

De diagnostische waarde van hoogcervicale instabiliteitstests: een systematische review

Auteurs:

drs. Nathan Hutting¹, dr. Wendy Scholten-Peeters², drs. Veerle Vijverman³, drs. Martin Keesenberg⁴, dr. Arianne Verhagen⁵

- 1) Fysiotherapeut / manueeltherapeut / wetenschappelijk onderzoeker / promovendus, Faculteit Geneeskunde en Farmacie, Vrije Universiteit Brussel, België en Lectoraat Arbeid & Gezondheid Hogeschool van Arnhem en Nijmegen. Email: nhutting@fysiotherapiehetcentrum.nl
- 2) Fysiotherapeut / manueeltherapeut / senior onderzoeker, Faculteit Geneeskunde en Farmacie, Vrije Universiteit Brussel, België en Lectoraat Diagnostiek, Avans Hogeschool en afdeling huisartsengeneeskunde Erasmus MC Rotterdam.
- 3) Fysiotherapeut / manueeltherapeut, Faculteit Geneeskunde en Farmacie, Vrije Universiteit Brussel, België en praktijk voor fysiotherapie in Geel, België.
- 4) Fysiotherapeut & klinisch gezondheidswetenschapper, Corpus Mentis, Centrum Fysiotherapie & Wetenschap in Leiden en bestuurslid Vereniging Fysiotherapie & Wetenschap
- 5) Fysiotherapeut / manueeltherapeut / lector, Lectoraat Diagnostiek, Avans Hogeschool en universitair hoofddocent afdeling huisartsengeneeskunde Erasmus MC Rotterdam.

Samenvatting

Doel: Evaluatie van de diagnostische waarde van de screeningstest voor hoog cervicale instabiliteit bij patiënten en gezonde personen.

Methode: Diagnostische systematische review volgens de standaarden van de Cochrane Collaboration. Twee onafhankelijke onderzoekers zochten, selecteerden en beoordeelde studies naar de diagnostische waarde van hoog cervicale instabiliteitstests bij patiënten of gezonde proefpersonen. In de studies moest de sensitiviteit en specificiteit zijn gerapporteerd of berekend kunnen worden met een 2x2 tabel. De methodologische kwaliteit is beoordeeld met de QUADAS-2. In het geval van klinische of statistische heterogeniteit werd een beschrijvende analyse uitgevoerd.

Resultaten: We vonden vijf studies waarin zeven test zijn onderzocht. De meeste tests zijn slecht één keer onderzocht, veelal in specifieke patiëntenpopulaties. Overall was de accuratesse matig.

Conclusie: Screenen voor hoogcervicale instabiliteit kan op dit moment niet valide worden gedaan.

Dit artikel is een vertaling van:

Hutting N, Scholten-Peeters GG, Vijverman V, Keesenberg MD, Verhagen AP. Diagnostic Accuracy of Upper Cervical Spine Instability Tests: A Systematic Review. Phys Ther. 2013.

Inleiding

De prevalentie van hoog cervicale instabiliteit is laag (0.6 %) en varieert tussen verschillende patiëntengroepen.¹ Hoog cervicale instabiliteit komt vaker voor bij patiënten met inflammatoire aandoeningen zoals reumatoïde artritis en de ziekte van Bechterew.^{2,3} Ook bij patiënten na een trauma en met congenitale afwijkingen (Syndroom van Down) komt hoog cervicale instabiliteit vaker voor.² Symptomen van hoog cervicale instabiliteit zijn wisselend en variëren van nekpijn, beperkte range of motion of torticollis tot neurologische verschijnselen.⁴

Hoofdpijn, nekpijn, en beperkte mobiliteit zijn veelvoorkomende redenen voor het toepassen van manuele- of fysiotherapeutische interventies.^{5,6} Echter, als er een verdenking is van (hoog)cervicale instabiliteit moeten patiënten niet behandeld worden met speciale technieken, zoals manipulaties, vanwege het mogelijke risico op complicaties. Om het risico op complicaties na behandeling zoveel mogelijk te beperken wordt aanbevolen om patiënten te screenen op de aanwezigheid van een hoog cervicale instabiliteit,^{5,7-11} met name bij patiënten met een verhoogd risico.⁸

In de literatuur worden verschillende klinische tests beschreven voor het screenen op hypermobiliteit en instabiliteit van de craniocervicale ligamenten,^{3,12} namelijk de Sharp-Purser test (SPT), de 'side-bending test', de hoogcervicale flexiekniktest en de laterale verschuifbaarheidstest. Deze tests zijn echter alleen bruikbaar als de betrouwbaarheid en validiteit van deze tests voldoende zijn.

