

Praktijkgericht Onderzoek

Welk classificatiesysteem is het meest effectief met betrekking tot de interventies ten aanzien van pijn en functioneren bij patiënten met aspecifieke lage rugpijn? Een systematische review.

10 juni 2016

Aspirant-onderzoekers:

S.C. Bekhuis
519018

M.G.J. Jansen
520915

F.A.M. Kuijpers
522616

A.J.M. Peek
523730

van Hogeschool Arnhem & Nijmegen te Nijmegen

Opdrachtgever:

Martin Oortwijn

Laurens Brummelhuis

van Arcus Fysiotherapie te Zutphen

Docentbegeleider:

Bart Staal

van Hogeschool Arnhem & Nijmegen te Nijmegen

SAMENVATTING

Achtergrond: Lage rugpijn is een veelvoorkomende klacht. Voor 90% tot 95% van alle patiënten met lage rugpijn is geen specifieke lichamelijke oorzaak aan te wijzen¹. Door de onduidelijkheid over de aard van de aandoening zijn er richtlijnen ontworpen waar behandeltrajecten beschreven staan². In de KNGF richtlijn³ worden de patiënten opgedeeld in patiëntenprofielen waar een behandeltraject aan gekoppeld is. Voor de therapeuten binnen Arcus Fysiotherapie is er geen duidelijke richtlijn voor het maken van de keuze welk behandeltraject het beste bij de patiënt past, waardoor de patiënt automatisch het behandeltraject doorloopt van de expertise van de therapeut waar de patiënt wordt geplaatst.

Doel: Het doel van dit onderzoek is om te onderzoeken welk classificatiesysteem, met betrekking tot interventies bij patiënten van 16-65 jaar met specifieke lage rugpijn, het meest effectief is ten aanzien van pijn en functioneren.

Studiedesign: Systematische review, meta-analyse en focusgroep.

Databases: Google Scholar, HANQuest, MEDLINE, PEDro en PubMed

Methode: Aan de hand van een trechtersvorm strategie zijn artikelen gezocht aan de hand van een zoekschema waarbij trefwoorden en synoniemen met behulp van Boleaanse operatoren zijn verbonden. De Randomised Controlled Trials zijn gerandomiseerd door een externe persoon middels loting over twee groepen en door iedere onderzoeker onafhankelijk van elkaar beoordeeld om zo het risico op bias te voorkomen. De Randomised Controlled Trials zijn geselecteerd op basis van de Cochrane RCT Checklist⁴ (te voldoen aan 7/10 criteria van 'Validiteit') en de TIDieR Checklist⁵ (te voldoen aan 8/12 criteria). De uitkomstmaten van de geïnccludeerde artikelen zijn weergegeven in de studiekarakteristieken ter verduidelijking voor de data-analyse. De uitkomstmaten betreffen pijn en functioneren op korte termijn (0-3 maanden), middellange termijn (3-6 maanden) en lange termijn (>12 maanden). De resultaten zijn gepoold en uitgezet in forest plots door middel van Review Manager⁶. De kwaliteit van bewijs is beoordeeld door middel van de GRADE methodiek⁷.

Resultaten: Totaal zijn n=111 artikelen geselecteerd aan de hand van titel/samenvatting en full text. Uiteindelijk zijn totaal n=11 RCT's meegenomen in de data-analyse. De artikelen zijn verdeeld over de classificatiesystemen n=3 RCT's van Treatment-Based Classification (TBC)⁸⁻¹⁰, n=3 RCT's van Mechanical Diagnosis and Treatment (MDT)¹¹⁻¹³, n=3 RCT's van

Movement System Impairment (MSI)¹⁴⁻¹⁶, n=1 RCT van Start Back Screening Tool (SBST)¹⁷ en n=1 RCT van O'Sullivan Classification (OSC)¹⁸.

Uit de meta-analyse geeft de Start Back Screening tool op korte termijn voor pijn het beste resultaat SMD=0.20; 95%-BI [0.04, 0.37]. Op lange termijn voor pijn geeft de Mechanical Diagnosis and Treatment het beste resultaat SMD=0.23; 95%-BI [0.02, 0.44]. Op korte termijn voor functioneren geeft de Mechanical Diagnosis and Treatment het beste resultaat SMD=0.17; 95%-BI [-0.01, 0.35]. Op lange termijn geeft de Treatment-Based Classification het beste resultaat SMD=0.22; 95%-BI [-0.35, 0.78]. Met een gemiddelde van 2 in de rangorde van de resultaten staat de Mechanical Diagnosis and Treatment op de hoogste plek vergeleken met andere classificatiesystemen. In de rangorde op basis van kwaliteit van bewijs scoren de Treatment-Based Classification en de Mechanical Diagnosis and Treatment in het niveau hoog. Doordat de Treatment-Based Classification meer gepoolde uitkomstmaten in het niveau hoog scoort, waar de Mechanical Diagnosis and Treatment in het niveau redelijk scoort, zal de Treatment-Based Classification sterker in bewijs zijn. Uit de focusgroep is gebleken dat een combinatie van de Treatment-Based Classification met het psychosociale aspect van de O'Sullivan Classification het beste geïmplementeerd kan worden bij Arcus Fysiotherapie op basis van de effectschatting, mate van kwaliteit en toepasbaarheid in de praktijk.

Conclusie: Uit het onderzoek wordt geconcludeerd dat de Treatment-Based Classification op basis van effectschatting en kwaliteit van bewijs het meest effectief is ten aanzien van pijn en functioneren bij patiënten van 16-65 jaar met specifieke lage rugpijn.

Aanbevelingen: Voor Arcus Fysiotherapie wordt, op basis van effectschatting, kwaliteit van bewijs, de focusgroep en toepasbaarheid in de praktijk, het samengestelde classificatiesysteem van de Treatment-Based Classification met het psychosociale aspect van de O'Sullivan Classification aanbevolen ter implementatie in de praktijk.

Er is verder onderzoek nodig naar het onderlinge effect van de classificatiesystemen en naar het samengestelde classificatiesysteem vergeleken met de reguliere fysiotherapie.

ACHTERGROND

Lage rugpijn is een veelvoorkomende klacht. Van de Nederlandse bevolking zal jaarlijks 20% een episode van lage rugpijn doormaken. Circa 6,8% van de Nederlandse bevolking komt terecht bij de huisarts en 1,4% wordt doorverwezen naar de paramedische zorg, voornamelijk naar de fysiotherapeut. Na invoering van de directe toegankelijkheid fysiotherapie (DTF) zijn deze percentages veranderd. Gedurende de periode van 2008 t/m

2010 kwam circa 15% van de patiënten met lage rugpijn bij de eerstelijns fysiotherapeuten voor een behandeling³. De incidentie stijgt met de leeftijd. De meeste kans op lage rugpijn is bij een leeftijd van 45 tot 64 jaar, vanaf 65 jaar neemt de incidentie af. De aandoening wordt zowel bij mannen als vrouwen even vaak vastgesteld¹. Bij patiënten met lage rugpijn gaat 25% gepaard met ziekteverzuim², waardoor lage rugpijn de belangrijkste reden is voor ziekteverzuim op het werk³. Het ontstaan van de klachten hangt samen met zowel de fysieke belasting van arbeid als de psychosociale factoren volgens Burdorf². Uit onderzoek van Burdorf² is gebleken dat persoonlijke kenmerken en klachtenprestatie invloed hebben op de prognose van het verzuim. Het kortdurend verzuim is werkgerelateerd in tegenstelling tot het langdurig verzuim wat beïnvloed wordt door psychosociale factoren².

Op welke criteria wordt de aandoening ingedeeld? Er wordt onderscheid gemaakt tussen specifieke en aspecifieke lage rugpijn. Specifieke lage rugpijn is een aandoening met een specifieke lichamelijke oorzaak die door middel van diagnostiek vastgesteld kan worden³. Wat de term aspecifieke lage rugpijn (ALRP) exact inhoudt, is nog onduidelijk. Men weet dat dit een aandoening is wanneer er geen valide specifieke pathologie gevonden kan worden. Nog steeds wordt er onderzoek gedaan om duidelijkheid te creëren. Bij 90% tot 95% van alle patiënten met lage rugpijn is geen specifieke lichamelijke oorzaak aan te wijzen¹. Doordat het een niet goed aantoonbare diagnose is, zal dit onbevredigend zijn voor de patiënt, omgeving en de maatschappij. Door onduidelijkheid over de aard van de aandoening zijn er richtlijnen ontworpen waar behandeltrajecten beschreven staan². In de KNGF richtlijn: Lage rugpijn³ worden patiënten opgedeeld in patiëntenprofielen gebaseerd op de gezondheidstoestand, acute/subacute/chronische fase, en de beïnvloedbaarheid van de belemmerende factoren van het gezondheidsprobleem. Aan de profielen zijn behandeltrajecten gekoppeld die als richtlijn gehanteerd kunnen worden³.

Tot op heden wordt bij Arcus Fysiotherapie het besluit van het behandeltraject van de patiënt gebaseerd op de richtlijn door de desbetreffende therapeut. Dit kan per therapeut verschillen vanwege zijn/haar expertise. De plaatsing van de patiënt wordt gebaseerd op welke behandelaar als eerste plek heeft, verwijzing met een voorkeur voor een behandelaar of door inschatting van de receptionisten. Doordat de patiënt automatisch het traject van de behandelaar volgt, betekent dit niet direct dat dit behandeltraject het beste bij de patiënt past. Arcus Fysiotherapie wil meer uniformiteit in het maken van de keuze welk behandeltraject de patiënt gaat doorlopen. Door middel van een classificatiesysteem met betrekking tot de interventies kan een richtlijn gecreëerd worden voor de therapeuten om te bepalen welk

behandeltraject het best bij de patiënt past, zodat de patiënt direct bij de goede therapeut terecht komt.

In dit onderzoek wordt aan de hand van een systematische review gebaseerd op bestaande classificatiesystemen een advies uitgebracht, met betrekking tot interventies bij patiënten van 16-65 jaar met aspecifieke lage rugpijn (ALRP), welk classificatiesysteem het meest effectief is ten aanzien van pijn en functioneren voor Arcus Fysiotherapie.

DOELSTELLING EN VRAAGSTELLING

Probleemstelling: Voor de therapeuten binnen Arcus Fysiotherapie is er geen duidelijke richtlijn voor het maken van de keuze welk behandeltraject het beste bij de patiënt past, waardoor de patiënt automatisch het behandeltraject doorloopt van de expertise van de therapeut waar de patiënt wordt geplaatst.

Doelstelling: Op 12 juni 2016 wordt aan de hand van een systematische review gebaseerd op bestaande classificatiesystemen een advies uitgebracht welk classificatiesysteem, met betrekking tot interventies bij patiënten van 16-65 jaar met aspecifieke lage rugpijn, het meest effectief is ten aanzien van pijn en functioneren voor Arcus Fysiotherapie.

Vraagstelling: Welk classificatiesysteem, met betrekking tot interventies bij patiënten van 16-65 jaar met aspecifieke lage rugpijn, is het meest effectief met betrekking tot pijn en functioneren voor Arcus Fysiotherapie?

Deelvragen

1. Welke classificatiesystemen voor lage rugpijn bestaan er in de literatuur?
2. Welk classificatiesysteem is het meest effectief ten aanzien van pijn en functioneren?
3. Welk classificatiesysteem vindt Arcus Fysiotherapie toepasbaar in hun praktijk?

PICO

Patient: Man of vrouw van 16 t/m 65 jaar met aspecifieke lage rugpijn

Intervention: classificatiesysteem

Comparison:

- alternatief classificatiesysteem
- reguliere fysiotherapie
- placebogroep

Outcome: het meest effectief ten aanzien van pijn en functioneren

PICO-vraag: Welk classificatiesysteem is het meest effectief ten aanzien van pijn en functioneren bij een man of vrouw van 16 t/m 65 jaar met aspecifieke lage rugpijn ten opzichte van een alternatief classificatiesysteem, reguliere fysiotherapie of een placebogroep?

METHODE

Selectiecriteria

In het oriënterend literatuuronderzoek zijn zowel systematische reviews als Randomised Controlled Trials (RCT) gebruikt als onderzoekseenheid. Uit het oriënterend literatuuronderzoek bleek welke classificatiesystemen voor aspecifieke lage rugpijn bestaan. Op basis van deze gegevens vond een literatuuronderzoek plaats waarbij de onderzoekseenheid expliciet uit RCT's bestond. De RCT's werden ten eerste geselecteerd op grond van de overeenkomst van de onderzoeksvraag met de titel/samenvatting. Vervolgens werd aan de hand van de full text bepaald of dit artikel daadwerkelijk bij de onderzoeksvraag paste. De artikelen die bij de onderzoeksvraag pasten, werden gescreend op de volgende selectiecriteria: aspecifieke lage rugpijn bij patiënten van 16-65 jaar, aanwezigheid van één van de controlegroepen: alternatief classificatiesysteem/reguliere fysiotherapie/placebogroep, daarnaast bevat het artikel één van de uitkomstmaten "pijn" en/of "functioneren" en het classificatiesysteem moet toepasbaar zijn in de eerstelijnspraktijk (*tabel 1*). Wanneer het artikel aan vier van de vijf criteria voldeed, werd deze meegenomen in het verdere onderzoek. In het literatuuronderzoek werden de artikelen geselecteerd op basis van de Cochrane RCT Checklist⁴ en de TIDieR Checklist⁵. De artikelen behoorden aan zeven van de tien criteria te voldoen van het kopje "*Validiteit*" van de Cochrane RCT Checklist en daarnaast moest er een betrouwbaarheidsinterval en/of standaarddeviatie aanwezig zijn voor de uitkomstmaten pijn en/of functioneren. De artikelen behoorden aan acht van de twaalf criteria te voldoen van de TIDieR Checklist om meegenomen te worden in het vervolgonderzoek.

Tabel 1: Selectiecriteria

Criteria	+	-	?
Aspecifieke lage rugpijn			
Mannen en vrouwen van 16-65 jaar			
Eén van de controlegroepen is aanwezig: alternatief classificatiesysteem/reguliere fysiotherapie/placebogroep			
Eén van de uitkomstmaten is aanwezig: pijn of functioneren			
Het classificatiesysteem is toepasbaar in de eerstelijns praktijk			

Zoekstrategie en data-extractie

Oriënterend literatuuronderzoek

Het oriënterend literatuuronderzoek is uitgevoerd om bestaande classificatiesystemen te inventariseren. Deze zijn gezocht door middel van het zoekschema van *tabel 6*. Trefwoorden en synoniemen werden met elkaar verbonden met behulp van Boleaanse operatoren. Door middel van een trechtersvorm strategie leverden de zoektermen/MeSh-terms van de trefwoorden 'stroomdiagram' en 'lage rugpijn' de meeste resultaten op. De databases die werden gebruikt zijn PEDro, Google Scholar, PubMed, MEDLINE en HANQuest. In dit literatuuronderzoek zijn zowel systematische reviews als RCT's meegenomen. Systematische reviews zijn gebruikt om tot RCT's te komen. Aan de hand van dit onderzoek zijn de artikelen ingedeeld op namen van classificatiesystemen en auteurs. Dit leidde tot een nieuw zoekschema voor het literatuuronderzoek (*tabel 7*).

