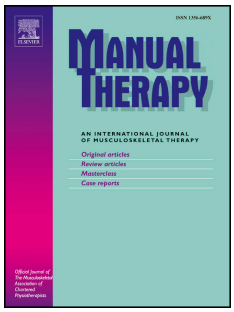
Clinimetric properties of illness perception questionnaire revised (IPQ-R) and brief illness perception questionnaire (Brief IPQ) in patients with musculoskeletal disorders : a systematic review

Reference:

Leysen Marijke, Nijs Jo, Meeus Mira, Van Wilgen C. Paul, Struyf Filip, Vermandel Alexandra, Kuppens An, Roussel Nathalie.- *Clinimetric properties of illness perception questionnaire revised (IPQ-R) and brief illness perception questionnaire (Brief IPQ) in patients with musculoskeletal disorders : a systematic review*
Manual therapy / **Manipulation Association of Chartered Physiotherapists** - ISSN 1356-689X - (2014), p. 1-35
DOI: http://dx.doi.org/doi:10.1016/j.math.2014.05.001

Accepted Manuscript



Clinimetric properties of illness perception questionnaire revised (IPQ-R) and brief illness perception questionnaire (Brief IPQ) in patients with musculoskeletal disorders: A systematic review

Marijke Leysen , Jo Nijs , Mira Meeus , C. Paul van Wilgen , Filip Struyf , Alexandra Vermandel , Kevin Kuppens , Nathalie Roussel

PII: S1356-689X(14)00084-8

DOI: 10.1016/j.math.2014.05.001

Reference: YMATH 1566

To appear in: *Manual Therapy*

Received Date: 19 December 2013

Revised Date: 4 April 2014

Accepted Date: 6 May 2014

Please cite this article as: Leysen M, Nijs J, Meeus M, Paul van Wilgen C, Struyf F, Vermandel A, Kuppens K, Roussel N, Clinimetric properties of illness perception questionnaire revised (IPQ-R) and brief illness perception questionnaire (Brief IPQ) in patients with musculoskeletal disorders: A systematic review, *Manual Therapy* (2014), doi: 10.1016/j.math.2014.05.001.

Clinimetric properties of illness perception questionnaire revised (IPQ-R) and brief illness perception questionnaire (Brief IPQ) in patients with musculoskeletal disorders: A systematic review

Marijke Leysen [ab], Jo Nijs [bce], Mira Meeus [abg], C. Paul van Wilgen [bd], Filip Struyf [abc], Alexandra Vermandel [af], Kevin Kuppens [abc], Nathalie Roussel [abc]

[a] Rehabilitation Sciences and Physiotherapy (REVAKI), Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium

[b] Pain in Motion Research Group (www.paininmotion.be)

[c] Departments of Human Physiology and Physiotherapy, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Belgium

[d] Transcare, Transdisciplinary Pain Management Centre, the Netherlands.

[e] Department of Physical Medicine and Physiotherapy, University Hospital Brussels, Belgium

[f] Department of Urology, University Hospital Antwerp, Belgium

[g] Ghent University Hospital (6K3) (REVAKI), Faculty of Medicine, Ghent University

Corresponding author: Nathalie Roussel;

Address of correspondence and reprints requests to Nathalie Roussel, Universiteit Antwerpen, Campus Drie Eiken D.S.121, Universiteitsplein 1, 2610 Wilrijk, Belgium (e-mail: Nathalie.Roussel@uantwerpen.be; phone: +3232652859; website: www.paininmotion.be

ABSTRACT

Several questionnaires are available to evaluate illness perceptions in patients, such as the illness perception questionnaire revised (IPQ-R) and the brief version (Brief IPQ). This study aims to systematically review the literature concerning the clinimetric properties of the IPQ-R and the Brief IPQ in patients with musculoskeletal pain. The electronic databases Web of Sciences and Pubmed were searched. Studies were included when the clinimetric properties of the IPQ-R or Brief IPQ were assessed in adults with musculoskeletal pain. Methodological quality was determined using the COSMIN checklist. Eight articles were included and evaluated. The methodological quality was good for 3 COSMIN boxes, fair for 11 and poor for 3 boxes. None of the articles obtained an excellent methodological score. The results of this review suggest that the IPQ-R is a reliable questionnaire, except for illness coherence. Internal consistency is good, except for the causal domain. The IPQ-R has good construct validity, but the factor structure is unstable. Hence, the IPQ-R appears to be a useful instrument for assessing illness perceptions, but care must be taken when generalizing the results of adapted versions of the questionnaires. The Brief IPQ shows moderate overall test-retest reliability. No articles examining the validity of the Brief IPQ were found. Further research should therefore focus on the content and criterion validity of the IPQ-R and the clinimetric properties of the Brief IPQ.