Ondanks internationale aanbevelingen op screening van hoog cervicale instabiliteit,^{5,7-11} bestaat in de praktijk geen consensus over het gebruik van screeningstests. Uit een Australische survey bleek dat 34,3% van de leden van de 'Australian Physiotherapy Association' zelden instabiliteitstests voor de hoog cervicale wervelkolom uitvoerden en 23% nooit, terwijl 12,3% van de therapeuten alleen hoog cervicale instabiliteitstests uitvoerden wanneer ze cervicale manipulaties als behandeloptie overwogen.¹²

Er is weinig bekend over de diagnostische waarde van de hoog cervicale instabiliteitstests en er zijn, voor zover bekend, geen systematische reviews uitgevoerd over dit onderwerp. Het doel van deze systematische review is dan ook de diagnostische waarde van de hoog cervicale instabiliteitstests, in

termen van sensitiviteit, specificiteit, voorspellende waarden en likelihood ratios, te evalueren bij patiënten en gezonde personen.

Methode

Zoekstrategie

Twee onafhankelijke onderzoekers (NH, MK) zochten naar relevante literatuur in de elektronische databases van PUBMED, CINAHL, EMBASE en RECAL Legacy (tot eind oktober 2012). Er is gebruik gemaakt van MeSH termen (PUBMED), Thesaurus (EMBASE, CINAHL) en vrije tekstwoorden. De zoektermen waren gerelateerd aan de diagnostische waarden, manipulatie en hoog cervicale instabiliteit. De volledige zoekstrategie kan worden opgevraagd bij de corresponderende auteur van dit artikel.

Studie selectie

Studies die de diagnostische waarde van hoog cervicale instabiliteitstests bij patiënten of gezonde proefpersonen onderzochten, waarbij sensitiviteit en specificiteit waren gerapporteerd of berekend konden worden met een 2x2 tabel zijn geïnccludeerd. Alle artikelen zijn op titel en abstract beoordeeld door twee onafhankelijke onderzoekers (NH, MK). Van de relevante artikelen zijn de fulltekst opgezocht en beoordeeld aan de hand van de selectie criteria. Er was geen restrictie wat betreft het jaartal van publicatie of taal. In geval van onenigheden tussen beide reviewers werd een derde persoon (WSP) geraadpleegd. De referenties van de geïnccludeerde artikelen werden ook handmatig gecontroleerd op relevantie studies, die niet in de elektronische databases waren gevonden.

Data-extractie

Twee onafhankelijke onderzoekers (NH, WSP) voerden de data-extractie uit door middel van een gestandaardiseerd formulier. De volgende data werd geëxtraheerd: auteur, jaar van publicatie, kenmerken van de studiepopulatie, de indextest, de referentietest en de waarden van sensitiviteit,

specificiteit, voorspellende waarden en likelihood ratios. Wanneer geen ruwe data werden gerapporteerd in het artikel, is de eerste auteur van het artikel per email benaderd voor aanvullende informatie.

Methodologische kwaliteit

De methodologische kwaliteit van de geïncludeerde artikelen is onderzocht met behulp van de QUADAS-2.¹³ De QUADAS-2 bestaat uit vier domeinen die vertekening (bias) in kaart brengt met betrekking tot de selectie van patiënten, index test, referentiestandaard, patiëntstroom en tijdstip van index en referentie test.¹³

Twee onderzoekers (NH, MK) scoorden onafhankelijk van elkaar alle items als 'laag', 'hoog' of 'onduidelijk'. Eventuele verschillen in beoordeling werden onderling besproken en wanneer geen consensus bereikt kon worden is een derde persoon (WSP) geraadpleegd.

