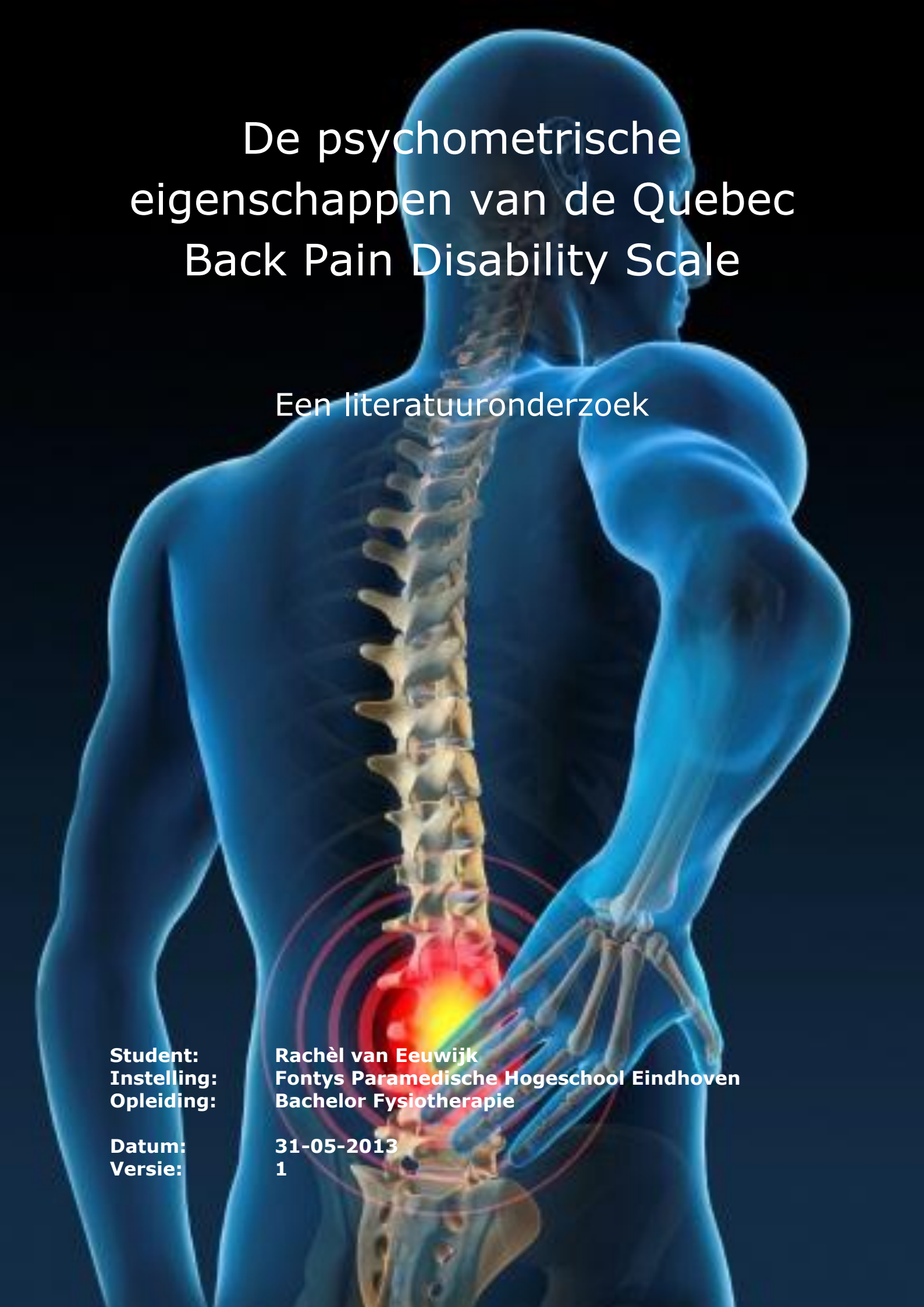




De psychometrische eigenschappen van de Quebec Back Pain Disability Scale

Een literatuuronderzoek

Student:
Instelling:
Opleiding:

Rachèl van Eeuwijk
Fontys Paramedische Hogeschool Eindhoven
Bachelor Fysiotherapie

Datum:
Versie:

31-05-2013
1

Personalia

Student:

Naam: Eeuwijk, van
Voorletters: R.G.H.
Voornaam: Rachèl
Adres: Papaverstraat 30
5482MD Schijndel
Mobiel: 0643484484
Telefoonnummer: -
Email: r.vaneeuwijk@student.fontys.nl

Gegevens opleiding:

Naam: Fontys Paramedische Hogeschool
Adres: Ds. Th. Fliednerstraat 2
5600 AH Eindhoven
Telefoonnummer: 0877-877011
Email: paramedisch@fontys.nl

Docentbegeleider:

Naam: Tjarco Koppedaal
Email: t.koppenaal@fontys.nl
Telefoonnummer: 0885-089760

Voorwoord

U staat op het punt mijn literatuuronderzoek te lezen, welke ik heb geschreven als afstudeeropdracht van de opleiding Fysiotherapie van de Fontys Paramedische Hogeschool te Eindhoven.

Deze studie gaat over de psychometrische eigenschappen van de Quebec Back Pain Disability Scale. Dit onderwerp is aangedragen vanuit de opleiding door Tjarco Koppenaar, docent en tevens begeleider van deze afstudeeropdracht.

Ik heb voor dit onderwerp gekozen, omdat de psychometrische eigenschappen van meetinstrumenten in de vorm van een vragenlijst niet aan bod zijn gekomen gedurende de opleiding. Gezien de ruime beschikbaarheid van vragenlijsten binnen de fysiotherapie, leek mij dit een geschikt onderzoek om meer over de psychometrische eigenschappen te leren en hoe je een vragenlijst kunt beoordelen op bruikbaarheid. Ik denk dat dit zeer relevant is en bruikbaar wanneer ik zelf als fysiotherapeut aan de slag ga in de praktijk.

Voordat ik mocht starten met mijn literatuuronderzoek, heb ik een projectplan moeten maken. Goedkeuring van dit projectplan vindt u in de bijlagen (*zie bijlage 1*).

Ten slotte wil ik graag een dankwoord uitbrengen naar de volgende personen:
Tjarco Koppenaar voor een prettige begeleiding en zijn fulltime beschikbaarheid.
Mijn medestudenten Patricia Meijs en Myou Coumans voor de feedback die ik heb mogen ontvangen van hen.
Cindy Vogels voor het controleren van mijn onderzoeksverslag op taal- en spelfouten.

Rachèl van Eeuwijk,

Schijndel, mei 2013

Samenvatting

Achtergrond

Aspecifieke lage rugklachten veroorzaken de hoogste kosten wat betreft arbeidsongeschiktheid en werkverzuim binnen Nederland. Deze waren 3,5 biljoen euro in 2007 en het aantal mensen met lage rugklachten neemt nog steeds toe. Er zijn weinig afstemmingsafspraken gemaakt binnen de zorgverlening met betrekking tot aspecifieke lage rugklachten, omdat er geen gouden standaard bestaat en een goed diagnostisch instrument ontbreekt. Hierdoor zijn er verschillende vragenlijsten ontwikkeld. Het Koninklijk Nederlands Genootschap voor Fysiotherapeuten adviseert binnen de richtlijn lage rugpijn de Quebec Back Pain Disability Scale.

Vraagstelling

Wat zijn de psychometrische eigenschappen van de Quebec Back Pain Disability Scale in het meten van beperkingen in activiteiten bij volwassen patiënten met subacute en chronische aspecifieke lage rugklachten?

Methode

Er is gezocht naar Nederlandse en Engelse literatuur in de databases Pubmed, MEDline en CINAHL door zoektermen betreft het klachtenbeeld, de vragenlijst en de psychometrische eigenschappen te combineren. De methodologische kwaliteit en uiteindelijke 'levels of evidence' zijn bepaald aan de hand van de methode, opgesteld door Consensus-based Standards for the selection of health Measurement INstruments.

Resultaten

De Quebec Back Pain Disability Scale heeft een goede interne consistentie (Cronbach's alpha 0.92-0.979), betrouwbaarheid (Intraclass Correlation Coefficient 0.86-0.96 en Kappa 0.86-0.92), meetfout (MIC 14 en > SDC), inhoudsvaliditeit, hypothese-testende eigenschap (correlaties van 0.63-0.857) responsiviteit (correlaties van 0.80-0.91) en is gevalideerd in het Engels en Frans.

Conclusie

De Quebec Back Pain Disability Scale heeft goede psychometrische eigenschappen betreft de interne consistentie, betrouwbaarheid, meetfout, inhoudsvaliditeit, hypothese-testen, cross-culturele validiteit van de Franse versie en responsiviteit.

Abstract

Background

Nonspecific low back pain causes the highest costs in disability and absenteeism in the Netherlands. These were 3,5 billion euro's in 2007 and the number of people with low back pain is still increasing. There are few coordination agreements within healthcare concerns nonspecific low back pain, because there is no golden standard and a good diagnostic instrument is missing. This is the reason several questionnaires have been developed. The Royal Dutch community for physiotherapists advises the Quebec Back Pain Disability Scale within the guideline low back pain.

Question

What are the psychometric properties of the Quebec Back Pain Disability Scale in measuring disability in activities of adult patients with sub-acute or chronic nonspecific low back pain?

Method

There has been searched for literature in Dutch and in the databases Pubmed, MEDline and CINAHL by combining search terms concerns the image of the complaint, questionnaire and the psychometric properties. The methodological quality and the final 'levels of evidence' are determined by the method developed by Consensus-based Standards for the selection of health Measurement INstruments.

Results

The Quebec Back Pain Disability Scale has a good internal consistency (Cronbach's alpha 0.92-0.979), reliability (Intraclass Correlation Coefficient 0.86-0.96 en Kappa 0.86-0.92), measurement error (MIC 14 en > SDC), content validity, hypothese-testing (correlations of 0.63-0.857), responsiveness (correlations of 0.80-0.91) and is validated in English and French.

Conclusion

De Quebec Back Pain Disability Scale has good psychometric properties contains the internal consistency, reliability, error of measurement, content validity, hypothese-testing, cross-cultural validity of the French version and responsiveness.