Key Words: musculoskeletal pain (MeSH), epidemiologic methods (MeSH), Questionnaires (MeSH), Perception (MeSH)

INTRODUCTION

Recent guidelines advise health care personal to evaluate and treat patients with musculoskeletal pain from a biopsychosocial perspective (Airaksinen et al., 2006; Tulder et al., 2006). In both medical and psychological literature, Leventhal's Common Sense Model (CSM) is often used as a theoretical framework for the evaluation and treatment of patients (Leventhal et al., 2003). According to this model, patients develop cognitions about their illness, based on former experiences, interpretation of symptoms and provided information. These cognitions are often referred to as illness perceptions.

These illness perceptions have been studied in several pathologies such as cardiovascular disorders (Schoormans et al., 2013), respiratory disorders (Kaptein et al., 2011) and musculoskeletal disorders e.g. fibromyalgia (van Wilgen et al., 2008), sports injuries (van Wilgen et al., 2010; Larmer et al., 2011), low back pain (Foster et al., 2008; van Wilgen et al., 2012), chronic fatigue syndrome and rheumatoid arthritis (Moss-Morris & Chalder, 2003). Especially when there is no clear diagnosis (e.g. no bodily cause of pain or medically unexplained symptoms), patients form their own interpretation of symptoms to explain the disorder. Illness perceptions will determine the patient's coping strategy. (Sumathipala et al., 2008). Some patients will typically develop negative beliefs about their illness (Stenner et al., 2000). These negative illness perceptions can include believing that the problem will last long, relating all symptoms to their illness or having weak beliefs about self-control and low confidence in performing activities despite their pain (Foster et al., 2008). In a large prospective study with acute, sub-acute and chronic low back pain patients, negative illness perceptions were better predictors of disability at 6 months than fear avoidance, catastrophizing or depression (Foster et al., 2008; Foster et al., 2010). In chronic pain patients, negative illness perceptions are associated with maladaptive illness

behaviour, dysfunction, poor treatment adherence and treatment outcome (Spinhoven et al., 2004; Edwards et al., 2006).

In order to evaluate illness perceptions, the Illness Perceptions Questionnaire (IPQ) (Weinman et al., 1996) was developed. Subsequent to publication of the IPQ, further evolvement of the tool was undertaken, leading to the creation of the IPQ-Revised (IPQ-R) (Moss-Morris et al., 2002). The IPQ-R measures 9 dimensions of illness perceptions and consists of 3 domains. In the first domain, called illness identity, the perceived symptoms and their possible relation to the illness are evaluated. The second domain, the beliefs domain, covers 7 dimensions: the acute/chronic timeline as well as the cyclical character of the illness represent the first and second dimension. Consequences, as the third dimension, include perceived short- and long-term effects on physical, psychological and social functioning. Controllability and curability refers to the extent to which a condition is perceived to be controllable or curable, while emotional representations, the sixth dimension, represent the emotions experienced as a result of their illness. Finally, illness coherence reflects an individual's understanding of their condition. For each dimension, responders rate their level of agreement on a five-point Likert scale, ranging from 'strongly disagree' to 'strongly agree'. The third domain lists 18 possible causes to which individuals attribute their condition, the degree to which individuals perceive themselves as responsible for the illness, as well as the responsibility individuals take for curing themselves. Again, patients rate their level of agreement on a five-point Likert scale, ranging from 'strongly disagree' to 'strongly agree' (Hill et al., 2007).

In 2006 Broadbent et al. constructed a briefer version from the IPQ-R, which is referred to as the Brief IPQ (Broadbent et al., 2006). The aim was to construct a very short and simple

measure of illness perceptions for clinical use and to provide an alternative for the 5-point Likert scale approach. The Brief IPQ is an eight-item instrument that measures the cognitive perceptions with respect to an illness on an ordinal scale (0–10). Eight areas are examined: consequences (item 1), timeline (item 2), personal control (item 3), treatment control (item 4), identity for describing the condition and symptoms (item 5), coherence (item 7), and concern and emotions (items 6 and 8). The maximal score on the Brief IPQ is 80, where higher scores reflect more negative perceptions.

Since the IPQ, IPQ-R and Brief IPQ are general questionnaires, researchers are allowed to substitute the term 'illness' with the name of the condition they are investigating (Weinman et al., 1996; Hill et al., 2007). Moreover, researchers should feel free to modify the causal and identity scales in order to suit particular illnesses, cultural settings or populations (Moss-Morris et al., 2002).