Datasynthese en analyse

De overeenkomst van de methodologische kwaliteit tussen beide reviewers werd berekend en weergegeven met kappa waarde (κ). Een kappa waarde werd beoordeeld als slechte (≤ 0.00), geringe (0.00-0.2), matige (0.21-0.40), redelijke (0.41-0.60), voldoende tot goede (0.61-0.80) of bijna perfecte overeenstemming (0.81-1.00).¹⁴

In principe worden de gegevens bij elkaar opgeteld (gepooled), maar in het geval van klinische (verschil in patiënten, index test of referentie test) of statistische heterogeniteit ($I^2 > 40\%$) werd een beschrijvende analyse uitgevoerd.

Wanneer de diagnostische waarden niet waren weergegeven in het oorspronkelijke artikel zijn de ruwe data gebruikt om de diagnostische waarden met behulp van een 2x2 tabel te berekenen. Bij lege cellen in de 2x2 tabel is de waarde 0.5 toegevoegd aan elk veld in de 2x2 tabel.¹⁵

Een sensitiviteit, specificiteit en voorspellende waarde van minimaal 80% is als voldoende beschouwd. Een positieve likelihood ratio van >10 en een negatieve likelihood ratio van <0.1 werd tevens beschouwd als voldoende.¹⁶

Resultaten

Studie selectie

De zoekstrategie in de verschillende databases leverde in totaal 773 potentiële artikelen op. Figuur 1 geeft de flowchart weer van het selectieproces. Na het verwijderen van dubbele referenties en selectie op basis van titel en abstract, bleven vier studies¹⁷⁻²⁰ over waarvan de fulltekst werd bekeken. Van deze artikelen zijn drie studies geïncludeerd.¹⁸⁻²⁰ Eén studie werd geëxcludeerd omdat deze studie geen diagnostische waarden of ruwe data rapporteerde.¹⁷ Via handmatig doorzoeken van de referenties van de geïncludeerde artikelen werden 2 extra studies getraceerd.^{21,22} Uiteindelijk voldeden 5 studies aan de selectiecriteria en zijn opgenomen in deze systematische review.³¹⁻³⁵

De eerste auteur van één studie is gecontacteerd met het verzoek tot aanlevering van ruwe data en data likelihood ratios.¹⁹ We ontvingen alleen de likelihood ratios van deze studie.

Beschrijving van de studies

Karakteristieken van de geïncludeerde studies zijn weergegeven in tabel 1. Vier studies zijn gepubliceerd tussen 1969 en 1999.^{18,19,21,22} Het gemiddeld aantal geïncludeerde personen was 115.5 (range 31-123). Slechts drie studies rapporteerden de gemiddelde leeftijd van de proefpersonen.^{19,20,22} Vier studies includeerden patiënten met reumatoïde artritis.^{18,19,21,22} In drie studies was de diagnose reumatoïde artritis gesteld op basis van de criteria van de American Rheumatism Association^{18,21,22} en in één studie waren de criteria niet duidelijk beschreven.¹⁹ Eén studie includeerde patiënten met whiplash associated disorders (WAD).²⁰ Twee studies includeerden ook gezonde proefpersonen^{20,21}, twee studies includeerden patiënten die het ziekenhuis bezochten^{18,22} en één studie includeerden patiënten die de afdeling fysiotherapie van een ziekenhuis bezochten.²¹ In een andere studie was de

setting niet duidelijk.¹⁹ De prevalentie van hoog cervicale instabiliteit in de geïnccludeerde studies varieerde tussen de 0.07 en 0.44 (%?).

Index tests

In totaal werden zeven verschillende tests geëvalueerd: de Sharp-Purser test (SPT),^{18,19,21,22} 'clunking test',²¹ 'palate sign',²¹ test van het ligament alare (TLA),²⁰ test van het ligament transversum atlantis (TLTA),²⁰ test van het tectoriaal membraan van het atlanto-axiale gewricht (TTM)²⁰ en test van het posterieur atlanto-occipitaal membraan (TPAM)²⁰ (zie tabel 1). Alleen de SPT werd geëvalueerd in meer dan één studie. De interbeoordelaars betrouwbaarheid weergegeven in een kappa waarde van de SPT varieerde tussen 0.06 en 0.67 (k) en de SPT werd als niet betrouwbaar beschouwd.^{19,23} Over de betrouwbaarheid van de andere geëvalueerde tests is geen informatie bekend.