Inhoudsopgave

Personalia	2
Voorwoord	3
Samenvatting	4
Abstract	5
Inhoudsopgave	6
Inleiding	7
Methode	9
Resultaten	11
Discussie	15
Conclusie	18
Literatuurlijst	19
Figuren	22
Tabellen	25
Bijlagen	35

Inleiding

Lage rugklachten behoren tot een van de meest voorkomende aandoeningen onder de Nederlandse bevolking.¹⁻² Ongeveer 60-90% krijgt ooit last van lage-rugpijn.³ Uit gegevens van 2003 blijkt dat voor fysiotherapeuten 'lage-rugpijn' de meest voorkomende verwijsdiagnose is.³⁻⁴ Zevenentwintig procent van alle patiënten die de fysiotherapeut bezoekt heeft lage-rugpijn.²⁻³

Lage rugklachten vormen een belangrijk economisch probleem in Nederland.^{2-3, 5-8} Van alle klachten aan het houdings- en bewegingsapparaat veroorzaken lage rugklachten de hoogste kosten wat betreft arbeidsongeschiktheid en werkverzuim.^{2-3, 5-8} Eén op de vijf Nederlanders met lage rugklachten moet als gevolg van lage rugklachten het werk verzuimen.⁹ In 2007 waren de kosten ten gevolge van lage rugklachten binnen Nederland 3,5 biljoen euro en het aantal mensen met lage rugklachten neemt nog steeds toe.⁸ De verwachtingen zijn dat het aantal mensen met lage rugklachten in 2025 met 7% zal zijn toegenomen ten opzichte van 2007.¹⁰

Bij ongeveer 90% van de mensen met lage rugklachten zijn dit aspecifieke lage rugklachten (ALR).¹¹ Dat wil zeggen dat er geen specifieke oorzaak gevonden kan worden voor de pijn en beperkingen met betrekking op de lage rugklachten.³ De variatie in diagnostiek en behandeling is als gevolg van het ontbreken van een specifieke oorzaak dan ook enorm, gezien wetenschappelijk bewijs voor de effectiviteit van diagnostiek en behandeling nog ontbreekt.² Ook is het zorgaanbod bij ALR zeer divers (bijvoorbeeld de huis- of bedrijfsarts, de fysio-, manueel-, Mensendieck- of Cesar therapeut en/of de psycholoog).² Ondanks de omvang van het probleem en de diversiteit in zorgaanbod zijn er met betrekking tot ALR weinig afstemmingsafspraken gemaakt binnen de zorgverlening.² Elke zorgaanbieder bekijkt ALR vanuit de eigen beroepsgroep en hanteert een eigen monodisciplinaire richtlijn.² Er worden dan ook verschillende classificatiesystemen gebruikt, omdat er geen gouden standaard bestaat om beperkingen bij ALR te meten en een goed diagnostisch instrument ontbreekt.^{2, 12}

Door de fysiotherapeut wordt het dagelijks functioneren van mensen met ALR bekeken vanuit drie perspectieven, namelijk op functie-, activiteiten-, en participatieniveau.³⁻⁴ Activiteiten als lang staan, zitten, wandelen, tillen, bukken, reiken en gooien zijn veel voorkomende problemen bij ALR.¹³⁻¹⁴ Deze enkelvoudige activiteiten zijn weer nodig om dagelijkse bezigheden zoals het huishouden en het uitvoeren van werk te kunnen doen.¹³⁻¹⁴ Wanneer mensen worden beperkt in deze activiteiten kan dat aanleiding geven tot werkverzuim.¹³⁻¹⁴

De belangrijkste taken van de fysiotherapeut bij ALR zijn het geven van voorlichting, helpen bij het herstel van functies en herpakken van de beperkte activiteiten en de patiënt activeren.³⁻⁴ Het is dus voor de fysiotherapeut essentieel om een goed inzicht te krijgen in de beperkingen van de patiënt met ALR. Er kan dan een persoonlijk behandelplan worden opgesteld, gericht op het voorkomen van werkverzuim of het herpakken van werk. In de richtlijn lage rugpijn van het Koninklijk Nederlands Genootschap voor Fysiotherapeuten (KNGF) wordt dan ook geadviseerd gebruik te maken van meetinstrumenten bij het diagnosticeren en behandelen van ALR om de beperkingen in activiteiten goed in kaart te kunnen brengen en verandering in functionele status te kunnen evalueren.³ Echter zijn er meerdere meetinstrumenten en vragenlijsten ontwikkeld.^{13,15} Het KNGF adviseert binnen de richtlijn lage rugpijn het gebruik van de Quebec Back Pain Disability Scale (QBPDS) (zie *bijlage 2*).³

De QBPDS is een evaluatief meetinstrument dat bestaat uit 20 items, ontwikkelt om door de patiënt zelf in te vullen en functionele beperkingen te meten bij patiënten met lage rugpijn.^{14,16} Het gaat hierbij om enkelvoudige fysieke activiteiten.¹⁶ De 20 items kunnen

worden gescoord tussen de 0 en de 5 waarbij 0 niet beperkt betekend en 5 niet in staat om het genoemde item uit te voeren. Er ontstaat een score tussen de 0 en de 100, wat aangeeft in welke mate iemand functioneel wordt beperkt.¹⁶ In de QBPDS wordt de definitie van 'beperking' gebruikt zoals deze door de World Health Organization (WHO) is omschreven.¹³⁻¹⁴ Deze definitie luidt 'enige beperking of gebrek aan vermogen om een activiteit uit te voeren op een wijze, of binnen de range beschouwd als normaal voor een mens'.¹⁷ Om als fysiotherapeut te weten of de QBPDS een goed en bruikbaar instrument is in het vaststellen van de beperkingen in activiteiten en deze ook te kunnen evalueren, moeten de psychometrische eigenschappen van de QBPDS bekend zijn.¹⁸

De psychometrische eigenschappen kunnen worden onderverdeeld in drie hoofdcategorieën, namelijk betrouwbaarheid, validiteit en responsiviteit.¹⁸ Onder de categorie betrouwbaarheid vallen de psychometrische eigenschappen interne consistentie, betrouwbaarheid en meetfout.¹⁸ Onder validiteit vallen de psychometrische eigenschappen inhoudsvaliditeit, constructvaliditeit, hypothese-testen, cross-culturele validiteit en criteriumvaliditeit.¹⁸ De categorie responsiviteit bevat enkel de psychometrische eigenschap responsiviteit (*zie figuur 1 en bijlage 3*).¹⁸

Consensus-based Standards for the selection of health Measurement INstruments (COSMIN) heeft een methode ontwikkeld om de kwaliteit van een vragenlijst aan de hand van de psychometrische eigenschappen te bepalen.¹⁸ Er is voor elke psychometrische eigenschap een box met vragen opgesteld om alle aspecten die meegenomen moeten worden in een studie naar die specifieke psychometrische eigenschap te scoren.¹⁸⁻¹⁹ Aan de hand van deze methode wordt er dus nauwkeurig gekeken per studie en per onderzochte psychometrische eigenschap.¹⁸⁻¹⁹

Er is tot op heden geen systematisch literatuuronderzoek gedaan naar de psychometrische eigenschappen van de QBPDS waarbij gebruik is gemaakt van de COSMIN-methode. Het is daarom onduidelijk of de QBPDS een kwalitatief goed meetinstrument is en bruikbaar in de praktijk voor de fysiotherapeut.

Het doel van dit onderzoek is dan ook door middel van een literatuurstudie de psychometrische eigenschappen van de QBPDS kritisch te beoordelen en te vergelijken, door het systematisch beoordelen van de methodologische kwaliteit en inhoud van studies die de psychometrische eigenschappen van de QBPDS onderzoeken en/of ontwikkelen.

Daarom wordt in dit systematisch literatuuronderzoek de volgende vraag onderzocht: 'Wat zijn de psychometrische eigenschappen van de Quebec Back Pain Disability Scale in het meten van beperkingen in activiteiten bij volwassen patiënten met subacute en chronische aspecifieke lage rugklachten?'

Methode

Zoekstrategie

Databases

Voor dit systematisch literatuuronderzoek werd gezocht in de databases Pubmed, MEDline en CINAHL.

Zoektermen

Om artikelen te vinden die onderzoek hebben gedaan naar de juiste patiëntenpopulatie is gezocht op: lumbar spine, back pain, back ache, low back pain, chronic back pain, aspecific back pain, aspecific low back pain, non specific back pain en non specific low back pain.

Om artikelen te vinden over de QBPDS werd gebruik gemaakt van de zoektermen: Quebec Back Pain Disability Scale, QBPDS en Quebec.

Om artikelen te vinden die ten minste één psychometrische eigenschap onderzochten of ontwikkelden werd gebruik gemaakt van de volgende zoektermen: measurement properties, psychometric characteristics, validity, content validity, face validity, construct validity, structural validity, hypotheses testing, cross cultural validity, criterion validity, reliability, internal consistency, measurement error en responsiveness.

Alle zoektermen zijn gecombineerd met OR. De zoekstrategieën voor het vinden van artikelen met de juiste patiëntenpopulatie, de QBPDS en de psychometrische eigenschappen zijn vervolgens gecombineerd met AND. De gehele zoekstrategie voor de drie databases is opgenomen in de bijlagen (*zie bijlage 4*).

In- en exclusiecriteria

Om geïnccludeerd te worden in dit systematisch literatuur onderzoek moesten de artikelen aan de volgende inclusiecriteria voldoen: in het Engels of Nederlands verkrijgbaar zijn, full tekst verkrijgbaar zijn, minimaal één psychometrische eigenschap onderzoeken of ontwikkelen van de QBPDS, de onderzoekspopulatie moest bestaan uit patiënten met specifieke lage rugpijn of lage rugklachten tussen de 18 en 80 jaar oud.

Artikelen werden geëxcludeerd van dit systematisch literatuur onderzoek als: het een systematische literatuuronderzoek over de psychometrische eigenschappen van de QBPDS betrof, het een onderzoekspopulatie bevatte van patiënten met specifieke lage rugpijn of rugklachten of met acute specifieke lage rugpijn of rugklachten.

Selectieprocedure

Aan de hand van de in- en exclusiecriteria zijn de artikelen eerst gescreend op titel. Vervolgens is het abstract van de overgebleven artikelen gescreend op basis van de in- en exclusiecriteria. De artikelen die na screening op titel en abstract zijn over gebleven werden full tekst gescreend op basis van de in- en exclusiecriteria en vervolgens is bepaald welke artikelen zijn meegenomen in het systematisch literatuuronderzoek. De resultaten van de selectieprocedure zullen overzichtelijk worden weergegeven in een stroomdiagram (*zie figuur 2*).

Data-extractie

De volgende gegevens zijn uit de geïnccludeerde artikelen geëxtraheerd: auteur(s), jaartal, land, taal, setting, studiepopulatie, geslacht en gemiddelde leeftijd van de studiepopulatie, in- en exclusiecriteria betreft het klachtenbeeld, onderzochte of ontwikkelde psychometrische eigenschap(en) en het resultaat van de onderzochte psychometrische eigenschap(en). Voor een goede vergelijking van de artikelen, is de belangrijke informatie uit de geïnccludeerde artikelen weergegeven in een data-extractietabel (zie *tabel 1*).

Kwaliteitsbeoordeling

De kwaliteitsbeoordeling is gedaan in 3 verschillende stappen. Bij de data-extractie is reeds vastgesteld welke psychometrische eigenschappen van de QBPDS in het artikel zijn onderzocht of ontwikkeld.

De eerste stap was het bepalen van de methodologische kwaliteit per onderzochte eigenschap van een studie. Dit werd gedaan aan de hand van de COSMIN checklist die bestaat uit de eerder genoemde boxen per psychometrische eigenschap, waarmee kan worden bepaald of de psychometrische eigenschap op de juiste manier is onderzocht.¹⁸⁻¹⁹ De vragen binnen elke box kunnen beantwoord worden met *poor*, *fair*, *good* en *excellent*. Voor het totaaloordeel van de box telde uiteindelijk het laagste oordeel.¹⁹

Nadat de methodologische kwaliteit van de studie naar een bepaalde psychometrische eigenschap is bepaald, werd het resultaat beoordeeld aan de hand van de criteria van Terwee et al.²⁰ (Zie *tabel 2*). Er kon worden gescoord met *positief*, wat wijst op een sterke psychometrische eigenschap. *Negatief*, wat wijst op een zwakke psychometrische eigenschap. En *onbepaald*, dat wil zeggen dat het onduidelijk is of de psychometrische eigenschap sterk of zwak aanwezig is.

Het totaal oordeel van de onderzochte psychometrische eigenschap per studie is uiteindelijk bepaald door de methodologische kwaliteit en het resultaat van de psychometrische eigenschap per studie te combineren. Dit is wederom gedaan aan de hand van de criteria van Terwee et al.²⁰ (zie *tabel 3*). Dit is beoordeeld met één tot drie plussen wat een voldoende, goede of uitstekende methodologische kwaliteit aangeeft in combinatie met een positief resultaat. Of met één tot drie minnen, wat eveneens voldoende, goede of uitstekende methodologische kwaliteit aangeeft in combinatie met een negatief resultaat. Of met een vraagteken, wat aangeeft dat de studie van slechte methodologische kwaliteit is of het resultaat is onbepaald.

