 31% van de patiënten bij een fysiotherapeut of huisarts om aan hun schouderklacht te worden geholpen (1) (2). De fysiotherapeutische interventies die het beste resultaat geven voor het verhelpen van schouderklachten kunnen gericht zijn op zowel oefentherapie als manuele therapie (3) (4).

Schouderklachten kunnen worden onderverdeeld in de driedeling subacromiaal pijnsyndroom (SAPS), glenohumerale schouderklachten en overige schouderklachten. Het subacromiale pijnsyndroom omvat subacromiale bursitis, (calcificerende) tendinopathie, (partiële dikte, volledige dikte of complete) ruptuur van een of meerdere pezen van de rotator cuff of ruptuur van de lange kop van de bicepspees. Het subacromiale pijnsyndroom heeft binnen het totaal aantal schouderklachten in de huisartsenpraktijk een prevalentie van 80% (5). Onder de glenohumerale schouderklachten worden voornamelijk frozen shoulder en artrose gezien. De frozen shoulder komt het meest frequent voor in de leeftijdsgroep 40 tot 60 jaar, terwijl artrose van de schouder een beeld is dat vooral gezien wordt bij de leeftijdsgroep boven de 60 jaar (5). De overige schouderklachten zijn een combinatie van glenohumerale instabiliteit, afwijkingen van het acromioclaviculaire gewricht, functiestoornissen van de cervicale wervelkolom en/of cervicothorocale overgang en scapula dyskinesie (5).

Binnen het subacromiale pijnsyndroom hebben tendinopathieën een prevalentie van 29% (5) en het beeld varieert hierbij in een ontsteking van de rotator cuff pezen als in een ontsteking van de caput longum bicepspees. Bij het gebruiken van de term tendinopathie moet er rekening gehouden worden met verschillende fases waarin de pees zich kan bevinden. De reactieve tendinopathie wordt gezien als de eerste fase. Dit is een proliferatie reactie in de extracellulaire matrix als gevolg van (acute) druk- of trekbelastingen. De cellen veranderen van vorm en er ontstaat een toename van proteïne productie door dat er meer cytoplasmatische organellen ontstaan (6). De tweede fase wordt gekenmerkt als zijnde pees verval. Hier staat een toename van het aantal cellen als chondrocyten en myofibroblasten, met als gevolg een verhoging van de proteïne productie, op de voorgrond. Deze toename leidt vervolgens tot afscheiding in de collageenvezels en tot desorganisatie in de extracellulaire matrix (6). Binnen de degeneratieve tendinopathie, ook wel de laatste fase, zijn er veel ruimtes met dode cellen en is er een gehele desorganisatie van de extracellulaire matrix. De ruimtes zijn gevuld met vezels en er is een kleinschalige productie van collageen wat de kans op pathofysiologische veranderingen verkleind (6).

Anatomisch gezien komt de bicepspees voort uit de lange kop van de M. Biceps Brachii (7). Ontsteking van de bicepspees in de sulcus intertubercularis wordt primaire biceps tendinopathie genoemd en komt voor 5% binnen de biceps tendinopathieën voor (8). De overige 95% van het tendinopathische beeld is een gevolg van onderliggende pathologie zoals rotator cuff klachten of een SLAP laesie (8). De caput longum biceps tendinopathie komt vaak voor in de patiëntenpopulatie 18-35 jaar die betrokken zijn bij sporten als zwemmen, gymnastiek en werpsporten (8). Verder kunnen degeneratie en overbelasting tijdens bovenhandse activiteiten ook zorgen voor het ontstaan van desbetreffende tendinopathie. Naast de tendinopathie van de lange kop van de M. Biceps Brachii pees spelen instabiliteit en traumatische oorzaken ook een rol in het pathologische beeld van de lange kop van de M. Biceps Brachii pees.

Binnen de diagnostiek kunnen er meerdere middelen worden ingezet om tussen deze pathologische beelden onderscheid te maken. Er wordt gebruik gemaakt van zowel beeldvormend onderzoek als van klinische tests (9). Om tussen het beeldvormende onderzoek en de klinische tests onderscheid te maken is het ontdekken van de ernst van de schouderpijn van belang. In de literatuur binnen de bicepspees aandoeningen wordt voornamelijk het algehele beeld beschreven van de pathologieën. Deze pathologieën van de bicepspees kunnen het gevolg zijn van een ontsteking, een traumatische oorzaak (met als gevolg een ruptuur) of instabiliteit (10).

Gezien het feit dat tendinopathieën ofwel geïsoleerd ofwel ten gevolge van onderliggende pathologie voorkomen moet de diagnostische standaard zo valide mogelijk zijn. De moeilijkheidsgraad van het diagnosticeren van de onderliggende schouderklacht ligt hoog en dit brengt met zich mee dat accuraatheid op zijn plaats is in het proces van diagnostiek (11). De waarde van de diagnose tendinopathie is onzeker en wordt momenteel gevormd door klinische tests, al dan niet in combinatie met beeldvormend onderzoek (11). De onderzoeksvraag die hieruit volgt luidt: wat zijn de meest valide diagnostische middelen om caput longum bicepspees tendinopathie zo accuraat mogelijk te kunnen stellen in de algehele patiëntenpopulatie? Hieruit voortkomend zou de accurate diagnose op basis van klinische tests/beeldvormend onderzoek de benadering van de klacht vereenvoudigen om vanuit daar een succesvol behandeltraject te kunnen laten starten aan de hand van fysiotherapeutische interventies.

# Methode

Om antwoord te krijgen op de onderzoeksvraag, waarin de validiteit van diagnostische middelen binnen het stellen van caput longum bicepspees tendinopathie centraal staat, is er gekozen voor een literatuurstudie. Het doel van het onderzoek betreft het definiëren van de diagnostische validiteit binnen de caput longum bicepspees tendinopathie. Hierbij is er gekeken naar de sensitiviteit, specificiteit, positief voorspellende waarde (PVW), negatief voorspellende waarde (NVW), de accuraatheid, de positieve likelihood ratio (LR+) en de negatieve likelihood ratio (LR-). De validiteit is onderzocht aan de hand van het gebruik van klinische tests dan wel fysieke tests als beeldvormend onderzoek. Omdat er sprake is van een literatuurstudie, is er volgens het stoomdiagram van WMO-toetsing geen ethische toetsing nodig.

## Databases en zoekstrategieën

De literatuur die is gebruikt voor deze literatuurstudie is gevonden in de databanken PubMed, Google Scholar en ScienceDirect. De databank PubMed is gekozen wegens het toegankelijke gebruik naar medische wetenschappelijke artikelen waar zowel de MEDLINE- als de NCIB-databank doorzocht wordt. Ook heeft PubMed een geavanceerde zoekoptie waardoor het aantal specifieke studies vergroot kan worden. Daarnaast is er voor Google Scholar, de speciale zoekmachine van Google voor wetenschappelijke artikelen, gekozen. Deze databank richt zich op alle wetenschappelijke disciplines in de zoekbalk. Tenslotte is er gekozen voor de databank ScienceDirect wegens de mogelijkheid om te zoeken naar internationale wetenschappelijke literatuur op alle wetenschappelijke gebieden, te gebruiken met geavanceerde zoekopties.

De in de inleiding genoemde onderzoeksvraag ‘wat zijn de meest valide diagnostische middelen om caput longum bicepspees tendinopathie in de algehele patiëntenpopulatie zo accuraat mogelijk te kunnen stellen?’ is in de zoektocht naar geschikte artikelen, verder uitgewerkt aan de hand van het PIO-principe. Het onderzoeksprobleem (P) staat voor de caput longum bicepspees tendinopathie. De interventie (I) staat voor de diagnostische middelen in het onderzoek. De uitkomst (O) is gericht op het stellen van een zo accuraat mogelijke diagnose. Er komt geen specifieke comparison (C) naar voren in de vraagstelling van deze literatuurstudie, hier is dan ook geen sprake van. Aan de hand van het genoemde PIO-principe is er verder gezocht naar vrije zoektermen en MeSH termen in het afzonderlijke PIO-deel. Voor het onderzoeksprobleem is er gezocht naar de MeSH term “tendinopathy”. Voor de interventie is er gezocht naar de MeSH termen “diagnosis”, “diagnostic techniques” en “diagnostic imaging”. Daarnaast is er gezocht naar de MeSH term “reproducibility of results” om binnen de diagnostische middelen zo specifiek mogelijk te kunnen zoeken. Binnen de uitkomstmaat (O) kon er geen gebruik worden gemaakt van een MeSH term, om deze reden zijn er vrije zoektermen opgezocht. Tevens werd het gebruik van deze vrije zoektermen in het gehele PIO-principe toegepast om de meest geschikte artikelen te vinden. De MeSH termen en vrije zoektermen zijn weergegeven in de bijlage 2. In het afzonderlijke PIO-deel zijn de MeSH termen en vrije zoektermen aan de booleaanse operator “OR” gekoppeld. Deze operator heeft als doel het vergroten van de zoekresultaten. Verder is boleaanse operator “AND” gebruikt om de losse zoekstrings met elkaar te combineren waardoor de zoekstrategie specifieker kan worden gemaakt. In bijlage 1 is de zoekstring per databank weergegeven.

## Selectiecriteria

Voor het selecteren van geschikte artikelen zijn de zoekstrings per databank ingevuld en gefilterd aan de hand van de in- en exclusiecriteria. De zoekstrings werden waar nodig aangepast aan de eisen van desbetreffende databank. De in- en exclusiecriteria zijn weergegeven in tabel 1.