Literatuuronderzoek

Aan de hand van het zoekschema (*tabel 7*) is het literatuuronderzoek uitgevoerd, gebaseerd op de PICO vraag. Trefwoorden en synoniemen zijn met elkaar verbonden met behulp van Boleaanse operatoren. De databases die werden gebruikt zijn PEDro, Google Scholar, PubMed, MEDLINE en HANQuest.

In dit onderzoek zijn expliciet RCT's meegenomen. De artikelen zijn eerst geselecteerd op de titel/samenvatting. Vervolgens zijn de artikelen gescreend op full text of ze bij de onderzoeksvraag passen. Ieder artikel werd door twee onderzoekers onafhankelijk van elkaar beoordeeld op de selectiecriteria van *tabel 1*. Dit is vervolgens besproken. Wanneer de twee onderzoekers het niet met elkaar eens waren, werd een derde aspirant-onderzoeker gevraagd om de doorslaggevende factor te zijn of een artikel wel of niet meegenomen werd in het vervolgonderzoek. De zoekresultaten zijn beschreven in *figuur 1*.

Risico op bias

De onderzoekers zijn opgesplitst in groepen, waarbij in iedere groep de onderzoekers onafhankelijk van elkaar de artikelen beoordeelden. Indien een derde persoon nodig was, werd deze gevraagd uit de andere groep. Groep 1 bestond uit S.C. Bekhuis en M.G.J. Jansen. Groep 2 bestond uit F.A.M. Kuijpers en A.J.M. Peek. Middels loting door een

externe persoon zijn de gevonden RCT's gerandomiseerd over de twee groepen.

Ieder artikel werd door twee onderzoekers onafhankelijk van elkaar beoordeeld aan de hand van de selectiecriteria (*tabel 1*). Doordat er gebruik is gemaakt van een externe persoon en doordat de onderzoekers onafhankelijk van elkaar hebben beoordeeld wordt voorkeur van artikelen, voorkeur van toewijzing van de artikelen en selectiebias uitgesloten.

De artikelen zijn door dezelfde onderzoekers beoordeeld aan de hand van de Cochrane RCT Checklist en de TIDieR Checklist. Hierbij is de aanwezigheid van de criteria beoordeeld op ja (+), nee (-) of onduidelijk (?).

Data-analyse

De gevonden artikelen uit het literatuuronderzoek zijn meegenomen in de meta-analyse. De artikelen zijn geanalyseerd wanneer de artikelen een gemiddelde en een spreidingsmaat (standaarddeviatie, betrouwbaarheidsinterval) bevatten van de uitkomstmaten pijn en/of functioneren. De artikelen zijn gepoold op de uitkomstmaten en follow-up. De studiekarakteristieken van de studies zijn weergegeven in *tabel 8*. Hierin is de studie, patiëntenpopulatie, interventie- en controlegroep, follow-up, uitkomstmaten, resultaten en de conclusie van de artikelen beschreven die bijdragen aan de meta-analyse.

De meta-analyse is uitgevoerd door middel van Review Manager⁶. Er is gebruik gemaakt van een random effect model, omdat tussen de groepen verschillen aanwezig zijn (behandeling, follow-up). Tevens is gebruik gemaakt van de standaard mean differences (SMD), omdat verschillende meetinstrumenten voor dezelfde uitkomstmaten zijn gebruikt, waardoor het gegeneraliseerd kan worden. Voor een grotere precisie van de effectschatting werd gebruik gemaakt van de 95% betrouwbaarheidsinterval. De resultaten zijn weergegeven middels een tabel met de gepoolde effectschatting en forest plots. Tevens is de rangorde, van 1. meest effectief t/m 5. minst effectief, van de classificatiesystemen per follow-up weergegeven in een tabel met daarbij de gemiddelde plek welk classificatiesysteem het meest effectief is. Van de RCT's is iedere uitkomstmaat tevens beoordeeld door middel van de GRADE-methode op kwaliteit van bewijs^{7,19} en ingedeeld op niveau van bewijs; hoog, redelijk, laag of zeer laag bewijs (*tabel 2*).

Tabel 2: Niveaus van kwaliteit van bewijs GRADE²²

Kwaliteit	Interpretatie
Hoog	Het werkelijk effect ligt dicht in de buurt van de schatting van het effect.
Matig	Het werkelijk effect ligt waarschijnlijk dicht bij de schatting van het effect, maar er is een mogelijkheid dat het hier substantieel van afwijkt.
Laag	Het werkelijke effect kan substantieel verschillend zijn van de schatting van het effect.
Zeer laag	Het werkelijke effect wijkt waarschijnlijk substantieel af van de schatting van het effect.

De volgende vijf factoren kunnen de kwaliteit van bewijs verlagen en worden per uitkomstmaat bepaald:

1. kans op vertekening door beperkingen in uitvoering en opzet van het onderzoek (risk of bias);
2. inconsistenties in de onderzoeksresultaten (inconsistentie);
3. indirect bewijs;
4. onnauwkeurigheid van het geschatte effect van de interventie;
5. publicatiebias (andere factoren)"²²

Aan de hand van deze factoren worden de resultaten van de meta-analyse versterkt of verzwakt. De onderzoekspopulatie ligt bij fysiotherapeutische studies vrijwel altijd onder de 300 patiënten, waardoor de nauwkeurigheid van een uitkomstmaat de gradering ernstig toebedeeld krijgt. Wanneer de spreiding van de betrouwbaarheidsindex tevens groot is, zal de uitkomstmaat de gradering zeer ernstig toebedeeld krijgen. Gevolg is dat de kwaliteit van bewijs al snel in niveau laag geplaatst wordt. De norm is dat een uitkomstmaat een goede kwaliteit van bewijs heeft wanneer deze in niveau hoog-redelijk valt. Wanneer twee systemen gelijk scoorden op niveau van kwaliteit van bewijs, werd gekeken naar de hoeveelheid studies die meegenomen zijn in de analyse en hoeveel uitkomstmaten in het niveau hoog scoren. Meerdere studies en meerdere uitkomstmaten in het niveau hoog zorgen voor een krachtiger bewijs.

Bovenstaande gegevens zijn meegenomen in de keuze van het classificatiesysteem.

Focusgroep

De conclusie uit het literatuuronderzoek is meegenomen in de focusgroep met Arcus Fysiotherapie. De focusgroep vond plaats met 6-12 personen, bestaande uit de opdrachtgevers Martin Oortwijn & Laurens Brummelhuis, andere collegae en de aspirant-onderzoekers. Zij uitten hun ideeën, belangen en meningen omtrent de classificatiesystemen. Ter voorbereiding op de focusgroep zijn vragen opgesteld en is informatie uit de systematische review geanalyseerd²⁰.

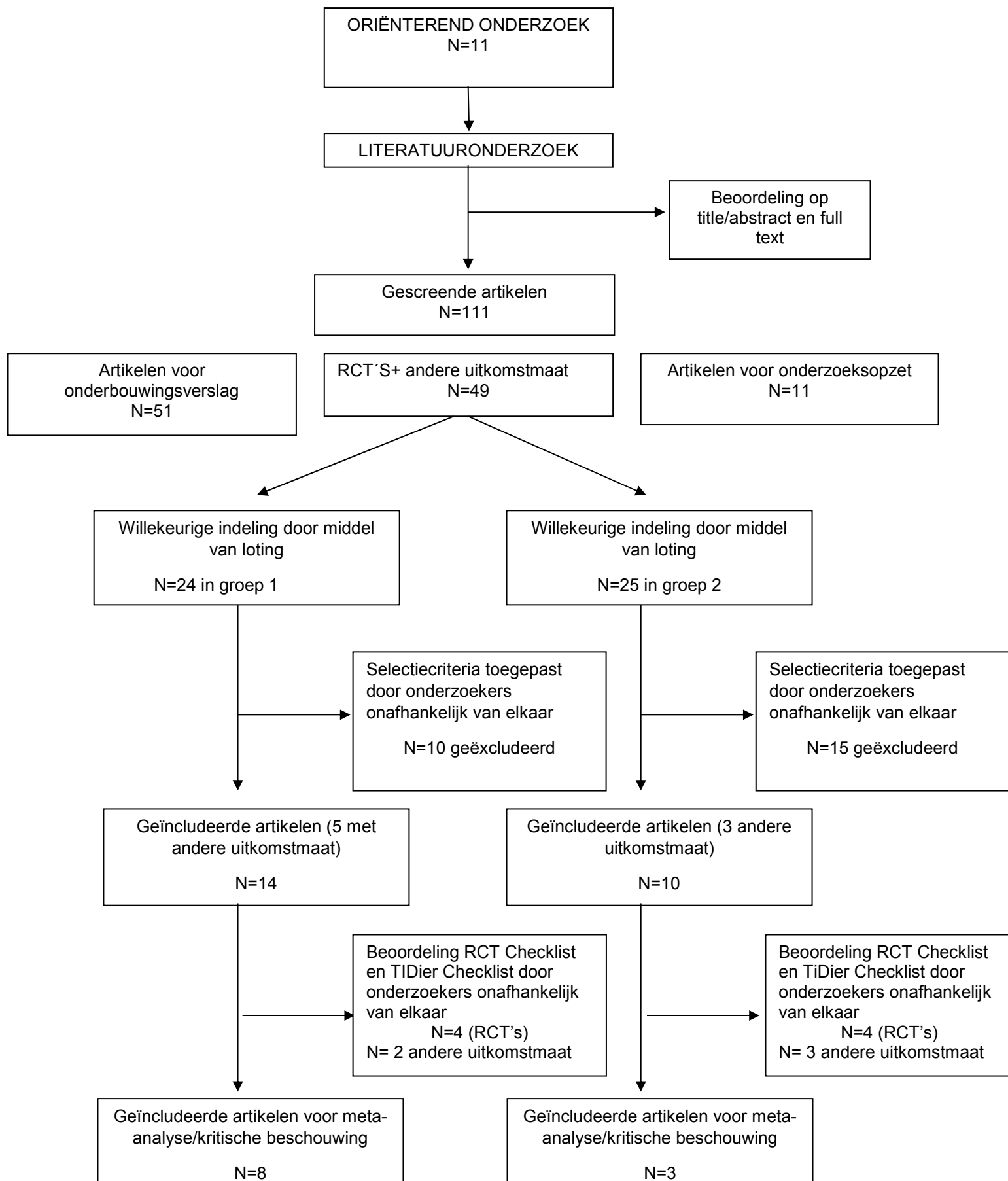
De focusgroep is opgenomen door middel van een recorder. Deze informatie is uitgeschreven in het onderbouwingsverslag²⁴. De conclusie uit het focusgroepinterview is het antwoord op de onderzoeksvraag. Dit classificatiesysteem zal in de toekomst bij Arcus Fysiotherapie geïmplementeerd kunnen worden.

RESULTATEN

Resultaten literatuuronderzoek

In totaal zijn n=111 artikelen geselecteerd op basis van titel/samenvatting en full text (*figuur 1*). Na de screening zijn alleen de RCT's en artikelen met een andere soort uitkomstmaat, die een gemiddelde en betrouwbaarheidsinterval/standaarddeviatie bevatten, meegenomen. N=62 artikelen zijn afgevalen, aangezien dit geen RCT's waren. N=49 artikelen zijn door middel van een loting door een externe persoon gerandomiseerd over groep 1 (n=24) en groep 2 (n= 25) artikelen. Ten eerste zijn de artikelen beoordeeld op de selectiecriteria (*tabel 1*). In groep 1 zijn n=10 artikelen geëxcludeerd (voldoen niet aan minimaal vier van de vijf selectiecriteria)²⁵⁻³⁴ en n=14 artikelen geïncludeerd, waarvan 4 artikelen met een andere uitkomstmaat en 10 RCT's. In groep 2 zijn n=15 artikelen geëxcludeerd (voldoen niet aan minimaal vier van de vijf inclusie criteria)³⁵⁻⁴⁹ en n=10 artikelen geïncludeerd, waarvan 3 met een andere uitkomstmaat en 7 RCT's. Na de beoordeling door middel van de Cochrane RCT Checklist en de TIDieR checklist zijn in groep 1 n=6 artikelen geëxcludeerd, waarvan n=4 RCT's en n=2 met een andere uitkomstmaat⁵⁰⁻⁵⁵ (voldoen niet aan minimaal acht van de twaalf criteria bij de TIDieR checklist en zeven van de tien bij de Cochrane RCT Checklist). In groep 2 n=7 artikelen geëxcludeerd, waarvan n=4 RCT's en n=3 andere uitkomstmaat⁵⁶⁻⁶¹ (voldoen niet aan minimaal acht van de twaalf criteria bij de TIDieR checklist en zeven van de tien bij de Cochrane RCT Checklist). In totaal zijn n=11 RCT's meegenomen in de meta-analyse in Review Manager en GRADE.

Figuur 1: studie flow-diagram²¹



Interventies

In totaal zijn n=11 artikelen meegenomen in de meta-analyse. De artikelen zijn verdeeld over verschillende classificatiesystemen, namelijk Treatment-Based Classification (TBC)⁸⁻¹⁰, Mechanical Diagnosis and Treatment (MDT)¹¹⁻¹³, Movement System Impairment (MSI)¹⁴⁻¹⁶, Start Back Screening Tool (SBST)¹⁷ en O'Sullivan Classification (OSC)¹⁸.

Uitkomstmaten en follow-up

De artikelen zijn ingedeeld op de follow-up op korte termijn (0-3 maanden), middellange termijn (3-6 maanden) en lange termijn (12 maanden of langer) en op de uitkomstmaten pijn en functioneren. Voor pijn op korte termijn deden in totaal 5 classificatiesystemen en 9 studies mee, voor pijn op lange termijn deden in totaal 3 classificatiesystemen en 3 studies mee, voor functioneren op korte termijn deden in totaal 5 classificatiesystemen en 11 studies mee en voor functioneren op lange termijn deden in totaal 3 classificatiesystemen en 5 studies mee. Er zijn geen studies voor pijn en/of functioneren op middellange termijn meegenomen.