Levels of evidence

Uiteindelijk is een totaal oordeel gevormd per psychometrische eigenschap van de QBPDS. Dit is gedaan op basis van alle geïnccludeerde artikelen die een bepaalde psychometrische eigenschap onderzoeken. De resultaten van alle artikelen die dezelfde psychometrische eigenschap onderzoeken zijn bij elkaar gelegd om de consistentie van de resultaten te bekijken. Er is gescoord met één tot drie plussen, wat een beperkt, matig of sterk bewijsniveau aangeeft voor sterke aanwezigheid van die psychometrische eigenschap. Met één tot drie minnen, wat een beperkt, matig of sterk bewijsniveau aangeeft voor zwakke aanwezigheid van die psychometrische eigenschap. Een plus-min, wat aangeeft dat er tegenstrijdige resultaten zijn gevonden voor een sterke of zwakke aanwezigheid van de psychometrische eigenschap. En de laatste scoringsmogelijkheid was een vraagteken, wat aangeeft dat de mate waarin een psychometrische eigenschap aanwezig is onbepaald is (zie *tabel 4*).

Resultaten

Zoekstrategie en selectieprocedure

Er is gezocht tot 15-04-2013 en er zijn na deze datum geen artikelen meer geïnccludeerd.

De zoekstrategie leverde in de database Pubmed 86 resultaten op. In de databases MEDline en CINAHL is tegelijkertijd gezocht en deze leverden 287 resultaten op. Na screening op titel aan de hand van de in- en exclusiecriteria bleven er van de resultaten in Pubmed 42 artikelen over. Van de databases MEDline en CINAHL waren dat er 47. Er was een overlap van 25 artikelen, waardoor er in totaal 46 artikelen overbleven voor screening op abstract.

Na screening op abstract aan de hand van de in- en exclusiecriteria vielen er nog 21 artikelen af, waardoor er 25 artikelen overbleven voor screening op full text eveneens aan de hand van de in- en exclusiecriteria.

Van de 25 overgebleven artikelen waren er 3 niet full text beschikbaar, twee artikelen bevatte een onderzoekspopulatie met een acuut klachtenbeeld, 1 artikel bevatte een onderzoekspopulatie met een hoog percentage patiënten ouder dan 80 jaar in combinatie met veel nevenpathologie en 2 artikelen waren systematische literatuuronderzoeken. Dit gaf 17 geïnccludeerde artikelen.

Deze resultaten zijn weergegeven in een stroomdiagram (*zie figuur 3*).

Data-extractie

De belangrijkste gegevens zijn uit de artikelen geëxtraheerd om zo de artikelen gemakkelijk te kunnen vergelijken. Bij het interpreteren van de aantallen moet men weten dat er één studie staat omschreven in twee artikelen.¹³⁻¹⁴

In veertien artikelen^{13-14, 22-26, 28-30, 32, 34-36} is er onderzoek gedaan naar de interne consistentie. In dertien artikelen^{13-14, 22-26, 28-30, 34-36} is er onderzoek gedaan naar de betrouwbaarheid. Twee artikelen¹³⁻¹⁴ omschreven onderzoek naar de meetfout. Zes artikelen^{13-14, 23, 27, 30, 32} bevatten onderzoek naar de inhoudsvaliditeit. In tien artikelen^{13-14, 22-27, 34-35} is er onderzoek gedaan naar de constructvaliditeit. Eveneens zijn er tien artikelen^{13-14, 25, 29-35} die onderzoek hebben gedaan naar de hypothese-testende eigenschap. In elf artikelen^{13-14, 22-27, 29-30, 36} is er onderzoek gedaan naar de cross-culturele validiteit. Zeven artikelen^{23-24, 26-27, 30-31, 36} omschreven onderzoek naar de criteriumvaliditeit en in vijf artikelen^{13-14, 28, 33-34} is er onderzoek gedaan naar de responsiviteit (*zie tabel 5*).

De onderzoekspopulatie varieerde van 32 patiënten tot 242 patiënten, waarbij de onderzoekspopulatie in de meeste gevallen gemengd was betreft geslacht. Eén studie³³ betrof enkel vrouwen en één studie²⁴ bestond voor 95% uit vrouwen. De gemiddelde leeftijd lag over het algemeen tussen de 30 en 50 jaar. Betreft het klachtenbeeld zijn er vier artikelen^{13-14, 22, 34} met een gemengde onderzoekspopulatie van zowel specifieke als aspecifieke klachten en twee van deze artikelen¹³⁻¹⁴ bevatten naast een subacuut- en chronisch klachtenbeeld ook patiënten met een acuut klachtenbeeld. In twee artikelen^{24, 33} staat het klachtenbeeld niet aangegeven en in één artikel²⁵ wordt wel aangegeven dat het om een chronisch klachtenbeeld gaat, maar niet of het een specifiek of aspecifiek klachtenbeeld betreft (*zie tabel 5*).

Het grootste gedeelte van de omschreven studies zijn uitgevoerd in een ziekenhuis, rugcentrum of revalidatiecentrum. In drie artikelen^{13-15, 27} is ook de particuliere praktijk meegenomen en in twee van deze artikelen¹³⁻¹⁴ is daarnaast ook een psychiatrische instelling en een familiale groepspraktijk meegenomen. In één artikel³⁴ is de studie uitgevoerd binnen een huisartsenpraktijk (zie tabel 5).

De gevonden resultaten in de artikelen worden besproken bij de resultaten van het bepalen van de methodologische kwaliteit.

Kwaliteitsbeoordeling

Aan de hand van de COSMIN-checklist¹⁸⁻¹⁹ is de methodologische kwaliteit van elke studie naar een psychometrische eigenschap, die in de artikelen staan omschreven vastgesteld. Vervolgens is bekeken of het resultaat positief, negatief of onbepaald was aan de hand van de criteria van Terwee et al.²⁰ Uiteindelijk zijn deze resultaten gecombineerd voor een totaal oordeel, eveneens aan de hand van de criteria van Terwee et al.²⁰ (Zie tabel 6 en 7).²⁰ Door de consistentie van de resultaten te bekijken van alle studies die onderzoek hebben gedaan naar één bepaalde psychometrische eigenschap zijn de uiteindelijke bewijsniveaus (levels of evidence) per psychometrische eigenschap bepaald.

Interne consistentie

Dertien studies^{13-14, 22-26, 28-30, 32, 34-36} hebben onderzoek gedaan naar de interne consistentie van de QBPDS voor het meten van de onderlinge correlaties van de items. Elf van deze studies^{13-14, 22-26, 29-30, 34-36} hebben dit uitgedrukt in Cronbach's alpha. Drie van de dertien studies^{13-14, 22, 29} hadden ten minste een voldoende methodologische kwaliteit en tien van deze studies^{23-26, 28, 30, 32, 34-36} waren van slechte methodologische kwaliteit (zie tabel 6). Alle studies gaven positieve resultaten weer (zie tabel 6). De Cronbach's alpha varieerde van 0.92 tot 0.979 (zie tabel 5). Door deze resultaten te combineren, bleven er drie positieve resultaten over uit de studies van ten minste voldoende methodologische kwaliteit. De overige resultaten zijn onbepaald wegens een slechte methodologische kwaliteit (Zie tabel 7). Wegens consistente positieve resultaten in drie studies^{13-14, 22, 27} van ten minste goede methodologische kwaliteit die de interne consistentie hebben onderzocht, is er een sterk bewijsniveau voor een goede interne consistentie van de QBPDS (zie tabel 8).

Betrouwbaarheid

De betrouwbaarheid van de QBPDS is in twaalf studies^{13-14, 22-27, 29-30, 34-36} onderzocht en uitgedrukt met Kappa (K) of de Intraclass Correlation Coefficient (ICC). Tien van deze studies^{13-14, 22-27, 29, 34, 36} waren van ten minste voldoende methodologische kwaliteit en twee studies^{30, 35} van slechte methodologische kwaliteit (zie tabel 6). Al deze studies gaven positieve resultaten weer die varieerden tussen K 0.86 en 0.92²³ en ICC 0.86 en 0.96^{13-14, 22, 24-27, 29, 34, 36} (zie tabel 5 en 6). Combinatie van deze resultaten laat nog tien positieve resultaten over en twee resultaten zijn onbepaald wegens een slechte methodologische kwaliteit (Zie tabel 7). Vijf studies^{13-14, 22, 26-27, 34} van goede methodologische kwaliteit die de betrouwbaarheid hebben onderzocht, vonden consistente positieve resultaten. Dit geeft een sterk bewijsniveau voor een goede betrouwbaarheid van de QBPDS (zie tabel 8).

Meetfout

Slechts één studie²² van goede methodologische kwaliteit heeft onderzoek gedaan naar de meetfout van de QBPDS weergegeven met de minimal important change (MIC) en de smallest detectable change (SDC) (zie tabel 6). De MIC was 14 punten en is groter bevonden dan de SDC²² (zie tabel 5), wat een positief resultaat is (zie tabel 6). Door deze resultaten te combineren blijft er een positief resultaat staan wegens de goede methodologische kwaliteit van de studie (zie tabel 7).

Er is één studie²² die de meetfout heeft onderzocht van goede methodologische kwaliteit met een positief resultaat. Dat geeft een matig bewijsniveau voor de meetfout van de QBPDS (zie tabel 8).

Inhoudsvaliditeit

Vijf studies^{13-14,23,27,30,32} hebben onderzoek gedaan naar de inhoudsvaliditeit van de QBPDS. Waarvan één studie¹³⁻¹⁴ van uitstekende methodologische kwaliteit, één studie²³ van voldoende methodologische kwaliteit en de overige drie studies^{27,30,32} van slechte methodologische kwaliteit (zie tabel 6).

Alle vijf de studies^{13-14,23,27,30,32} constateerden een goede inhoudsvaliditeit van de QBPDS, wat wil zeggen dat de items van de vragenlijst relevant zijn bevonden met betrekking tot de doelgroep en dat de vragenlijst is beschouwd als compleet. Dit zijn dus enkel positieve resultaten (zie tabel 6).

Na combinatie van de resultaten blijven er twee positieve resultaten staan en drie onbepaalde resultaten wegens slechte methodologische kwaliteit van de studies (zie tabel 7).

Eén studie¹³⁻¹⁴ van uitstekende methodologische kwaliteit met een positief resultaat onderzocht de inhoudsvaliditeit. Dit geeft een sterk bewijsniveau voor een goede inhoudsvaliditeit van de QBPDS (zie tabel 8).

Construct validiteit

Er zijn vijf studies^{13-14, 22, 25, 27, 34} gevonden die onderzoek hebben gedaan naar de constructvaliditeit van de QBPDS waarbij de variantie in score met een meetinstrument dat hetzelfde construct meet is verklaard, of terwijl de 'explained variance' (zie tabel 6). Deze studies^{13-14, 22, 25, 27, 34} waren allemaal van ten minste voldoende methodologische kwaliteit (zie tabel 6).

Echter zijn er tegenstrijdige resultaten die varieerden tussen <50% en 81,537%.^{13-14, 22, 25, 27, 34} Eén studie²⁵ gaf een negatief resultaat weer van <50% (zie tabel 5 en 6).

Gezien alle studies van ten minste voldoende methodologische kwaliteit waren is er één negatief resultaat blijven staan en vier positieve resultaten (zie tabel 7).

Er waren vijf studies^{13-14, 22, 26-27, 34} gevonden van ten minste goede methodologische kwaliteit met tegenstrijdige bevindingen die de constructvaliditeit onderzochten. Dit geeft dan ook een tegenstrijdig bewijsniveau voor een goede constructvaliditeit van de QBPDS (zie tabel 8).