Because illness perceptions are measured in a variety of disorders, the questionnaires can be adapted in function of each condition, such as fibromyalgia (Van Wilgen et al., 2008) and hand injury (Chan et al., 2009). However, information regarding the clinimetric properties of the (adapted versions of the) IPQ-R and Brief IPQ is lacking. The clinimetric approach is directed at the development of instruments to measure multiple constructs with a single index (Fayers & Hand, 2002), which is often the case in clinical practice (Vet et al., 2003). It is associated with rating scales that are used to describe or measure symptoms, physical signs and other distinctly clinical phenomena (Feinstein, 1983; 1987). A summary of the quality of the studies that have investigated IPQ-R or Brief IPQ will give perspective on how these articles can assist in directing approaches in clinical practice. Therefore, the aim of the present literature overview was to systematically review

the clinimetric properties of the IPQ-R and the Brief IPQ in patients with musculoskeletal disorders.

METHODS

Search strategy

Full details of the search strategy can be found in the addendum. In brief, alongside adherence to the PRISMA guidelines, the PICOS model was used to list three groups of keywords: (P) patients with musculoskeletal pain, (I) IPQ-R or Brief IPQ and (O) clinimetric properties. No limits were added.

Methodological quality of the included articles

The methodological quality of the included articles was reviewed using the COSMIN checklist with 4-point rating scale, representing excellent, good, fair and poor methodological quality (Mokkink et al., 2010a). The COSMIN checklist is a standardized tool for assessing the methodological quality of studies on measurement properties. It contains a generalizability box and 9 separate boxes, each dealing with one measurement property, with 5-18 items per box about the design and statistical methods. This incorporates potential bias of individual studies. Two researchers independently scored the selected studies. After reviewing the articles, the results of both researchers were compared and differences were discussed until consensus was obtained. Subsequently, a methodological quality score per box is obtained by taking the lowest rating of any item in a box (Terwee et al., 2012). The results were evaluated using the quality criteria for measurement properties of health status questionnaires described by Terwee et al. (Terwee et al., 2007).

Outcome measurements

For the purpose of this study reliability was analysed in terms of internal consistency and test-retest reliability (Lohr et al., 1996). ***Internal consistency*** is a measure of the extent to

104 which items in a subscale are correlated, thus measuring the same concept (Terwee et al.,
105 2007). To express the internal consistency of the different items in the domains of the IPQ-
106 R, Cronbach's alphas can be calculated. A Cronbach's alpha above 0.80 is considered to be
107 acceptable (Dijkers et al., 2002). ***Reproducibility or test–retest reliability*** over a period of
108 time can be calculated using an intraclass correlation coefficient (ICC), a weighted kappa
109 or Pearson correlation. To interpret the kappa statistics, values above 0.60 are considered
110 substantial agreement (Landis & Koch, 1977). For ICC, the threshold value of 0.75 for
111 good reliability was used (Portney & Watkins, 2000). For Pearson's correlations, critical
112 values are subject to the number of correlated items (Fisher & Yates, 1974; Portney &
113 Watkins, 2000).
114
115 Validity will be presented as construct-, content- and criterion-related validity (Lohr et al.,
116 1996; Mokkink et al., 2010b). Construct validity refers to the ability of an instrument to
117 measure a concept or construct. Convergence, discrimination, factor analysis, hypothesis
118 testing and known groups method are procedures to gather information about the construct
119 (Portney & Watkins, 2000). According to the COSMIN taxonomy, ***construct validity*** is
120 divided into hypotheses testing, structural validity and cross-cultural validity (Mokkink et
121 al., 2010b). ***Content validity*** is the degree to which the content of an instrument is an
122 adequate reflection of the construct to be measured (Mokkink et al., 2010b). ***Concurrent***
123 ***validity*** is an aspect of criterion validity and measures the agreement between the results
124 obtained by the IPQ-R and the results obtained by another instrument within the same
125 population at the same time.

RESULTS

Search strategy

The initial search strategy identified 75 unique abstracts from the PubMed and Web of Science databases. Two articles were included by hand search. Based on the inclusion criteria, 65 abstracts were excluded. Figure 1 presents a flowchart of the search strategy. A detailed overview of the included articles is presented in Table 1. The full text version of all papers that met the inclusion criteria was retrieved for quality assessment and data extraction.

Eight studies were included (Table 1) and scored for their methodological quality (Table 2). The methodological quality of the different items of the studies varied from good (van Ittersum et al., 2009; Nicholls et al., 2013) to fair (Moss-Morris et al., 2002; van Wilgen et al., 2008; Chan et al., 2009; Glattacker et al., 2009; Albert et al., 2013; Hallegraeff et al., 2013) to poor (Chan et al., 2009; Hallegraeff et al., 2013).