Referentie tests

Vier studies vergeleken één of meer instabiliteitstests met röntgenonderzoek als referentietest.^{18,19,21,22} De atlas dens interval (ADI) werd gemeten op een laterale röntgenfoto.^{18,19,21,22} In één studie werd de ADI gemeten in cervicale flexie en in extensie,¹⁸ in drie andere studies alleen in flexie.^{19,21,22} In vier studies werd een ADI ≥ 3 mm geïnclassificeerd als abnormaal.^{18,19,21,22} Eén studie berekende ook sensitiviteit en specificiteit met een ADI ≥ 4 mm.¹⁸ Eén studie vergeleek de TLA, TLTA, TTM en TPAM met Magnetic Resonance Imaging (MRI).²⁰

Methodologische kwaliteit

De interbeoordelaars betrouwbaarheid van de beoordeling van de methodologische kwaliteit met behulp van de QUADAS-2 was kappa (k) = 0.58 (Betrouwbaarheidsinterval 0.32-0.83), wat een redelijke betrouwbaarheid aangeeft. De totale overeenstemming tussen de reviewers was 80%. Alle meningsverschillen werden opgelost tijdens een consensus bijeenkomst.

De verschillende studies vertoonden verschillende typen vertekening van de resultaten (bias) en bij ongeveer de helft de items werd de optie 'onduidelijk' gescoord (zie tabel 2). Dit was met name het geval bij de patiëntselectie en de index- en referentietest.

Diagnostische waarden

De sensitiviteit, specificiteit, voorspellende waarden en likelihood ratios van de verschillende tests zijn weergegeven in tabel 3. We berekenden diagnostische waarden met de ruwe data uit drie studies.^{18,21,22}

Het statistisch poolen van de data van de drie studies die de SPT onderzochten was niet mogelijk vanwege statistische heterogeniteit (sensitiviteit $I^2 = 87.8\%$, specificiteit $I^2 = 91.6\%$).

De sensitiviteit varieerde van 0.19 tot 0.96 en de specificiteit varieerden van 0.91 tot 1.00 (zie tabel 3).

De positief voorspellende waarden van de 'clunking test' en het 'palate sign' waren niet voldoende en die van de SPT waren variabel. De negatief voorspellende waarden van de 'clunking test' en het 'palate sign' waren voldoende en die van de SPT variabel. In het algemeen waren de voorspellende waarden van de TLA, TLTA, TTM en TPAM voldoende.

Discussie

Belangrijkste bevindingen

De diagnostische waarden van de SPT, de enige test die in meerdere studies werd onderzocht, waren over het algemeen niet voldoende. De sensitiviteit van nagenoeg alle tests was onvoldoende; dit betekent dat de test niet gebruikt kan worden om hoog cervicale instabiliteit uit te sluiten. De betrouwbaarheidsintervallen van de likelihood ratios waren echter groot, hetgeen duidt op een gemis aan precisie. De methodologische kwaliteit van de meeste studies was slecht tot matig, waardoor de resultaten mogelijk vertekent zijn.

Klinische implicaties

Met het screenen van patiënten op hoog cervicale instabiliteit, willen we voorkomen dat deze patiënten risicovol behandeld worden, waardoor ze kans lopen op een complicatie. Het is daarom belangrijk om vals-negatieve resultaten te voorkomen, dat wil zeggen dat de patiënt toch een behandeling van de hoog cervicale wervelkolom ondergaat terwijl er sprake is van instabiliteit. Om dit te bereiken moet de sensitiviteit van de tests hoog zijn.²⁴ De sensitiviteit van de meeste tests is onvoldoende om hoog cervicale instabiliteit uit te sluiten. De klinische waarde van deze tests is dan ook laag.

Voldoende specificiteit geeft de mate aan waarmee een test in staat is om fout-positieve resultaten te voorkomen. Specificiteit is in dit kader minder belangrijk dan sensitiviteit omdat een fout-positief resultaat de patiënt niet kan schaden.²⁴ Echter, een in potentie effectieve behandelmethodede kan de patiënt hierdoor worden onthouden.