Hypothese-testen

Acht studies^{13-14, 25, 29-32, 34-35} deden onderzoek naar de hypothese-testende eigenschap van de QBPDS, waarvan in zeven studies^{13-14, 25, 29-31, 34-35} de correlatie is bepaald met een meetinstrument dat hetzelfde construct meet. Zes van deze studies^{13-14, 25, 29-31, 34} waren van ten minste voldoende methodologische kwaliteit en één van deze studies was van slechte methodologische kwaliteit³⁵ (Zie tabel 6).

Er zijn positieve resultaten gevonden van correlaties tussen de 0.63 en 0.857 met de Oswestry Questionnaire (OQ)¹³⁻¹⁴, de Roaland Morris Disability Questionnaire (RMDQ)^{25, 29}, de Short Form Health Survey 36 (SF-36)³⁰, de Oswestry Disability Index (ODI)³¹, de Dallas Pain Questionnaire (DPI)³², de Roland Disability Questionnaire (RDI)³⁴ en de Pain Disability Index (PDI)³⁵. Van één studie³² is het resultaat onbepaald, omdat deze de correlatie heeft bepaald met een meetinstrument dat een ander construct meet (Zie tabel 5 en 6).

Zo blijven er uiteindelijk zes positieve resultaten over en twee onbepaalde resultaten wegens een slechte methodologische kwaliteit of een onbepaald resultaat (zie tabel 7). Vijf studies^{13-14, 29-31, 34} van goede methodologische kwaliteit onderzochten de hypothese-testende eigenschap en vonden consistente positieve resultaten. Dit geeft een sterk bewijsniveau voor een goede hypothese-testende eigenschap van de QBPDS (zie tabel 8).

Cross-culturele validiteit

Tien studies^{13-14,22-27,29-30,36} hebben onderzoek gedaan naar de cross-culturele validiteit van de QBPDS. Negen van deze studies^{13-14,22, 24-27,29-30,36} waren van ten minste reële methodologische kwaliteit en hebben het vertalingsproces goed uitgevoerd. Eén studie²³ was van slechte methodologische kwaliteit en heeft het vertalingsproces onjuist uitgevoerd, dit betreft de Arabische versie van de QBPDS (zie tabel 6).

Slechts in één studie¹³⁻¹⁴ welke de Franse versie van de QBPDS betreft, is de 'original factor structure' bevestigd en een positief resultaat vastgesteld. In de overige studies^{22-27,29-30,36} is het resultaat onbepaald (zie tabel 6).

Bij het combineren van de resultaten wordt dus één positief resultaat behouden en de overige resultaten blijven onbepaald (zie tabel 7).

Slechts één studie¹³⁻¹⁴ (welke onderzoek deed naar de Franse versie) was van goede methodologische kwaliteit en had een positief resultaat vastgesteld. Dit geeft een matig bewijsniveau voor de cross-culturele validiteit van de QBPDS (zie tabel 8).

De QBPDS is vertaald vanuit het Engels naar het Frans,^{13-14,32} Hongaars,²² Arabisch,²³ Tswana,²⁴ Marokkaans,²⁵ Koreaans,²⁶ Grieks,²⁷ Braziliaans,²⁹ Perzisch,³⁰ Nederlands,³⁴ en Turks.³⁵⁻³⁶ De studie die het vertalingsproces van de Nederlandse versie bevat ontbreekt in mijn resultaten.

Criteriumvaliditeit

Zeven studies^{23-24,26-27,30-31,36} hebben onderzoek gedaan naar de criteriumvaliditeit van de QBPDS door de correlatie te bepalen met een meetinstrument dat hetzelfde construct meet. Echter wordt de criteriumvaliditeit bepaald aan de hand van een 'gouden standaard' en deze bestaat niet voor het meten van beperkingen. Al deze studies zijn daarom van slechte methodologische kwaliteit (zie tabel 6).

De resultaten varieerden van correlaties tussen de 0.67 en de 0.85 met de ODI^{23, 26, 31, 36} en de Disability Rating Index (DRI)²⁴. In twee studies^{27, 30} is het resultaat onjuist of niet weergegeven (zie tabel 5). Echter blijven alle resultaten onbepaald volgens de criteria van Terwee et al.²⁰, omdat het resultaat alleen kan worden bepaald aan de hand van een 'gouden standaard' (zie tabel 6).

Zo blijft het uiteindelijke resultaat voor criteriumvaliditeit onbepaald (zie tabel 7).

Gezien de criteriumvaliditeit van de QBPDS niet kan worden vastgesteld wegens het ontbreken van een 'gouden standaard', waren er enkel studies^{23-24,26-27,30-31,36} beschikbaar van slechte methodologische kwaliteit. Dit geeft een onbekend bewijsniveau voor de criteriumvaliditeit van de QBPDS (zie tabel 8).

Responsiviteit

Vier studies^{13-14, 28, 33-34} hebben de responsiviteit van de QBPDS onderzocht, maar slechts twee studies hebben dit gedaan door het bepalen van de correlatie met een meetinstrument dat hetzelfde construct meet over langere termijn.^{13-14,34} Deze studies^{13-14,34} zijn beide van goede methodologische kwaliteit en de overige twee^{28, 33} zijn van slechte methodologische kwaliteit (zie tabel 6).

De twee studies^{13-14,34} die wel een correlatie hebben bepaald vonden positieve resultaten van 0.80 en 0.91 met de OQ en de RDQ (zie tabel 5 en 6).

Vanwege de methodologische kwaliteit en de vastgestelde- of juist niet vastgestelde resultaten zijn uiteindelijk twee positieve resultaten blijven staan en twee onbepaalde resultaten (zie tabel 7).

Twee studies^{13-14, 34} van goede methodologische kwaliteit onderzochten de responsiviteit en vonden consistente positieve bevindingen. Dit geeft een sterk bewijsniveau voor een goede responsiviteit van de QBPDS (zie tabel 8).

Er is een sterk bewijsniveau gevonden voor de psychometrische eigenschappen interne consistentie, betrouwbaarheid, inhoudsvaliditeit, hypothese-testen en responsiviteit.

Discussie

De vraagstelling van deze literatuurstudie was: 'Wat zijn de psychometrische eigenschappen van de Quebec Back Pain Disability Scale in het meten van beperkingen in activiteiten bij volwassen patiënten met subacute en chronische aspecifieke lage rugklachten?'

In de resultaten is gevonden dat de QBPDS een goede interne consistentie (Cronbach's alpha 0.92-0.979)^{13-14, 22-26, 29-30, 34-36}, betrouwbaarheid (Intraclass Correlation Coefficient 0.86-0.96 en Kappa 0.86-0.92)^{13-14, 22-27, 29-30, 34-36}, meetfout (MIC 14 en > SDC)²², inhoudsvaliditeit^{13-14, 23, 27, 30, 32}, hypothese-testende eigenschap (correlaties van 0.63-0.857)^{13-14, 25, 29-31, 34-35}, en responsiviteit (correlaties van 0.80-0.91)^{13-14, 34} heeft. De QBPDS is gevalideerd in het Engels¹³⁻¹⁴ en Frans^{13-14, 32}. Daarnaast is de QBPDS vertaald vanuit het Engels naar het Hongaars,²² Arabisch,²³ Tswana,²⁴ Marokkaans,²⁵ Koreaans,²⁶ Grieks,²⁷ Braziliaans,²⁹ Perzisch,³⁰ Nederlands,³⁴ en Turks.³⁵⁻³⁶ Betreft de constructvaliditeit zijn er tegenstrijdige resultaten gevonden en de criteriumvaliditeit van de QBPDS kan niet worden vastgesteld wegens het ontbreken van een 'gouden standaard'.

Sterke- en zwakke punten

Het is een sterk punt van dit literatuuronderzoek dat de uiteindelijke 'levels of evidence' zijn bepaald door middel van de COSMIN-methode^{18, 19}, waarbij er eerst stapsgewijs en nauwkeurig wordt gekeken naar de methodologische kwaliteit aan de hand van de COSMIN-checklist¹⁸⁻¹⁹ en de resultaten worden beoordeeld aan de hand van de criteria van Terwee et al.²⁰ Door deze twee stappen te combineren is het oordeel per psychometrische eigenschap van de QBPDS volledig.

Echter is dit literatuuronderzoek uitgevoerd door één onderzoeker (R. v. E.). De methodologische kwaliteit en de uiteindelijke 'levels of evidence' per psychometrische eigenschap zijn dus vastgesteld door één onderzoeker. Mogelijk zijn er bepaalde onderdelen van de COSMIN-checklist^{18, 19} of de criteria van Terwee et al.²⁰ verkeerd geïnterpreteerd of over het hoofd gezien. Dit kan gevolgen hebben gehad op het oordeel van de methodologische kwaliteit en de uiteindelijke 'levels of evidence' per psychometrische eigenschap.

Echter zijn er meerdere studies met consistente positieve bevindingen gevonden voor de psychometrische eigenschappen interne consistentie, betrouwbaarheid, hypothese-testen en responsiviteit. Het is aannemelijk dat de uiteindelijke bewijsniveaus van deze psychometrische eigenschappen ook daadwerkelijk kloppen, ondanks dat dit literatuuronderzoek door één onderzoeker is uitgevoerd. Echter is het aanbevolen de psychometrische eigenschappen van de QBPDS aan de hand van de COSMIN-methode¹⁸⁻¹⁹ nogmaals te onderzoeken met minimaal twee onderzoekers.

Voor het vinden van artikelen die psychometrische eigenschappen onderzoeken is in Pubmed al een zoekstrategie ontwikkeld.³⁷ Echter is hier in dit onderzoek geen gebruik van gemaakt. De in dit systematisch literatuuronderzoek gebruikte zoekstrategie is een door de onderzoeker opgestelde zoekstrategie die behoorlijk volledig lijkt te zijn. Er zijn namelijk veel artikelen gevonden die aan de inclusiecriteria voldoen en veel geschikte resultaten.

Het kan als een beperking worden gezien dat er in dit literatuuronderzoek geen artikelen geïnccludeerd zijn die studies omschreven met een onderzoekspopulatie van mensen met acute ALR. De resultaten zijn dan ook niet van toepassing zijn op patiënten met dit klachtenbeeld. Kanttekening is echter, dat een afwijkend beloop van de klachten pas een indicatie geeft voor fysiotherapie.³ Mensen met een acuut klachtenbeeld hebben op dat moment in elk geval nog een normaal beloop van de klachten en zouden niet of

nauwelijks bij de fysiotherapeut komen. Dit maakt dat deze resultaten ook helemaal niet van belang zouden zijn.

Toch is er één studie¹³⁻¹⁴ geïnccludeerd die een onderzoekspopulatie bevatte met zowel een chronisch-, subacuut-, en acuut klachtenbeeld. Daarnaast zijn er nog drie studies^{13-14, 22, 34} die een onderzoekspopulatie bevatte van zowel specifieke als aspecifieke klachten. En er zijn nog drie studies^{24-25, 33} waarin niet of niet volledig staat aangegeven wat het klachtenbeeld is. Eén van deze studies³³ betrof een onderzoekspopulatie van enkel vrouwen met posterieure bekkenpijn na zwangerschap. Deze aantallen zijn vrij klein (zie tabel 5) en niet elk resultaat is meegenomen na controle aan de hand van de COSMIN-checklist¹⁸⁻¹⁹ of aan de hand van de criteria van Terwee et al.²⁰ (zie tabel 6 en 7). Daarom wordt niet verwacht dat de uiteindelijke 'levels of evidence' hiermee sterk beïnvloed zijn en dus toepasbaar op de patiëntengroep, zoals omschreven in de onderzoeksvraag.