Seven studies analysed the clinimetric properties of the IPQ-R (Moss-Morris et al., 2002; van Wilgen et al., 2008; Chan et al., 2009; Glattacker et al., 2009; van Ittersum et al., 2009; Albert et al., 2013; Nicholls et al., 2013). Only one study administered the Brief IPQ (Hallegraeff et al., 2013). To target a specific patient population, the IPQ-R was adapted in each article. These changes are presented in table 3.

Methodological quality of the included articles

The assessment of methodological quality of the included articles is shown in Table 2. Agreement between the two researchers was 83%. Consensus was obtained on all items. The answers on the generalizability box of the COSMIN checklist of each article are presented in Table 1. The items with poor methodological quality will not be further discussed.

Reliability

The Pearson correlations for test-retest reliability varied between 0.50 and 0.87 for the beliefs domain, except for cyclical timeline, where a lower correlation was observed (0.35). For illness identity and the causal domain, the correlations varied between 0.24-0.57 and 0.53-0.85, respectively (Table 4). The ICC varied between 0.55 and 0.87 (Glattacker et al., 2009). The test-retest reliability of the Brief IPQ over a one-week period was acceptable (ICC 0.72, 95% CI:0.53-0.82) (Hallegraeff et al., 2013).

Internal consistency of the beliefs domain of the IPQ-R among different patient populations was satisfactory, ranging between 0.51 and 0.87 (table 4). Of the sub-domains within the causal domain, only psychological attributions presented an alpha ≥0.82. The sub-domain 'accident or chance' showed a very low internal consistency. No studies examined the internal consistency of the Brief IPQ.

The measurement error was evaluated in the Brief IPQ only (Hallegraeff et al., 2013). Limits of agreement ranged from -25.3 to 17.1. No systematic trend was visible in the Bland-Altman plot. The standard error of the mean was 1.17 and the smallest detectable change was 42, compared to a maximum score of 80 (Hallegraeff et al., 2013).

Validity

Three articles tested different hypotheses on the construct validity of the IPQ-R (van Wilgen et al., 2008; Chan et al., 2009; Albert et al., 2013) (table 5).

Three studies established structural validity of the IPQ-R as an aspect of construct validity (Moss-Morris et al., 2002; van Ittersum et al., 2009; Nicholls et al., 2013). Moss-Morris et al. used an independent samples t-test to explore known group validity within acute versus chronic patients (Moss-Morris et al., 2002). Chronic pain patients were significantly different from acute patients on all dimensions of the IPQ-R ($p<.001$), except for risk factor attributions ($p<.01$).

Two studies performed a factor analysis: one study used both an exploratory and confirmatory factor analysis (Nicholls et al., 2013) while the other used confirmatory factor analysis only (van Ittersum et al., 2009). Results are presented in table 6.

No studies assessed the validity of the Brief IPQ.

DISCUSSION

The results of this review suggest that the IPQ-R is a reliable questionnaire, except for the illness coherence, with good internal consistency, except for the causal domain. The IPQ-R demonstrates good construct validity, but the factor structure is unstable. The Brief IPQ shows moderate overall test-retest reliability. There is a lack of articles studying the validity of the Brief IPQ used in musculoskeletal conditions.

Methodological quality of the included articles

The methodological quality of the different items of the included studies ranged from poor (N=3) to good (N=3). Methodological problems included an insufficient sample size, selection bias (e.g. convenience sampling), lack of description of handling with missing data or the lack of a priori formulated hypotheses. The items with poor methodological quality were eliminated from this literature review, since the precision of the results in these articles is doubtful. None of the selected articles obtained an excellent methodological score, implying that all included studies had methodological flaws.

Test-retest reliability

The results of the present study suggest that test-retest reliability of the IPQ-R and Brief IPQ is acceptable in the observed patient populations. Two out of three articles only calculated Pearson correlations (Moss-Morris et al., 2002; van Wilgen et al., 2008). Pearson correlation coefficients are less accurate to measure reliability than ICC, because systematic differences are not taken into account (Streiner & Norman, 2003). The moderate ICC in one study evaluating orthopaedic patients (Glattacker et al., 2009) suggests that further research is necessary to improve the test-retest reliability.

To measure test-retest reliability, it is important to ensure the stability of the illness perceptions of the patients within the time frame. Therefore, it must be questioned whether illness perceptions remain stable over time if symptoms are fluctuating. The differences in test-retest reliability across studies might be explained by the time interval between the consecutive measurements, which was much longer (6 months) in the study by Moss-Morris et al. (2002) compared to the 3 weeks (van Wilgen et al., 2008) or 4 days-time interval (Glattacker et al., 2009) in other studies.