**Tabel 1: in- en exclusiecriteria**

|  |  |
| --- | --- |
| Inclusiecriteria | Exclusiecriteria |
| Artikelen waarin gekeken wordt naar diagnostische middelen omtrent bicepspees tendinopathie | Artikelen die niet in het Engels of Nederlands verkrijgbaar zijn |
| Artikelen gericht op fysieke tests als interventie | Artikelen eerder gepubliceerd dan 2011 |
| Artikelen gericht op beeldvormend onderzoek als interventie | Artikelen met de aanwezigheid van postoperatieve schouder patiëntenpopulatie |
| Artikelen met als het doel het meten van de validiteit (accuraatheid, sensitiviteit, specificiteit, PVW en NVW, LR+ en LR-) | Artikelen niet full-tekst beschikbaar |

## Methodologische kwaliteitsbepaling

Voor het beoordelen van de geïncludeerde systematische review is er gekozen voor de AMSTAR-checklist. De AMSTAR-checklist is, ten opzichte van andere beoordelingschecklisten die zich richten op systematische reviews, met een gebruiksduur van gemiddeld 10-15 minuten makkelijker en sneller in gebruik. Het doel van de AMSTAR-checklist is om gebruikers te helpen in het herkennen van kwalitatief hoge systematische reviews (12). De checklist bestaat uit 16 meerkeuzevragen die beantwoord kunnen worden met ja, gedeeltelijk ja of nee. De mogelijke scores zijn verdeeld met 1,0 punt voor “ja”, met een 0,5 punt voor “gedeeltelijk ja” en met 0 punten voor “nee”.

**Tabel 2: Scorelijst methodologische kwaliteit AMSTAR checklist**

|  |  |
| --- | --- |
| Methodologische kwaliteit | Score |
| Hoge kwaliteit | 13-16 |
| Gemiddelde kwaliteit | 9-12 |
| Lage kwaliteit | 0-8 |

Voor het beoordelen van zowel de geïncludeerde cross-sectionele studie als de geïncludeerde cohortstudies als de case-control studie is er gekozen voor de STROBE-checklist. De STROBE is een te gebruiken checklist voor meerdere vormen van observationeel onderzoek. De checklist betreft aanbevelingen die gedaan kunnen worden over een nauwkeurige en complete rapportage van een observationeel onderzoek (13).

De STROBE-checklist bestaat uit 22 aanbevelingen die betrekking hebben op de traditionele onderdelen van een wetenschappelijk artikel: de titel en de samenvatting (aanbeveling 1), de inleiding (aanbevelingen 2 en 3), de beschrijving van de methode (aanbevelingen 4-12), de weergave van de resultaten (aanbevelingen 13-17), de beschouwing (aanbevelingen 18-21) en de vermelding van de rol van de financiers (aanbeveling 22). Per aanbeveling kunnen er een aantal extra items behaald worden. De aanbevelingen worden beoordeeld op “aanwezig” of “niet-aanwezig”. Totaal kunnen er 34 items in de STROBE-checklist worden behaald.

|  |  |
| --- | --- |
| Methodologische kwaliteit | Score |
| Hoge kwaliteit | 27-34 |
| Gemiddelde kwaliteit | 18-26 |
| Lage kwaliteit | 0-17 |
|  |  |

**Tabel 3: Scorelijst methodologische kwaliteit STROBE checklist**

## Data-extractie

Om antwoord te krijgen op de onderzoeksvraag, is na het beoordelen van de methodologische kwaliteit de benodigde data geëxtraheerd. Van elke studie werden de auteurs, het jaar van uitgave, de afdeling van het te verrichten onderzoek en het onderzoeksdesign genoteerd. Verder is de studie populatie, het aantal participanten, het aantal mannen en vrouwen en de gemiddelde leeftijd genoteerd. Daarnaast zijn de uitgevoerde interventies, de referentie tests en de beoordelaar van interventies geëxtraheerd om vervolgens de resultaten te extraheren. Ten slotte werden de sensitiviteit, specificiteit, PVW, NVW, accuraatheid, LR+ en LR- als uitkomst binnen de gebruikte interventies geëxtraheerd.

## Data-analyse

Voor het meten van de validiteit bij het gebruik van diagnostische middelen binnen de diagnose caput longum bicepspees tendinopathie is er gekeken naar de eerder genoemde zevental factoren, deze zullen hieronder worden toegelicht. De sensitiviteit is het percentage correct geclassificeerde positief geteste patiënten (14). De sensitiviteit heeft de mogelijkheid om het resultaat van een uitgevoerde test uit te sluiten voor de diagnose van een ziekte of aandoening (15). De specificiteit is het percentage correct niet-geclassificeerde negatief geteste patiënten (14). De specificiteit heeft de mogelijkheid om het resultaat van een uitgevoerde test in te sluiten voor de diagnose van een ziekte of aandoening (15). De positief voorspellende waarde is het percentage positief geteste patiënten met de diagnose van een ziekte of aandoening (14). De negatief voorspellende waarde is het percentage negatief geteste patiënten zonder de diagnose van een ziekte of aandoening (14). De likelihood ratio’s meten de waarschijnlijkheid van een test op hoe aannemelijk het test resultaat is bij een patient met of zonder de diagnose van een ziekte of aandoening (15). Hierbij geeft de positieve likelihood ratio aan hoe sterk de gestelde diagnose van een ziekte of aandoening wordt vergroot. Een negatieve likelihood ratio geeft aan hoe sterk de niet gestelde

diagnose van een ziekte of aandoening wordt verkleind (16). De accuraatheid is het percentage, over het vermogen van een diagnostisch middel, om de diagnose ziekte of aandoening te voorspellen of te onderscheiden (17).

**Tabel 4: Interpretatie LR+ en LR**

|  |  |  |
| --- | --- | --- |
| LR+ | LR- | Interpretatie |
| > 10 | < 0,1 | Heeft grote invloed op de te genereren diagnose |
| 5-10 | 0,1-0,2 | Belangrijke test, duidelijke invloed op diagnose |
| 2-5 | 0,2-0,5 | Geringe maar klinisch soms belangrijke invloed |
| 1-2 | 0,5-1 | Weinig relevante test |

Voor de interpretatie van de likelihood ratio’s zijn de waarden in de tabel 4 weergegeven (18).

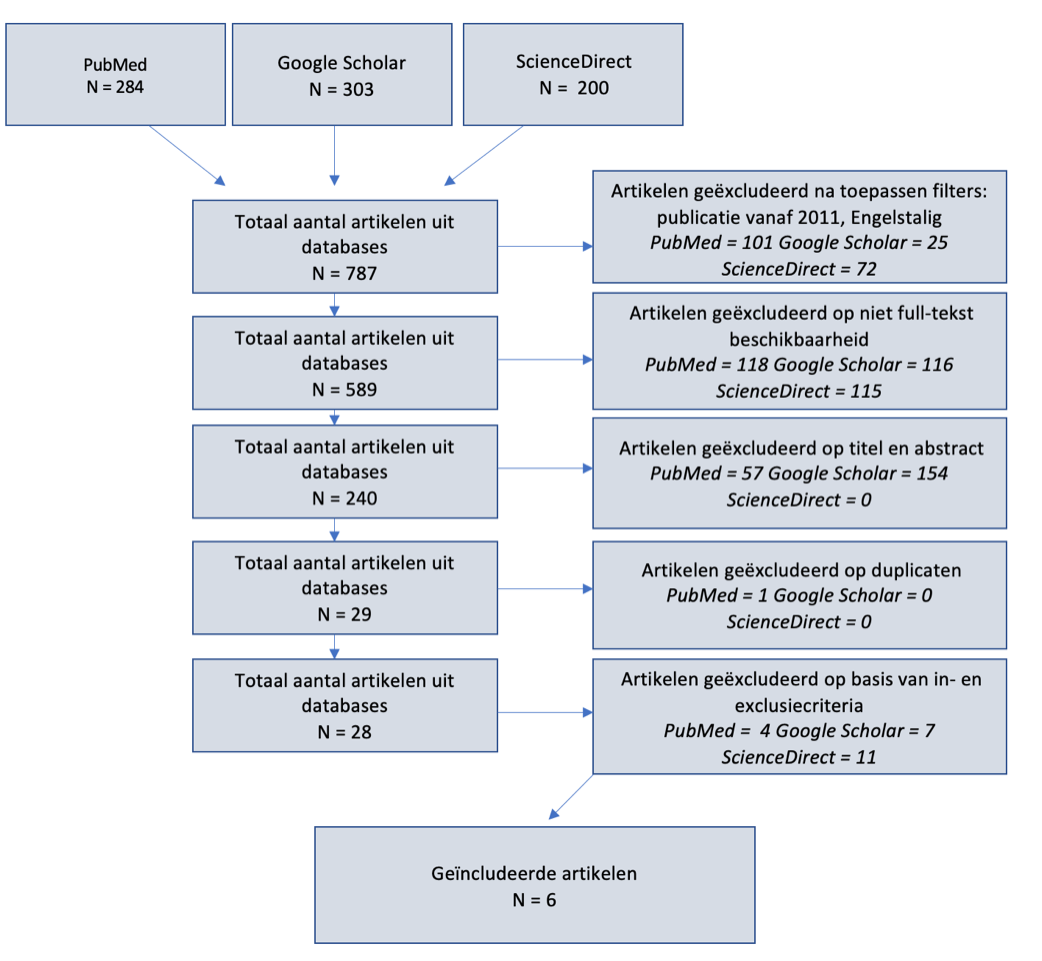
Voor het toepassen van de variabelen zijn diagnostische tests met een hoge sensitiviteit en een lage LR- het meest bruikbaar voor het uitsluiten van een ziekte of aandoening. Diagnostische tests met een hoge specificiteit en een hoge LR+ zijn het meest bruikbaar om een ziekte of aandoening in te sluiten (19).