De uitkomstmaten van de artikelen per classificatiesysteem zijn te vinden in *tabel 8*.

Resultaten meta-analyse

Resultaten pijn

De resultaten van de meta-analyse voor pijn zijn weergegeven in *tabel 3 en figuur 2a en 2b*. Op korte termijn geeft de SBST het beste resultaat SMD=0.20; 95%-BI [0.04, 0.37]. De OSC gaf het minste resultaat SMD=-1.18; 95%-BI [-1.79, -0.57]. Op korte termijn scoort de kwaliteit van bewijs van de MDT en TBC beiden in niveau redelijk. De gradering van de korte termijn van de MDT wordt toegekend wegens ernstige onnauwkeurigheid in het artikel van Garcia¹¹. De TBC scoort op korte termijn in niveau redelijk, vanwege ernstige onnauwkeurigheid in het artikel van Apeldoorn⁸. De MDT heeft meer gepoolde uitkomstmaten en wordt als hoogste in de rangorde geplaatst. De SBST heeft de laagste kwaliteit van bewijs en scoort in het niveau zeer laag. Deze gradering wordt toegekend, omdat zeer ernstig risico op bias en ernstige onnauwkeurigheid aanwezig is in het artikel van Murphy¹⁷ (*tabel 4, 5, 9.1-9.5*).

Op lange termijn geeft de MDT het beste resultaat SMD=0.23, 95%-BI [0.02, 0.44]. De MSI geeft op lange termijn het minste resultaat SMD=0.25; 95%-BI [-0.20, 0.71]. Op lange termijn scoort de MDT op kwaliteit van bewijs in niveau hoog en wordt als hoogste in de rangorde geplaatst. De MSI scoort op de kwaliteit van bewijs in niveau laag en wordt als laatste in de rangorde geplaatst. Deze gradering is toegekend wegens zeer ernstige onnauwkeurigheid in het artikel van Henry¹⁴ (*tabel 4, 5, 9.1-9.5*).

Tabel 3: Rangorde resultaten subgroepen

Classificatiesysteem /follow-up	Pijn korte termijn	Pijn lange termijn	Functioneren korte termijn	Functioneren lange termijn	Gemiddelde*
Treatment-Based Classification	2	3	3	1	9/4 = 2.25
Mechanical Diagnosis and Treatment	4	1	1	2	8/4 = 2
Movement System Impairment	3	2	2	3	10/4 = 2.5
Start Back Screening Tool	1	-	4	-	5/2 = 2.5
O'Sullivan Classification	5	-	5	-	10/2 = 5

*1 meest effectief t/m 5 minst effectief

Tabel 4: Rangorde kwaliteit van bewijs subgroepen

Classificatieschema/ follow-up	Pijn korte termijn	Pijn lange termijn	Functioneren korte termijn	Functioneren lange termijn	Gemiddelde*
Treatment-Based Classification	2	2	1	1	6/4 = 1.5**
Mechanical Diagnosis and Treatment	1	1	2	2	6/4 = 1.5**
Movement System Impairment	3	3	4	3	13/4 = 3.25
Start Back Screening Tool	5	-	5	-	10/2 = 5
O'Sullivan Classification	4	-	3	-	7/2 = 3.5

*1 meest effectief t/m 5 minst effectief

**De TBC en de MDT scoren gelijk op de geanalyseerde uitkomstmaten. Er is gekozen om de TBC op een eerste plaats te zetten, omdat deze met meerdere gepoolde uitkomstmaten in niveau hoog scoort, waar de MDT met deze uitkomstmaten in niveau redelijk scoort.

Tabel 5: Conclusies kwaliteit van bewijs

Classificatiesysteem	Conclusie
Mechanical Diagnosis and Treatment	De algemene kwaliteit van bewijs voor pijn en functioneren wordt als hoog-redelijk ingeschat. De korte termijn follow-ups van pijn en functioneren scoren beiden in niveau redelijk. Deze gradering is toegekend wegens ernstige onnauwkeurigheid ¹¹ . De lange termijn follow-ups van pijn en functioneren scoren beiden in niveau hoog. Deze resultaten bevatten geen factoren die de kwaliteit van bewijs verminderen.
Movement System Impairment	De algemene kwaliteit van bewijs voor pijn en functioneren wordt als laag ingeschat. Dit is voor alle follow-ups en uitkomstmaten geanalyseerd. Reden ter degradatie is zeer ernstige onnauwkeurigheid ¹⁴ .
O'Sullivan Classification	De algemene kwaliteit van bewijs voor pijn en functioneren wordt als redelijk-laag ingeschat. De follow-ups vinden alleen op korte termijn plaats. De uitkomstmaat functioneren scoort in niveau redelijk, wegens ernstige onnauwkeurigheid ¹⁸ . De uitkomstmaat pijn scoort in niveau laag, wegens zeer ernstige onnauwkeurigheid ¹⁸ .
Start Back Screening Tool	De Start Back Screening Tool scoort een volledige kwaliteit van bewijs in niveau zeer laag. Redenen ter degradatie zijn zeer ernstig risico op bias ¹⁷ en ernstige onnauwkeurigheid ¹⁷ .
Treatment-Based Classification	De algemene kwaliteit van bewijs voor pijn en functioneren wordt als redelijk-hoog ingeschat. De korte en lange termijn follow-ups van pijn scoren beiden in niveau redelijk. Deze gradering is toegekend wegens ernstige onnauwkeurigheid ⁸ . De korte en lange termijn follow-ups van functioneren scoren beiden in niveau hoog. Deze resultaten bevatten geen factoren die de kwaliteit van bewijs verminderen.

Resultaten functioneren

De resultaten van de meta-analyse voor functioneren zijn weergegeven in *tabel 3* en *figuur 2c* en *2d*.

Op korte termijn geeft de MDT het beste resultaat SMD=0.17; 95%-BI [-0.01, 0.35]. Het minste resultaat geeft de OSC SMD=-1.10; 95%-BI [-1.70, -0.49]. Op korte termijn scoren de TBC, MDT en OSC volledig in het niveau hoog-redelijk op kwaliteit van bewijs. De TBC scoort in niveau hoog en heeft de hoogste plaats in de rangorde. De resultaten van de TBC bevatten geen factoren die de kwaliteit van bewijs verminderen. De MDT en OSC scoren beiden in niveau redelijk, de MDT wordt hoger geplaatst in de rangorde omdat dit classificatiesysteem meerdere gepoolde uitkomstmaten heeft. De SBST heeft de laagste kwaliteit van bewijs en scoort in niveau zeer laag. Deze gradering wordt toegekend, omdat er zeer ernstig risico op bias en ernstige onnauwkeurigheid aanwezig is in het artikel van Murphy¹⁷ (*tabel 4, 5, 9.1-9.5*).

Op lange termijn geeft de TBC het beste resultaat SMD=0.22; 95%-BI [-0.35, 0.78]. De MSI geeft het minste resultaat SMD=-0.08; 95%-BI [-0.53, 0.37]. Op lange termijn scoren de TBC en MDT beide in het niveau hoog op kwaliteit van bewijs. De TBC wordt hoger geplaatst in de rangorde, omdat dit classificatiesysteem met meerdere gepoolde uitkomstmaten in niveau

hoog scoort, waar de MDT één uitkomstmaat meeneemt. Meerdere gepoolde uitkomstmaten zorgen voor een hogere kwaliteit van bewijs. De MSI wordt op een derde en laatste plaats gezet, deze uitkomstmaat scoort in niveau laag wegens zeer ernstige onnauwkeurigheid in het artikel van Henry¹⁴ (tabel 4, 5, 9.1-9.5).

Resultaten subgroepen

De classificatiesystemen zijn onderling met elkaar vergeleken per uitkomstmaat/follow-up. Er is een rangorde genomen per follow-up/uitkomstmaat en een algemeen gemiddelde rangorde. De rangorde is aangegeven van 1. meest effectief t/m 5. minst effectief. Voor de uitkomstmaat pijn op korte termijn is de SBST het meest effectief en staat daardoor op een eerste plaats in de rangorde van de meta-analyse. In de gemiddelde rangorde staat de SBST op een vierde plaats. Tevens is de kwaliteit van bewijs zeer laag, als gevolg neemt de effectgrootte en dus de klinische relevantie af.

Uit de gemiddelde rangorde scoort de MDT het hoogst en is dus het meest effectief. De effectgroottes bevinden zich gemiddeld meer aan de positieve kant van de forest plot. De TBC volgt de MDT in effectgrootte. Op basis van kwaliteit van bewijs is de TBC sterker van kwaliteit van bewijs dan de MDT, doordat de TBC meer gepoolde uitkomstmaten in niveau hoog scoort waar de MDT in het niveau redelijk scoort. Dit zal de effectgrootte van de TBC versterken en is de TBC klinisch relevanter (tabel 3, 4, 5).

Resultaten focusgroep

In de focusgroep is het advies uitgebracht op basis van de resultaten van de meta-analyse en kwaliteit van bewijs. Uit de meta-analyse geeft de Mechanical Diagnosis and Treatment de beste resultaten. De Treatment-Based Classification staat op de tweede plaats in de rangorde van de resultaten, maar deze scoort het hoogste op kwaliteit van bewijs en is daardoor sterker dan de Mechanical Diagnosis and Treatment (tabel 3, 4). Vergeleken met de Mechanical Diagnosis and Treatment is de Treatment-Based Classification overzichtelijker in gebruik, waardoor het beter toepasbaar is in de praktijk. Daarnaast neemt de Treatment-Based Classification meerdere vormen van therapie mee in vergelijking met de Mechanical Diagnosis and Treatment.

In de focusgroep zijn de verschillende classificatiesystemen vergeleken. Voor Arcus Fysiotherapie is het belangrijk dat er verschillende vormen van therapie wordt meegenomen en dat het makkelijk toepasbaar is in de praktijk. Hierdoor sloten ze zich aan bij het advies. Echter mist het psychosociale aspect in dit classificatiesysteem. Het psychosociale aspect van de O'Sullivan Classification sprak hen het meeste aan, ondanks dat dit classificatiesysteem niet effectief is volgens de meta-analyse en kwaliteit van bewijs. Uit de focusgroep is daardoor gebleken dat een combinatie van de Treatment-Based Classification

met een psychosociaal aspect van de O'Sullivan Classification kan worden geïmplementeerd in de praktijk van Arcus Fysiotherapie²⁴.

DISCUSSIE

Binnen de fysiotherapie bestaan meerdere classificatiesystemen die aangeven hoe een patiënt met specifieke lage rugpijn behandeld kan worden. Er is onderzoek gedaan waarbij de classificatiesystemen met reguliere fysiotherapie worden vergeleken, maar het onderling effect tussen de classificatiesystemen ten aanzien van pijnvermindering en toename van functioneren is nog niet onderzocht. In deze review wordt onderzoek gedaan naar welk classificatiesysteem het meest effectief is ten aanzien van pijn en functioneren bij patiënten met specifieke lage rugpijn.

De classificatiesystemen zijn vergeleken op basis van de resultaten uit de meta-analyse met behulp van Review Manager⁶ en de uitkomsten zijn op de kwaliteit van bewijs beoordeeld met behulp van GRADE⁷. De resultaten van de classificatiesystemen worden vergeleken op pijn en functioneren op korte en lange termijn. Ieder classificatiesysteem krijgt een plek van 1: het meest effectief t/m 5: het minst effectief. Per classificatiesysteem wordt in de meta-analyse naar zowel de gepoolde resultaten als de effectgrootte per artikel gekeken. Wanneer het betrouwbaarheidsinterval aan de positieve kant van de forest plot ligt, betekent dit dat de interventiegroep effectiever is dan de controlegroep.

Voor pijn op korte termijn is de Start Back Screening Tool¹⁷ (SBST) het meest effectief vergeleken met de rest van de artikelen, omdat er een significant verschil aanwezig is, $p=0.02$, en het betrouwbaarheidsinterval als enige volledig aan de positieve kant van de forest plot staat (*figuur 2a*). Dit betekent dat de SBST effectiever is dan de controlegroep. Echter is de kwaliteit van de SBST zeer laag, omdat het risico op bias zeer ernstig is en een ernstige onnauwkeurigheid aanwezig is in het artikel van Murphy¹⁷. Hierdoor neemt de klinische relevantie van dit classificatiesysteem af (*tabel 4, 5, 9.1-9.5*). De Treatment-Based Classification⁸⁻¹⁰ (TBC) volgt de SBST op basis van de effectgrootte op korte termijn, omdat het betrouwbaarheidsinterval grotendeels aan de positieve kant ligt vergeleken met de andere classificatiesystemen. Het gemiddelde van de Movement System Impairment¹⁴⁻¹⁶ (MSI) ligt net aan de positieve kant en ligt meer in het negatieve dan de TBC, waardoor het op de derde plek in de rangorde van de resultaten komt. De O'Sullivan Classification¹⁸ (OSC) staat volledig aan de negatieve kant, waardoor deze het minst effectief is vergeleken met de rest. Bij de overige artikelen is geen significant verschil zichtbaar (*tabel 3, figuur 2a*). Op basis van de kwaliteit van bewijs scoren de Mechanical Diagnosis and Treatment¹¹⁻¹³ (MDT) en de Treatment-Based Classification beide in niveau redelijk. De gradering van de MDT wordt toegekend vanwege ernstige onnauwkeurigheid in het artikel van Garcia¹¹. De

gradering van de TBC wordt toegekend vanwege ernstige onnauwkeurigheid in het artikel van Apeldoorn⁸. De MDT heeft daarentegen meer gepoolde uitkomstmaten, waardoor deze als hoogste in de rangorde van kwaliteit van bewijs wordt ingedeeld. Uit bovenstaande analyses blijkt dat de TBC zowel op effectgrootte als op kwaliteit van bewijs het beste scoort. (tabel 3, 4, 5, 9.1-9.5)

Voor pijn op lange termijn scoort de Mechanical Diagnosis and Treatment (MDT) het beste, omdat er een significant verschil zichtbaar is, $p=0.03$, en het betrouwbaarheidsinterval volledig aan de positieve kant ligt (figuur 2b). Tevens scoort de MDT hoog op kwaliteit van bewijs en staat daardoor op een eerste plaats in de rangorde voor kwaliteit van bewijs (tabel 3). Zowel de MSI als de TBC geeft geen significant verschil. Het betrouwbaarheidsinterval van de MSI ligt grotendeels aan de positieve kant. De TBC daarentegen ligt grotendeels aan de negatieve kant. Daarom volgt de MSI de MDT en geeft de TBC het minste effect voor pijn op lange termijn (tabel 3, figuur 2b). Voor pijn op lange termijn scoort de TBC hoger dan de MSI op kwaliteit van bewijs. De onnauwkeurigheid bij de MSI is zeer ernstig en bij de TBC is dit ernstig, waardoor de TBC sterker is op kwaliteit van bewijs. De TBC zal hoger in de rangorde van kwaliteit van bewijs staan (tabel 4, 5, tabel 9.1-9.5)