De likelihood ratio is een alternatief om de diagnostische waarde van een test weer te geven.²⁵

Likelihood ratios ≥ 10 en ≤ 0.1 worden verondersteld sterk bewijs te leveren voor het aantonen of uitsluiten van een aandoening.¹⁶ De likelihood ratios van de TTM en TPAM waren voldoende, hetgeen aangeeft dat deze tests gebruikt zouden kunnen worden om hoog cervicale instabiliteit uit te sluiten. Echter, de betrouwbaarheidsintervallen waren grootmogelijk door de kleine onderzoekspopulatie (lage sample size). Daarnaast werden deze tests slechts één keer geëvalueerd. Niettemin zijn deze tests het meest veelbelovend en toekomstige studies zullen de klinische waarde van deze tests verder aan dienen te tonen.

Een diagnostische test is van belang wanneer deze tevens een acceptabele betrouwbaarheid heeft. De hoog cervicale flexiekniktest toonde een acceptabele intra- en interbetrouwbaarheid bij kinderen met het Syndroom van Down.²³ Dit is echter niet generaliseerbaar naar de patiënt met nekpijn in de praktijk. De SPT en de laterale verschuifbaarheidstest waren niet betrouwbaar.^{19,23}

Sterke en zwakke punten van de studie

Dit is de eerste systematische review die de diagnostische waarde van de hoog cervicale instabiliteitstests heeft onderzocht. Helaas werden de meeste tests slechts eenmaal onderzocht en geen van de studies onderzocht de waarde van de tests als screeningtest in de gewone fysiotherapie praktijk. Dit betekent dat de generaliseerbaarheid van de resultaten naar de praktijk laag is. De prevalentie van hoog cervicale instabiliteit is laag, maar is in de tweede lijn vaak hoger dan bij patiënten in de eerste lijn.^{20,26}

De prevalentie van hoog cervicale instabiliteit zoals aangetoond door de referentietest tussen de geïncludeerde studies varieerde. Deze varieerde zelfs in één studie die verschillende tests onderzocht.^{18,20,21} Het is vreemd dat de prevalentie varieerde omdat dezelfde populatie gebruikt werd. Er was in de geïncludeerde studies onvoldoende rapportage van gegevens, mogelijk omdat een aantal studies relatief oud zijn. Dit kan leiden tot verschillende typen bias en dit beperkt de kracht van onze conclusies.

De interbeoordelaars betrouwbaarheid van het scoren van de methodologische kwaliteit met de QUADAS-2¹³ was redelijk. Verschillen tussen de onderzoekers werden voornamelijk veroorzaakt door leesfouten en verschillen in interpretatie. Beide beoordelaars hadden moeilijkheden met het scoren met de QUADAS-2¹³ vanwege vaak onduidelijke of onvolledige informatie in de studies. Het wordt aanbevolen het rapporteren in de toekomst te verbeteren door gebruik te maken van de STARD richtlijnen.²⁷

Vier studies^{18,19,21,22} gebruikten het ADI op een röntgenopname om instabiliteit vast te stellen. Ondanks dat deze methode frequent wordt gebruikt bij het vaststellen van hoog cervicale instabiliteit, heeft deze methode enkele beperkingen. Omdat de test-hertest betrouwbaarheid van een röntgenopname niet voldoende is,²⁸ is dit niet de perfecte referentie test. Volgens de 'American Association of Radiologists', zijn de referentiewaarden voor het vaststellen van instabiliteit een ADI > 2.5-3 mm voor volwassenen en een ADI > 4.5-5 mm voor kinderen.²³ Ongeveer 95% van de gezonde personen heeft een ADI van 0.3-1.8 mm in flexie, van 0.4-0.2 mm in de neutrale positie en van 0.3-2.2 mm in extensie.⁴ Omdat de ADI is klein is, dienen beoordelaars bewust zijn op de kans op meetfouten.⁴ Enkele auteurs

hebben gesuggereerd dat de diameter van het spinale kanaal een beter diagnostisch criterium is.^{4,29} Anderen geven aan dat Computed Tomography (CT) en MRI de voorkeur hebben als referentietest voor het diagnosticeren van hoog cervicale instabiliteit.⁴ Een studie vergeleek het verlies van collageen integriteit, beoordeeld op MRI (graad 0-3), met de mate van vergrootte mobiliteit beoordeeld middels klinisch onderzoek (categorie 0-3).²⁰ Deze auteurs gingen na of laesies van een specifieke nekstructuur invloed hadden op de mobiliteit van die structuur. Tot op heden is wordt MRI gezien als de beste manier om collageen integriteit van weke delen structuren te visualiseren,³⁰ en wordt gezien als de gouden 'standaard'.²⁰ Eén studie onderzocht of de resultaten van een klinische test (TLA, TLTA, TTM, TPAM) correspondeerde met tekenen van lichamelijk letsel, beoordeeld op MRI.²⁰ Bij letsel van weke delen kan in deze regio een vergrote mobiliteit worden verwacht.²⁰ Helaas classificeerden de auteurs deze vergrote beweeglijkheid niet als instabiliteit.