# Resultaten

## Selectie van de studies

De zoekresultaten van de uit de databanken gebruikte zoekstrengen zijn beschreven in figuur 1. Uit deze drie gebruikte databanken PubMed, Google Scholar en ScienceDirect zijn in totaal 787 mogelijk te gebruiken artikelen naar voren gekomen. Als eerst stap zijn de artikelen met een publicatiedatum voor 2011 en niet-Engelstalig geëxcludeerd, waardoor er 589 mogelijk te gebruiken artikelen overbleven. Na het excluderen van de artikelen op full-tekst beschikbaarheid, titel en abstract en duplicaten bleven er 28 mogelijk te gebruiken artikelen over. Deze 28 artikelen werden door middel van de vooraf opgestelde in- en exclusiecriteria gescreend, dit heeft geresulteerd in 6 relevante artikelen. Het volledige zoekproces is in figuur 1 beschreven.

**Figuur 1: Flowchart**



## Methodologische kwaliteit van artikelen

De studie Cardoso et al. (2019) scoorde een gemiddelde kwaliteit. Er werd geen beschrijving gegeven van de matchingscriteria. Voor de statistische methode werd er geen beschrijving gegeven over het verlies van data en van follow up. Naast de subgroepen werden er ook geen eventueel te verrichten sensitiviteitsanalyses beschreven. Binnen de te behalen resultaten werd er over de verloren data en de follow up time voor de participanten niets vermeld. Het absolute risico en de continue variabelen werden niet gecategoriseerd. Ten slotte werd er geen beschrijving gegeven over de rol en de bron van de financiering.

De studie Chen et al. (2011) scoorde een gemiddelde kwaliteit. In de inleiding werden er geen vooraf opgestelde gespecificeerde hypotheses opgesteld. Daarnaast werden de matchingscriteria niet benoemd. Binnen de statistische methode werd er geen verlies van follow up en van verloren data beschreven. Er werd geen beschrijving gegeven van de aanwezigheid van subgroepen. Ook de eventueel te verrichten sensitiviteitsanalyses werden niet beschreven. Binnen de resultaten sectie werd er geen flow diagram met karakteristieke eigenschappen van de participanten beschreven. Verder ontbraken de absolute risico’s en de continue variabelen. Ten slotte werd de bron en de rol van de financiering niet vermeld.

De studie Teixeira et al. (2019) scoorde een gemiddelde kwaliteit. In de inleiding werden er geen vooraf gespecificeerde hypotheses opgesteld. Verder werden er geen matchingscriteria benoemd. Bij het gebruik van de statistische methodes werd er niet aangegeven hoe er werd omgegaan met verloren data en verloren follow op methodes. De eventueel te verrichten sensitiviteitsanalyses zijn niet aanwezig. Daarnaast werd er voor de participanten niet vermeld hoe er met verloren data werd omgegaan, werden er geen karakteristieke eigenschappen vermeld, was er geen flow diagram bijgevoegd en werd er geen uitleg gegeven over de niet-deelnemende participanten. De continue variabelen en het absolute risico werden niet beschreven. Ten slotte werd de rol en de bron van de financiering niet vermeld.

De studie Nourissat et al. (2014) scoorde een gemiddelde kwaliteit. Er was geen vermelding van de vooraf opgestelde gespecificeerde hypotheses en er werden geen matchingscriteria benoemd. Binnen de statistische methode werd er geen beschrijving gegeven over de aanwezigheid van eventuele subgroepen, de verloren data en de beschrijving over al dan niet aanwezige sensitiviteitsanalyses. Bij de resultatensectie ontbraken naast de kenmerken van de participanten met het gebruik van een flowchart ook de bijbehorende variabelen. De continue variabelen en het absolute risico werden binnen de studie niet vermeld. Ten slotte werd de rol en de bron van de financiering niet besproken.

De studie Lasbleiz et al. (2014) scoorde een gemiddelde kwaliteit. In de inleiding werden er geen vooraf gespecificeerde hypotheses opgesteld. Daarnaast werd het verlies van follow up niet uitgelegd. Ook werden de eventueel te verrichten sensitiviteitsanalyses niet beschreven. Binnen de resultatensectie ontbrak het naast het flow diagram ook aan de beschrijving van de data. De continue variabelen en absolute risico’s werden niet vermeld. Ten slotte werd de rol en de bron van de financiering niet besproken.

De studie Bélanger et al. (2019) scoorde een gemiddelde kwaliteit. De review auteurs hebben niet gerapporteerd over de mogelijke financieringsbronnen binnen de in de review geïncludeerde studies. Daarnaast werd er bij het uitvoeren van de kwantitatieve synthese geen adequaat onderzoek naar publicatiebias verricht met daarbij horend de bespreking van de waarschijnlijke impact op de resultaten.

**Tabel 5: Overzicht methodologische kwaliteit per studie**

|  |  |  |  |
| --- | --- | --- | --- |
| Studie | Design | Checklist | Score methodologische kwaliteit |
| Cardoso et al. (2019) | Cohort | STROBE | 23/34 |
| Chen et al. (2011) | Cohort | STROBE | 18/34 |
| Teixeira et al. (2019) | Cohort | STROBE | 22/34 |
| Nourissat et al. (2014) | Case-control | STROBE | 19/33\* |
| Lasbleiz et al. (2014) | Cross-sectioneel | STROBE | 22/32\* |
| Bélanger et al. (2019) | Systematische review met meta-analyse | AMSTAR | 11/16 |

\*maximale score van 34 niet mogelijk

## Beschrijving van de studies

Totaal zijn er 6 studies geïncludeerd, waarvan 2 prospectieve cohort studies, 1 retrospectieve cohort studie, een prospectieve case-controle studie, een cross-sectionele studie en een systematische review. De gebruikte geëxtraheerde data zoals beschreven in de methode zal voor de geïncludeerde studies in de tabel worden weergegeven. Hieronder volgt een overzicht van de kenmerken van de onderzoekspopulatie, de interventie, de uitkomstmaten en de resultaten.

## Onderzoekspopulatie

Het totaal aantal participanten voor de geïncludeerde studies kwam neer op 5600. Het aandeel van de systematisch review is hierin aanzienlijk groter dan die van de andere studies, gezien een patiëntenpopulatie van 5190. Binnen Bélanger et al. (2019) wordt er verder geen onderscheid gemaakt in geslacht, gemiddelde leeftijd en diagnosegroep. Van het totaal aantal participanten exclusief de systematische review waren er 225 vrouwen en 185 mannen. De gemiddelde leeftijd ligt tussen de 45.7 jaar en 55.1 jaar, waarbij er binnen een studie geen gemiddelde leeftijd werd vermeld. Verder varieerde de leeftijd van het aantal participanten tussen de 16 en 78 jaar. Binnen twee studies werd geen leeftijd spreiding vermeld.

De onderzoekspopulatie die vermeld werd bij Cardoso et al. (2019) bestond uit 46 vrouwen en 19 mannen. De gemiddelde leeftijd was 55.1 jaar en de leeftijd spreiding werd niet verduidelijkt. In Chen et al. (2011) was er een man/vrouw verdeling van 69 vrouwen en 56 mannen. Totaal werden er 143 schouders onderzocht, waarvan 18 bilateraal. De gemiddelde leeftijd en de leeftijd spreiding werd niet vermeld binnen deze studie. In Teixeira et al. (2019) bestond de man/vrouw verdeling uit 43 vrouwen en 55 mannen. Hierbij was de gemiddelde leeftijd 54.8 jaar en de leeftijd spreiding varieerde tussen de 16 en 77 jaar. Bij Nourissat et al. (2014) waren er 40 vrouwen en 47 mannen. De gemiddelde leeftijd was 45.7 jaar en de leeftijd spreiding varieerde van 17 tot 78 jaar. In Lasbleiz et al. (2014) bestond de verhouding uit 27 vrouwen en 8 mannen. Totaal werden er 39 schouders onderzocht, waarvan 4 bilateraal. De gemiddelde leeftijd was 59 jaar en voor de leeftijd spreiding werd er niets vermeld.

## Interventie

De interventie die werd ingezet bij Cardoso et al. (2019) bestond uit een klinisch onderzoek van 5 fysieke tests voordat de geïncludeerde participanten in de periode april 2017 tot augustus 2018 een artroscopie voor de schouder ondergingen. Er werd gebruik gemaakt van de Speed’s test, de Yergason’s test, de Uppercut test, de Biceps resisted flexie test en de gemodificeerde Biceps resisted flexie test. Het klinische onderzoek werd gemiddeld 2.1 dagen voor de artroscopie uitgevoerd. Aan het eind van de artroscopie beschreef de chirurg de gevonden beschadigingen van het weefsel.

De interventie die werd ingezet bij Chen et al. (2011) bestond uit het beoordelen van echografie als gouden standaard in combinatie met de uitvoering van de Yergason’s test, Speed’s test en palpatie van de caput longum bicepspees in de sulcus intertubercularis als fysieke tests. Het apparaat dat werd gebruikt voor de echografie was de Medison ACCUVIX V20 met een lineaire array 7.5-13 MHz transducer. De lange kop van de bicepspees werd op zowel de lange- als de korte as geobserveerd.

De interventie die werd ingezet bij Teixeira et al. (2019) bestond uit het, door een radioloog gekozen, beoordelen van participanten die tussen januari 2014 en december 2016 een schouder CT-artrografie gevolgd door een artroscopie ondergingen. De artroscopische beelden en de operatie verslagen werden door een orthopedisch schouder chirurg bekeken. De tijd tussen de CT-artrografie en artroscopie was gemiddeld 87.7 dagen. Het apparaat dat werd gebruikt voor de CT-artrografie was de Aquilon One, Canon Medical Systems Corporation. Bij het uitvoeren van de CT-artrografie lagen de participanten in rug lig met de te onderzoeken arm in lichte exorotatie. De beelden van de CT-artrografie werden geëvalueerd op de PACS werkplaats door twee radiologen.

