

Systematische review en meta-analyse

Validiteit en betrouwbaarheid van palpatie bij het SI-gewricht

Bron: Ribeiro RP, Guerrero FG, Camargo EN, Beraldo LM, Candotti CT. Validity and Reliability of Palpatory Clinical Tests of Sacroiliac Joint Mobility: A Systematic Review and Meta-analysis. J Body Move Ther. 2021;44(4);307-318

Lage rugpijn (LRP) is het meest voorkomende gezondheidsprobleem in de wereld. Het sacro-iliacaal gewricht (SIG) kan hierbij een bron van pijn zijn. Mogelijke bewegingen in het SIG bestaan uit rotaties en translaties. Bij het onderzoek van het SIG maken manueel therapeuten gebruik van klinische tests onderverdeeld in drie categorieën: positietests, pijn-provocatietests en mobiliteit-tests.

Tekst: Ilse Swinkels

Bij de mobiliteit-tests worden vaak routinematig palpatietests gebruikt voor het vaststellen van een dysfunctie van het SIG. De betrouwbaarheid van deze tests was echter ontmoedigend en de validiteit is nog niet vastgesteld. Systematische reviews naar SIG-tests zijn meer dan twintig jaar oud en mogelijk zijn er sindsdien nieuwe tests ontwikkeld of zijn er aanpassingen van oude tests. Vandaar dat een update naar deze palpatietests nodig is.

Doel studie

1. Onderzoek naar de betrouwbaarheid en validiteit van palpatietests voor de mobiliteit van het SIG.
2. Inventarisatie van deze palpatietests van het SIG.

Hypothese

Er bestaan veel palpatietests voor het SIG, maar slechts weinige met voldoende betrouwbaarheid of validiteit.

Methode en materiaal

Design: Systematische review en meta-analyse

De zoektocht werd verricht in vijf databases: PubMed, Embase, Scopus, Medline en de Physiotherapy Evidence Database. Er waren geen restricties betreffende

populatie, publicatiejaar of onderzoeksdesign. Alleen publicaties in de Engelse en Portugese taal werden geïnccludeerd. Studies naar pijnprovocatietests van het SIG werden geëxcludeerd. Twee auteurs includeerden de publicaties en verrichten onderzoek naar de methodologische kwaliteit van de studies. Bij disconsensus beoordeelde een derde auteur de publicaties. Dezelfde twee auteurs verrichtten de data-extractie, zoals grootte van de populatie, design, onderzochte tests, data betreffende betrouwbaarheid, aantal onderzoekers/testers, concurrente validiteit, of en welke gouden standaard was gebruikt en de statistische tests die werden gebruikt. Om te worden geïnccludeerd in de meta-analyse van de betrouwbaarheid van een test moesten er minstens twee publicaties betreffende die test zijn met minimaal een matige betrouwbaarheid (kappa van meer dan 0,40; $k > 0,40$).

Resultaten

Van de 74 potentieel relevante studies waren er uiteindelijk 28 die voldeden aan de inclusiecriteria. Van de 14 studies naar de betrouwbaarheid van palpatietests van het SIG waren er 3 van lage methodologische kwaliteit. Bij de methodologische beoordeling bleken vooral de randomisatie van klinische tests en de tijdsinterval tussen twee tests laag te scoren.

In totaal werden er 15 verschillende palpatietests voor de beoordeling van de mobiliteit van het SIG besproken. Hierbij werden de 'standing flexion test' (STFT), de 'sitting flexion test' (SIFT) en de Gillet-test het meest frequent onderzocht. 14 studies analyseerden de interbeoordelaarsbetrouwbaarheid en 8 analyseerden tevens de intrabeoordelaarsbetrouwbaarheid.

Er was geen enkele studie die de concurrente validiteit heeft onderzocht.

De betrouwbaarheid van tests worden in een tabel weergegeven. Hierbij wordt soms de percentage of agreement (PoA) gegeven, soms de kappa en bij een aantal studies werden beide berekend. De studie van Arab et al. heeft de betrouwbaarheid van vier SIG-tests onderzocht voor zowel intra- als interbeoordelaars betrouwbaarheid.¹ Dit waren de STFT, de SIFT, de Gillet-test en de Prone knee flexion test. In deze studie scoorden alle vier de tests boven een kappa van 0,40 op beide maten van betrouwbaarheid. Bij de overige studies bleek de interbeoordelaarsbetrouwbaarheid te variëren tussen -0,05 en 0,40 op één studie na waar de kappa 0,77 was. Bij slechts drie studies wordt een betrouwbaarheidsinterval gegeven.

De STFT, SIFT en Gillet-test waren de enige tests in de meta-analyse. Er werd geen meta-analyse van de interbeoordelaarsbetrouwbaarheid verricht vanwege het gebrek aan studies met een kappa $> 0,40$. Uit de meta-analyse van de intrabeoordelaarsbetrouwbaarheid bleek dat deze voor de STFT en Gillet-test niet significant was. Alleen de intrabeoordelaarsbetrouwbaarheid van de SIFT-test bleek significant.

Discussie

De interbeoordelaarsbetrouwbaarheid is de overeenkomst tussen verschillende beoordelaars die dezelfde schaal, classificatie, hetzelfde meetinstrument, dezelfde procedure en evaluatie van een object of subject gebruiken. De intrabeoordelaarsbetrouwbaarheid, ook wel test-hertest genoemd, is de overeenkomst binnen een en dezelfde beoordelaar die dezelfde schaal, classificatie, hetzelfde meetinstrument, dezelfde

procedure en evaluatie van een object of subject gebruikt op verschillende tijdstippen. De betrouwbaarheid van een meet-instrument is belangrijker dan de validiteit. Een test heeft immers geen waarde wanneer deze niet betrouwbaar is.