Voor functioneren op korte termijn is bij geen van de classificatiesystemen een significant verschil zichtbaar. De MDT is het meest effectief, omdat het betrouwbaarheidsinterval grotendeels aan de positieve kant ligt vergeleken met de andere artikelen. Vervolgens komt de MSI, omdat dit betrouwbaarheidsinterval kleiner is en het minst aan de negatieve kant staat ten opzichte van de TBC. Hoe kleiner het interval, des te betrouwbaarder het effect. Doordat het betrouwbaarheidsinterval van de MSI minder aan de negatieve kant ligt dan de TBC, is de MSI effectiever dan de TBC (tabel 3, figuur 2c). Echter scoort de TBC in het niveau hoog en zijn er geen factoren die de kwaliteit van het bewijs verminderen. De MDT scoort op basis van ernstige onnauwkeurigheid in het niveau redelijk, de MSI scoort op basis van zeer ernstige onnauwkeurigheid in het niveau laag. (tabel 4, 5, 9.1-9.5). Hierdoor neemt de klinische relevantie van de MSI en MDT af. Omdat de TBC een effectgrootte heeft die voornamelijk aan de positieve kant ligt en sterk is in kwaliteit van bewijs, neemt de klinische relevantie toe. De SBST heeft een kleiner betrouwbaarheidsinterval vergeleken met de TBC, maar het gemiddelde van de SBST ligt dicht bij de 'nul' waardoor de kans groter is op toeval vergeleken met de TBC (figuur 2c). De OSC ligt volledig aan de negatieve kant, waardoor de controlegroep effectiever is dan de interventiegroep. De OSC is daardoor vergeleken met de andere artikelen het minst effectief (figuur 2c). Daarnaast scoort de OSC op kwaliteit van bewijs in het niveau redelijk, omdat het een ernstige onnauwkeurigheid heeft. Hierdoor is de kans dat de OSC daadwerkelijk minder effectief is dan de controlegroep groter (tabel 4, 5, 9.1-9.5).

Voor functioneren op lange termijn is de TBC het meest effectief. In principe ligt het betrouwbaarheidsinterval van de TBC deels aan de negatieve kant en de MDT volledig aan de positieve kant. Desondanks heeft de TBC een grotere effectgrootte, omdat hierbij één artikel⁸ een negatief effect heeft en de andere twee artikelen^{9,10} volledig aan de positieve kant liggen. De MDT heeft één artikel¹³, waardoor dit minder zwaar weegt dan de artikelen van de TBC (*tabel 3, figuur 2d*). Op kwaliteit van bewijs scoort de TBC in het niveau hoog en zijn er geen factoren die de kwaliteit van bewijs verminderen, als gevolg neemt de klinische relevantie van de TBC toe. De MDT scoort in het niveau redelijk en scoort dus lager dan de TBC. Dit versterkt de klinische relevantie van de TBC (*tabel 4, 5, 9.1-9.5*). Het betrouwbaarheidsinterval van de MSI ligt grotendeels aan de negatieve kant van de forest plot, waardoor deze het minst effectief is vergeleken met de TBC en MDT (*figuur 2d*). Tevens scoort de MSI het laagst op kwaliteit van bewijs in het niveau laag op basis van zeer ernstige onnauwkeurigheid (*tabel 4, 5, 9.1-9.5*).

De classificatiesystemen zijn onderling vergeleken met een gemiddelde rangorde op basis van de meta-analyse en op kwaliteit van bewijs. In de rangorde van de meta-analyse is de Mechanical Diagnosis and Treatment het meest effectief, gevolgd door de Treatment-Based Classification. De O'Sullivan Classification is het minst effectief (*tabel 3*). In de rangorde op basis van kwaliteit van bewijs is de Treatment-Based Classification het sterkst in bewijs samen met de Mechanical Diagnosis and Treatment. Omdat de Treatment-Based Classification met meer gepoolde uitkomstmaten in niveau hoog scoort, waar de Mechanical Diagnosis and Treatment in het niveau redelijk scoort, zal het bewijs van de Treatment-Based Classification sterker zijn. De effectgrootte is daardoor bij de Treatment-Based Classification groter op basis van de meta-analyse en kwaliteit van bewijs groter en is dus klinisch relevant.

Uit de meta-analyse is de O'Sullivan Classification het minst effectief. De gepoolde resultaten zijn in vergelijking met de rest van de classificatiesystemen groot en staan tevens volledig aan de negatieve kant van de forest plot (*figuur 2a-d*). De controlegroep is effectiever dan de interventiegroep. Daarnaast scoort de O'Sullivan Classification¹⁸ voornamelijk in het niveau redelijk-laag, gebaseerd op een ernstig-zeer ernstige onnauwkeurigheid door een kleine steekproef en een groot betrouwbaarheidsinterval (*tabel 4, 5, 9.1-9.5*).

In de Focusgroep is het advies uitgebracht dat de Treatment-Based Classification het meest effectief is ten aanzien van pijn en functioneren op basis van de meta-analyse, kwaliteit van bewijs en de toepasbaarheid van het classificatiesysteem in de praktijk. Het classificatiesysteem is gemakkelijker toe te passen en overzichtelijker vergeleken met de

andere classificatiesystemen. Uit de focusgroep blijkt dat de wensen en behoeften van Arcus Fysiotherapie het beste overeenkomen met de Treatment-Based Classification. De Start Back Screening Tool is minder gericht op de behandeling, de Mechanical Diagnosis and Treatment beschrijft alleen mechanische aspecten en de O'Sullivan Classification en Movement System Impairment scoren het laagst in de meta-analyse. Echter mist de TBC het psychosociale aspect in de classificatie. Arcus Fysiotherapie is zoekende naar een classificatiesysteem waarin mechanische en psychosociale aspecten zijn verwerkt om zo tot het beste behandeltraject voor patiënten te komen. In de O'Sullivan Classification worden psychosociale aspecten meegenomen in tegenstelling tot de andere classificatiesystemen. Om bovenstaande redenen wordt de combinatie van de TBC met het psychosociale aspect van de OSC aanbevolen ter implementatie in de praktijk van Arcus Fysiotherapie.

Een discussiepunt in deze review is dat twee verschillende artikelen^{8,23} van dezelfde auteur over dezelfde patiëntenpopulatie door de selectiecriteria zijn gekomen. Het artikel van Apeldoorn²³ is afgevalen, omdat de kosteneffectiviteit is beschreven en dit niet past bij het doel van deze review. Tevens kan dit artikel niet meegenomen worden in de analyse. De leeftijden van de patiënten in de geïncludeerde studies vallen binnen de selectiecriteria. Echter zijn in Garcia¹¹ de patiënten onderverdeeld in twee subgroepen wat betreft de leeftijdscategorie. De subgroepen bestaan uit patiënten jonger of ouder dan 54 jaar. Het is niet bekend hoeveel patiënten er per subgroep aanwezig zijn. De aspirant-onderzoekers hebben ervoor gekozen om de subgroep te analyseren met patiënten onder de 54 jaar, deze komt beter overeen met de gemiddelde leeftijd van de andere studies.

Een belangrijk discussiepunt is dat de aspirant-onderzoekers ervoor hebben gekozen om de verschillende uitkomstmaten van zowel pijn als functioneren gelijk te stellen en deze met elkaar te vergelijken. Zo zijn de NPRS en VAS gelijk gesteld, zodat deze als één uitkomstmaat geanalyseerd kon worden. Dit zelfde geldt voor de ODI en RMDQ. Naast het gelijkstellen van de uitkomstmaten zijn ook de data van de follow-ups samengevoegd. De follow-ups van drie, vier, zeven en acht weken vallen binnen de korte termijn follow-up (0-3 maanden). Door het samenvoegen kan de data geanalyseerd worden. Door het generaliseren kunnen de resultaten negatief worden beïnvloedt. Daarom is gekozen voor een random effect model bij de meta-analyse. De follow-ups van de lange termijn worden gehouden op twaalf maanden. Hierdoor zijn deze goed met elkaar te vergelijken zonder dat dit de resultaten negatief beïnvloedt. Eén artikel⁸ heeft een follow-up die binnen de middellange termijn valt. Deze follow-up is niet meegenomen in de analyse, omdat er geen andere studie is waar de resultaten mee vergeleken kunnen worden.

Een volgend discussiepunt is het berekenen van de standaarddeviatie en betrouwbaarheidsinterval in Review Manager. Niet ieder artikel heeft een standaarddeviatie of betrouwbaarheidsinterval tussen de groepen beschreven. Met behulp van Review Manager zijn de ontbrekende betrouwbaarheidsintervallen en standaarddeviaties automatisch berekend. In GRADE is het betrouwbaarheidsinterval nodig voor het invullen van de tabel. Hiervoor zijn de gepoolde betrouwbaarheidsintervallen van de meta-analyse gebruikt om zo homogeniteit te creëren tussen de analyses. Niet ieder classificatiesysteem bevat meerdere artikelen, waardoor de gepoolde betrouwbaarheidsintervallen gelijk staan aan het betrouwbaarheidsinterval van dat artikel. Bij sommige artikelen komt het berekende betrouwbaarheidsinterval niet overeen met het betrouwbaarheidsinterval in het artikel. In GRADE heeft het invullen betrouwbaarheidsinterval geen invloed op de uiteindelijk toebedeelde kwaliteit van bewijs. De spreiding van het betrouwbaarheidsinterval wordt wel geanalyseerd bij de beoordeling van de factor "Onnauwkeurigheid". Tevens heeft het betrouwbaarheidsinterval geen invloed op de resultaten van de meta-analyse. Wanneer in Review Manager het betrouwbaarheidsinterval van het artikel wordt gebruikt in plaats van het berekende betrouwbaarheidsinterval, veranderen de resultaten niet.

Ten slotte hebben de aspirant-onderzoekers weinig ervaring met het beoordelen van artikelen die geschreven zijn in de Engelse taal, waardoor woorden anders geïnterpreteerd kunnen zijn.

Daarentegen wordt de kracht van bewijs van deze review vergroot, doordat de gevonden artikelen uit verschillende databases komen. Daarnaast komen de patiëntenpopulaties wat betreft het klachtenbeeld overeen. De meeste studies hebben minder dan 75 geïnccludeerde patiënten in de interventie- en controlegroep. De sterkte van bewijs verminderd met deze kleine patiëntenpopulatie. De sterkte van het bewijs neemt toe bij de statistische pooling, aangezien bij statistische pooling de patiëntenpopulatie boven de 300 patiënten is bij de meeste classificatiesystemen. Daarnaast is het belangrijk dat de artikelen een interventie- en controlegroep hebben zodat er een analyse kan worden uitgevoerd. Dit is bij de geïnccludeerde artikelen het geval.

De randomisatie voor de te beoordelen artikelen is geblindeerd uitgevoerd door een extern persoon. Tevens zijn de artikelen door de aspirant onderzoekers onafhankelijk van elkaar beoordeeld. Hierdoor is voorkeur voor artikelen voorkomen (selectiebias). De aspirant-onderzoekers hebben zich volledig aan de methode gehouden tijdens het uitvoeren van dit onderzoek.

Aspecifieke lage rugpijn is een aandoening waarbij geen specifieke pathologie gevonden kan worden¹. In de KNGF richtlijn³ worden patiënten ingedeeld in patiëntenprofielen. Per profiel wordt beschreven waar de behandeling op gericht moet worden. Echter is er geen

duidelijkheid welke behandeling moet worden gegeven. In dit onderzoek worden classificatiesystemen vergeleken met verschillende soorten behandeltrajecten, zoals oefentherapie en manuele verrichtingen. Hierdoor is er een duidelijkere richtlijn voor therapeuten welk behandeltraject een patiënt het beste kan volgen en bij welke therapeut. Arcus Fysiotherapie heeft diverse expertises, waardoor er niet volgens één duidelijke richtlijn gehandeld wordt. Deze review heeft onderzocht welk classificatiesysteem het meest effectief is volgens de literatuur en welk classificatiesysteem het beste toepasbaar is in de praktijk van Arcus Fysiotherapie. Voor Arcus Fysiotherapie is gekozen, op basis van bovenstaande gegevens, voor de Treatment-Based Classification in combinatie met het psychosociale aspect van O'Sullivan Classification. De classificatiesystemen zijn in de literatuur vergeleken met reguliere fysiotherapie, maar onderling zijn ze niet vergeleken. Om deze reden moet nog verder onderzoek worden gedaan naar welk classificatiesysteem onderling het beste effect geeft. Er is geen classificatiesysteem die een combinatie heeft van de Treatment-Based Classification met het psychosociale aspect van de O'Sullivan Classification. Er moet verder onderzoek worden gedaan naar het effect van dit samengestelde classificatiesysteem vergeleken met reguliere fysiotherapie (*figuur 3*).

CONCLUSIE

De classificatiesystemen die in dit onderzoek zijn meegenomen en vergeleken zijn met reguliere fysiotherapie betreffen de Treatment-Based Classification, Mechanical Diagnosis and Treatment, Movement System Impairment, Start Back Screening Tool en O'Sullivan Classification. De Treatment-Based Classification bereikt op een na hoogste resultaten op de uitkomstmaten pijn en functioneren en het hoogste op kwaliteit van bewijs. Arcus Fysiotherapie wenst zich een classificatiesysteem die hoog scoort in de analyse en gemakkelijk toe te passen is in de praktijk, waarbij psychosociale aspecten worden meegenomen.

Het advies is om de Treatment-Based Classification te combineren met de psychosociale aspecten van de O'Sullivan Classification (*figuur 3*).

AANBEVELINGEN

Er is verder onderzoek nodig naar het onderlinge effect van de classificatiesystemen. Tevens moet verder onderzoek worden gedaan het effect van het samengestelde classificatiesysteem vergeleken met reguliere fysiotherapie (*figuur 3*).

DANKWOORD

Hierbij dank aan Bart Staal, docentbegeleider, lector aan de Hogeschool van Arnhem en Nijmegen en onderzoeker aan het Radboud UMC, voor zijn enthousiaste en inhoudelijke bijdrage gedurende het onderzoek. Daarnaast bedanken we Martin Oortwijn en Laurens Brummelhuis, opdrachtgevers namens Arcus Fysiotherapie te Zutphen, voor de ondersteuning en mogelijkheid tot onderzoeken. Ten slotte bedanken wij onze familie, vrienden en studiegenoten voor hun ondersteuning tijdens het onderzoeksproces.