De interventie die werd ingezet bij Nourissat et al. (2014) bestond uit een artroscopische ingreep bij participanten die preoperatief een CT-artrografie of MRI-scan ondergingen. Voor de MRI-scan werd de Magnetom AR1, 1.5 T gebruikt. Het doel van de artroscopie was het benoemen van de term tendinopathie door macroscopische veranderingen van de peesstructuur aan te kunnen tonen, en het beoordelen van het intra-articulaire gedeelte van de lange kop van de bicepspees. Dit werd rond het glenoïd gedaan, ook wel het mediale insertie deel, en verder rond het deel voor de sulcus intertubercularis, ook wel het horizontale deel gedaan. De radioloog observeerde de lange kop van de bicepspees voor een insertie tendinopathie of een horizontale tendinopathie.

De interventie die werd ingezet bij Lasbleiz et al. (2014) bestond uit een klinisch onderzoek van twee fysieke tests met echografie als referentiestandaard. Het doel hiervan was het kunnen inschalen van verandering op pathologisch gebied van de lange kop van de bicepspees. De fysieke tests die werden ingezet waren de Speed’s test en de Yergason’s test. Voor de echografie werd er gebruik gemaakt van een 7- tot 15- MHz transducer.

De interventie die werd ingezet bij Bélanger et al. (2019) bestond uit de hoge resolutie echografie met als referentie standaard een MRI-scan of een operatieve ingreep. Verder bestond de interventie uit fysieke tests als de Speed’s test, de Yergason’s test en de palpatie test. Tussen de uitvoering van de indextest echografie en de uitvoering van de referentiestandaard MRI-scan of operatieve ingreep zaten 4 weken.

## Uitkomstmaten

De uitkomstmaat tendinopathie werd binnen de geïncludeerde studies voornamelijk gebruikt. Toch waren er een aantal verschillen in definities op te merken. In Cardoso et al. (2019) werden de termen totale ruptuur en intra-articulaire laesie gebruikt. De intra-articulaire laesie kon worden onderverdeeld in tendinopathie en een verstoring van de peesdikte van meer of minder dan 50%. In Chen et al. (2011) werd de term biceps tendinitis gedefinieerd als een combinatie van het doppler signaal met een verdikking van de caput longum bicepspees, het bewijs van synoviale verdikking van meer dan 3 millimeter, het bewijs van vloeistofophoping van meer dan 3 millimeter in de caput longum bicepspeesschede of een scheurende, hypoechoïsche aantasting van de caput longum bicepspees . Binnen Teixeira et al. (2019) worden de gebruikte uitkomsten geclassificeerd als normaal, tendinopathie inclusief partiële ruptuur en als totale ruptuur. In Nourissat et al. (2014) werd de caput longum bicepspees beoordeeld als normaal, ontstoken, degeneratief of gescheurd. De term tendinopathie werd gebruikt bij macroscopische abnormaliteiten van de caput longum bicepspees. Binnen Lasbleiz et al. (2014) werden de normale pees, tendinopathie inclusief partiële ruptuur en totale ruptuur als uitkomstmaten gebruikt. In Bélanger et al. (2019) werd alleen de term tendinopathie aangehouden.

## Resultaten

De validiteit van de Yergason’s test werd in Cardoso et al. (2019), Chen et al. (2011), Lasbleiz et al. (2014) en Bélanger et al. (2019) onderzocht, waarbij er een gemiddelde sensitiviteit van 45,4% (32-66,7%), specificiteit van 74,3% (54-83%), PVW van 57% (40-79%), NVW van 59,3%, accuraatheid van 60,8% (49-79,5%), LR+ van 2.10 (1.01-3.70) en een LR- van 0.76 gevonden werd. Bij Bélanger et al. (2019) werd er een gemiddelde sensitiviteit en specificiteit variërend van 32% tot 86% en 74% tot 82% gevonden. Binnen desbetreffende studie werden de variabelen PVW, NVW, accuraatheid, LR+ en LR- niet onderzocht. In Chen et al. (2011) werd de variabele accuraatheid niet benoemd.

De validiteit van de Speed’s test werd in Cardoso et al. (2019), Chen et al. (2011), Lasbleiz et al. (2014) en Bélanger et al. (2019) onderzocht, waarbij er een gemiddelde sensitiviteit van 72,5% (61-83,3%), specificiteit van 50,1% (33-71%), PVW van 77,3% (19,2-78%), NVW van 67,1% (52-92,3%), accuraatheid van 57,9% (43,6-65%), LR+ van 1.55 (1.24-2.09) en een LR- van 0.54 (0.46-0.63) gevonden werd. De sensitiviteit en specificiteit in Bélanger et al. (2019) varieerde van 47% tot 83% en 36% tot 75%. De accuraatheid werd bij Chen et al. (2011) eveneens niet onderzocht.

De validiteit van de Upper cut test werd in Cardoso et al. (2019) onderzocht, waarbij er binnen de strict en de loose interpretatie een sensitiviteit van 61% en 90%, specificiteit van 63% en 38%, PVW van 74% en 71%, NVW van 48% en 69%, accuraatheid van 62% en 69%, LR+ van 1.63 en 1.44 en een LR- van 0.62 en 0.26 gevonden werd.

De validiteit van de mBRF test werd in Cardoso et al. (2019) onderzocht, waarbij er binnen de strict en de loose interpretatie een sensitiviteit van 34% en 49%, specificiteit van 75% en 71%, PVW van 70% en 74%, NVW van 40% en 45%, accuraatheid van 49% en 57%, LR+ van 1.36 en 1.67 en een LR- van 0.88 en 0.72 werd gevonden.

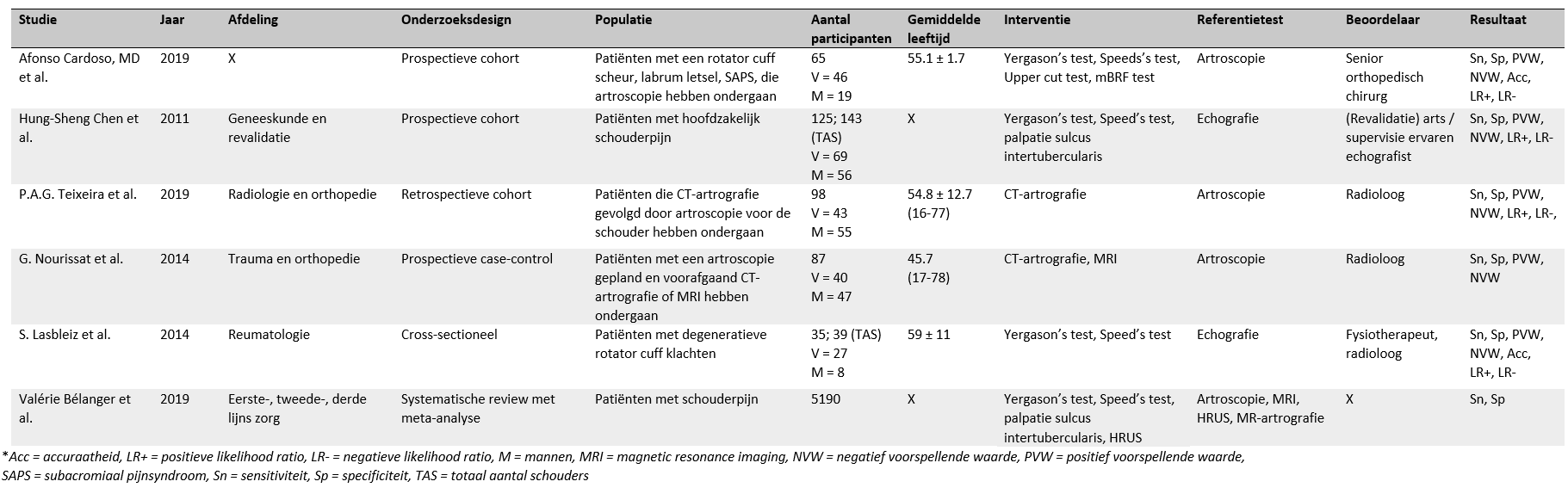
De validiteit van de sulcus intertubercularis palpatie werd in Chen et al. (2011) en Bélanger et al. (2019) onderzocht, waarbij er binnen Chen et al. (2011) een sensitiviteit van 57,1%, specificiteit van 72,4%, PVW van 57%, NVW van 72%, LR+ van 2.07 en een LR- van 0.59 werd gevonden. De accuraatheid werd niet onderzocht. In Bélanger et al. (2019) werd er een sensitiviteit van 57% tot 85% en specificiteit van 49% tot 72% gevonden.

De validiteit van de CT-artrografie werd Teixeira et al. (2019) en Nourissat et al. (2014) onderzocht, waarbij er een gemiddelde sensitiviteit van 74,8% (71,4-79%), specificiteit van 96% (93-100%), PVW van 96,3% (93-100%), NVW van 71,9% (67,8-76%), accuraatheid van 84% (82-86%), LR+ van 12.97 (10.14-15.80) en een LR- van 0.25 (0.22-0.28) gevonden werd. Binnen Nourissat et al. (2019) werd er geen onderzoek gedaan naar de accuraatheid, LR+ en LR-.