Binnen het domein betrouwbaarheid is de psychometrische eigenschap meetfout van de QBPDS slechts in één studie²² onderzocht. Deze psychometrische eigenschap is dus bepaald aan de hand van slechts één resultaat. Echter was deze studie was wel van goede methodologische kwaliteit, wat het resultaat meer betrouwbaar maakt. De geconstateerde uitkomst is dat de score op de QBPDS minimaal 14 punten veranderd moet zijn, om er zeker van te zijn dat er sprake is van een verandering in functionele status en geen sprake is van een meetfout. Om er zeker van te zijn als fysiotherapeut dat de functionele status van de patiënt is verbeterd of verslechterd, moet deze dus minimaal 14 punten meer of minder scoren op de QBPDS. Toch wordt er aanbeveling gedaan de psychometrische eigenschap meetfout van de QBPDS verder te onderzoeken.

Binnen het domein validiteit is onder andere de psychometrische eigenschap cross-culturele validiteit onderzocht. Met uitzondering van één studie¹³⁻¹⁴ blijft het uiteindelijke resultaat onbepaald, omdat er niet wordt aangegeven of de 'original factor structure' is behouden of er geen Differential Item Functioning (DIF) is bepaald. Dat wil zeggen dat de vertaalde versie niet is vergeleken met de oorspronkelijke versie of de scoring daarin eenduidig was. Echter is het belangrijkste om te weten, hoe het vertaalproces is uitgevoerd. Wanneer dit goed is gedaan, kan men aannemen dat de scoring van de vertaalde versie eenduidig zal zijn. Alleen de studie naar de ontwikkeling van de Arabische versie²³ heeft dit proces niet goed uitgevoerd. Dat wil zeggen dat de QBPDS in elk geval te gebruiken is in het engels¹³⁻¹⁴, Frans^{13-14,32}, Hongaars²², Tswana²⁴, Marokkaans²⁵, Koreaans,²⁶ Grieks²⁷, Braziliaans²⁹, Perzisch³⁰, en Turks.³⁵⁻³⁶ Onderzoek ter verbetering van de Arabische versie²³ van de QBPDS is aanbevolen.

De criteriumvaliditeit van de QBPDS kan niet worden vastgesteld. Er is namelijk geen 'gouden standaard' voor het meten van beperkingen in activiteiten bij ALR.¹² Dit is dan ook de rede waarom er zoveel vragenlijsten zijn ontwikkeld die dit construct beogen te meten^{13,15} en het juist zo belangrijk is dat de psychometrische eigenschappen van een vragenlijst bekend zijn¹⁸. Uiteindelijk was dat de aanleiding van dit literatuuronderzoek.

Het derde domein was responsiviteit met de psychometrische eigenschap responsiviteit. Er waren twee studies^{13-14, 34} die positieve resultaten hadden bevonden betreffende de responsiviteit van de QBPDS. Dit wil zeggen dat de QBPDS in staat is verandering te meten over de langere termijn en dus geschikt is als een evaluatief meetinstrument binnen de fysiotherapie. Echter is verder onderzoek naar de responsiviteit van de QBPDS aanbevolen, gezien het beperkte aantal studies.

Vergelijking met de literatuur

Er zijn drie literatuuronderzoeken³⁸⁻⁴⁰ met overeenkomstige resultaten betreft interne consistentie, betrouwbaarheid, cross-culturele validiteit en responsiviteit van de QBPDS. Vier literatuuronderzoeken³⁸⁻⁴¹ bevatten overeenkomstige resultaten betreft de inhoudsvaliditeit van de QBPDS. Twee literatuuronderzoeken^{38, 41} bevatten overeenkomstige resultaten betreft de meetfout van de QBPDS en één literatuuronderzoek³⁹ bevat overeenkomstige resultaten betreft de constructvaliditeit van de QBPDS. Ondanks dat dit literatuuronderzoek door één onderzoeker is uitgevoerd, was er voldoende kennis en vaardigheid aanwezig om overeenkomstige resultaten te vinden. Dit versterkt de vastgestelde bewijsniveaus per psychometrische eigenschap. Ondanks de overeenkomstige resultaten, zijn er enkele noemenswaardige verschillen in het vaststellen van deze resultaten. Hierdoor ontbreken er soms gegevens in de genoemde literatuur.

In één literatuuronderzoek⁴¹ worden enkele aandachtspunten omschreven betreft de inhoudsvaliditeit van de QBPDS. Namelijk dat wanneer de QBPDS niet in combinatie wordt gebruikt met een meetinstrument voor pijn, patiënten automatisch hun pijn mee kunnen nemen bij het scoren van de items. Verder dat een patiënt eventuele nevenpathologie mee kan nemen in het scoren van bepaalde items (bijvoorbeeld het gooien van een bal bij schouderpathologie) en dat het item autorijden storing kan geven wanneer iemand geen rijbewijs heeft. Echter is niet terug te vinden in het artikel hoe de inhoudsvaliditeit van de QBPDS is onderzocht en waar deze aandachtspunten op zijn gebaseerd. Dit pleit voor het toepassen van de COSMIN-methode¹⁸⁻¹⁹, zodat resultaten altijd te herleiden zijn. Echter kan men deze aandachtspunten binnen de praktijk toch meenemen en de patiënt erop wijzen dat de QBPDS ingevuld moet worden op basis van de ALR. Het toepassen van een meetinstrument voor pijn wordt in meerdere literatuuronderzoeken^{38-39, 41} geconcludeerd en wordt ook geadviseerd vanuit de KNGF richtlijn lage rugpijn³. De QBPDS meet ten slotte de beperkingen in activiteiten bij patiënten met ALR en is geen meetinstrument voor het meten van pijn.

Betreft de constructvaliditeit zijn er ook volgens Müller et al.³⁸ tegenstrijdige resultaten. Echter is in dit literatuuronderzoek de constructvaliditeit van de QBPDS alleen vergeleken met andere meetinstrumenten die hetzelfde construct meten. De verschillende aspecten die horen bij een studie naar de constructvaliditeit volgens de COSMIN-checklist¹⁹ zijn niet bekeken en beoordeeld. Dit maakt dat er geen uitspraak kan worden gedaan over de constructvaliditeit van de QBPDS en geen betrouwbare uitspraak kan worden gedaan over de vergelijking van de constructvaliditeit van de QBPDS met andere meetinstrumenten. Het construct 'beperkingen in activiteiten bij ALR' is dan ook een zeer complex construct en zeer divers, wat het juist zo lastig te meten maakt. Het is dan ook sterk aanbevolen meer onderzoek te verrichten naar de constructvaliditeit van de QBPDS, maar dan wel volgens de COSMIN-methode¹⁸⁻¹⁹.

Conclusie

De psychometrische eigenschappen van de QBPDS in het meten van beperkingen bij volwassenen met subacute en chronische ALR zijn een goede interne consistentie, een goede betrouwbaarheid, een positief resultaat betreft de meetfout, een goede inhoudsvaliditeit, een goede hypothese-testende eigenschap, een goede cross-culturele validiteit betreft de Franse versie van de QBPDS en een goede responsiviteit. Dit maakt de QBPDS een bruikbaar meetinstrument voor het meten en evalueren van beperkingen in activiteiten bij volwassenen met subacute en chronische aspecifieke lage rugklachten.

Aanbevelingen

- Het systematisch beoordelen en bepalen van de methodologische kwaliteit van studies die psychometrische eigenschappen onderzoeken of ontwikkelen volgens de COSMIN-methode¹⁸⁻¹⁹, uitgevoerd door minimaal twee personen.
- Het gebruik van de QBPDS binnen de praktijk als evaluatief meetinstrument in combinatie met een meetinstrument voor pijn.

Literatuurlijst

- (1) van der Lucht F, JJ Polder. Kernrapport van de Volksgezondheid Toekomst Verkenning 2010 Van gezond naar beter. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu; 2010. 94.
- (2) Baan CA, Hutten JBF, Rijken PM. Afstemming in de zorg Een achtergrondstudie naar de zorg voor mensen met een chronische aandoening. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu; 2003. 564.
- (3) Bekkering GE, Hendriks HJM, Koes BW, Oostendorp AB, Ostelo RJWG, Thomassen J, van Tulder MW. KNGF richtlijn voor lage rugklachten. Nederlands tijdschrift voor fysiotherapie 2001; 111: 1-7.
- (4) Faber E, Custers JWH, Van Ederen C, Bout J, Cinjee G, Kolnaar BGM, Schotsman R, Spinnewijn WEM, Staal JB, Ten Cate A, Wildervanck-Dekker CMJ. Landelijke Eerstelijns Samenwerkings Afspraak Aspecifieke Lagerugpijn. Huisarts Wet 2008; 51: S5-S9
- (5) Gast GCM, Spijkerman AMW, Schoenmaker CG. Preventie van chronische ziekten in risicogroepen. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu; 2011. 91.
- (6) Swinkels ICS, Leemrijse C, De Bakker D. Jaarboek LiPZ 2004 Deel 2: Beroepsgroep oefentherapie Cesar . Utrecht: NIVEL, 2005.
- (7) Swinkels ICS, Leemrijse C, De Bakker D. Jaarboek LiPZ 2004 Deel 3: Beroepsgroep oefentherapie Mensendieck. Utrecht: NIVEL, 2006.
- (8) Lambeek LC, van Tulder MW, Swinkels ICS, Koppes LL, Anema JR, van Mechelen W. The trend in total cost of back pain in The Netherlands in the period 2002-2007. Spine 2011; 36: 1050-1058.
- (9) Picavet HSJ, van Gils HWV, Schoutem JSAG. Klachten van het bewegingsapparaat in de Nederlandse bevolking. Prevalenties, consequenties en risicogroepen. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu/ Centraal Bureau voor Statistiek; 2000. 60.
- (10) Koes BW, van Tulder MW, Poos MJJC. Neemt het aantal mensen met nek- en rugklachten toe of af? In: Volksgezondheid Toekomst Verkenning, National Kompas Volksgezondheid 2006.
- (11) Deyo RA. Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil 1988; 69: 1044-1053.
- (12) Calmels P, Béthoux F, Condemine A, Fayolle-Minon I. Low back pain disability assessment tools. Ann Readapt Med Phys 2005; 6: 97-288.
- (13) Jacek A, Kopec, John M, Esdaile, Abrahamowicz M, Abenheim L, Wood-Dauphinee S, Donna L, Lamping, Williams J. The Quebec Back Pain Disability Scale measurement properties. Spine 1995; 20: 341-352.
- (14) Jacek A, Kopec, John M, Esdaile, Abrahamowicz M, Abenheim L, Wood-Dauphinee S, Donna L, Lamping, Williams J. The Quebec Back Pain Disability Scale conceptualization and development. J Clin Epidemiol 1996; 49:151-161.
- (15) Kopec JA, Esdaile JM. Functional disability scales for back pain. Spine 1995; 20: 1943-1949
- (16) Caroline B, Terwee A, Sandra DM, Bota, Michael R, de Boera B, Daniëlle AWM, van der Windta C, Dirk L, Knola D, Dekkera JE, Lex M, Bouter LM, Henrica CW, de Veta. Accepted Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology 2007; 60: 34-42.
- (17) International Classification of Impairments, Disabilities and Handicaps—A manual of Classification Relating to the Consequences of Disease. Geneva: World Health Organization, 1980.
- (18) Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012; 21: 651-657.
- (19) Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCH. The COSMIN checklist for assessing the methodological quality of studies