The single study examining test-retest reliability of the Brief IPQ (Hallegraeff et al., 2013) suggests an acceptable test-retest reliability. In that study, the smallest detectable change was 42, which means that a change in the Brief IPQ overall score must exceed a value of 42 in order to reflect a true difference between test and retest scores. With a maximum overall score of 80, it can be suggested that the Brief IPQ is not suitable for detecting real individual changes. However, it can also be questioned if an overall score can be calculated in the Brief IPQ, for each question measures a different dimension of illness perceptions.

Internal consistency

The Cronbach's alphas for the beliefs domain of the IPQ-R showed good internal consistency (0.75-0.82). Two studies had lower scores on some of the subscales (van Wilgen et al., 2008; Albert et al., 2013). This may be related to the smaller sample size in comparison to the third study (van Ittersum et al., 2009). The latter had a good methodological quality. Furthermore, Albert et al. created a virtually new questionnaire by adding 26 items to the beliefs domain, making it hazardous to compare.

Illness identity consists of disparate symptoms, such as pain, fatigue, nausea and stiff joints. Some symptoms may be more relevant to particular illnesses than other symptoms (e.g. stiff

joints is common for fibromyalgia, but less common for low back pain (van Wilgen et al., 2008; van Ittersum et al., 2009; van Wilgen et al., 2012)). Therefore, the internal consistency of this scale is less relevant than in the other subscales. Symptoms and their frequency are presented as a checklist, therefore they are not supposed to measure a certain construct.

Within the causal domain, internal consistency is very good for the psychological attribution (0.82-0.90). The Cronbach's alphas for the other subscales in the causal domain are moderate (0.47-0.62), except for accident or chance, which are very low (0.00-0.14). By analogy with symptoms, causes can be very diverse between different pathologies. Again, some causes may be more relevant to particular illnesses than other (e.g. 'hereditary' is often cited as a cause in fibromyalgia, whereas it is not mentioned frequently by patients with low back pain (van Ittersum et al., 2009; van Wilgen et al., 2012)). This is supported by the unstable factor structure of the causal domain (Nicholls et al., 2013). It is suggested that a satisfactory factor solution could be found if the list of causal items is sufficiently modified to relate more clearly to musculoskeletal pain patients, by removing items or including new items (Nicholls et al., 2013).

Construct validity

The significant differences in test results between acute and chronic patients on all dimensions reflect clear known group validity (Moss-Morris et al., 2002). In patients with fibromyalgia, catastrophizing showed a negative relationship with illness coherence and a positive association with emotional representations and cyclical timeline (van Wilgen et al., 2008), suggesting that patients who do not have a clear understanding of their situation have the tendency to catastrophize. This indicates that education and information play a key role in the treatment process.

However, pain intensity proves to be unrelated to the subscales of the IPQ-R in patients with musculoskeletal disorders which are absent from work (Albert et al., 2013). In this particular patient population, pain intensity might be of less importance compared to functional limitations. This is reflected in the fact that a high illness identity endorsed by participants is more strongly associated with psychological distress than with pain intensity (Albert et al., 2013).

Structural validity of the IPQ-R was assessed in two articles with good methodological quality (van Ittersum et al., 2009; Nicholls et al., 2013). The factor structure of the beliefs domain as suggested in the original IPQ-R (Moss-Morris et al., 2002) could not be completely affirmed, nor could the causal domain. The factor structure of the original IPQ-R was calculated in 711 patients with a variety of disorders, such as rheumatoid arthritis, type II diabetes, asthma, chronic pain, acute pain, multiple sclerosis, myocardial infarction and HIV (Moss-Morris et al., 2002). Comparison of the clinimetric properties of the questionnaires should ideally be calculated in a homogeneous patient group. For the causal domain, this may be even more important, as attributions are probably disease specific. Another potential reason why the seven-factor model of the beliefs domain does not generally provide a good fit could be related to the presentation of the items. A mixture of positively and negatively worded items may confuse some respondents. There is some evidence that positively worded items are more highly correlated with each other than negatively worded items, and vice-versa (Nicholls et al., 2013).

There is a lack of studies with good methodological quality examining the measurement error and predictive validity of the IPQ-R. This would favour the use of this type of questionnaires

in clinical practice. Furthermore, no studies with good methodological quality examined the criterion validity or content validity of the IPQ-R. Concerning the Brief IPQ, only one article met the inclusion criteria (Hallegraeff et al., 2013). This suggests the need of future research to study the clinimetric properties of the Brief IPQ within musculoskeletal patients more closely.