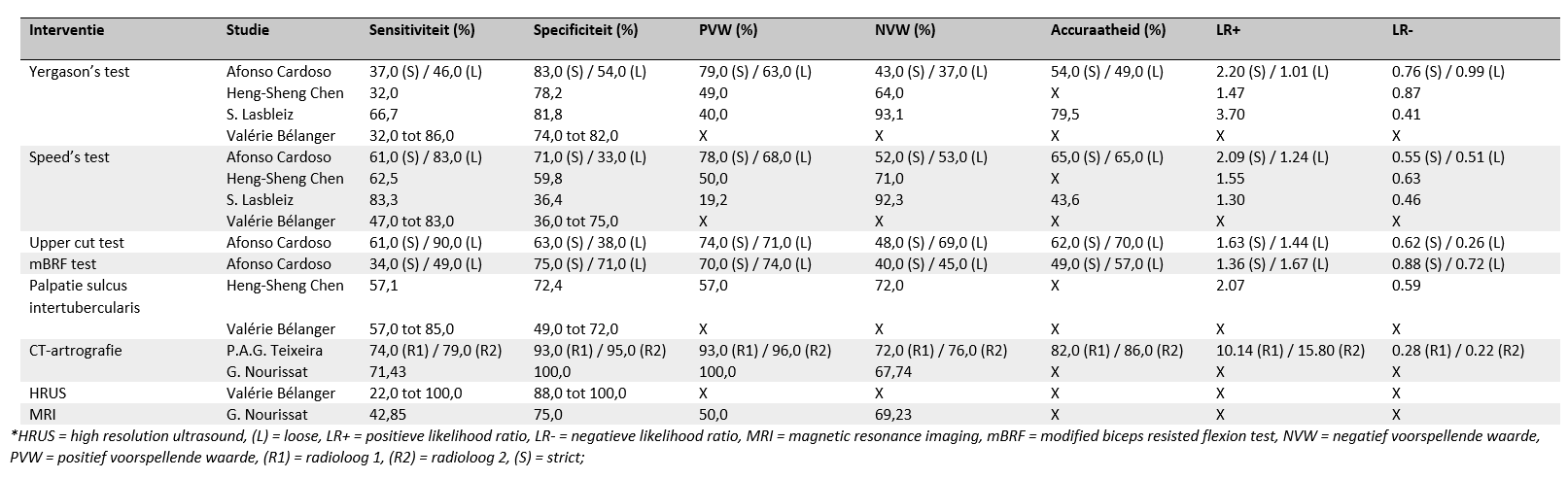
De validiteit van de HRUS werd in Bélanger et al. (2019) onderzocht, waarbij er een sensitiviteit van 72% tot 100% en specificiteit van 88% tot 100% werd gevonden.

De validiteit van de MRI werd in Nourissat et al. (2014) onderzocht, waarbij er een sensitiviteit van 42,9%, specificiteit van 75%, PVW van 50% en een NVW van 69,2% gevonden werd. Naar de variabelen accuraatheid, LR+ en LR- werd geen onderzoek verricht.

**Tabel 6: Karakteristieken van geïncludeerde studies**



**Tabel 7: Resultaten van uitgevoerde interventies**



# Discussie

Doordat de moeilijkheidsgraad van de diagnostiek binnen de caput longum bicepspees tendinopathie hoog ligt, werd de onderzoeksvraag als volgt opgesteld: ‘wat zijn de meest valide diagnostische middelen om caput longum bicepspees tendinopathie zo accuraat mogelijk te kunnen stellen in de algehele patiëntenpopulatie?’. Het doel van deze literatuurstudie was het definiëren van de diagnostische validiteit om zo antwoord te geven op de onderzoeksvraag en daarmee de benadering van de klacht te vereenvoudigen.

## Belangrijkste resultaten studies

Valide data kan worden geëxtraheerd uit de zes geïncludeerde studies. Niet alle gekozen variabelen (sensitiviteit, specificiteit, PVW, NVW, accuraatheid, LR+ en LR-) komen echter terug in ieder individuele studie. Desalniettemin zijn er twee factoren die in iedere studie zijn bepaald, namelijk de sensitiviteit en de specificiteit. Zodoende kunnen aan deze variabelen een duidelijke conclusie worden verbonden. De LR+ en LR- geven belangrijke toegevoegde informatie en komen terug in 4 studies, hiermee zijn deze variabelen eveneens van belang voor het stellen van de conclusie. Indien in een artikel laatstgenoemde variabelen niet zijn bepaald zal, waar mogelijk, worden gekeken naar de PVW en NVW. De sensitiviteit en specificiteit vormen de basis, andere variabelen zijn van secundair belang in het bepalen van de validiteit.

Door het gebruik van de ‘strict’ en de ‘loose’ interpretatie van de pijn, lijkt er een verschil in diagnose te zijn in *Cardoso et al. (2019)*. De Yergason’s test en de Speed’s test zijn meer waardevol in het insluiten van de diagnose tendinopathie, met respectievelijk een specificiteit van 83% en 71% en een LR+ van 2.20 en 2.09. De Upper cut is met een ‘loose’ interpretatie geschikt voor het uitsluiten van de diagnose tendinopathie, hierbij ligt de sensitiviteit op 90% en de LR- op 0.26. De mBRF test heeft geen diagnostische betrouwbaarheid in het correct kunnen diagnosticeren van de tendinopathie. Op basis van de diagnostische uitkomsten met echografie als referentiestandaard zijn de Yergason’s test en palpatie van de sulcus intertubercularis in *Chen et al. (2011)* specifieker in het bevestigen van diagnose tendinitis. In *Teixeira et al. (2019)* heeft CT-artrografie een sterke evidentie voor het insluiten van de diagnose tendinopathie. Hierbij geldt een specificiteit van 95% en een LR+ van 15.80. Verder dient CT-artrografie, met een sensitiviteit van 79% en een LR- van 0.22, als een geschikt middel voor het uitsluiten van de diagnose tendinopathie. Binnen het ontdekken van de diagnose tendinopathie is aan de hand van beeldvormend onderzoek een verschil op te merken in *Nourissat et al. (2014)*. Ten opzichte van de MRI kan CT-artrografie, met een specificiteit van 100% en een PVW van 100%, als een zeer geschikt diagnose middel worden gezien voor het insluiten van de diagnose. De Yergason’s test en de Speed’s test zijn als fysieke tests onderzocht in *Lasbleiz et al. (2014).* Met een specificiteit van 81,8% en een LR+ van 3.7 heeft de Yergason’s test een meerwaarde in het insluiten van de diagnose tendinopathie. Daarentegen heeft de Speed’s test met een sensitiviteit van 83,3%, een LR- van 0.46 en een NVW van 92,3% een meerwaarde in het uitsluiten van de diagnose tendinopathie. Ten slotte is in *Bélanger et al. (2019)* voor het insluiten van de diagnose tendinopathie zowel het beeldvormende onderzoek echografie als de Yergason’s test een geschikt diagnose middel. Hierbij varieert de specificiteit respectievelijk van 88% tot 100% en van 74% tot 82%. De palpatie van de sulcus intertubercularis en het inzetten van de Speed’s test zijn beide met een sensitiviteit van 85% en 83% geschikt voor het uitsluiten van de diagnose tendinopathie.

## Kritische beschouwing studies

De resultaten uit *Cardoso et al. (2019)* zijn door middel van de ‘strict’ en de ‘loose’ interpretatie uitgewerkt. Het verschil tussen deze termen is dat de ‘strict’ interpretatie bij een positief of negatief testresultaat uitgaat van pijn in de sulcus intertubercularis en dat de ‘loose’ interpretatie uitgaat van zowel pijn in de sulcus intertubercularis als in een andere regio. Het effect hiervan is dat fysieke tests specifieker worden bij de ‘strict’ interpretatie. Het tegenovergestelde doet zich voor bij de ‘loose’ interpretatie waarbij de sensitiviteit toeneemt en de specificiteit afneemt. Voor de uitkomstmaat caput longum bicepspees laesie is er geen onderscheid gemaakt tussen tendinopathie en ruptuur, met als gevolg dat de benadering van uitkomst verandert. In de studie is er gekozen voor een operatieve patiëntenpopulatie die zich beperkt tot artroscopie als referentiestandaard. Een referentiestandaard voor de niet-operatieve patiëntenpopulatie ontbreekt, resulterend in een verminderde mate van generaliseerbaarheid van de geselecteerde populatie in relatie tot de algehele. Ten slotte zijn de fysieke tests door één onderzoeker uitgevoerd waardoor de intra- en interbeoordelaar analyses zijn uitgesloten.

De resultaten uit *Chen et al. (2011)* laten een overeenkomst in specificiteit zien tussen de Yergason’s test en de palpatie van de sulcus intertubercularis. Binnen de uitgevoerde interventies behaalt de Speed’s test daarentegen de hoogste sensitiviteit. Echografie wordt in de studie als een nieuw geïntroduceerde referentiestandaard gebruikt. Vergeleken met MRI en artroscopie heeft echografie de beperking dat het niet volledig accuraat is in het stellen van de diagnose tendinopathie in samenhang met een extra aandoening, zoals een SLAP laesie. De correlatie tussen de fysieke tests en echografie zijn lager uitgevallen wegens een aantal factoren, zoals het onbekende pijn patroon van de bicepspees en de moeilijk te palperen caput longum bicepspees door de diepe ligging in de schouder met bedekking van overige schouderstructuren.

De resultaten uit *Teixeira et al. (2019)* laten zien dat CT-artrografie een geschikt middel is in het diagnosticeren van caput longum bicepspees tendinopathie. CT-artrografie laat daarentegen een brede spreiding van sensitiviteit en specificiteit zien. Een verstoord signaal rond de sulcus intertubercularis wordt veroorzaakt door anatomische veranderingen van de caput longum bicepspees en de verschillende scanposities. Daarnaast wordt de uitkomstmaat tendinopathie samengevoegd met partiële rupturen van de caput longum bicepspees, hierdoor is er geen eenduidige benaming van de term tendinopathie te gebruiken. Ten slotte heeft de invloed van bestraling een rol gespeeld in het minder brede bereik van het diagnosticeren van de tendinopathie.

De resultaten uit *Nourissat et al. (2014)* laten zien dat CT-artrografie als beeldvormend superieur is aan MRI in het ontdekken van caput longum bicepspees tendinopathie. De beperking van CT-artrografie heeft wederom betrekking op de invloed van de bestraling en het daardoor smallere bereik. Daarnaast zijn de bevindingen eenmalig bekeken door de radioloog, waardoor de intra- en interbeoordelaar analyses zijn uitgesloten. In de studie is de geïncludeerde patiëntenpopulatie wat betreft bicepspees aandoeningen ongelijk. Het onderzoek zou effectiever zijn bij geïsoleerde caput longum bicepspees tendinopathie als patiëntenpopulatie.