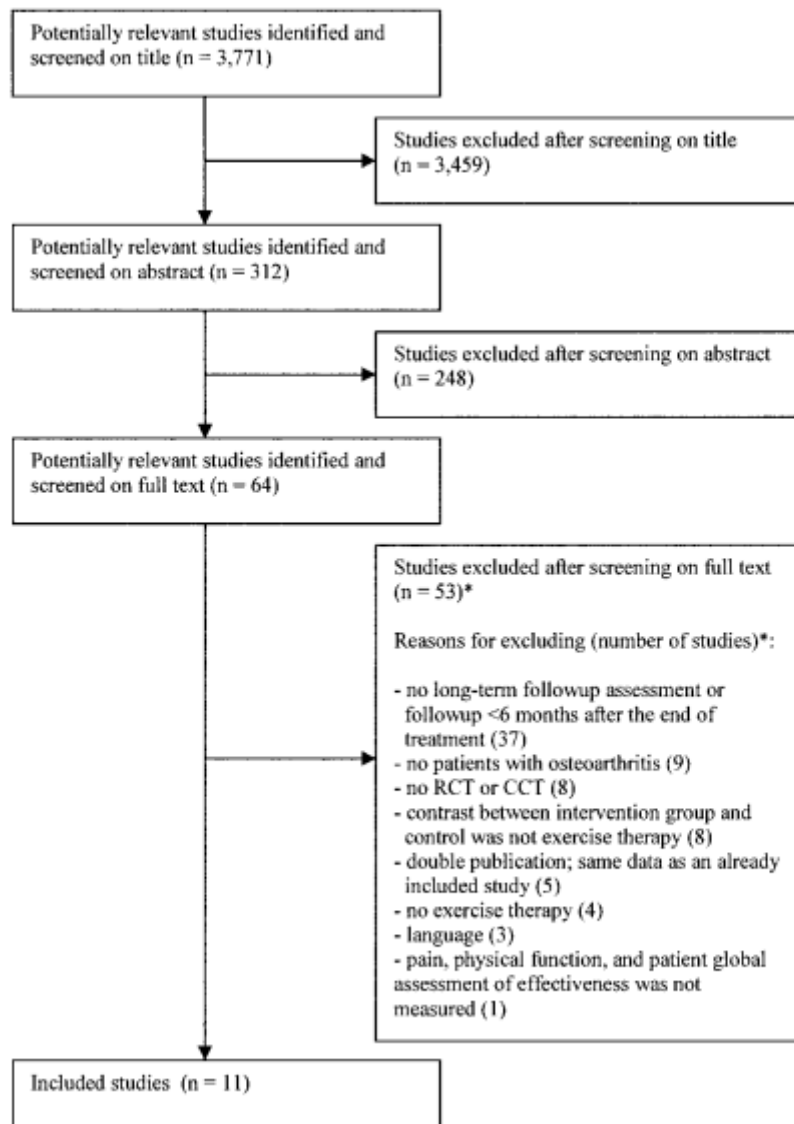
- on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research* 2010; 19: 539–549.
- (20) Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34-42
 - (21) van Tulder M, Furlan A, Bombardier C, Bouter L. Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003; 28: 1290-1299.
 - (22) Valasek T, Varga PP, Szövérfi Z, Kümin M, Fairbank J, Lazary A. Reliability and validity study on the Hungarian versions of the Oswestry Disability Index and the Quebec Back Pain Disability Scale. *Eur Spine J*. 2013; 22: 1010-8.
 - (23) Validation of the Arabic version of the quebec Back Pain Disability Scale. Alnahhal A, May S. Validation of the Arabic version of the quebec Back Pain Disability Scale. *Spine* 2012; 37: E1645-50.
 - (24) De Beer N, Stewart A, Becker P. Validation of the Tswana versions of the Roland-Morris Disability Questionnaire, Quebec Disability Scale and Waddell Disability Index . *South African Journal of Physiotherapy* 2008; 64: 23-30.
 - (25) Bendeddouche I, Rostom S, Bahiri R, Boudali A, Srifi N, Mawani N, Mengat M, El Badri D, Lazrak N, Abouqal R, Allali F, Hajjaj-Hassouni N. Translation, adaptation and validation of the Moroccan version of the Quebec Back Pain Disability Scale. *Clin Rheumatol* 2012; 31: 943-9.
 - (26) Suh KT, Kim JI, Lim JM, Goh TS, Lee JS. Validation of the Korean version of the Quebec Back Pain Disability Scale. *J Spinal Disord Tech* 2012; 25: 447-50.
 - (27) Christakou A, Andriopoulou M, Asimakopoulos PJ. Validity and reliability of the Greek version of the Quebec Back Pain Disability Scale. *Back Musculoskelet Rehabil* 2011; 24: 145-54.
 - (28) Demoulin C, Ostelo R, Knottnerus JA, Smeets RJ. Quebec Back Pain Disability Scale was responsive and showed reasonable interpretability after a multidisciplinary treatment. *J Clin Epidemiol* 2010; 63: 1249-55.
 - (29) Rodrigues MF, Michel-Crosato E, Cardoso JR, Traebert J. Psychometric properties and cross-cultural adaptation of the Brazilian Quebec back pain disability scale questionnaire. *Spine* 2009; 34: E459-64.
 - (30) Mousavi SJ, Parnianpour M, Mehdian H, Montazeri A, Mobini B. The Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Quebec Back Pain Disability Scale: translation and validation studies of the Iranian versions. *Spine* 2006; 31: E454-9.
 - (31) Reneman MF, Jorritsma W, Schellekens JM, Göeken LN. Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain. *J Occup Rehabil*. 2002; 12: 119-29.
 - (32) Yvanes-Thomas M, Calmels P, Béthoux F, Richard A, Nayme P, Payre D, Laurent B. Validity of the French-language version of the Quebec back pain disability scale in low back pain patients in France. *Spine*. 2002; 69: 397-405.
 - (33) Mens JM, Vleeming A, Snijders CJ, Ronchetti I, Ginai AZ, Stam HJ. Responsiveness of outcome measurements in rehabilitation of patients with posterior pelvic pain since pregnancy. *Spine* 2002; 27: 1110-5.
 - (34) Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Phys Ther* 1996; 76: 268-75.
 - (35) Bicer A, Yazici A, Camdeviren H, Milcan A, Erdogan C. Assessment of pain and disability in patients with chronic low back pain: reliability and construct validity of the Turkish version of the Quebec Back Pain Disability Scale and Pain Disability Index. *Journal of Back & Musculoskeletal Rehabilitation* 2005; 18: 37-44.
 - (36) Melikoglu MA, Kocabas H, Sezer I, Bilgilişoy M, Tuncer T. Validation of the Turkish version of the Quebec back pain disability scale for patients with low back pain. *Spine* 2009; 34: E219-24.

- (37) Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009; 18: 1115–1123.
- (38) Müller U, Röder C, Greenough CG. Back related outcome assessment instruments. *Spine* 2006; 15: S25-31.
- (39) Müller U, Duetz MS, Roeder C, Greenough CG. Condition-specific outcome measures for low back pain. Part I: validation. *Spine* 2004; 13: 301-13.
- (40) Rocchi MB, Sisti D, Benedetti P, Valentini M, Bellagamba S, Federici A. *Eura Medicophys*. Critical comparison of nine different self-administered questionnaires for the evaluation of disability caused by low back pain. 2005; 41: 275-81.
- (41) Müller U, Roeder C, Dubs L, Duetz MS, Greenough CG. Condition-specific outcome measures for low back pain. Part II: scale construction. *Spine* 2004; 13: 314-24.

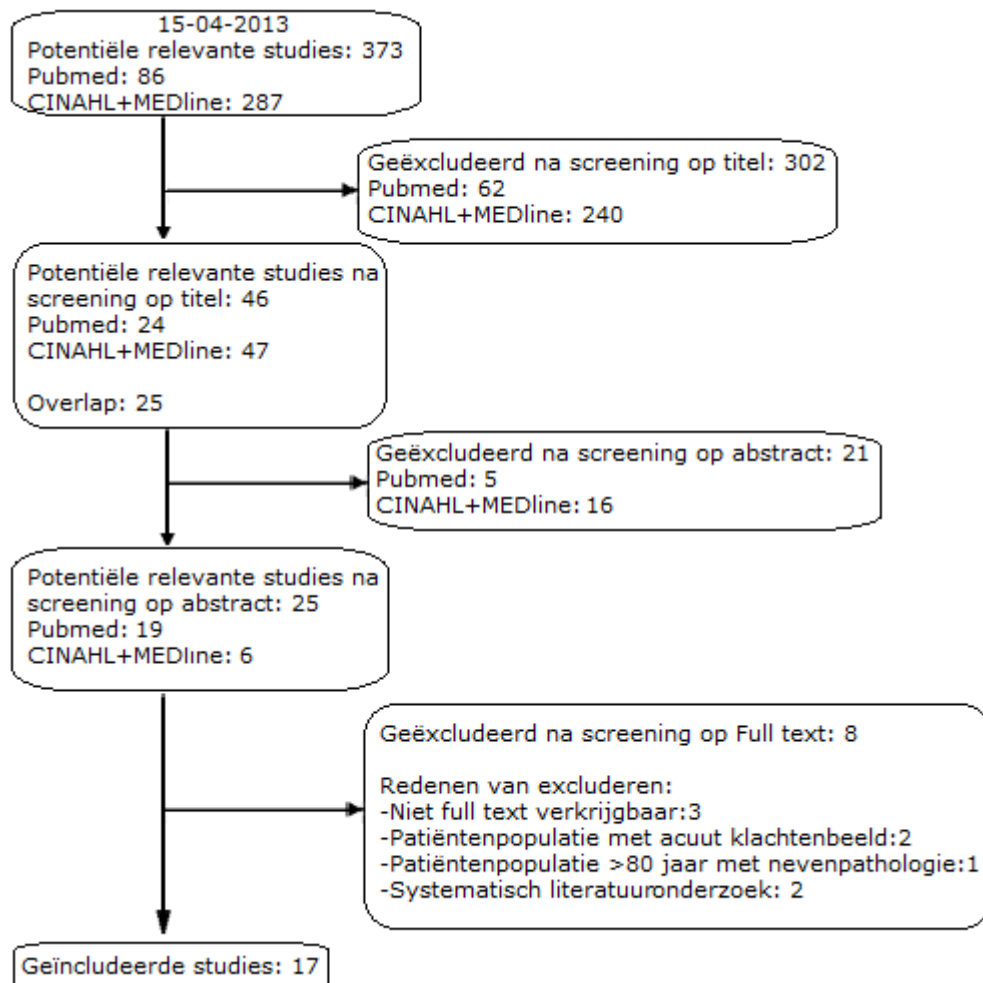
Figuren



Figuur 1 Psychometrische eigenschappen¹⁸



Figuur 2 Stroomdiagram



Figuur 3 Resultaten selectieprocedure

Tabellen

Tabel 1 – Data-extractietabel

Artikel	Land	Taal	Setting	Onderzoeks- populatie	Ge- slacht	Gemiddelde leeftijd	Klachtenbeeld	Psychometrische eigenschap(pen)	Resultaat