BRONNENLIJST

1. Chavannes, A., Mens, J., Koes, B., Lubbers, W., Ostelo, R. & Spinnewijn, W. (2005). NHG-standaard aspecifieke lagerugpijn. Gedownload op 23 februari 2016, van <https://www.nhg.org/standaarden/volledig/nhg-standaard-aspecifieke-lagerugpijn>
2. Crul, B. J. P., Houdenhove, B., van, P., R.S.G.M., Vissers, K., & Wit, R. d. (2005). Pijn info: Aspecifieke lage rugklachten. Houten: Bohn Stafleu van Loghum.
3. Staal, J. B., Hendriks, E.J.M., Heijnmans, M., Kiers, H., Lutgers-Boomsma, A.M., Rutten, G. et al. (2013). KNGF richtlijn: Lage rugpijn verantwoording & toelichting. Gedownload op 23 februari 2016, van https://www.fysionet-evidencebased.nl/images/pdfs/richtlijnen/lage_rugpijn_2013/lage_rugpijn_verantwoording_en_toelichting.pdf
4. Cochrane Netherlands (2014). RCT Checklist. Gedownload op 15 maart 2016, van <http://netherlands.cochrane.org/sites/netherlands.cochrane.org/files/uploads/4.4%20Checklist%20Beoordeling%20randomsed%20controlled%20trial%20%28RCT%29.pdf>
5. Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., ... & Lamb, S. E. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj*, 348, g1687. Gedownload op 14 maart 2016, van <http://www.equator-network.org/wp-content/uploads/2014/03/TIDieR-Checklist-PDF.pdf>
6. RevMan 5.3 [software voor het maken van Cochrane Reviews] (2014, 13 juni). Nederland: Cochrane. Geraadpleegd van 25 april 2016 tot en met 3 mei 2016, van <http://tech.cochrane.org/revman>
7. GRADE Working Group, the Cochrane Collaboration and Members of McMaster University. *GRADEpro*. Geraadpleegd van 2 mei 2016 tot 25 mei 2016, van <http://gradepro.org/>
8. Apeldoorn, A. T., Ostelo, R. W., van Helvoirt, H., Fritz, J. M., Knol, D. L., van Tulder, M. W., et al. (2012). A randomized controlled trial on the effectiveness of a classification-based system for subacute and chronic low back pain. *Spine* (03622436), 37(16), 1347-1356 10p.
9. Brennan, G. P., Fritz, J. M., Hunter, S. J., Thackeray, A., Delitto, A., & Erhard, R. E. (2006). Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: Results of a randomized clinical trial. *Spine*, 31(6), 623-631.
10. Fritz, J. M., Delitto, A., & Erhard, R. E. (2003). Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: A randomized clinical trial. *Spine* (03622436), 28(13), 1363-1372 10p.
11. Garcia, A. N., Costa, L. d. C. M., Hancock, M., & Costa, L. O. P. (2015). Identifying patients with chronic low back pain who respond best to mechanical diagnosis and therapy: Secondary analysis of a randomized controlled trial. *Physical Therapy*,
12. Machado, L. A. C., Maher, C. G., Herbert, R. D., Clare, H., & McAuley, J. H. (2010). The effectiveness of the McKenzie method in addition to first-line care for acute low back pain: A randomized controlled trial. *BMC Medicine*, 8, 10-19.

13. Petersen, T., Larsen, K., Nordsteen, J., Olsen, S., Fournier, G., & Jacobsen, S. (2011). The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: A randomized controlled trial. *Spine*, 36(24), 1999-2010.
14. Henry, S. M., Van Dillen, L. R., Ouellette-Morton, R. H., Hitt, J. R., Lomond, K. v., DeSarno, M. J., et al. (2014). Outcomes are not different for patient-matched vs. non-matched treatment in subjects with chronic, recurrent low back pain: A randomized clinical trial. *Spine J*, 14(12), 2799-810.
15. Jacobs, J. V., Lomond, K. V., Hitt, J. R., DeSarno, M. J., Bunn, J. Y., & Henry, S. M. (2016). Original article: Effects of low back pain and of stabilization or movement-system-impairment treatments on induced postural responses: A planned secondary analysis of a randomised controlled trial. *Manual Therapy*, 21, 210-219.
16. Lomond, K. V., Jacobs, J. V., Hitt, J. R., DeSarno, M. J., Bunn, J. Y., & Henry, S. M. (2015). Effects of low back pain stabilization or movement system impairment treatments on voluntary postural adjustments: A randomized controlled trial. *The Spine Journal*, 15(4), 596-606.
17. Murphy, S. E., Blake, C., Power, C. K., & Fullen, B. M. (2015). Comparison of a stratified group intervention (STarT back) with usual group care in patients with low back pain: A non-randomised controlled trial. *Spine*
18. Sheeran, L., van Deursen, R., Caterson, B., & Sparkes, V. (2013). Classification-guided versus generalized postural intervention in subgroups of nonspecific chronic low back pain: A pragmatic randomized controlled study. *Spine* (03622436), 38(19), 1613-1625 13p.
19. Schoemaker, C. (2011). Grade zet bewijs om in concreet advies - Vaak is van een richtlijnadvies onduidelijk hoe bindend dat is. Een internationale groep richtlijnmakers formuleert adviezen in gewonemensentaal, direct te gebruiken in de spreekkamer. *Medisch Contact*, 66(45), 2756.
20. Kwaliteitsinstituut voor de gezondheidszorg CBO (2004). *Handleiding focusgroepen*. Gedownload op 15 maart 2016, van <http://www.communicerenmetarmen.be/sites/default/files/HandleidingFocusgroepenCBO200409.pdf>
21. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & The PRISMA Group. (2009). Prisma flow-diagram. Gedownload op 22 februari 2016, van <http://www.prisma-statement.org/documents/PRISMA%202009%20flow%20diagram.pdf>
22. Nederlands Huisartsen Genootschap. (2015). Ontwikkelen van NHG-standaarden. Gedownload op 26 april 2016, van https://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwiS5lTlmtTMAhWrKsAKHSr6DHMQFggjMAE&url=https%3A%2F%2Fwww.nhg.org%2Fsites%2Fdefault%2Ffiles%2Fcontent%2Fnhg_org%2Fuploads%2Fhandleiding_standarden_2015.pdf&usg=AFQjCNHerTn2z444kRsZA1hpgGrssjsfMw&sig2=HT2ph0-3EW0c4EF6_8BDNw
23. Apeldoorn, A. T., Bosmans, J. E., Ostelo, R. W., de Vet, H. C., & van Tulder, M. W. (2012). Cost-effectiveness of a classification-based system for sub-acute and chronic low back pain. *European Spine Journal*, 21(7), 1290-1300.

24. Onderbouwingsverslag

25. Clare, H. A., Adams, R., & Maher, C. G. (2005). Reliability of McKenzie classification of patients with cervical or lumbar pain. *Journal of Manipulative & Physiological Therapeutics*, 28(2), 122-127 6p.
26. Hefford, C. (2008). McKenzie classification of mechanical spinal pain: Profile of syndromes and directions of preference
27. Kent, P., & Keating, J. L. (2005). Classification in nonspecific low back pain: What methods do primary care clinicians currently use? *Spine* (03622436), 30(12), 1433-1440 8p.
28. O'Sullivan, P. (2005). Diagnosis and classification of chronic low back pain disorders: Maladaptive movement and motor control impairments as underlying mechanism. *Manual Therapy*, 10(4), 242-255 14p.
29. Petersen, T., Olsen, S., Laslett, M., Thorsen, H., Manniche, C., Ekdahl, C., et al. (2004). Inter-tester reliability of a new diagnostic classification system for patients with non-specific low back pain. *Australian Journal of Physiotherapy*, 50(2), 85-94.
30. Razmjou, H., Kramer, J. F., & Yamada, R. (2000). Intertester reliability of the McKenzie evaluation in assessing patients with mechanical low-back pain. *Journal of Orthopaedic & Sports Physical Therapy*, 30(7), 368-389 22p.
31. Stanton, T. R., Fritz, J. M., Hancock, M. J., Latimer, J., Maher, C. G., Wand, B. M., et al. (2011). Evaluation of a treatment-based classification algorithm for low back pain: A cross-sectional study. *Physical Therapy*, 91(4), 496.
32. Widerström, B., Olofsson, N., Arvidsson, I., Harms-Ringdahl, K., & Larsson, U. E. (2012). Inter-examiner reliability of a proposed decision-making treatment based classification system for low back pain patients. *Manual Therapy*, 17, 164-171.
33. Widerström, B., Olofsson, N., Boström, C., & Rasmussen-Barr, E. (2016). Feasibility of the subgroup criteria included in the treatment-strategy-based (TREST) classification system (CS) for patients with non-specific low back pain (NSLBP). *Manual Therapy*
34. Wilson, L., Hall, H., McIntosh, G., & Milles, T. (1999). Intertester reliability of a low back pain classification system. *Spine*, 24(3), 248-254.
35. Delitto, A., Erhard, R. E., & Bowling, R. W. (1995). A treatment-based classification approach to low back syndrome: Identifying and staging patients for conservative treatment. *Physical Therapy*, 75(6), 470-485.
36. George, S. Z., Zeppieri, G., Cere, A. L., Cere, M. R., Borut, M. S., Hodges, M. J., et al. (2008). A randomized trial of behavioral physical therapy interventions for acute and sub-acute low back pain (NCT00373867). *Pain*, 140(1), 145-157.
37. Hall, H., McIntosh, G., & Boyle, C. (2009). Effectiveness of a low back pain classification system. *The Spine Journal*, 9, 648-657.
38. Harris-Hayes, M., & Van Dillen, L. (2009). The inter-tester reliability of physical therapists classifying low back pain problems based on the movement system impairment classification system
39. Hill, J. C., Dunn, K. M., Main, C. J., & Hay, E. M. (2010). Subgrouping low back pain: A comparison of the STarT back tool with the örebro musculoskeletal pain screening

questionnaire. *European Journal of Pain*, 14(1), 83-89.

40. Hill, J. C., Whitehurst, D. G., Lewis, M., Bryan, S., Dunn, K. M., Foster, N. E., et al. (2011). Comparison of stratified primary care management for low back pain with current best practice (STarT back): A randomised controlled trial. *The Lancet*, 378(9802), 1560-1571.
41. Hill, J. C., Vohora, K., Dunn, K. M., Main, C. J., & Hay, E. M. (2010). Comparing the STarT back screening tool's subgroup allocation of individual patients with that of independent clinical experts. *The Clinical Journal of Pain*, 26, 783-787.
42. Kilby, J., Stigant, M., & Roberts, A. (1990). The reliability of back pain assessment by physiotherapists, using a 'McKenzie algorithm'. *Physiotherapy*, 76(9), 579-583 5p.
43. Lamb, S. E., Lall, R. S., Hansen, Z., Castelnovo, E., Withers, E. J., Nichols, V., et al. (2010). A multicentred randomised controlled trial of a primary care-based cognitive behavioural programme for low back pain: The back skills training (BeST) trial. *Health Technology Assessment*, 14(41), 1-281.
44. Murphy, E. S., Blake, C., Power, K. C., & Fullen, M. B. (2013). The effectiveness of a stratified group intervention using the STarTBack screening tool in patients with LBP - a non randomised controlled trial. *BMC Musculoskeletal Disorders*, 14
45. Petersen, T., Laslett, M., Thorsen, H., Manniche, C., Ekdahl, C., & Jacobsen, S. (2003). Diagnostic classification of non-specific low back pain. A new system integrating patho-anatomic and clinical categories. *Physiotherapy Theory & Practice*, 19(4), 213-237 25p.
46. Petersen, T., Thorsen, H., Manniche, C., & Ekdahl, C. (1999). Classification of non-specific low back pain: A review of the literature on classifications systems relevant to physiotherapy. *Physical Therapy Reviews*, 4(4), 265-281 17p.
47. Riddle, D. L., & Rothstein, J. M. (1993). Intertester reliability of McKenzie's classifications of the syndrome types present in patients with low back pain. *Spine (03622436)*, 18(10), 1333-1344 12p.
48. Trudelle-Jackson, E., Sarvaiya-Shah, S. A., & Wang, S. S. (2008). Interrater reliability of a movement impairment-based classification system for lumbar spine syndromes in patients with chronic low back pain. *Journal of Orthopaedic & Sports Physical Therapy*, 38(6), 371-376 6p.
49. Van Dillen, L. R., Sahrmann, S. A., Norton, B. J., Caldwell, C. A., Fleming, D. A., McDonnell, M. K., et al. (1998). Reliability of physical examination items used for classification of patients with low back pain. *Physical Therapy*, 78(9), 979-988 10p.
50. Fritz, J. M., & George, S. (2000). The use of a classification approach to identify subgroups of patients with acute low back pain: Interrater reliability and short-term treatment outcomes. *Spine (03622436)*, 25(1), 106-114 9p.
51. George, S. Z., & Delitto, A. (2005). Clinical examination variables discriminate among treatment-based classification groups: A study of construct validity in patients with acute low back pain. *Physical Therapy*, 85(4), 306-314 9p.
52. Henry, S. M., Fritz, J. M., Trombley, A. R., & Bunn, J. Y. (2012). Reliability of a treatment-based classification system for subgrouping people with low back pain. *Journal of Orthopaedic & Sports Physical Therapy*, 42(9), 797-805 9p.