De resultaten uit *Lasbleiz et al. (2014)* laten zien dan de Yergason’s test en de Speed’s test effectief zijn in het in- of uitsluiten van de diagnose tendinopathie. Volgens de studie is de klinische relevantie van de Speed’s test laag doordat de LR+ niet gelijk of meer dan 2 is. Daarnaast heeft echografie als referentiestandaard geen onderscheid gemaakt tussen partiële rupturen en tendinopathie. De geïncludeerde patiëntenpopulatie had betrekking tot conservatieve behandeling. De resultaten zijn daardoor niet te generaliseren over de gehele patiëntenpopulatie met schouder pijn.

De resultaten uit *Bélanger et al. (2019)* laten zien dat zowel fysieke tests als beeldvormend onderzoek kunnen worden ingezet voor het in- of uitsluiten van de diagnose tendinopathie. De resultaten zijn daarentegen niet geheel representatief voor de gehele patiëntenpopulatie wegens het gebrek aan demografische gegevens. Daarnaast zijn de resultaten met een spreiding weergeven, in plaats van eenduidig per studie. Hierdoor is het interpreteren van de resultaten bemoeilijkt.

## Vergelijking resultaten met bestaande literatuur

De resultaten uit de huidige literatuurstudie komen overeen met de resultaten uit *Rosas et al. (2017)*. Voor de Yergason’s test en de palpatie van de sulcus intertubercularis geldt een hogere specificiteit dan sensitiviteit. Deze gegevens duiden op het kunnen insluiten van de klacht of aandoening. Echter laat de LR+ geen duidelijke verbetering zien ten opzichte van de huidige literatuurstudie. Dit is wel het geval met de Speed’s test, waarin een verbeterde specificiteit en LR+ naar voren komen, wat duidt op het insluiten van de klacht of aandoening. In de huidige literatuurstudie laat de Speed’s test eerder een hogere sensitiviteit dan specificiteit zien en een meer lage LR+. De keerzijde van de resultaten uit *Rosas et al. (2017)* is dat alleen het begrip caput longum bicepspees pathologie als verzamelnaam wordt gebruikt. De uitkomstmaat tendinopathie komt niet naar voren.

In de systematische review *Hegedus (2012)* worden de Yergason’s test, Speed’s test en palpatie van de sulcus intertubercularis als interventie uitgewerkt. Bij de gemiddelde sensitiviteit en specificiteit van de geïncludeerde studies die betrekking hebben op caput longum bicepspees tendinopathie, laat de Yergason’s test een hogere specificiteit dan sensitiviteit zien. Het verschil van de Speed’s test ten opzichte van de huidige literatuurstudie is dat de specificiteit hoger is, en de sensitiviteit lager. Desbetreffende test kan dus worden gebruikt voor het insluiten van de diagnose caput longum bicepspees tendinopathie. Het resultaat van de sulcus intertubercularis palpatie komt overeen met de huidige literatuurstudie, waarin de specificiteit hoger is dan de sensitiviteit.

In *Read et al. (1998)* komen de resultaten overeen met de resultaten uit de huidige literatuurstudie. Zowel voor de sensitiviteit als specificiteit kan geconcludeerd worden dat echografie een toegevoegde waarde heeft in zowel het insluiten als uitsluiten van de diagnose caput longum bicepspees tendinopathie. Het verschil met de huidige literatuurstudie is dat de PVW en NVW in *Read et al. (1998)* worden toegelicht. Hierin geldt een PVW van 100% en een NVW van 94%.

## Sterke punten literatuurstudie

Een sterk punt van deze literatuurstudie is het includeren van studies waarin het gebruik van verschillende diagnosemiddelen naar voren komen. Binnen de geïncludeerde studies is beeldvormend onderzoek driemalig als interventie ingezet. De verschillende fysieke tests zijn viermaal als interventie ingezet. Bovendien zijn er totaal acht diagnostische middelen gebruikt, waardoor er een brede kijk is op de validiteit van de mogelijk te stellen diagnose caput longum bicepspees tendinopathie.

Een ander sterk punt van deze literatuurstudie is het gebruik van 3 verschillende databanken, waarin eventueel te gebruiken literatuur gezocht en geselecteerd is. Deze databanken hebben elk hun eigen kenmerken. Het voordeel van het gebruik van verschillende databanken is het kunnen specificeren van de gestelde onderzoeksvraag.

Verder is er gezocht naar de meest recente literatuur over de validiteit van de diagnostische middelen binnen de caput longum bicepspees tendinopathie. Hierdoor is er geschikt gefilterd op de meest bruikbare informatie omtrent de diagnostiek.

## Zwakke punten literatuurstudie

Een zwak punt van deze literatuurstudie is de beperkte ervaring met het gebruik van zowel de AMSTAR checklist als de STROBE checklist. Het classificeren en definiëren van de uitkomsten van de geïncludeerde studies is hierdoor niet geheel reproduceerbaar tot het interpreteren bij andere studies.

Een zwak punt wat hier op aansluit is de methodologische kwaliteit van de geïncludeerde studies. Van desbetreffende studies is de methodologische kwaliteit gemiddeld. Hierdoor kunnen de resultaten in de huidige literatuurstudie als niet geheel wetenschappelijk onderbouwd worden gezien. Daarnaast is er geen volledige score toegedeeld binnen *Nourissat et al. (2014)* en *Lasbleiz et al. (2014)*.

Binnen de huidige literatuurstudie zijn geen randomised controlled trials gebruikt. Het niveau van bewijs binnen de studie valt lager uit aangezien randomised controlled trials de beste klinische resultaten geven, wegens het feit van daadwerkelijk uitgevoerd onderzoek. (28)

Een ander zwak punt is de beperkte geëxtraheerde data uit de geïncludeerde studies. Niet alle gekozen variabelen konden worden gebruikt om tot een uitspraak te komen over de validiteit van de diagnostische middelen omtrent caput longum bicepspees tendinopathie. Daarnaast is de beoogde uitkomstmaat verschillend geformuleerd in de geïncludeerde studie. De vooropgestelde probleemstelling ‘tendinopathie’ zal als uitkomstmaat in twijfel moeten worden getrokken.

## Aanbeveling voor vervolgonderzoek

Voor het vergroten van de klinische relevantie is vervolgonderzoek van belang. Ten eerste is het belangrijk om naar echografie als interventie te kijken. Desbetreffende interventie is in de huidige literatuurstudie éénmalig naar voren gekomen en zal voor een betere validiteit vaker moeten worden onderzocht, zowel in een operatieve setting als in een fysiotherapeutische setting. Daarnaast heeft de patiëntenpopulatie van de geïncludeerde studies voornamelijk betrekking op zowel de operatieve- als de revalidatie setting. Dit heeft als gevolg dat de patiëntenpopulatie niet geheel representatief is voor de algehele patiëntenpopulatie met schouderpijn en mogelijke caput longum bicepspees tendinopathie. Ten slotte wordt er aanbevolen om een cluster van fysieke tests te combineren, dit zou de diagnostische accuraatheid bekrachtigen. In de huidige literatuurstudie zijn de resultaten van de fysieke tests binnen de sensitiviteit en specificiteit verschillend. Toekomstig onderzoek zal gericht zijn op het ontwikkelen van clusters waarin fysieke tests met een hoge sensitiviteit dan wel hoge specificiteit gecombineerd worden om de diagnostische accuraatheid van caput longum bicepspees tendinopathie te versterken.

## Aanbeveling voor de praktijk

Fysieke tests dan wel beeldvormend onderzoek kunnen worden ingezet als diagnostisch middel voor het diagnosticeren van caput longum bicepspees tendinopathie. Voor de fysiotherapeutische setting dient er rekening te worden gehouden met het verschil in sensitiviteit en specificiteit van verschillende fysieke tests. De fysieke tests waarbij een hogere sensitiviteit naar voren is gekomen, kunnen worden gebruikt als screeningstest om vanuit daar de mogelijke tendinopathie uit te sluiten. Daarentegen dienen fysieke tests met een hogere specificiteit ervoor om de mogelijke tendinopathie in te sluiten. Zoals eerder aangegeven wordt vervolgonderzoek voor desbetreffende situatie aanbevolen. Daarnaast is het voor een fysiotherapeut van belang om waakzaam te zijn op de klinische symptomen tijdens de anamnestische fase. Relevante informatie moet worden gefilterd om gerichter onderzoek uit te kunnen voeren. Ten slotte is het effect van beeldvormend onderzoek gunstig in het accuraat kunnen diagnosticeren van caput longum bicepspees tendinopathie. Echografie kan, uit gaande van de resultaten van desbetreffende literatuurstudie, worden ingezet als diagnostisch middel. Zoals eerder vermeld dient er vervolgonderzoek te worden uitgevoerd voor de validiteit binnen de fysiotherapeutische setting. Bovendien dient er binnen de beleidsbepaling overwogen te worden of er aan de hand van MRI en/of CT-artrografie een diagnose moet worden gesteld, reden hiervoor is dat zowel MRI als CT-artrografie kosten met zich meebrengen.