Tabel 2 – Criteria van Terwee et al.²⁰

Eigenschap	Waardering	Kwaliteitscriteria
Betrouwbaarheid		
Interne consistentie	+	Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha niet omschreven
	-	Cronbach's alpha(s) < 0.70
Betrouwbaarheid	+	ICC / Kappa ≥ 0.70 of Pearsons product-momentcorrelatiecoëfficiënt ≥ 0.80
	?	Geen ICC / Kappa, Pearsons product-momentcorrelatiecoëfficiënt omschreven
	-	ICC / Kappa < 0.70 of Pearsons product-momentcorrelatiecoëfficiënt < 0.80
Meetfout	+	MIC $>$ SDC of MIC buiten de LOA
	?	MIC niet omschreven
	-	MIC $\leq$ SDC of MIC gelijkgesteld of binnen de LOA
Validiteit		
Inhoudsvaliditeit	+	Alle items in de vragenlijst zijn relevant met betrekking tot de doelgroep en de vragenlijst is compleet.
	?	Er is geen doelgroep bij betrokken
	-	De items in de vragenlijst zijn niet relevant of de vragenlijst is niet compleet.
Construct validiteit	+	Ten minste 50% van de 'explained variance' wordt uitgelegd
	?	'Explained variance' wordt niet genoemd
	-	Minder dan 50% van de 'explained variance' wordt uitgelegd
Testen van hypothesen	+	Er is een correlatie met een ander instrument dat hetzelfde meet, deze waarde is > 0.50 , of tenminste 75% van de resultaten komt overeen met de hypothesen, of de AUC waarde is > 0.70 én er is een hogere waarde zichtbaar met 'related constructs' dan met 'unrelated constructs'.
	?	Er zijn uitsluitend correlaties vastgesteld met 'unrelated constructs'
	-	Er is een correlatie met een instrument dat hetzelfde meet, maar deze waarde is < 0.50 , of minder dan 75% van de hypothesen komt overeen met de resultaten, of AUC waarde is < 0.70 , of de correlatie met 'related constructs' is lager dan met 'unrelated constructs'
Cross-culturele validiteit	+	De 'original factor structure' wordt bevestigd of er wordt geen 'DIF' gevonden tussen de verschillende taalversies.
	?	Er is geen 'confirmatory factor analysis' toegepast of de 'DIF' is niet bepaald.
	-	'Original factor structure' wordt niet bevestigd of er is een belangrijke 'DIF' gevonden tussen de verschillende taalversies.
Criterium	+	Er zijn overtuigende argumenten dat de gouden standaard

validiteit		'goud' is en de correlatie > 0.70 bedraagt.
	?	Er zijn geen overtuigende argumenten dat de gouden standard 'goud' is, of er is een twijfelachtige inleiding of methode aanwezig.
	-	De correlatie met de gouden standaard is <0.70, ondanks een goede inleiding en methode.
Responsiviteit		
Responsiviteit	+	Er is een correlatie met een ander instrument dat hetzelfde meet, deze waarde is > 0.50, of tenminste 75% van de resultaten komt overeen met de hypothesen, of de AUC waarde is >0.70 én er is een hogere waarde zichtbaar met 'related constructs' dan met 'unrelated constructs'.
	?	Er zijn uitsluitend correlaties vastgesteld met 'unrelated constructs'
	-	Er is een correlatie met een instrument dat hetzelfde meet, maar deze waarde is < 0.50, of minder dan 75% van de hypothesen komt overeen met de resultaten, of AUC waarde is <0.70, of de correlatie met 'related constructs' is lager dan met 'unrelated constructs'

(..) = reference number, MIC = minimal important change, SDC = smallest detectable change, LoA = limits of agreement, ICC = intraclass correlation coefficient, DIF = differential item functioning, AUC = area under the curve

[†] + = positieve score, ? = onbepaalde score, - = negatieve score

Tabel 3 – Criteria van Terwee et al.²⁰

Score [†]	Kwaliteitscriteria
+++	Positief resultaat en methodologische kwaliteit uitstekend
++	Positief resultaat en methodologische kwaliteit goed
+	Positief resultaat en methodologische kwaliteit reëel
?	Resultaat onbepaald of methodologische kwaliteit slecht
-	Negatief resultaat en methodologische kwaliteit voldoende
--	Negatief resultaat en methodologische kwaliteit goed
---	Negatief resultaat en methodologische kwaliteit uitstekend

[†] + = positieve score, ? = onbepaalde score, - = negatieve score

Tabel 4 – Levels of evidence²¹

Level	Score [†]	Criteria
sterk	+++ or ---	Consistente bevindingen in meerdere studies van goede methodologische kwaliteit of in één studie van uitstekende methodologische kwaliteit
matig	++ or --	Consistente bevindingen in meerdere studies van reële methodologische kwaliteit of in één studie van goede methodologische kwaliteit
beperkt	+ or -	Één studie van reële methodologische kwaliteit
Tegenstrijdig	+/-	Tegenstrijdige bevindingen
onbekend	?	Alleen studies van slechte methodologische kwaliteit

[†] + = positieve score, ? = onbepaalde score, - = negatieve score

Tabel 5 – Data-extractie van de geïncludeerde studies

Artikel	Land	Taal	Setting	Onderzoeks- populatie	Ge- slacht	Gemiddelde leeftijd	Klachten- beeld	Psychometrische eigenschap(pen)	Resultaat
Kopec et al. ¹⁴	Canada	Engels/Frans	Part Pr/Zkh	153	M 122	NA	S/As	Interne consistentie	Ca 0.96
			Ps instelling	54	V 120	18-29 46	Ac/Sub/Chr	Betrouwbaarheid	ICC 0.92
			Fam GrPr	30		30-39 59		Inhoudsvaliditeit	
			Overig	NA		40-49 63		Construct validiteit	Ev 53%
Kopec et al. ¹⁴				T:242		50-59 33		Hypothese-testen	r=0.80 (OQ)
						>60 41		Cross-culturele validiteit	
								Responsiviteit	r=0.80 (OQ)
Valasek et al. ²²	Hongarije	Hongaars	Polikliniek van Na- tionaal centrum voor rugproble- matiek	T: 133	M 56 V 77	48.1-SD 15.3	S/As Chr	Interne consistentie	Ca 0.95
								Betrouwbaarheid	ICC 0.923
								Meetfout	MIC>SDC
								Construct validiteit	Ev 71%
								Cross-culturele validiteit	
Alnahhal A en May S ²³	Palestina	Arabisch	Zkh	T:148	M 99 V 49	33.4-SD 9.2	As Sub/Chr	Interne consistentie	Ca 0.92
								Betrouwbaarheid	K 0.86-0.92
								Inhoudsvaliditeit	
								Construct validiteit	Ev NA
								Cross-culturele validiteit	
								Criterium validiteit	r=0.67 (ODI)
De Beer et al. ²⁴	Zuid-Afrika	Tswana	Zkh	T:100	M 5 V 95	42-SD 9.13	NA NA	Interne consistentie	Ca 0.95
								Betrouwbaarheid	ICC 0.91
								Construct validiteit	Ev
								Cross-culturele validiteit	
								Criterium validiteit	r=0.85 (DRI)
Bendeddouche et al. ²⁵	Marokko	Marrokaans	Zkh	T:64	M 24 V 40	47.6-SD 12.3	NA Chr	Interne consistentie	Ca 0.979
								Betrouwbaarheid	ICC 0.959
								Construct validiteit	Ev <50%

								Hypothese-testen Cross-culturele validiteit	r=0,644 (RMDQ)
Kuen Tak Suh et al. ²⁶	Korea	Koreaans	Rugcentrum	T:80	M 28 V 52	48-SD 7.5	As Chr	Interne consistentie Betrouwbaarheid Construct validiteit Cross-culturele validiteit Criterium validiteit	Ca 0.9319 ICC 0.9094 Ev r=0.718 (ODI)
Christakou et al. ²⁷	Griekenland	Grieks	Part Pr/Rev Kliniek	T:130	M 60 V 70	41.1-SD 11.6	As Chr	Interne consistentie Betrouwbaarheid Inhoudsvaliditeit Construct validiteit Cross-culturele validiteit Criterium validiteit	Ca NA ICC 0.92 Ev 81.537% r NA
Demoulin et al. ²⁸	Nederland	Nederlands	Rev centrum	T:223	M 118 V 105	43-SD NA	As Chr	Responsiviteit	r NA
Marcelo et al. ²⁹	Brazilië	Braziliaans (Portugees)	Rugcentrum	T:54	M 16 V 34	44.3-SD 12.1	As Chr	Interne consistentie Betrouwbaarheid Hypothese-testen Cross-culturele validiteit	Ca 0.97 ICC 0.96 r=0.857 (RMQ)
Sayed Javad Mousavi et al. ³⁰	Iran	Perzisch	Zkh	T:100	M 45 V 55	40.1-SD 11.59	Aspecifiek Chr	Interne consistentie Betrouwbaarheid Inhoudsvaliditeit Hypothese-testen Cross-culturele validiteit Criteriumvaliditeit	Ca 0.92 ICC 0.86 r=0.69 (SF-36) r NA
Reneman et al. ³¹	Nederland	Nederlands	Polikliniek van universiteit	T:64	M 54 V 10	38-SD 8.9	As Chr	Hypothese-testen Criteriumvaliditeit	r=0.74 (ODI) r=0.74 (ODI)

Yvanes-Thomas et al. ³²	Frankrijk	Frans	Zkh	T:32	M 21 V 11	42-SD 16.2	As Chr	Interne consistentie Inhoudsvaliditeit Hypothese-testend	Ca NA r=>0.50 (DPQ)
Mens et al. ³³	Nederland	Nederlands	Rev Centrum	T:100	M 0 V 100	31.7 -SD 3.2 33.5-SD 4.9	NA NA	Responsiviteit	r NA
Schoppink et al. ³⁴	Nederland	Nederlands	Huisartsenpraktijk	T:115	M 69 V 46	39.7-SD 10.4	A/As Chr	Interne consistentie Betrouwbaarheid Construct validiteit Hypothese-testen Responsiviteit	Ca 0.96 ICC 0.90 Ev 81% r=0.80 (RDQ) r=0.91 (RDQ)
Bicer et al. ³⁵	Turkije	Turks	Rev centrum	T:83	M 20 V 63	43,6-SD 10.146	As Chr	Interne consistentie Betrouwbaarheid Construct validiteit Hypothese-testen	Ca 0.93 ICC 0.90 Ev r=0.630 (PDI)
Meltem Alkan Melikoglu et al. ³⁶	Turkije	Turks	Rev centrum van universiteit	T:100	M 26 V 74	45,4-SD 15.05	As Sub/Chr	Interne consistentie Betrouwbaarheid Inhoudsvaliditeit Cross-culturele validiteit Criterium validiteit	Ca 0,9537 ICC 0.9221 r=0.681 (ODI)

NA	Niet Aangegeven
M	Mannen
V	Vrouwen
T	Totaal
Part Pr	Particuliere Praktijk
Zkh	Ziekenhuis
Ps	Psychiatrische Familliaire
Fam GrPr	Groepspraktijk
Rev	Revalidatie
SD	Standaard Deviatie

S	Specifiek
As	Aspecifiek
Ac	Acuut
Sub	Subacuut
Chr	Chronisch
Ca	Cronbach's alpha
ICC	Intraclass Correlation Coefficient
Ev	Explained variance
r	Correlatie (berekend met)
MIC	Minimal Important Change
SDC	Smallest Detectable Change
K	Kappa

OQ	Oswestry Questionnaire
ODI	Oswestry Disability Index
DRI	Disability Rating Index
RDQ	Roland Disability Questionnaire
EigenV	Eigenvalues
RMQ	Roland Morris Questionnaire
SF-36	Short Form Health Survey 36
DPQ	Dallas Pain Questionnaire
PDI	Pain Disability Index

Tabel 6 – Methodologische bepaald aan de hand van de COSMIN-checklist¹⁸⁻¹⁹ en resultaat bepaald aan de hand van de criteria van Terwee et al.²⁰

	Interne consistentie	Betrouwbaarheid	Meetfout	Inhoudsvaliditeit	Construct- validiteit	Hypothese- testen	Cross-culturele validiteit	Criterium- validiteit	Responsiviteit
Kopec et al.(14)	Excellent +	Good +		Excellent +	Excellent +	Good +	Good +		Good +
Kopec et al.(13)									
Valasek et al.(22)	Excellent +	Good +	Good +		Excellent +		Excellent ?		
Alnahhal A en May S(23)	Poor +	Fair +		Fair +	Fair ?		Poor ?	Poor ?	
De Beer et al.(24)	Poor +	Fair +			Fair ?		Fair ?	Poor ?	
Bendeddouche et al.(25)	Poor +	Fair +			Fair -	Fair +	Good ?		
Kuen Tak Suh et al.(26)	Poor +	Good +			Good ?		Good ?	Poor ?	
Christakou et al.(27)	Good ?	Good +		Poor +	Good +		Fair ?	Poor ?	
Demoulin et al.(28)									Poor ?
Marcelo et al.(29)	Fair +	Fair +				Good +	Good ?		
Sayed Javad Mousavi et al.(30)	Poor +	Poor +		Poor +		Good +	Good ?	Poor ?	
Reneman et al.(31)						Good +		Poor ?	
Yvanes-Thomas et al.(32)	Poor ?			Poor +		Poor ?			
Mens et al.(33)									Poor ?
Schoppink et al.(34)	Poor +	Good +			Good +	Good +			Good +
Bicer et al.(35)	Poor +	Poor +			Poor ?	Poor +			
Meltem Alkan Melikoglu et al.(36)	Poor +	Fair +					Fair ?	Poor ?	