53. Mbada, C. E., Ayanniyi, O., Ogunlade, S. O., Orimolade, E. A., Oladiran, A. B., & Ogundele, A. O. (2014). Influence of mckenzie protocol and two modes of endurance exercises on health-related quality of life of patients with long-term mechanical low-back pain. *The Pan African Medical Journal*, 17 Suppl 1, 5-5.
54. Moncelon, S., & Otero, J. (2015). Savoirs: Méthode McKenzie et lombalgiques chroniques avec préférence directionnelle. *Kinésithérapie, La Revue*, 15, 31-37.
55. Petersen, T., Larsen, K., Nordsteen, J., Olsen, S., Fournier, G., & Jacobsen, S. (2011). The McKenzie method compared with manipulation when used adjunctive to information and advice in low back pain patients presenting with centralization or peripheralization: A randomized controlled trial. *Spine*, 36(24), 1999-2010.
56. Apeldoorn, A. T., Bosmans, J. E., Ostelo, R. W., de Vet, H. C., & van Tulder, M. W. (2012). Cost-effectiveness of a classification-based system for sub-acute and chronic low back pain. *European Spine Journal*, 21(7), 1290-1300.
57. Camara Azevedo, D., Van Dillen, L. R., de, O. S., Ribeiro Oliveira, D., Ferreira, P. H., & Oliveira, P. C. (2015). Movement system impairment-based classification versus general exercise for chronic low back pain: Protocol of a randomized controlled trial. *Physical Therapy*, 95(9), 1287-1294.
58. Chou, R. (2012). Risk-stratified primary care management of low back pain reduced disability. *ACP Journal Club*, 156(2), 6-7.
59. Foster, N. E., Mullis, R., Hill, J. C., Lewis, M., Whitehurst, D. G. T., Doyle, C., et al. (2014). Effect of stratified care for low back pain in family practice (IMPACT back): A prospective population-based sequential comparison. *The Annals of Family Medicine*, 12(2), 102-11.
60. Fritz, J. M., Brennan, G. P., Clifford, S. N., Hunter, S. J., & Thackeray, A. (2006). An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. *Spine* (03622436), 31(1), 77-82 6p.
61. Vibe Fersum, K., O'Sullivan, P. B., Kvåle, A., & Skouen, J. S. (2009). Inter-examiner reliability of a classification system for patients with non-specific low back pain. *Manual Therapy*, 14, 555-561.

BIJLAGEN

Tabel 6: Zoekschema oriënterend literatuuronderzoek

Term (ongeacht taal)	Classificatiesysteem	Lage rugpijn	Pijn	Patiënteigenschappen	Behandeling	Praktijk
Synoniem	- Beslisboom - Flowchart - Decision tree	- Lumbale klachten	- Pain	- Man - Vrouw	- Interventie - Therapy - Intervention	-Eerstelijnspraktijk
Verwante term	- Flow diagram - Flow sheet	- Lower back pain - Lumbar spine pain - Back pain - Lumbago - Lower back ache - Low back pain	- Ache	- Man - Male - Woman - Female	- Therapeutics management - Treatment - Treatment strategies - Manuele therapie - Manual therapy - Dry needling - Exercise therapy - Massage - Psychosomatisch/psychosomatic	- Algemene praktijk
Bredere term	- Stratified pathway(s) - Step care strategies - Stepped care - Treatment-based classification - Classification systems - Classification algorithm	- Aspecifieke lage rugklachten - Non-specific back-pain - Lower back pain - LBP - Back ache	- Sensation - Discomfort - Gevoel	- Middle-aged - Subgroups - Chronic - Chronical - Acute	- Cure - Healing - Physical therapy	- Algemene praktijk

Nauwere term	- Stratified care - Step care	- Acute lage rugklachten - Chronische lage rugklachten - CLRP - Lower back pain, mechanical - Lower back pain, postural - Lower back pain, posterior compartment - Lower back pain, recurrent	- Injury	- 16-65 jaar	- Remedy	- General practice
---------------------	--	---	----------	--------------	----------	--------------------

Tabel 7: Zoekschema literatuuronderzoek

Classificatiesystemen gevonden in oriënterend literatuuronderzoek:

1. Mechanical Diagnosis and Therapy (MDT): McKenzie R, May S.
2. Movement System Impairment (MSI): Sahrman SA
3. O'Sullivan Classification (OSC): O'Sullivan P
4. Pathoanatomic Based Classification (PBC): Petersen T, Laslett M, Thorsen H, Manniche C, Ekdahl C, Jacobsen S
5. Treatment Based Classification (TBC): Delitto A, Erhard RE, Bowling RW, DeRosa CP, Greathouse DG / Fritz JM, Cleland JA, Childs J / Alrwaily M, Timko M, Schneider M, Stevans J, Bise C, Hariharan K, Delitto A
6. *Versie uit 1995, 2006 en 2015 worden meegenomen en samengevoegd.*
7. Start Back Screening Tool: Hill
8. Impact Back
9. Treatment-Strategy-Based Classification System: Widerström

PICO-vraag

Welk classificatiesysteem is het meest effectief ten aanzien van pijn en functioneren bij een man of vrouw van 16 t/m 65 jaar met specifieke lage rugpijn ten opzichte van een alternatief classificatiesysteem, reguliere fysiotherapie of een placebogroep?

	Patient	Intervention	Comparison	Outcome
Term	Man Vrouw	1. Mechanical Diagnosis and Therapy 2. Movement System Impairment 3. O'Sullivan Classification 4. Pathoanatomic Based Classification 5. Treatment Based Classification 6. Start Back Screening Tool 7. Impact Back 8. Treatment-Strategy-Based Classification System	Alternatief classificatiesysteem (= een van de classificatiesystemen uit de Intervention) Reguliere fysiotherapie Placebo groep	Effectief Pijn Functioneren
Synoniem	Man Male Woman Female	1. MDT, McKenzie method 2. MSI 3. OSC 4. PBC 5. TBC 6. SBST	Synoniem van een ander classificatie uit de Intervention Regular physiotherapy Regular physical therapy Normal physiotherapy	Effective Useful Efficient Pain Ache

		7. IB 8. TREST, decision-making algorithm	Normal physical therapy Traditional physiotherapy Traditional physical therapy Standard physiotherapy Standard physical therapy Placebo group Placebo	Physical suffering
Verwante term		1. McKenzie R, May S 2. Sahrmann SA 3. O'Sullivan P 4. Petersen T, Laslett M, Thorsen H, Manniche C, Ekdahl C, Jacobsen S 5. Delitto A, Erhard RE, Bowling RW, DeRosa CP, Greathouse DG / Fritz JM, Cleland JA, Childs JD / Alrwaily M, Timko M, Schneider M, Stevans J, Bise C, Hariharan K, Delitto A 6. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM 7. Auteur niet bekend 8. Widerström B, Olofsson N, Boström C, Rasmussen-Barr E	Auteurs van een ander classificatie uit de Intervention	
Bredere term	Middle-aged		Other therapy No physiotherapy No physical therapy	
Nauwere term	16-65 jaar 16-65 years old			Practical Valid Range of capabilities

Tabel 8: Studiekarakteristieken

Auteur (jaar)	Patiëntenpopulatie	Interventiegroep	Controlegroep	Follow-up	Uitkomst (meetinstrument)	Gemiddelde resultaten	Conclusie auteur
Apeldoorn et al (2012)	Aspecifieke lage rugklachten (n=156) Gemiddelde leeftijd: I: 43.2 (11.7) C: 42.0 (10.9)	Treatment-Based Classification. Oefentherapie specifieke bewegingen, manipulaties en stabilisatietraining (n=74)	Reguliere fysiotherapie volgens de KNGF richtlijn aspecifieke lage rugklachten (n=82)	8 weken 26 weken 52 weken	Pijn (NPRS) Functioneren (ODI)	NPRS / ODI <u>Interventiegroep:</u> 4.04 / 13.66 3.14 / 12.09 3.24 / 10.25 <u>Controlegroep:</u> 3.61 / 16.67 3.20 / 13.45 3.41 / 14.59 <u>P (CI 95%) NRS</u> -0.49 (-1.34-0.37) 0.01 (-0.95 – 0.98) 0.13 (-0.83-1.10) <u>P (CI 95%) ODI</u> -0.48 (-4.59-3.63) -2.13 (-6.92-2.66) 0.86 (-3.61-5.33)	Er is meer verbetering waarneembaar wanneer de Classification-Based Treatment wordt gebruikt. Verder onderzoek is nodig.
Brennan et al (2006)	Aspecifieke lage rugklachten (n=123) Gemiddelde leeftijd: I: 38.4 (10.0) C: 37.2 (11.1)	Treatment-Based Classification. Bewust geclassificeerd in oefentherapie specifieke bewegingen, manipulaties of stabilisatietraining (n=50)	Treatment-Based Classification Random geclassificeerd in oefentherapie specifiek bewegen, manipulaties of stabilisatietraining (n=73)	4 weken 12 weken	Functioneren (ODI)	ODI Interventiegroep: 29.9 (15.6*) 27.9 (16.0*) Controlegroep: 23.3 (16.6*) 19.6 (16.0*) P=0.03 (4 weken) P=0.006 (1 jaar)	Betere resultaten kunnen worden behaald wanneer de patiënten volgens de classificatie systeem worden behandeld.
Fritz et al	Aspecifieke lage	Treatment-Based	Reguliere	4 weken	Functioneren (ODI)	ODI	De geclassificeerde

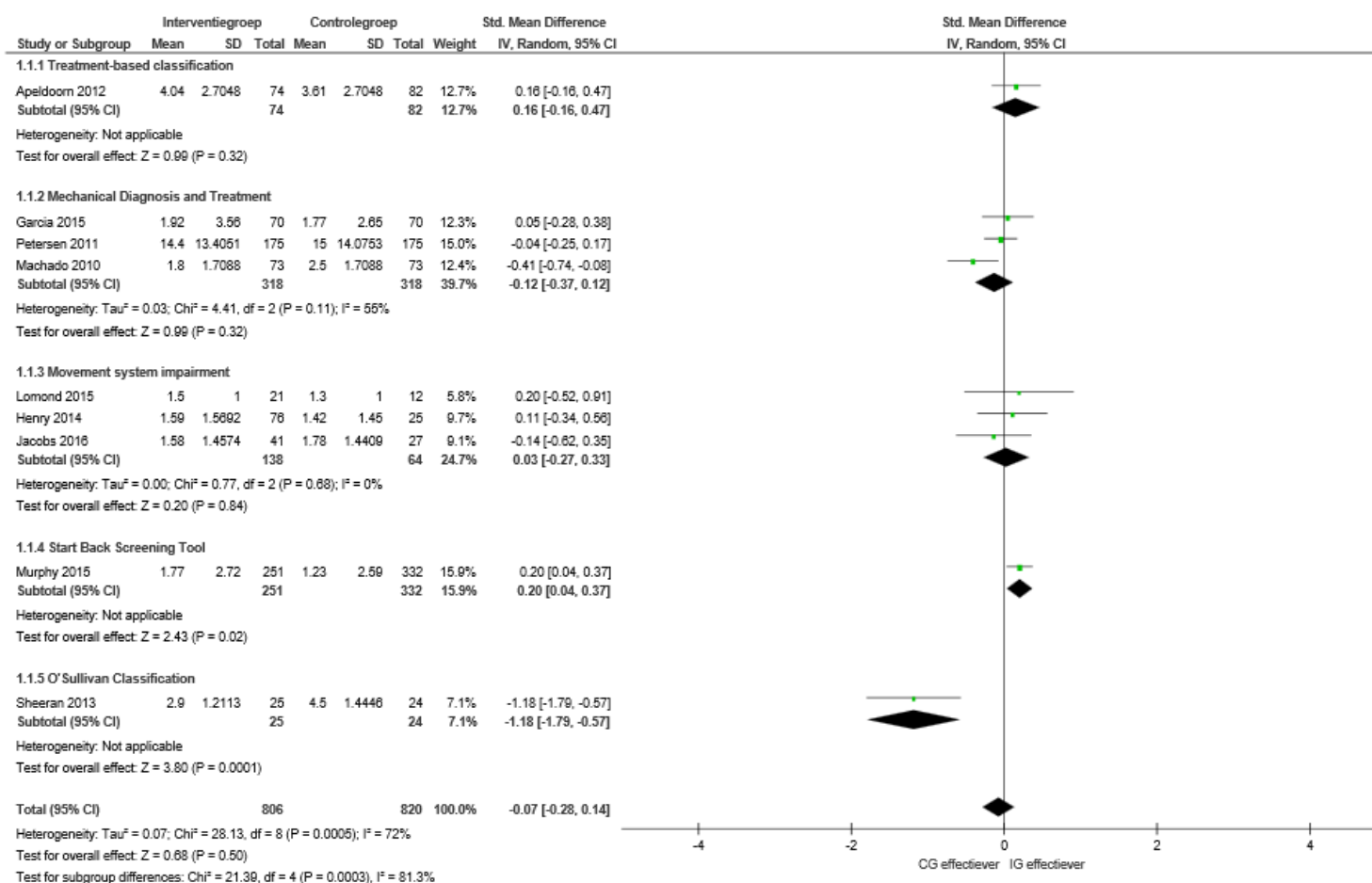
(2003)	rugklachten (n=78) Gemiddelde leeftijd: I: 35.9 (10.3) C: 39.1 (10.3)	Classification. Oefentherapie specifieke bewegingen, manipulaties en stabilisatietraining (n=41)	fysiotherapie volgens de Health Care Policy and Research richtlijnen (n=37)	52 weken		<u>Interventiegroep:</u> 22.5 (19.3*) 26.4 (16.9*) <u>Controlegroep:</u> 11.6 (18.1*) 17.4 (20.6*) 10.9 (1.9 – 19.9) P=0.018 9.0 (0.30 – 17.7) P=0.044	behandelingen geven een beter resultaat, maar verder onderzoek is nodig.
Petersen et al (2011)	Aspecifieke lage rugklachten (n=350) Gemiddelde leeftijd: I: 38 (10.4) C: 37 (9.4)	Mechanical Diagnosis and Treatment. McKenzie methode en advies over een individueel oefenprogramma. Zelfmanagement, mobilisatie, stabilisatie, rekken en versterkende oefeningen (n=175)	Manipulatie en advies over een individueel oefenprogramma. Zelfmanagement, mobilisatie, stabilisatie, rekken en versterkende oefeningen (n=175)	8 weken 52 weken	Pain reduction Functioneren (RMDQ)	RMDQ <u>Interventiegroep:</u> 6.7 (5.8-7.6) 7.1 (6.1-8.0) <u>Controlegroep:</u> 5.2 (4.3-6.1) 5.6 (4.6-6.6) P=0.022 95%: 1.5 (0.2-2.8) P=0.030: 95% 1.5 (0.2-2.9) Pain reduction <u>Interventiegroep:</u> 14.4 (12.4-16.4) 15.0 (12.9-17.1) <u>Controlegroep:</u> 13.0 (11.1-14.9) 12.2 (10.1-14.3) P=0.309 95% 1.4 (-1.3-4.1) P=0.063 95% 2.8 (-0.2-5.8)	McKenzie methode is effectiever dan spinale manipulatie met voorlichting en advies, maar het is nog de vraag of het klinisch relevant is. Reguliere fysiotherapie kan beter worden toegepast.
Machado et al (2010)	Aspecifieke lage rugklachten (n=146)	Mechanical Diagnosis and Treatment. Bewust gegroepeerd	Reguliere fysiotherapie Advies activiteiten en	1 week 3 weken	Pijn (NPRS) Functioneren (RMDQ)	NPRS <u>Interventiegroep:</u> 3.6 (0.2**)	Een behandeling volgens de McKenzie methode