Study limitations

Since the aim of present study was to identify clinimetric properties of the IPQ-R or Brief IPQ within musculoskeletal patients, the results of this review are only applicable to the included populations. Furthermore, it is uncertain whether clinimetric qualities of translated versions can be generalized to the original version. The results of the present study are therefore only applicable to the questionnaire and language used in a particular study (table 1). It has to be noted that none of the included articles had an excellent score on the COSMIN checklist for methodological quality. Therefore the results of the articles should not be rejected, but one must be attentive to the interpretation. As the first and third domain (i.e. illness identity and causal domain) are adjustable by researchers, care must be taken when comparing or generalizing the results of adapted questionnaires. In the last question of the IPQ-R, patients are asked to describe the three most important causes for their illness. With this open-ended format, a wealth of information is obtained from the patients, but due to the design it is very difficult to objectify, measure or compare these results. Nevertheless, the latter is very interesting for clinical practice, given the fact that negative illness perceptions influence behaviour (Leventhal et al., 2003) and predict disability in low back pain patients (Foster et al., 2008; Foster et al., 2010).

Conclusion

The results of the present systematic review confirm that the IPQ-R is an appropriate instrument to explore illness beliefs in patients with musculoskeletal disorders. Since the questionnaire can be adapted to target a specific patient population, the factor structure remains a delicate issue. Further research should be conducted to optimise the clinimetric properties of the Brief IPQ in patients with musculoskeletal disorders.

REFERENCES

Airaksinen, O, Brox, J, Cedraschi, C, Hildebrandt, J, Klaber-Moffett, J, Kovacs, F, Mannion, A, Reis, S, Staal, J, Ursin, H, Zanoli, G. Chapter 4: European guidelines for the management of chronic nonspecific low back pain Eur Spine J. 2006; (15): Suppl 2:S192-300.

Albert, V, Coutu, MF, Durand, MJ. Internal consistency and construct validity of the Revised Illness Perception Questionnaire adapted for work disability following a musculoskeletal disorder. Disabil Rehabil 2013; 35(7): 557-65.

Broadbent, E, Petrie, KJ, Main, J, Weinman, J. The brief illness perception questionnaire. J Psychosom Res 2006; 60(6): 631-7.

Chan, JCY, Ong, JCY, Avalos, G, Regan, PJ, McCann, J, Groarke, A, Kelly, JL. Illness representations in patients with hand injury. Journal of plastic, reconstructuve and aesthetic surgery 2009; 62: 927-932.

Dijkers, MP, Kropp, GC, Esper, RM, Yavuzer, G, Cullen, N, Bakdalieh, Y. Reporting on reliability and validity of outcome measures in medical rehabilitation research. Disabil Rehabil 2002; 24(16): 819-27.

Edwards, RR, Bingham, CO, 3rd, Bathon, J, Haythornthwaite, JA. Catastrophizing and pain in arthritis, fibromyalgia, and other rheumatic diseases. Arthritis Rheum 2006; 55(2): 325-32.

Fayers, PM, Hand, DJ. Causal variables, indicator variables and measurement scales: an example from quality of life. Journal of the Royal Statistical Society: Series A (Statistics in Society) 2002; 165(2).

Feinstein, AR. An additional science for clinical medicine: IV. The development of clinimetrics. Ann Intern Med 1983; 99: 843-848.

Feinstein, AR. Clinimetrics New Haven, Connecticut: Yale University Press, 1987;

Fisher, RA, Yates, F. Statistical tables for biological, agricultural and medical research, 6th revised and enlarged ed. UK: Longman, 1974;

Foster, NE, Bishop, A, Thomas, E, Main, C, Horne, R, Weinman, J, Hay, E. Illness perceptions of low back pain patients in primary care: what are they, do they change and are they associated with outcome? Pain 2008; 136(1-2): 177-87.

Foster, NE, Thomas, E, Bishop, A, Dunn, KM, Main, CJ. Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. Pain 2010; 148(3): 398-406.

Glattacker, M, Bengel, J, W.H., J. German version of the Illness perception questionnaire-revised. Zeitschrift für Gesundheitspsychologie 2009; 17(4): 158-169.

Hallegraeff, JM, van der Schans, CP, Krijnen, WP, de Greef, MH. Measurement of acute nonspecific low back pain perception in primary care physical therapy: reliability and validity of the brief illness perception questionnaire. BMC Musculoskelet Disord 2013; 14: 53.

Hill, S, Dziedzic, K, Thomas, E, Baker, SR, Croft, P. The illness perceptions associated with health and behavioural outcomes in people with musculoskeletal hand problems: findings from the North Staffordshire Osteoarthritis Project (NorStOP). Rheumatology (Oxford) 2007; 46(6): 944-51.