Tabel 7 – Combinatie methodologische kwaliteit en het resultaat aan de hand van de criteria van Terwee et al.²⁰

	Interne consistentie	Betrouwbaarheid	Meetfout	Inhoudsvaliditeit	Construct-validiteit	Hypothese-testen	Cross-culturele validiteit	Criterium-validiteit	Responsiviteit
Kopec et al.(14)	+++	++		+++	+++	++	++		++
Kopec et al.(13)									
Valasek et al.(22)	+++	++	++		+++		?		
Alnahhal A en May S(23)	?	+		+	?		?	?	
De Beer et al.(24)	?	+			?		?	?	
Bendeddouche et al.(25)	?	+			-	+	?		
Kuen Tak Suh et al.(26)	?	++			?		?	?	
Christakou et al.(27)	?	++		?	++		?	?	
Demoulin et al.(28)									?
Marcelo et al.(29)	+	+				++	?		
Sayed Javad Mousavi et al.(30)	?	?		?		++	?	?	
Reneman et al.(31)						++		?	
Yvanes-Thomas et al.(32)	?			?		?			
Mens et al.(33)									?
Schoppink et al.(34)	?	++			++	++			++
Bicer et al.(35)	?	?			?	?			
Meltem Alkan Melikoglu et al.(36)	?	+					?	?	

Tabel 8 – Levels of evidence

Interne consistentie	Betrouwbaarheid	Meetfout	Inhoudsvaliditeit	Construct-validiteit	Hypothese-tester	Cross-culturele validiteit	Criterium-validiteit	Responsiviteit
+++	+++	++	+++	+/-	+++	++	?	+++

Bijlagen

Bijlage 1 – Goedkeuring projectplan

Beoordelingsformulier projectplan

Naam student: Rachél van Eeuwijk

Datum: 12 april 2013

Diversen

- Het projectplan is opgesteld volgens het format **ja / nee**
- Spelling en taalgebruik zijn correct **ja / nee**

Probleemomschrijving en probleemstelling (inleiding)

- De probleemomschrijving is voldoende helder geformuleerd **ja / nee**
- Uit de probleemomschrijving, blijkt de maatschappelijke en paramedische relevantie **ja / nee**
- Op basis van de probleemstelling wordt een concrete en relevante vraagstelling (of meerdere) geformuleerd met eventueel sub vragen **ja / nee**

Doelstelling

De doelstelling is:

- Voldoende helder en concreet geformuleerd **ja / nee**
- Relevant voor een gekozen doelgroep binnen de (paramedische) beroepspraktijk **ja / nee**
- Praktisch uitvoerbaar **ja / nee**
- Haalbaar binnen de tijd **ja / nee**

Projectproduct

Het projectproduct:

- Sluit aan bij de probleemstelling, vraagstelling en doelstelling **ja / nee**
- Is bruikbaar voor de gekozen doelgroep **ja / nee**
- Sluit aan bij de wens van de opdrachtgever **ja / nee**
- De producteisen zijn nauwkeurig omschreven **ja / nee**

Activiteiten/methode

Er wordt voldoende inzicht gegeven in soorten activiteiten en soorten bronnen voor het uitvoeren van het onderzoek en het tot stand komen van het product

ja / nee

Tijdpad

- Het tijdpad geeft voor het project als geheel een globale fasering en tijdbesteding en voor de eerstkomende weken een steeds gedetailleerdere invulling **ja / nee**
- In de tabel zijn belangrijke momenten (typografisch opvallend) vastgelegd, (bv. contactmomenten, inlevermomenten e.d.) **ja / nee**
- In het tijdpad wordt al een globale invulling gegeven van de taakverdeling bij de geplande activiteiten **ja / nee**

Begrote kosten

Er wordt een helder inzicht gegeven in:

- De te verwachten soorten kosten qua geld en uren **ja / nee**
- De verdeling van deze kosten (projectleider, student, opleiding) **ja / nee**

Literatuur

- Gebruikte en geplande literatuur is specifiek en in voldoende omvang genoemd **ja / nee**
- Er wordt verwezen naar relevante en recente literatuur **ja / nee**
- Literatuurverwijzingen, in lopende tekst en in literatuurlijst, worden gegeven volgens de Schrijfwijzer (Wouters 2012) **ja / nee**

Toelichting:

Structuur van de inleiding is aanzienlijk beter geworden. Trechtersvorm is aanwezig en er wordt gewerkt naar een vraagstelling. Diepgang mag bij het uiteindelijke verslag nog beter worden. Er ontbreekt een inzicht in het probleem bij lage rugklachten die chronisch worden en zorgen voor kosten, uitval en werkverzuim. Er wordt daarnaast weinig inzicht gegeven in het mogelijk verklarende mechanisme waarom rugklachten deze problemen geven. En tot slot komt daarom de relevantie voor de fysiotherapeut en het gebruik van meetinstrumenten in de behandeling van deze doelgroep onvoldoende sterk naar voren. De aanleiding mag dus versterkt worden en dat geeft een versterking van de relevantie en het nut voor de praktijk.

De methode is reproduceerbaar geschreven. Op zinsbouw en alinea bouw kan het nog sterker worden. Nu ontbreekt door de schrijfwijze soms de duidelijkheid in wat er nu gedaan is. Door met name soms de volgorde in de tekst van een alinea te veranderen wordt de leesbaarheid aanzienlijk vergroot. Bij de kwaliteitsbeoordeling en levels of evidence moet echt elke stap perfect duidelijk zijn: wat wordt er gedaan, op welke manier en hoe mag het resultaat geïnterpreteerd worden. Inhoud behoeft alleen aanscherping, maar de methode bevat de noodzakelijke elementen. Aanvullende feedback staat in het projectplan.

2^e beoordelaar

Projectplan bevat de voorgeschreven onderdelen die beter zijn uitgewerkt tov het eerste plan.

In de methodesectie best de zoektermen duidelijker naar voor laten komen door ze in een tabel te plaatsen.

Over het geheel van de scriptie moet door de studenten meer aandacht besteed worden aan taal, schrijffouten, referentielijst en lay-out.

Alle punten onder B3.1 tot en met B3.8 moeten met 'ja' beantwoord zijn om een GO voor het project te krijgen. De begeleider bespreekt met de student op welke punten wijzigingen nodig zijn.

ALGEMEEN:

GO

Naam beoordelaar:

Datum + Handtekening

Tjarco Koppenaal

12 april 2013

Steven Onkelinx

Bijlage 2 – The Quebec Back Pain Disability Scale

Naam:

Geb.datum:

Datum:

www.fysiovrageenlijst.nl

Quebec Back Pain Disability Scale

Kopec 1995

Nederlandse vertaling Schoppink EM 1996

Onderstaande vragenlijst gaat over de manier waarop uw rugklachten uw dagelijks leven beïnvloeden. Mensen met rugklachten kunnen moeite hebben met het uitvoeren van sommige dagelijkse activiteiten. Wij willen graag weten of u moeite heeft met het uitvoeren van onderstaande activiteiten **vanwege uw rugklachten**. Voor elke activiteit is er een schaal van 0 tot 5. Wilt u bij iedere activiteit één antwoord kiezen (**geen activiteit overslaan**), en het daarbij behorende cijfer omcirkelen.

Heeft u **vandaag** moeite om de volgende activiteiten uit te voeren vanwege uw rugklachten?

	Totaal geen moeite	Nauwelijks moeite	Enige moeite	Veel moeite	Zeer veel moeite	Niet in staat
Opstaan uit bed	0	1	2	3	4	5
De hele nacht slapen	0	1	2	3	4	5
Omdraaien in bed	0	1	2	3	4	5
Auto rijden	0	1	2	3	4	5
20 tot 30 minuten (achter elkaar) staan	0	1	2	3	4	5
Enkele uren in een stoel zitten	0	1	2	3	4	5
Een trap oplopen	0	1	2	3	4	5
Een klein eindje lopen (300- 400 m)	0	1	2	3	4	5
Enkele kilometers lopen	0	1	2	3	4	5
Naar een hoge plank reiken	0	1	2	3	4	5
Een bal werpen	0	1	2	3	4	5
Een eindje hardlopen (100m)	0	1	2	3	4	5
Iets uit de koelkast pakken	0	1	2	3	4	5
Het bed opmaken	0	1	2	3	4	5
Sokken (of panty) aantrekken	0	1	2	3	4	5
Voorover buigen om bijv. badkuip of W.C. schoon te maken	0	1	2	3	4	5
Een stoel verplaatsen	0	1	2	3	4	5
Een zware deur opentrekken of openduwen	0	1	2	3	4	5
Dragen van 2 tassen met boodschappen	0	1	2	3	4	5

Een zware koffer optillen en dragen	0	1	2	3	4	5
-------------------------------------	---	---	---	---	---	---

Interpretatie:

Per vraag zijn er 6 antwoordcategorieën.

De eerste categorie (0) geeft geen beperkingen aan, de laatste categorie (score 5) betekent de meeste beperkingen.

De totaalscore is de som van de 20 vragen. De eindscore varieert van 0 (geen beperking) tot 100 (volledig beperkt)

Bijlage 3 – Psychometrische eigenschappen¹⁸

COSMIN definitions of domains, measurement properties, and aspects of measurement properties

Term			Definition
Domain	Measurement property	Aspect of a measurement property	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HR-PRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true' [†] differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (<i>for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups</i>) based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'
Responsiveness			The ability of an HR-PRO instrument to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability*			Interpretability is the degree to which one can assign qualitative meaning - that is, clinical or commonly understood connotations – to an instrument's quantitative scores or change in scores.

[†] The word 'true' must be seen in the context of the CTT, which states that any observation is composed of two components – a true score and error associated with the observation. 'True' is the average score that would be obtained if the scale were given an infinite number of times. It refers only to the consistency of the score, and not to its accuracy (ref Streiner & Norman)

* Interpretability is not considered a measurement property, but an important characteristic of a measurement instrument

Bijlage 4 – Zoekstrategie

((((low back pain) OR (backache) OR (back pain) OR (chronic back pain) OR (lumbar spine) OR (non specific back pain) OR (aspecific low back pain) OR (non specific low back pain) OR (aspecific backpain)) AND ((QBPDS) OR (quebec pain disability scale) OR (Quebec)) AND ((measurement properties) OR (psychometric characteristics) OR (validity) OR (content validity) OR (face validity) OR (construct validity) OR (structural validity) OR (hypotheses testing) OR (cross cultural validity) OR (criterion validity) OR (reliability) OR (internal consistency) OR (measurement error) OR (responsiveness))) al