# Conclusie

Het doel van de huidige literatuurstudie was het definiëren van de diagnostische validiteit binnen de caput longum bicepspees tendinopathie in de algehele patiëntenpopulatie. Uit de resultaten is gebleken dat CT-artrografie het meest valide diagnostische middel is voor zowel het insluiten als het uitsluiten van de diagnose caput longum bicepspees tendinopathie. Kijkend naar de sensitiviteit en specificiteit binnen de fysieke tests, komt de Speed’s test naar voren als geschikt middel voor het uitsluiten van de diagnose en de Yergason’s test als geschikt middel voor het insluiten van de diagnose. Aanbevolen wordt om vervolgonderzoek uit te voeren naar de effectiviteit van echografie in de fysiotherapeutische setting, de validiteit van de diagnostische middelen binnen de fysiotherapeutische setting en de validiteit van een combinatie van fysieke tests. Ten slotte wordt voor de praktijk aanbevolen om waakzaam te zijn op klinische symptomen tijdens de anamnestische fase om vanuit daar efficiënter onderzoek uit te voeren aan de hand van een screenings, dan wel een diagnostische test.

# Referentielijst

1. Huisstede BM, Wijnhoven HA, Bierma-Zeinstra SM, Koes BW, Verhaar JA, Picavet S. Prevalence and characteristics of complaints of the arm, neck, and/or shoulder (cans) in the open population. Clin J Pain 2008; 24: 253-9.
2. Picavet HS, Schouten JS. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC (3)- study. Pain 2003; 102: 167-78.
3. Ho C-YC, Sole G, Munn J. The effectiveness of manual therapy in the management of musculoskeletal disorders of the shoulder: a systematic review. *Man Ther* (Internet). 2009 Oct cited 2013 Feb 27; 14(5): 463–74
4. Lee N, Juliann M, Dalton D, Charles C. The effectiveness of therapeutic exercise for painful shoulder conditions: a meta-analysis. Journal of Shoulder and Elbow Surgery. Volume 20, Issue 8, December 2011, Pages 1351-1359
5. Damen, GJ., Koel, G., Kuijpers, T., Ottenheijm, RP., Schellingerhout, JM., Van den Donk, M., Winters, JC., Wittenberg, J. 2019. Schouderklachten. Beschikbaar via: <http://schoudernetwerk.nl/SNNbestanden/NHG_standaard_Schouderklachten/01.M08%20NHG%20-Standaard%20Schouderklachten.2020.02.pdf> – geraadpleegd op 17-05-2021
6. Cook JL, Rio E, Purdam CR, Docking SI. Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research? British Journal of Sports Medicine. 2016; 50: 1187-91.
7. Hsu D, Anand P, Chang K. Biceps tendon rupture. StatPearls (Internet). StatPearls publishing; 2021-jan
8. Catherina A, Churgay, MD, ST. Luke’s Hospital/The University of Toledo, Family Medicin Residency Program, Toledo, Ohio. Diagnosis and Treatment of Biceps Tendinitis and Tendinosis. Am Fam Physician. 2009 September 1;80(5): 470-476
9. Bélanger V, Dupuis F, LeBlond J, Roy JS. Accuracy of examination of the long head of the biceps tendon in the clinical setting: a systematic review. Accepted May 3, 2019. J Rehabil Med 2019; 51: 479-491
10. Sarmento M. Long head of biceps: from anatomy to treatment. Acta Reumatal Port. Jan-Mar 2015; 40 (1): 26-33
11. Shane J, Eric J, Brett A, Matthew T, Augustus D, Nikhil N, Anthony A. Long Head of the Biceps Tendinopathy: Diagnosis and Management. American Academy of Orthopaedic Surgeon: November 2010 – Volume 18 – Issue 11 – p 645-656
12. Shea, BJ., Reeves CB., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., Henry, DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. 2017. Sep 21; 358:j4000
13. Stehouwer CDA. Rapportage van observationeel onderzoek: nuttige en welkome aanbevelingen ter verbetering. 2008. Beschikbaar via: <https://www.ntvg.nl/system/files/publications/2008101820001a.pdf>

Geraadpleegd op: 08-06-2021

1. Parikh, R., Mathai, A., Parikh, S., Sekhar, GC., Thomas, R. Understanding and using sensitivity, specificity and predictive values. 2008. Jan-Feb; 56(1): 45-50
2. Pewsner, D., Battagli, M., Minder, C., Marx, A., Bucher, HC., Egger, M. Ruling a diagnosis in or out with “SpPIn” and “SnNOut: a note of caution. 2004. Jul 24; 329 (7459): 209-213
3. Van der Berg P. Tests: voorafkans, achterafkans en likelihood ratio. 2015. Beschikbaar via: <https://www.henw.org/artikelen/tests-voorafkans-achterafkans-en-likelihood-ratio>

Geraadpleegd op: 11-06-2021

1. Eusebi P. Diagnostic accuracy measures. 2013. 36(4): 267-72
2. Koel G. FYSIOTHERAPIE en het behandelen van patiënten met SCHOUDERKLACHTEN. 2017. Beschikbaar via: <http://www.schoudernetwerk.nl/SNNbestanden/Cursussen/KoelCombiCursus/Blok.03.pdf>

Geraadpleegd op: 11-06-2021

1. Cardoso, A., Amaro, P., Barbosa, L., Coelho, MA., Alonso, R., Pires, L. Diagnostic accuracy of clinical tests directed to the long head of biceps tendon in a surgical population: a combination of old and new tests. 2019. 1-7
2. Chen, HS., Lin, HS., Hsu, YH., Chen, SC., Kang, JH. A comparison of physical examinations with musculoskeletal ultrasound in the diagnosis of biceps long head tendinitis. 2011. Ultrasound in Med & Biol.Vol 37, No 9. Pp 1392-1398
3. Teixeira, PAG., Jaquet, P., Bakour, O., Jacquot, A., Molé, D., Sirveaux, F., Blum, A. CT arthrography of the intra-articular long head of the biceps tendon: Diagnostic performance outside the labral-biciptal complex. 2019.Diagnostic and Interventional Imaging 100. 437-444
4. Nourissat, G., Tribot-Laspiere, Q., Aim, F., Radier, C. Contribution of MRI and CT arthrography to the diagnosis of inta-articular tendinopathy of the long head of the biceps. 2014. Orthopaedics & Traumatology: Surgery & Research S391-S394
5. Lasbleiz, S., Quintero, N., Ea, K., Petrover, D., Aout, M., Laredo, JD., Vicaut, E., Bardin, T., Orcel, P., Beaudreuil, J. 2014. Annals of physical and Rehabilitation Medicine 57 - 228-243
6. Bélanger, V., Dupuis, F., Leblond, J., Roy, JS. Accuracy of clinical examination of the long head of the biceps tendon in the clinical setting: a systematic review. 2019. J Rehabil Med. 479-491
7. Rosas, S., Krill, MK., Achampong, KA., Kwon, K., Nwachukwu, BU., McCormick, F. A practical, evidence-based, comprehensive (PEC) physical examinations for diagnosing pathology of the long head of the biceps. 2019. J Shoulder Elbow Surg. Aug; 26 (8); 1484-1492.
8. Hegedus J. Which physical examinations tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. 2012. 7 July.
9. Read, JW., Perko, M., Nest, C. Shoulder ultrasound: diagnostic accuracy for impingement syndrome, rotator cuff tear, and biceps tendon pathology. 1998. Journal of shoulder and elbow Surgery, 7 (3), 264-271
10. Burns, PB., Rohrich, RJ., Chung, KC. The levels of Evidence and their role in Evidence-Based Medicine. 2011. Jul; 128(1): 305-310

# Bijlagen

## Bijlage 1 Zoekstring

|  |  |  |
| --- | --- | --- |
| **Databank** | **Zoekstreng** | **Resultaten** |
| PubMed | “Long head of the biceps tendinopathy” OR “biceps tendinopathy” AND “diagnosis” OR “diagnostic techniques” OR “diagnostic imaging” AND “reproducibility of results” OR “validity” OR “sensitivity” OR “specificity” OR “accuracy” AND “diagnosis long head of the biceps” | 284 |
| Google Scholar | “Long head biceps tendinopathy” OR “long head biceps tendinitis” OR “biceps tendinopathy” OR “biceps tendinitis” AND “diagnosis” OR “diagnostic imaging” OR “clinical tests” OR “physical tests” AND “diagnosis LHBT” OR “accuracy diagnosis LHBT” AND “validity” OR “specificity” OR “sensitivity” | 303 |
| ScienceDirect | “Long head of the biceps tendinopathy” AND “diagnosis” AND “clinical tests” AND “diagnostic imaging” AND physical tests” AND “accuracy” | 200 |

## Bijlage 2 Vrije zoektermen en MeSH term

|  |  |
| --- | --- |
| **PICO** | **Vrije zoektermen** |
| P (caput longum bicepspees tendinopathie) | “Tendinitis”, “tendinosis” |
| I (diagnostische middelen) | “Clinical tests, “physical tests”, “MRI”, CT-artrography”, “ultrasound” |
| O (accuratie positieve diagnose) | “Accuracy”, “validity”, “sensitivity”, “specificity” |

|  |  |
| --- | --- |
| **Vrije zoektermen** | **MeSH** |
| “Tendinitis”, “tendinosis” | Tendinopathy |
| “Clinical tests, “physical tests”, “MRI”, CT-artrography”, “ultrasound” | Diagnosis  Diagnostic techniques  Diagnostic imaging |
| “Accuracy”, “validity”, “sensitivity”, “specificity” | Reproducibility of results |

|  |  |  |  |
| --- | --- | --- | --- |
| Bijlage 3 scorelijst STROBE Cardoso et al. (2019) | **Item No** | **Recommendation** | **Page No** |
| **Title and abstract** | 1 | (*a*) Indicate the study’s design with a commonly used term in the title or the abstract | V |
| (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | V |
| **Introduction** | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | V |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | V |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper | V |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | V |
| Participants | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants  (*b*)*Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study*—For matched studies, give matching criteria and the number of controls per case | V  X |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | V |
| Data sources/ measurement | 8\* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | V |
| Bias | 9 | Describe any efforts to address potential sources of bias | V |
| Study size | 10 | Explain how the study size was arrived at | V |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | V |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | V |
| (*b*) Describe any methods used to examine subgroups and interactions | X |
| (*c*) Explain how missing data were addressed | X |
| (*d*) If applicable, describe analytical methods taking account of sampling strategy | X |
| (*e*) Describe any sensitivity analyses | X |
| **Results** | | | |
| Participants | 13\* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | V |
| (b) Give reasons for non-participation at each stage | V |
| (c) Consider use of a flow diagram | V |
| Descriptive data | 14\* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | V |
| (b) Indicate number of participants with missing data for each variable of interest  (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) | X  X |
| Outcome data | 15\* | *Cohort study*—Report numbers of outcome events or summary measures over time  *Case-control study—*Report numbers in each exposure category, or summary measures of exposure  *Cross-sectional study—*Report numbers of outcome events or summary measures | V |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | V |
| (*b*) Report category boundaries when continuous variables were categorized | X |
| (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | X |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | X |
| **Discussion** | | | |
| Key results | 18 | Summarise key results with reference to study objectives | V |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | V |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | V |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | V |
| **Other information** | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | X |