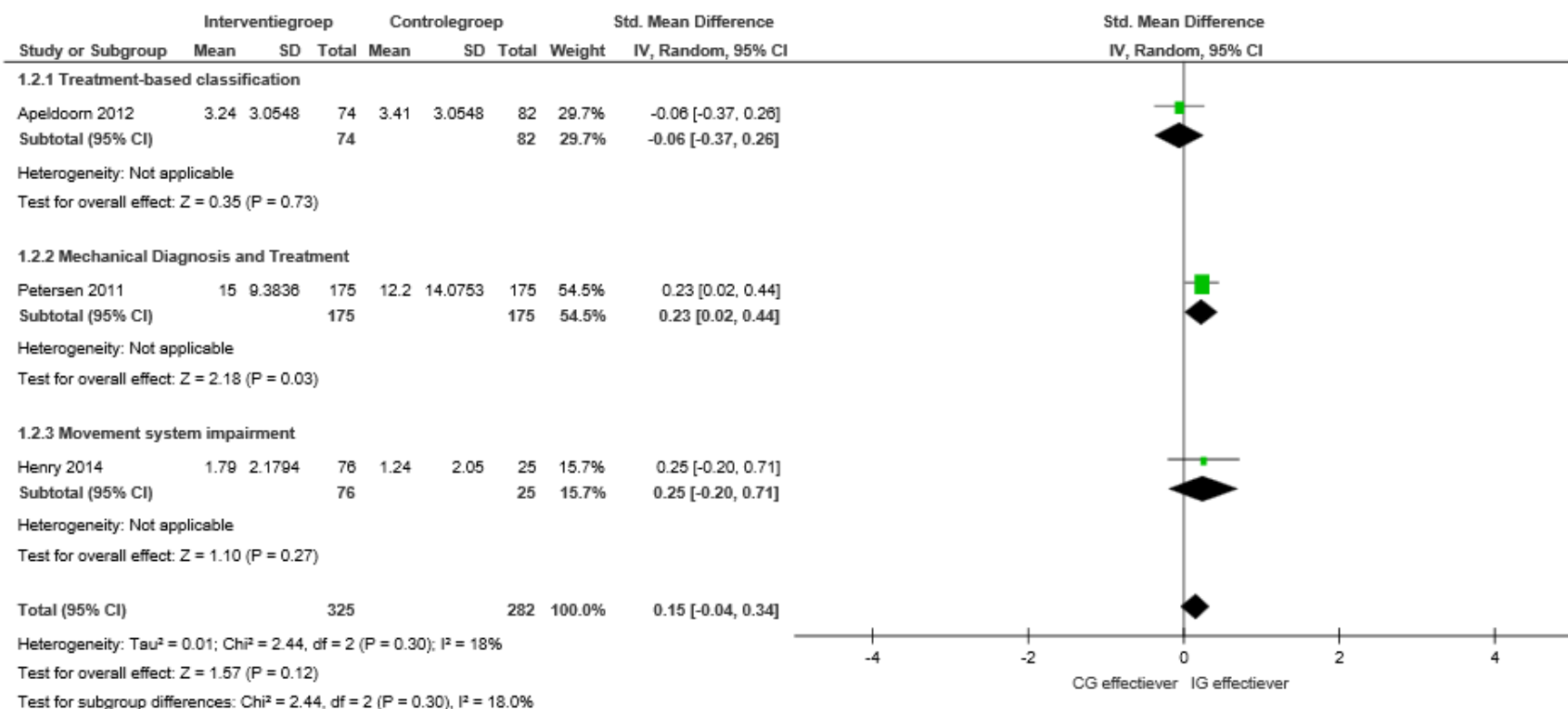
	Gemiddelde leeftijd: I: 47.5 (14.4) C: 45.9 (14.9)	in de McKenzie methode (n=73)	vermijden van bedrust (n=73)			1.8 (0.2**) <u>Controlegroep:</u> 4.0 (0.2**) 2.5 (0.2**) P=0.02 RMDQ <u>Interventiegroep:</u> 8.0 (0.5**) 4.8 (0.7**) <u>Controlegroep:</u> 8.2 (0.5**) 5.1 (0.7**) P=0.74	is niet effectiever dan een behandeling volgens de reguliere fysiotherapie in de eerstelij. Reguliere fysiotherapie kan het beste worden gebruikt.
Garcia et al (2015)	Aspecifieke lage rugklachten (n=140) Gemiddelde leeftijd: I: 53.5 (13.3) C: 54.8 (13.5)	Mechanical Diagnosis and Treatment en volgen van het boek 'Treat your own Back' (n=70)	Back school methode (n=70)	4 weken	Pijn (NPRS) Functioneren (RMDQ)	NPRS > 54 jaar <u>Interventiegroep:</u> 3.77 (3.06*) <u>Controlegroep:</u> 2.35 (2.75*) 1.42 (0.02;2.83) NPRS < 54 jaar <u>Interventiegroep:</u> 1.92 (3.56*) <u>Controlegroep:</u> 1.77 (2.65*) 0.15 (-1.38; 1.68) RMDQ > 54 jaar <u>Interventiegroep:</u> 4.97 (4.78*) <u>Controlegroep:</u> 1.46 (3.85*) 3.51 (1.42; 5.60) RMDQ < 54 jaar	Ouder dan 54 jaar reageert beter op MDT dan op Back school.

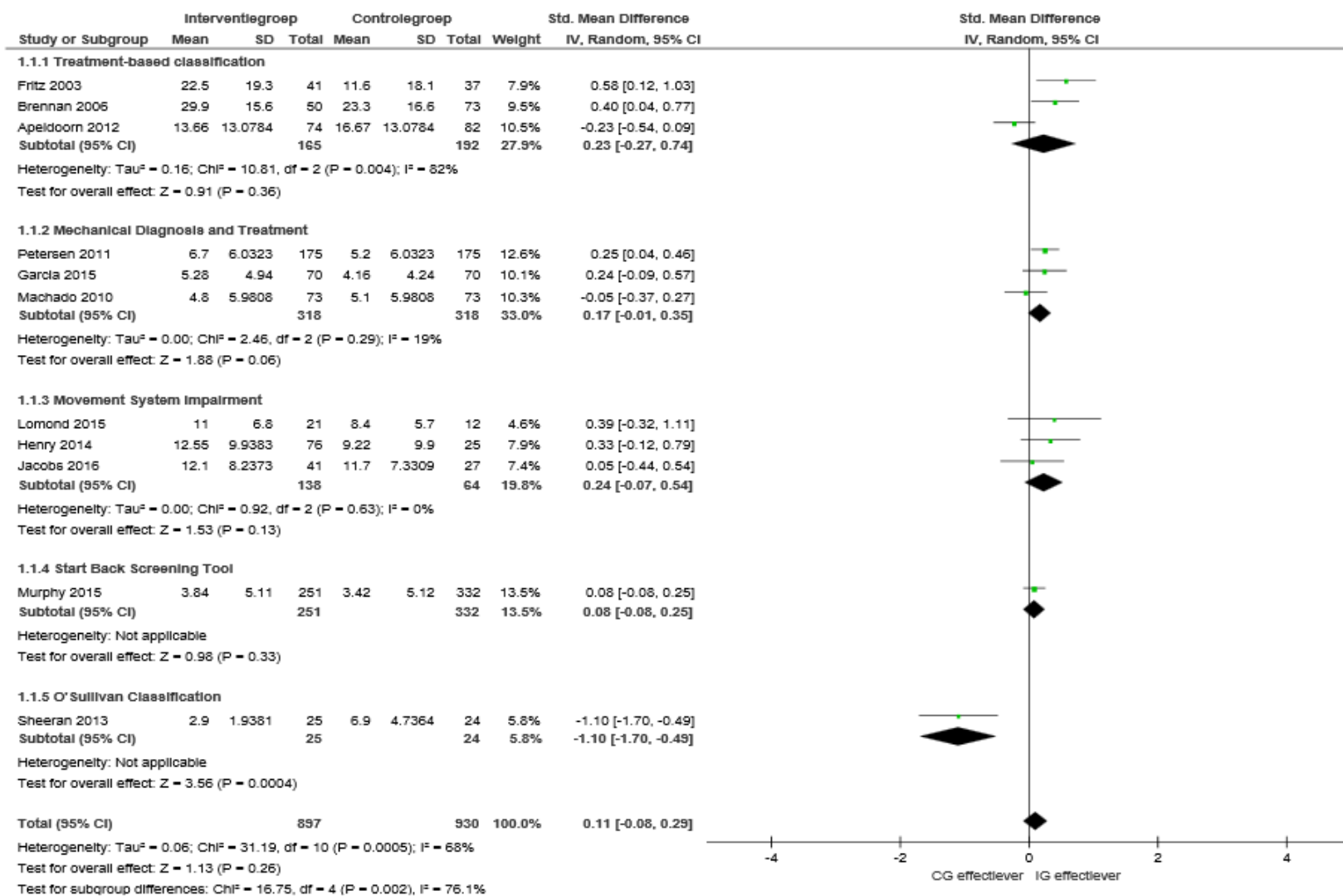
						<u>Interventiegroep:</u> 5.28 (4.94*) <u>Controlegroep:</u> 4.16 (4.24*) 1.12 (-1.11; 3.35)	
Jacobs et al (2016)	Aspecifieke lage rugklachten (n=68) Gemiddelde leeftijd: I: 42.6 (3.2) C: 39.6 (4.3)	Movement System Impairment behandelprotocol met huiswerkoefeningen (n=41)	Protocol voor stabilisatie oefeningen (n=27)	7 weken	Pijn (NPRS) Functioneren (ODI)	ODI <u>Interventiegroep:</u> 19.4% (16.7-22.0%) = pre 12.1% (9.5-14.7%) = post <u>Controlegroep:</u> 18.6% (14.9-22.3%) 11.7% (8.7-14.6%) P=<0.0001 NPRS <u>Interventiegroep:</u> 2.68 (2.10-3.26) 1.58 (1.11-2.04) <u>Controlegroep:</u> 2.37 (1.63-3.11) 1.78 (1.20-2.35) P=0.00015	De Oswestry en NPRS zijn in dezelfde maten verbeterd, zowel bij de interventiegroep als de controlegroep.
Lomond et al (2015)	Aspecifieke lage rugklachten (n=33) Gemiddelde leeftijd: I: 41.6 (10.9) C: 43.1 (11.9)	Movement System Impairment behandelprotocol met huiswerkoefeningen (n=21)	Protocol voor stabilisatie oefeningen (n=12)	7 weken	Pijn (NPRS) Functioneren (ODI)	ODI <u>Interventiegroep:</u> 11 (6.8*) <u>Controlegroep:</u> 8.4 (5.7*) P=0.229 NPRS <u>Interventiegroep:</u>	Er is geen verschil te zien tussen de interventiegroep of de controlegroep, beide groepen zijn in dezelfde maat verbeterd.

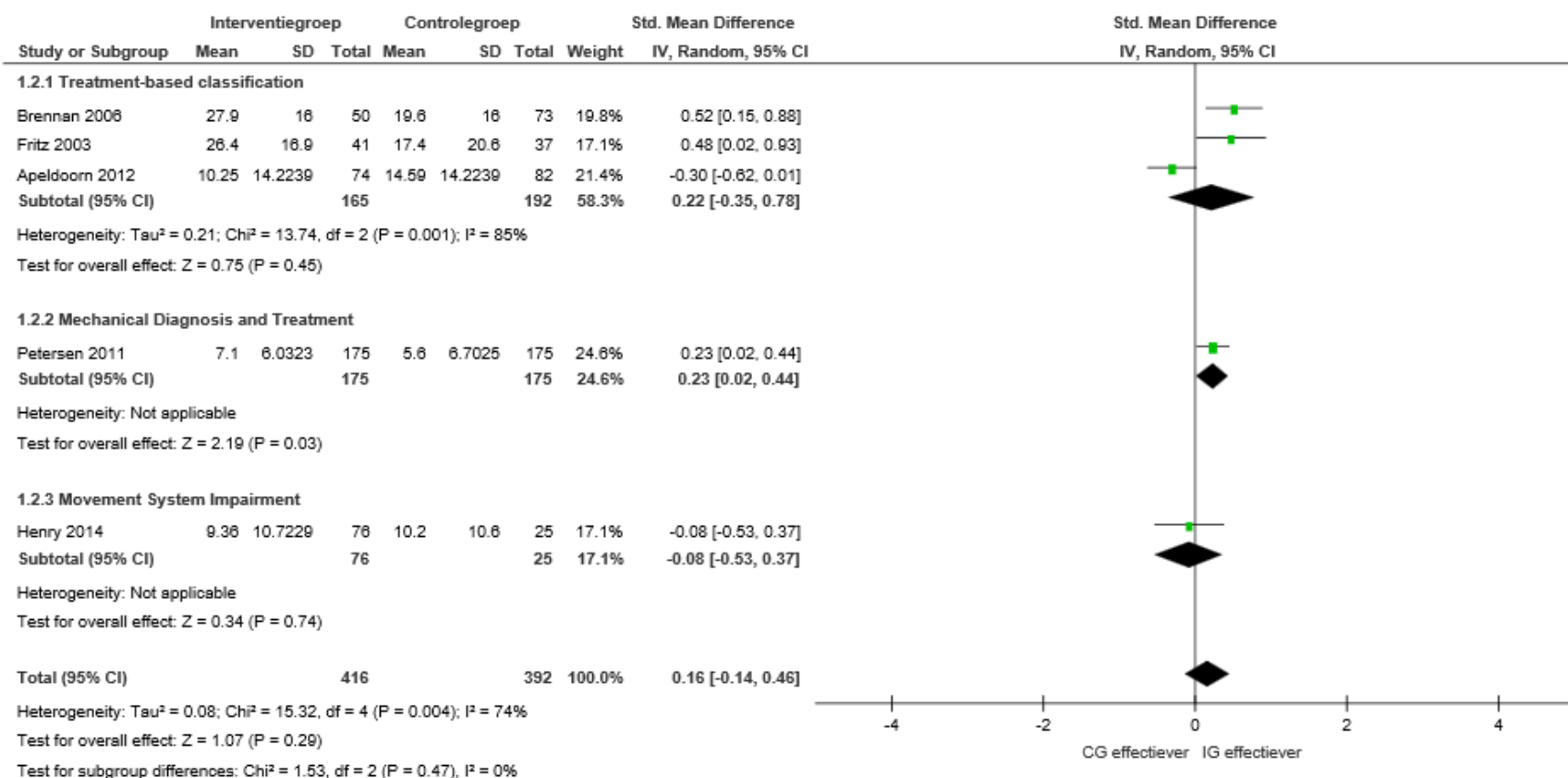
						1.5 (1.0*) <u>Controlegroep:</u> 1.3 (1.0*) P=0.232	
Henry et al (2014)	Aspecifieke lage rugklachten (n=101) Gemiddelde leeftijd: I: 40.15 (11.1) C: 45.88 (8.8)	Movement System Impairment en Treatment-Based classification. Stabilisatie oefeningen of movement system impairment gerichte oefeningen Bewust ingedeeld in een van deze classificatiesystemen (n=76)	Movement System Impairment en Treatment-Based Classification Stabilisatie oefeningen of Movement System Impairment gerichte oefeningen Onbewust ingedeeld in een van deze classificatiesystemen (n=25)	7 weken 52 weken	Pijn (NPRS) Functioneren (ODI)	ODI <u>Interventiegroep:</u> 12.55 (1.14**) 9.36 (1.23**) <u>Controlegroep:</u> 9.22 (1.98**) 10.20 (2.12**) P <0.0001 NPRS <u>Interventiegroep:</u> 1.59 (0.18**) 1.79 (0.25**) <u>Controlegroep:</u> 1.42 (0.29**) 1.24 (0.41**) P<0.0001	Er zijn geen verschillen tussen de patiënten die geclassificeerde behandelingen kregen en patiënten die random waren ingedeeld.
Murphy et al (2015)	Aspecifieke lage rugklachten (n=583) Gemiddelde leeftijd: I: 43.4 (11.4) C: 43.1 (10.7)	Geclassificeerde zorg door Start Back screening tool. Voorlichting en oefentherapie (n=251)	Niet geclassificeerde zorg, voorlichting en oefentherapie (n=332)	Direct na de behandeling 12 weken	Pijn (VAS) Functioneren (RMDQ)	RMDQ <u>Interventiegroep:</u> 6.31 (4.97*) 3.84 (5.11*) <u>Controlegroep:</u> 7.57 (5.75*) 3.42 (5.12*) P=0.028 VAS <u>Interventiegroep:</u> 3.31 (2.38*) 1.77 (2.72*) <u>Controlegroep:</u> 3.57 (2.57*)	Dit model van geclassificeerde zorg, zorgt voor een vroegtijdig en doeltreffend model in het management van chronische ziekte bij LBP patiënten en wordt aanbevolen om de patiënt resultaten te maximaliseren.