Conclusies

De geïncludeerde studies vertonen verschillende vormen van bias en de sensitiviteit is variabel. Wij concluderen dat screenen voor hoogcervicale instabiliteit op dit moment niet valide kan worden gedaan. De TTM en TPAM vertonen de beste accuratesse bij WAD patiënten in slechts één kleine studie.

References

1. Beck RW, Holt KR, Fox MA, Hurtgen-Grace KL. Radiographic anomalies that may alter chiropractic intervention strategies found in a New Zealand population. *J Manipulative Physiol Ther.* 2004;27(9):554-559.
2. Swinkels R, Beeton K, Alltree J. Pathogenesis of upper cervical instability. *Man Ther.* 1996;1(3):127-132.
3. Aspinnall W. Clinical testing for the craniovertebral hypermobility syndrome. *J Orthop Sports Phys Ther.* 1990;12(2):47-54.
4. Swinkels RA, Oostendorp RA. Upper cervical instability: fact or fiction? *J Manipulative Physiol Ther.* 1996;19(3):185-194.
5. Rivett DA, Thomas L, Bolton P. Pre-manipulative testing: where do we go from here? *NZ journal of physiotherapy.* 2005;33(3):78-84.
6. Colloca CJ, Pickar JG, Slosberg M. Special focus on spinal manipulation. *J Electromyogr Kinesiol.* 2012;22(5):629-631.
7. Refshauge KM, Parry S, Shirley D, Larsen D, Rivett DA, Boland R. Professional responsibility in relation to cervical spine manipulation. *Aust.J Physiother.* 2002;48(3):171-179.
8. Mintken PE, Metrick L, Flynn TW. Upper cervical ligament testing in a patient with os odontoideum presenting with headaches. *J Orthop Sports Phys Ther.* 2008;38(8):465-475.
9. Di Fabio RP. Manipulation of the cervical spine: risks and benefits. *Phys Ther.* 1999;79(1):50-65.
10. Hing WA, Reid DA, Monaghan M. Manipulation of the cervical spine. *Man Ther.* 2003;8(1):2-9.
11. Rushton A RD, Carlesso L, Flynn T, Hing W, Kerry R. *International Framework for Examination of the Cervical Region for potential of Cervical Arterial Dysfunction prior to Orthopaedic Manual Therapy Intervention: International Federation of Orthopaedic Manipulative Physical Therapists; 2012.*
12. Osmotherly PG, Rivett DA. Knowledge and use of craniovertebral instability testing by Australian physiotherapists. *Man Ther.* 2011;16(4):357-363.
13. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
15. Altman D, Machin D, Bryant T, Gardner MJ. *Statistics with confidence: Confidence intervals and statistical guidelines.* London: BMJ Books; 2000.
16. Guyatt G, Rennie D, Meade M, Cook M. *Users' Guides to the Medical Literature.* Chicago: McGraw-Hill Professional; 2008.
17. Olson-KA, Paris-SV, Spohr-. Radiographic assessment and reliability study of the craniovertebral sidebending test. 1998.
18. Uitvlugt G, Indenbaum S. Clinical assessment of atlantoaxial instability using the Sharp-Purser test. *Arthritis Rheum.* 1988;31(7):918-922.
19. Forrester G, Barlas P. Reliability and validity of the Sharp-Purser test in the assessment of atlanto-axial instability in patients with rheumatoid arthritis. *Physiotherapy.* 1999;85(7):376-376.
20. Kaale BR, Krakenes J, Albrektsen G, Wester K. Clinical assessment techniques for detecting ligament and membrane injuries in the upper cervical spine region--a comparison with MRI results. *Man Ther.* 2008;13(5):397-403.
21. Mathews JA. Atlanto-axial subluxation in rheumatoid arthritis. *Ann.Rheum.Dis.* 1969;28(3):260-266.
22. Stevens JC, Cartlidge NE, Saunders M, Appleby A, Hall M, Shaw DA. Atlanto-axial subluxation and cervical myelopathy in rheumatoid arthritis. *Q.J Med.* 1971;40(159):391-408.
23. Cattrysse E, Swinkels R, Oostendorp R, Duquet W. Upper cervical instability: are clinical tests reliable? *Man Ther.* 1997;2(2):91-97.