Kaptein, AA, Yamaoka, K, Snoei, L, Kobayashi, K, Uchida, Y, van der Kloot, WA, Tabei, T, Kleijn, WC, Koster, M, Wijnands, G, Kaajan, H, Tran, T, Inoue, K, van Klink, R, van Dooren-Coppens, E, Dik, H, Hayashi, F, Willems, L, Annema-Schmidt, D, Annema, J, van der Maat, B, van Kralingen, K, Meirink, C, Ogoshi, K, Aaronson, N, Nortier, H, Rabe, K. Illness perceptions and quality of life in Japanese and Dutch patients with non-small-cell lung cancer. Lung Cancer 2011; 72(3): 384-90.

366 Landis, JR, Koch, GG. The measurement of observer agreement for categorical data.
367 Biometrics 1977; 33(1): 159-74.
368 Larmer, PJ, McNair, PJ, Smythe, L, Williams, M. Ankle sprains: patient perceptions of
369 function and performance of physical tasks. A mixed methods approach. Disabil
370 Rehabil 2011; 33(23-24): 2299-304.
371 Leventhal, H, Brissette, I, Leventhal, E. The common-sense model of self-regulation of health
372 and illness., In: Cameron LD, Leventhal H, editors The self-regulation of health and
373 illness behavior New York: Routledge, 2003.
374 Lohr, KN, Aaronson, NK, Alonso, J, Burnam, MA, Patrick, DL, Perrin, EB, Roberts, JS.
375 Evaluating quality-of-life and health status instruments: development of scientific
376 review criteria. Clin Ther 1996; 18(5): 979-92.
377 Mokkink, LB, Terwee, CB, Patrick, DL, Alonso, J, Stratford, PW, Knol, DL, Bouter, LM, de
378 Vet, HC. The COSMIN checklist for assessing the methodological quality of studies
379 on measurement properties of health status measurement instruments: an international
380 Delphi study. Qual Life Res 2010a; 19(4): 539-49.
381 Mokkink, LB, Terwee, CB, Patrick, DL, Alonso, J, Stratford, PW, Knol, DL, Bouter, LM, de
382 Vet, HC. The COSMIN study reached international consensus on taxonomy,
383 terminology, and definitions of measurement properties for health-related patient-
384 reported outcomes. J Clin Epidemiol 2010b; 63(7): 737-45.
385 Moss-Morris, R, Chalder, T. Illness perceptions and levels of disability in patients with
386 chronic fatigue syndrome and rheumatoid arthritis. J Psychosom Res 2003; 55(4):
387 305-8.
388 Moss-Morris, R, Weinman, J, Petrie, KJ, Horne, R, Cameron, LD, Buick, D. The revised
389 illness perception questionnaire (IPQ-R). Psychology and Health 2002; 17(1): 1-16.
390 Nicholls, EE, Hill, S, Foster, NE. Musculoskeletal pain illness perceptions: factor structure of
391 the Illness Perceptions Questionnaire-Revised. Psychol Health 2013; 28(1): 84-102.
392 Portney, LG, Watkins, MP. Foundations of clinical research: applications to practice, 2 ed.:
393 Prentice-Hall, 2000; p759
394 Schoormans, D, Mulder, BJ, van Melle, JP, Pieper, PG, van Dijk, AP, Sieswerda, GT,
395 Hulsbergen-Zwarts, MS, Plokker, TH, Brunninkhuis, LG, Vliegen, HW, Sprangers,
396 MA. Illness perceptions of adults with congenital heart disease and their predictive
397 value for quality of life two years later. Eur J Cardiovasc Nurs 2013.
398 Spinhoven, P, Ter Kuile, M, Kole-Snijders, AM, Hutten Mansfeld, M, Den Ouden, DJ,
399 Vlaeyen, JW. Catastrophizing and internal pain control as mediators of outcome in the
400 multidisciplinary treatment of chronic low back pain. Eur J Pain 2004; 8(3): 211-9.
401 Stenner, PH, Dancey, CP, Watts, S. The understanding of their illness amongst people with
402 irritable bowel syndrome: a Q methodological study. Soc Sci Med 2000; 51(3): 439-
403 52.
404 Streiner, DL, Norman, GR. Health measurements scales. A practical guide to their
405 development and use. New York: Oxford University Press, 2003;
406 Sumathipala, A, Siribaddana, S, Hewege, S, Sumathipala, K, Prince, M, Mann, A.
407 Understanding the explanatory model of the patient on their medically unexplained
408 symptoms and its implication on treatment development research: a Sri Lanka Study.
409 BMC Psychiatry 2008; 8: 54.
410 Terwee, CB, Bot, SD, de Boer, MR, van der Windt, DA, Knol, DL, Dekker, J, Bouter, LM, de
411 Vet, HC. Quality criteria were proposed for measurement properties of health status
412 questionnaires. J Clin Epidemiol 2007; 60(1): 34-42.
413 Terwee, CB, Mokkink, LB, Knol, DL, Ostelo, RW, Bouter, LM, de Vet, HC. Rating the
414 methodological quality in systematic reviews of studies on measurement properties: a
415 scoring system for the COSMIN checklist. Qual Life Res 2012; 21(4): 651-7.