In het geval van de STFT-test is de intrabeoordelaarsbetrouwbaarheid ($0,46 < k < 0,70$) beter dan de interbeoordelaarsbetrouwbaarheid ($0,05 < k < 0,51$). Dit was ook het geval bij de overige tests. Er bleek überhaupt een grote heterogeniteit in de berekende kappa-scores van de verschillende tests en studies, vooral bij de berekening van de interbeoordelaarsbetrouwbaarheid, wat vraagt om meer onderzoek. Daarnaast lijkt het dus voor de herevaluatie van een patiënt beter te zijn dat dezelfde onderzoeker de patiënt evalueert, dan dat het door verschillende onderzoekers gebeurt.

De meta-analyse toonde dat alleen de SIFT-test een statistisch significante intrabeoordelaarsbetrouwbaarheid heeft. Maar dit resultaat werd berekend vanuit slechts twee studies die overigens wel een goed methodologische kwaliteit bezaten.

Bij LRP is er veel onderzoek gedaan naar Clinical Diagnostic Rules (CDR). Deze lijken een hoge mate van specificiteit te hebben bij LRP veroorzaakt door spondylolisthesis, discaal veroorzaakte radiculopathie en spinale stenose. De auteurs onderschrijven het belang van CDR en stellen voor dat in de toekomst vervolgstudies zouden kunnen focussen op een CDR van mobiliteit tests van het SIG waarbij dan vooral gekeken kan worden naar de SIFT, STFT en Gillet-test omdat naar deze tests al het meeste onderzoek is gedaan en deze tests de beste overeenkomstniveaus hadden.

Beperkingen

Voor de meta-analyse werden slechts 2 of 3 studies gebruikt wat erg weinig is. Daarnaast was het een beperking dat alleen de methodologische kwaliteit van die studies waarin ook een statische analyse was verricht is onderzocht.

Conclusie

Er werden 15 palpatietests voor SIG-mobiliteit gevonden. Geen enkele studie evalueerde de validiteit van deze tests. Uit de meta-analyse bleek alleen de SIFT-tests voldoende betrouwbaar en alleen voor wat betreft de intrabeoordelaarsbetrouwbaarheid.

Bespreking

De auteurs hebben onderzoek gedaan naar de betrouwbaarheid en validiteit van palpatietests voor de beoordeling van de mobiliteit van het SIG. Ze geven aan dat de laatste review uit 2000 is. Dit was correct, maar net voor de publicatie van deze review (2021) is in 2020 een vergelijkbare review naar de klinimetrische eigenschappen van mobiliteitstests van het SIG gepubliceerd.² Hoewel er in deze review niet volledig dezelfde tests werden beoordeeld als in onderhavig review is de conclusie wel hetzelfde. Namelijk dat de betrouwbaarheid van het onderzoeken van de mobiliteit van het SIG niet is vastgesteld en dat men zich zou moeten afvragen wat men eigenlijk onderzoekt wanneer men deze tests gebruikt, omdat de validiteit niet is onderzocht/vastgesteld. Dit houdt in dat men voorzichtig moet zijn om diagnostische dan wel evaluatieve conclusies te trekken wanneer bij een patiënt de mobiliteit van het SIG wordt onderzocht.

Bij deze review werd de methodologische kwaliteit onderzocht door middel van een lijst volgens Brink en Louw.³ De laatste jaren wordt hiervoor echter vaak de COSMIN-checklist gebruikt, wat een goed onderzocht en inmiddels veelvuldig gebruikt instrument is. Reviews zouden makkelijker vergelijkbaar zijn wanneer auteurs dezelfde lijst zouden gebruiken bij de beoordeling van de methodologische kwaliteit van studies.

Wat verder opvalt is dat bij de meta-analyse alleen studies met een kappa-score boven 0,40 worden geïnccludeerd. Wat hiervoor de reden is wordt door de auteurs niet beschreven. Door het vooraf al uitsluiten van studies die een lage score op betrouwbaarheid hebben wordt de uitslag van de meta-analyse beïnvloed. Die valt dan hoger uit dan in werkelijkheid het geval is.

Wanneer naar de resultaten van de 14 studies wordt gekeken valt op dat er één studie is die op alle tests die in die studie werden onderzocht hoog scoorde, hoger dan alle andere studies en zowel voor de inter- als intrabeoordelaarsbetrouwbaarheid. Dit is opvallend en wordt door de auteurs niet benoemd noch verklaard. De betreffende studie is door die hoge kappa-scores sterk aanwezig in de meta-analyse.

Al met al lijkt me de conclusie gerechtvaardigd dat het evalueren van de mobiliteit van het SIG met terughoudendheid

moet worden gedaan en dus ook de consequenties die men hieruit trekt voor wat betreft de in te zetten/ingezette interventie.

De door de auteurs van deze review vooraf gegeven hypothese “Er bestaan veel palpatietests voor het SIG, maar slechts weinigen met voldoende betrouwbaarheid of validiteit” dient derhalve te worden aangenomen.



Dr. Ilse Swinkels-Meewisse,
Amsterdam UMC, locatie VUmc



swinky@xs4all.nl



Literatuur:
www.kngf.nl/fysiopraxis