24. Hutting NV, A.P.; Vijverman, V.; Keesenberg, D.M.; Dixon, G.; Scholten-Peeters, G.G.M. Diagnostic accuracy of premanipulative vertebrobasilar insufficiency tests: a systematic review. *Man Ther.* 2013;18:177-182.
25. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ.* 2004;329(7458):168-169.
26. Neva MH, Hakkinen A, Makinen H, Hannonen P, Kauppi M, Sokka T. High prevalence of asymptomatic cervical spine subluxation in patients with rheumatoid arthritis waiting for orthopaedic surgery. *Ann.Rheum.Dis.* 2006;65(7):884-888.
27. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem.* 2003;49(1):7-18.
28. Gajdosik CG, Ostertag S. Cervical instability and down syndrome: Review of the literature and implications for physical therapists. *Pediatr Phys Ther.* 1996(8):31-36.
29. White KS, Ball WS, Prenger EC, Patterson BJ, Kirks DR. Evaluation of the craniocervical junction in Down syndrome: correlation of measurements obtained with radiography and MR imaging. *Radiology.* 1993;186(2):377-382.
30. Erickson SJ. High-resolution imaging of the musculoskeletal system. *Radiology.* 1997;205(3):593-618.

Tabellen




Tabel 1: Kenmerken van de geïncludeerde studies

| Eerste auteur, jaar van publicatie | Populatie (n) | Gemiddelde leeftijd (spreiding) in jaren | Index test | Referentietest |
|------------------------------------|---------------------|--|--|--|
| Forrester, 1999 ¹⁹ | RA (31) | 57.5 (29-74) | -Sharp-Purser test | Röntgenfoto (lateraal, flexie) |
| Kaale, 2008 ²⁰ | WAD (92) GP (30) | 39.1 45.3 | -Test van het ligament alare -Test van het ligament transversum atlantis -Test van het tectoriaal membraan van het atlanto-axiale gewricht -Test van het posterieur atlanto-occipitaal membraan | MRI |
| Mathews, 1969 ²¹ | RA (76) GP (28) | | - Sharp-Purser test - 'Clunking test' - 'Palate sign' | Röntgenfoto (lateraal, flexie) |
| Stevens, 1971 ²² | RA (100) | 54.2 (21-75) | -Sharp-Purser test | Röntgenfoto (lateraal, flexie) |
| Uitvlugt, 1988 ¹⁸ | RA (123) | | -Sharp-Purser test | Röntgenfoto (lateraal, flexie, extensie) |

RA=reumatoïde artritis, WAD=whiplash associated disorders, GP=gezonde personen, MRI=magnetic resonance imaging

Tabel 2: Beoordeling methodologische kwaliteit met de QUADAS-2

| Studie | Risico op bias | | | | Zorgen mbt toepassing | | |
|-------------------------------|------------------|------------|----------------------|--------------------------|-----------------------|------------|----------------------|
| | Patiënt Selectie | Index Test | Referentie-standaard | Patiënt stroom en timing | Patiënt Selectie | Index Test | Referentie-standaard |
| Forrester, 1999 ¹⁹ | ? | ? | ? | ? | 😊 | 😊 | ? |
| Kaale, 2008 ²⁰ | 😊 | 😊 | 😊 | 😊 | 😊 | ? | 😊 |
| Mathews, 1969 ²¹ | 😞 | 😊 | ? | 😊 | 😊 | 😊 | ? |
| Stevens, 1971 ²² | 😊 | ? | ? | 😊 | 😊 | 😊 | ? |
| Uitvlugt, 1988 ¹⁸ | 😊 | ? | ? | ? | 😊 | 😊 | ? |