						1.23 (2.59*) P=0.015	
Sheeran et al (2013)	Aspecifieke lage rugklachten (n=49) Gemiddelde leeftijd: I: 35.9 (10.1) C: 37.1 (11.1)	O'Sullivan classification, behandeling gericht op houding (n=25)	Reguliere interventies gericht op houding (n=24)	4 weken	Pijn (VAS) Functioneren (RMDQ)	RMDQ <u>Interventiegroep:</u> 2.9 (2.3-3.7) <u>Controlegroep:</u> 6.9 (4.9-8.9) 2.2 (1.5-3.0)95% P=0.001 VAS <u>Interventiegroep:</u> 2.9 (2.5-3.4) <u>Controle:</u> 4.5 (3.9-5.11) 3.8(2.2-5.5)95% P=0.001	Geclassificeerde behandelingen zijn effectiever als het gaat over een behandeling van actie extensie en flexie patronen.
*: Standaarddeviatie (SD) **: Standard error							

Figuur 2: Meta-analyse**Figuur 2a: Pijn korte termijn**

Figuur 2b: Pijn lange termijn


Figuur 2c: Functioneren korte termijn

Figuur 2d: Functioneren lange termijn


Tabel 9 Resultaten kwaliteit van bewijs**Tabel 9.1: Mechanical Diagnosis and Treatment**

Vraagstelling: Mechanical Diagnosis and Treatment vergeleken met reguliere fysiotherapie bij patiënten met aspecifieke lage rugpijn.

Beoordeling kwaliteit van bewijs							Aantal patiënten		Effect		Kwaliteit van bewijs
Aantal studies	Studieopzet	Risk of bias	Inconsistentie	Indirect bewijs	Onnauwkeurigheid	Andere factoren	Mechanical Diagnosis and Therapy	usual group care	Relatief (95% CI)	Absoluut (95% CI*)	
Functionele status (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: RMDQ; Scale from: 0 tot 24)											
3 ¹¹⁻¹³	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	ernstig ¹	niet gevonden	318	318	-	mean 0.17 hoger (0.01 lager tot 0.35 hoger)	⊕⊕⊕○ REDELIJK
Functionele status (follow up lange termijn: range 6 maanden en langer; vastgesteld met: RMDQ; Scale from: 0 tot 24)											
1 ¹³	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	niet ernstig	niet gevonden	175	175	-	mean 0.23 hoger (0.02 hoger tot 0.44 hoger)	⊕⊕⊕⊕ HOOG
Pijnintensiteit (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: NPRS en Pain Reduction; Scale from: 0 tot 10)											
3 ¹¹⁻¹³	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	ernstig ¹	niet gevonden	318	318	-	mean 0.12 lager (0.37 lager tot 0.12 hoger)	⊕⊕⊕○ REDELIJK
Pijnintensiteit (follow up lange termijn: range 6 maanden en langer ; vastgesteld met: Pain Reduction; Scale from: 0 tot 60)											

1 ¹³	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	niet ernstig	niet gevonden	175	175	-	mean 0.23 hoger (0.02 hoger tot 0.44 hoger)	⊕⊕⊕⊕ HOOG
-----------------	------------------------	--------------	--------------	--------------	--------------	---------------	-----	-----	---	---	--------------

***CI:** Confidence interval

1: Twee groepen: ouder en jonger dan 54, niet bekend hoeveel personen in welke groep zitten.

Conclusie: De algemene kwaliteit van bewijs voor pijn en functioneren wordt als hoog-redelijk ingeschat. De korte termijn follow-ups van pijn en functioneren scoren beiden in niveau redelijk. Deze gradering is toegekend wegens ernstige onnauwkeurigheid¹¹. De lange termijn follow-ups van pijn en functioneren scoren beiden in niveau hoog. Deze resultaten bevatten geen factoren die de kwaliteit van bewijs verminderen.

Tabel 9.2: Movement System Impairment**Vraagstelling:** Movement system impairment vergeleken met reguliere fysiotherapie bij patiënten met aspecifieke lage rugpijn.

Beoordeling kwaliteit van bewijs							Aantal patiënten		Effect		Kwaliteit van bewijs
Aantal studies	Studieopzet	Risk of bias	Inconsistentie	Indirect bewijs	Onnauwkeurigheid	Andere factoren	movement system impairment-based treatment	usual physical therapy	Relatief (95% CI)	Absoluut (95% CI*)	
Functionele status (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: ODI; Scale from: 0 tot 100)											
3 ¹⁴⁻¹⁶	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	zeer ernstig ¹	niet gevonden	138	64	-	mean 0.24 hoger (0.07 lager tot 0.54 hoger)	⊕⊕○○ LAAG
Functionele status (follow up lange termijn: 6 maanden en langer; vastgesteld met: ODI; Scale from: 0 tot 100)											
1 ¹⁴	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	zeer ernstig ¹	niet gevonden	76	25	-	mean 0.08 lager (0.53 lager tot 0.37 hoger)	⊕⊕○○ LAAG
Pijnintensiteit (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: NPRS; Scale from: 0 tot 10)											
3 ¹⁴⁻¹⁶	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	zeer ernstig ¹	niet gevonden	138	64	-	mean 0.03 hoger (0.27 lager tot 0.33 hoger)	⊕⊕○○ LAAG
Pijnintensiteit (follow up lange termijn: 6 maanden en langer; vastgesteld met: NPRS; Scale from: 0 tot 10)											

1 ¹⁴	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	zeer ernstig ¹	niet gevonden	76	25	-	mean 0.25 hoger (0.2 lager tot 0.71 hoger)	⊕⊕○○ LAAG
-----------------	------------------------	--------------	--------------	--------------	---------------------------	---------------	----	----	---	--	--------------

***CI:** Confidence interval

1: Kleine steekproef. Geen betrouwbaarheidsindex tussen groepen beschikbaar.¹⁴

Conclusie: De algemene kwaliteit van bewijs voor pijn en functioneren wordt als laag ingeschat. Dit is voor alle follow-ups en uitkomstmaten geanalyseerd. Reden ter degradatie is zeer ernstige onnauwkeurigheid¹⁴.

Tabel 9.3: O'Sullivan Classification

Vraagstelling: O'Sullivan Classification vergeleken met reguliere fysiotherapie bij patiënten met aspecifieke lage rugpijn.

Beoordeling kwaliteit van bewijs							Aantal patiënten		Effect		Kwaliteit van bewijs
Aantal studies	Studieopzet	Risk of bias	Inconsistentie	Indirect bewijs	Onnauwkeurigheid	Andere factoren	classification system guided postural intervention (CSPI)	generalized postural intervention	Relatief (95% CI)	Absoluut (95% CI*)	
Functionele status (follow up korte termijn: 0 maanden tot 3 maanden; vastgesteld met: RMDQ; Scale from: 0 tot 24)											
1 ¹⁸	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	ernstig ¹	niet gevonden	25	24	-	mean 1.1 lager (1.7 lager tot 0.49 lager)	⊕⊕⊕○ REDELIJK
Pijnintensiteit (follow up korte termijn: 0 maanden tot 3 maanden; vastgesteld met: VAS; Scale from: 0 tot 10)											
1 ¹⁸	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	zeer ernstig ^{1,2}	niet gevonden	25	24	-	mean 1.18 lager (1.79 lager tot 0.57 lager)	⊕⊕○○ LAAG

***CI:** Confidence interval

1: Kleine steekproef

2: Grote betrouwbaarheidsinterval

Conclusie: De algemene kwaliteit van bewijs voor pijn en functioneren wordt als redelijk-laag ingeschat. De follow-ups vinden alleen op korte termijn plaats. De uitkomstmaat functioneren scoort in niveau redelijk, wegens ernstige onnauwkeurigheid¹⁸. De uitkomstmaat pijn scoort in niveau laag, wegens zeer ernstige onnauwkeurigheid¹⁸.

Tabel 9.4: Start Back Screening Tool**Vraagstelling:** Start Back Screening Tool vergeleken met reguliere fysiotherapie bij patiënten met aspecifieke lage rugpijn

Beoordeling kwaliteit van bewijs							Aantal patiënten		Effect		Kwaliteit van bewijs
Aantal studies	Studieopzet	Risk of bias	Inconsistentie	Indirect bewijs	Onnauwkeurigheid	Andere factoren	a stratified group intervention (STarT Back)	usual group care	Relatief (95% CI)	Absoluut (95% CI*)	
Functionele status (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: RMDQ; Scale from: 0 tot 24)											
1 ¹⁷	gerandomiseerde trials	zeer ernstig ¹	niet ernstig	niet ernstig	ernstig ²	niet gevonden	251	332	-	mean 0.08 hoger (0.08 lager tot 0.25 hoger)	⊕○○○ ZEER LAAG
Pijnintensiteit (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: VAS; Scale from: 0 tot 10)											
1 ¹⁷	gerandomiseerde trials	zeer ernstig ¹	niet ernstig	niet ernstig	ernstig ²	niet gevonden	251	332	-	mean 0.2 hoger (0.04 hoger tot 0.37 hoger)	⊕○○○ ZEER LAAG

***CI:** Confidence interval

1: Onvoldoende proportie van follow-up beschikbaar, reden niet bekend.

2: Geen betrouwbaarheidsindex tussen groepen beschikbaar vanuit het artikel

Conclusie: De Start Back Screening Tool scoort een volledige kwaliteit van bewijs in niveau zeer laag. Redenen ter degradatie zijn zeer ernstig risico op bias¹⁷ en ernstige onnauwkeurigheid¹⁷.

Tabel 9.5: Treatment-Based Classification**Vraagstelling:** Treatment-Based Classification vergeleken met reguliere fysiotherapie bij patiënten met aspecifieke lage rugpijn.

Beoordeling kwaliteit van bewijs							Aantal patiënten		Effect		Kwaliteit van bewijs
Aantal studies	Studieopzet	Risk of bias	Inconsistentie	Indirect bewijs	Onnauwkeurigheid	Andere factoren	a modified version of Delitto's treatment-based classification approach	usual physical therapy	Relatief (95% CI)	Absoluut (95% CI)	
Functionele status (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: ODI; Scale from: 0 tot 100)											
3 ^{8,9,10}	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	niet ernstig	niet gevonden	165	192	-	mean 0.23 hoger (0.27 lager tot 0.74 hoger)	⊕⊕⊕⊕ HOOG
Functionele status (follow up lange termijn: range 6 maanden en langer; vastgesteld met: ODI; Scale from: 0 tot 100)											
3 ^{8,9,10}	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	niet ernstig	niet gevonden	165	192	-	mean 0.22 hoger (0.35 lager tot 0.78 hoger)	⊕⊕⊕⊕ HOOG
Pijnintensiteit (follow up korte termijn: range 0 maanden tot 3 maanden; vastgesteld met: NPRS; Scale from: 0 tot 10)											
1 ⁸	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	ernstig ¹	niet gevonden	74	82	-	mean 0.16 hoger (0.16 lager tot 0.47 hoger)	⊕⊕⊕○ REDELIJK
Pijnintensiteit (follow up lange termijn: range 6 maanden en langer; vastgesteld met: NPRS; Scale from: 0 tot 10)											

1. ⁸	gerandomiseerde trials	niet ernstig	niet ernstig	niet ernstig	ernstig ¹	niet gevonden	74	82	-	mean 0.06 lager (0.37 lager tot 0.26 hoger)	⊕⊕⊕○ REDELIJK
-----------------	------------------------	--------------	--------------	--------------	----------------------	---------------	----	----	---	---	------------------

CI: Confidence interval

Kleine steekproef⁸

Conclusie: De algemene kwaliteit van bewijs voor pijn en functioneren wordt als redelijk-hoog ingeschat. De korte en lange termijn follow-ups van pijn scoren beiden in niveau redelijk. Deze gradering is toegekend wegens ernstige onnauwkeurigheid⁸. De korte en lange termijn follow-ups van functioneren scoren beiden in niveau hoog. Deze resultaten bevatten geen factoren die de kwaliteit van bewijs verminderen.

Figuur 3: Samengesteld Classificatiesysteem op basis van Treatment-Based Classification en O'Sullivan Classification