--

|  |  |  |  |
| --- | --- | --- | --- |
| Bijlage 4 scorelijst STROBE ***Chen et al. (2011)*** | **Item No** | **Recommendation** | **Page No** |
| **Title and abstract** | 1 | (*a*) Indicate the study’s design with a commonly used term in the title or the abstract | V |
| (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | V |
| **Introduction** | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | V |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | X |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper | X |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | V |
| Participants | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants  (*b*)*Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study*—For matched studies, give matching criteria and the number of controls per case | V  X |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | V |
| Data sources/ measurement | 8\* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | V |
| Bias | 9 | Describe any efforts to address potential sources of bias | V |
| Study size | 10 | Explain how the study size was arrived at | V |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | V |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | V |
| (*b*) Describe any methods used to examine subgroups and interactions | X |
| (*c*) Explain how missing data were addressed | X |
| (*d*) If applicable, describe analytical methods taking account of sampling strategy | X |
| (*e*) Describe any sensitivity analyses | X |
| **Results** | | | |
| Participants | 13\* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | V |
| (b) Give reasons for non-participation at each stage | X |
| (c) Consider use of a flow diagram | X |
| Descriptive data | 14\* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | X |
| (b) Indicate number of participants with missing data for each variable of interest  (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) | X  X |
| Outcome data | 15\* | *Cohort study*—Report numbers of outcome events or summary measures over time  *Case-control study—*Report numbers in each exposure category, or summary measures of exposure  *Cross-sectional study—*Report numbers of outcome events or summary measures | V |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | V |
| (*b*) Report category boundaries when continuous variables were categorized | X |
| (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | X |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | X |
| **Discussion** | | | |
| Key results | 18 | Summarise key results with reference to study objectives | V |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | V |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | V |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | V |
| **Other information** | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | X |

--

|  |  |  |  |
| --- | --- | --- | --- |
| Bijlage 5 scorelijst STROBE ***Teixeira et al. (2019)*** | **Item No** | **Recommendation** | **Page No** |
| **Title and abstract** | 1 | (*a*) Indicate the study’s design with a commonly used term in the title or the abstract | V |
| (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | V |
| **Introduction** | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | V |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | X |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper | V |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | V |
| Participants | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants  (*b*)*Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study*—For matched studies, give matching criteria and the number of controls per case | V  X |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | V |
| Data sources/ measurement | 8\* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | V |
| Bias | 9 | Describe any efforts to address potential sources of bias | V |
| Study size | 10 | Explain how the study size was arrived at | V |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | V |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | V |
| (*b*) Describe any methods used to examine subgroups and interactions | V |
| (*c*) Explain how missing data were addressed | X |
| (*d*) If applicable, describe analytical methods taking account of sampling strategy | X |
| (*e*) Describe any sensitivity analyses | X |
| **Results** | | | |
| Participants | 13\* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | V |
| (b) Give reasons for non-participation at each stage | X |
| (c) Consider use of a flow diagram | X |
| Descriptive data | 14\* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | X |
| (b) Indicate number of participants with missing data for each variable of interest  (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) | X  V |
| Outcome data | 15\* | *Cohort study*—Report numbers of outcome events or summary measures over time  *Case-control study—*Report numbers in each exposure category, or summary measures of exposure  *Cross-sectional study—*Report numbers of outcome events or summary measures | V |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | V |
| (*b*) Report category boundaries when continuous variables were categorized | X |
| (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | X |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | V |
| **Discussion** | | | |
| Key results | 18 | Summarise key results with reference to study objectives | V |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | V |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | V |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | V |
| **Other information** | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | X |

--

|  |  |  |  |
| --- | --- | --- | --- |
| Bijlage 6 scorelijst STROBE ***Nourissat et al. (2014)*** | **Item No** | **Recommendation** | **Page No** |
| **Title and abstract** | 1 | (*a*) Indicate the study’s design with a commonly used term in the title or the abstract | V |
| (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | V |
| **Introduction** | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | V |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | X |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper | V |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | V |
| Participants | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants  (*b*)*Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study*—For matched studies, give matching criteria and the number of controls per case | V  X |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | V |
| Data sources/ measurement | 8\* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | V |
| Bias | 9 | Describe any efforts to address potential sources of bias | V |
| Study size | 10 | Explain how the study size was arrived at | V |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | V |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | V |
| (*b*) Describe any methods used to examine subgroups and interactions | X |
| (*c*) Explain how missing data were addressed | X |
| (*d*) If applicable, describe analytical methods taking account of sampling strategy | X |
| (*e*) Describe any sensitivity analyses | X |
| **Results** | | | |
| Participants | 13\* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | V |
| (b) Give reasons for non-participation at each stage | X |
| (c) Consider use of a flow diagram | X |
| Descriptive data | 14\* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | X |
| (b) Indicate number of participants with missing data for each variable of interest  (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) | X |
| Outcome data | 15\* | *Cohort study*—Report numbers of outcome events or summary measures over time  *Case-control study—*Report numbers in each exposure category, or summary measures of exposure  *Cross-sectional study—*Report numbers of outcome events or summary measures | V |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | V |
| (*b*) Report category boundaries when continuous variables were categorized | X |
| (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | X |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | X |
| **Discussion** | | | |
| Key results | 18 | Summarise key results with reference to study objectives | V |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | V |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | V |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | V |
| **Other information** | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | X |

--

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Bijlage 7 scorelijst STROBE ***Lasbleiz et al. (2014)*** | | **Item No** | **Recommendation** | **Page No** |
| **Title and abstract** | | 1 | (*a*) Indicate the study’s design with a commonly used term in the title or the abstract | V |
| (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | V |
| **Introduction** | | | | |
| Background/rationale | | 2 | Explain the scientific background and rationale for the investigation being reported | V |
| Objectives | | 3 | State specific objectives, including any prespecified hypotheses | X |
| **Methods** | | | | |
| Study design | | 4 | Present key elements of study design early in the paper | V |
| Setting | | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | V |
| Participants | | 6 | (*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls  *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants  (*b*)*Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed  *Case-control study*—For matched studies, give matching criteria and the number of controls per case | V |
| Variables | | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | V |
| Data sources/ measurement | | 8\* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | V |
| Bias | | 9 | Describe any efforts to address potential sources of bias | V |
| Study size | | 10 | Explain how the study size was arrived at | V |
| Quantitative variables | | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | V |
| Statistical methods | | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | V |
| (*b*) Describe any methods used to examine subgroups and interactions | V |
| (*c*) Explain how missing data were addressed | V |
| (*d*) If applicable, describe analytical methods taking account of sampling strategy | X |
| (*e*) Describe any sensitivity analyses | X |
| **Results** | | | | |
| Participants | | 13\* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | V |
| (b) Give reasons for non-participation at each stage | V |
| (c) Consider use of a flow diagram | X |
| Descriptive data | | 14\* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | X |
| (b) Indicate number of participants with missing data for each variable of interest  (c) *Cohort study*—Summarise follow-up time (eg, average and total amount) | X |
| Outcome data | | 15\* | *Cohort study*—Report numbers of outcome events or summary measures over time  *Case-control study—*Report numbers in each exposure category, or summary measures of exposure  *Cross-sectional study—*Report numbers of outcome events or summary measures | V |
| Main results | | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | V |
| (*b*) Report category boundaries when continuous variables were categorized | X |
| (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | X |
| Other analyses | | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | X |
| **Discussion** | | | | |
| Key results | | 18 | Summarise key results with reference to study objectives | V |
| Limitations | | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | V |
| Interpretation | | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | V |
| Generalisability | | 21 | Discuss the generalisability (external validity) of the study results | V |
| |  |  |  |  | | --- | --- | --- | --- | | Bijlage 8 scorelijst AMSTAR ***Bélanger et al. (2019)*** | **Ja** | **Gedeeltelijk ja** | **Nee** | | 1. Did the research questions and inclusion criteria for the review include the components of PICO? | X |  |  | | 2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify and significant deviations from the protocol? |  | X |  | | 3. Did the review authors explain their selection of the study designs for inclusion in the review? | X |  |  | | 4. Did the review authors use a comprehensive literature search strategy? |  | X |  | | 5. Did the review authors perform study selection in duplicate? | X |  |  | | 6. Did the review authors perform data extraction in duplicate? | X |  |  | | 7. Did the review authors provide a list of excluded studies and justify the exclusions? |  | X |  | | 8. Did the review authors describe the included studies in adequate detail? |  | X |  | | 9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review? |  | X |  | | 10. Did the review authors report on the sources of funding for the studies included in the review? |  |  | X | | 11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? | X |  |  | | 12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis? | X |  |  | | 13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review? | X |  |  | | 14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? | X |  |  | | 15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review? |  |  | X | | 16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review? | X |  |  |   **Other information** | | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | | X |