 Laag risico
  Hoog risico
  Onduidelijk risico

Tabel 3: Diagnostische waarden van de geïnccludeerde studies

| Eerste auteur, jaar van publicatie | Test | Prevalentie | Sensitiviteit (95% BI) | Specificiteit (95% BI) | Positief voorspellende waarde (95% BI) | Negatief voorspellende waarde (95% BI) | Positieve likelihood ratio (95% BI) | Negatieve likelihood ratio (95% BI) |
|-------------------------------------|-------|-------------|------------------------|------------------------|--|--|-------------------------------------|-------------------------------------|
| Forrester, 1999¹⁹ | SPT | ? | 0.43 | 0.77 | 0.34 | 0.89 | 1.88 (1.14–3.09) | 0.74 (0.53–1.02) |
| Kaale, 2008²⁰ | TLA-R | 0.44 | 0.69 (0.56-0.81) | 1.00 (1.00-1.00) | 1.00 (1.00-1.00) | 0.80 (0.71-0.89) | 94.09 (5.91-1498.02) | 0.33(0.22-0.47) |
| | TLA-L | 0.44 | 0.72 (0.60-0.84) | 0.96 (0.91-1.00) | 0.93 (0.85-1.01) | 0.81 (0.73-0.90) | 16.37 (5.35-50.10) | 0.29 (0.19-0.48) |
| | TLTA | 0.35 | 0.65 (0.51-0.79) | 0.99 (0.96-1.01) | 0.97 (0.90-1.03) | 0.84 (0.76-0.91) | 51.44 (7.25-365.12) | 0.35 (0.12-0.53) |
| | TTM | 0.13 | 0.94 (0.82-1.06) | 0.99 (0.97-1.01) | 0.94 (0.82-1.06) | 0.99 (0.97-1.01) | 98.44 (13.94-695.15) | 0.06 (0.01-0.42) |
| | TPAM | 0.19 | 0.96 (0.87-1.04) | 1.00 (1.00-1.00) | 1.00 (1.00-1.00) | 0.99 (0.97-1.01) | 185.6 (11.67-2952.63) | 0.04 (0.01-0.30) |
| Mathews, 1969²¹ | SPT | 0.41 | 0.19 (0.08–0.38) | 0.71 (0.55–0.83) | 0.32 (0.14–0.57) | 0.56 (0.42–0.70) | 0.67 (0.29–1.57) | 1.13 (0.94–1.38) |
| | CL | 0.08 | 0.33 (0.06–0.76) | 0.76 (0.64–0.85) | 0.11 (0.02–0.35) | 0.93 (0.82–0.98) | 1.37 (0.41–4.58) | 0.88 (0.50–1.56) |
| | PS | 0.09 | 0.71 (0.30–0.95) | 0.80 (0.68–0.88) | 0.26 (0.10–0.51) | 0.96 (0.87–0.99) | 3.52 (1.82–6.82) | 0.36 (0.11–1.16) |
| Stevens, 1971²² | SPT | 0.36 | 0.44 (0.28–0.62) | 0.98 (0.90–1.00) | 0.94 (0.70–1.00) | 0.76 (0.65–0.84) | 28.44 (3.93–205.72) | 0.56 (0.42–0.76) |
| Uitvlugt, 1988¹⁸ | SPT | 0.26 | 0.69 (0.50–0.83) | 0.96 (0.89–0.99) | 0.85 (0.64–0.95) | 0.90 (0.81–0.95) | 15.64 (5.83–41.94) | 0.33 (0.20–0.55) |
| | SPT* | 0.20 | 0.88 (0.68–0.97) | 0.96 (0.89–0.99) | 0.85 (0.64–0.95) | 0.97 (0.91–0.99) | 21.56 (8.17–56.91) | 0.13 (0.04–0.36) |

BI=betrouwbaarheidsinterval, SPT=Sharp-Purser Test, CL='clunking test', PS='palate sign', TLA-R=test van het rechter ligament alare, TLA-L=test van het linker ligament alare, TLTA=test van het ligament transversum atlantis, TTM=test van het tectoriaal membraan van het atlanto-axiale gewricht, TPAM=test van het posterieur atlanto-occipitaal membraan

*met atlas dens interval >4 mm

Figuren

Figuur 1: Flowchart van de geïncludeerde studies