Tulder, Mv, Becker, A, Bekkering, T, Breen, A, Real, Md, Hutchinson, A, Koes, B, Laerum, E, Malmivaara, A. Chapter 3: European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J. 2006; (15): Suppl 2:S169-91.

van Ittersum, MW, van Wilgen, CP, Hilberdink, WK, Groothoff, JW, van der Schans, CP. Illness perceptions in patients with fibromyalgia. Patient Educ Couns 2009; 74(1): 53-60.

van Wilgen, CP, Kaptein, AA, Brink, MS. Illness perceptions and mood states are associated with injury-related outcomes in athletes. Disabil Rehabil 2010; 32(19): 1576-85.

van Wilgen, CP, van Ittersum, MW, Kaptein, AA. Do illness perceptions of people with chronic low back pain differ from people without chronic low back pain? Physiotherapy 2012; 99(1): 27-32.

van Wilgen, CP, van Ittersum, MW, Kaptein, AA, van Wijhe, M. Illness perceptions in patients with fibromyalgia and their relationship to quality of life and catastrophizing. Arthritis Rheum 2008; 58(11): 3618-26.

Vet, HCWd, Terwee, CB, Bouter, LM. Clinimetrics and psychometrics: two sides of the same coin. Journal of Clinical Epidemiology 2003; 56: 1146–1147.

Weinman, J, Petrie, KJ, Moss-Morris, R. The illness perception questionnaire: a new method for assessing the cognitive representation of illness. Psychology and Health 1996; 11.

ADDENDUM: SEARCH STRATEGY

Using the PRISMA guidelines (Liberati et al., 2009), a systematic search strategy was performed via the electronic databases PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and Web of Science (http://isiwebofknowledge.com) from their inception up till April 2013. Several key word combinations were made to ensure that no relevant articles were missed. Using the PICOS (Population, Intervention, Comparison, Outcome, Study design) model, three groups of keywords were listed: (P) patients with musculoskeletal pain, (I) IPQ-R or Brief IPQ and (O) clinimetric properties. The keywords from the three groups were combined. No limits were used during the search strategy.

To identify relevant articles, all titles and/or abstracts of the selected articles were screened for inclusion. Articles were eligible for this review if they fulfilled the following criteria: 1) the study consisted of a prospective, population-based cohort or case-control design investigating the clinimetric properties of the IPQ-R or Brief IPQ, 2) subjects of the study were adult patients (18 years and older) with musculoskeletal complaints, and 3) the studies were written in English, German, French or Dutch. Articles were excluded from this systematic literature research if they were letters to the editor or reviews, abstracts, hypotheses or papers without scientific data or if they included only healthy subjects. In case of doubt of the eligibility of the article based on the content of the title and abstract, the full text version was retrieved and evaluated against the selection criteria as mentioned above. Literature was screened by the first and last author.

Liberati, A, Altman, DG, Tetzlaff, J, Mulrow, C, Gotzsche, PC, Ioannidis, JP, Clarke, M, Devereaux, PJ, Kleijnen, J, Moher, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009; 6(7): e1000100.

There are no conflicts of interest to report. The study was funded by the research council of Artesis University College, Antwerp, Belgium.

Table 1: Included studies

Author **Country**	**Patient Population** **Setting**	**N**	**Mean age** **(%male)**	**Questionnaire**	**Clinimetric Outcome**
Moss-Morris (2002) **New Zealand**	RA hospital outpatient clinics Chronic pain (> 3months) hospital based chronic pain clinics Acute pain (< 6 weeks) private PT practice	76 63 35	59,0 (24%) 53,9 (41%) 35,7 (57%)	IPQ-R (English) PANAS Ambulatory Index SIP Fatigue Severity Scale	Test-retest reliability (RA) - Pearson's correlations Construct validity: Known group method (acute vs chronic) - independent samples t-test
Van Ittersum (2009) **The Netherlands**	FM PT treatment centre	196	49 (12%)	IPQ-R-FM (Dutch) VAS IPQ-R (English)	Internal consistency - Cronbach's α Construct validity: structural validity - MGM (CFA)
Van Wilgen (2008) **The Netherlands**	FM Dutch FM patient association	51	44 (8%)	IPQ-R-FM (Dutch) with 8 FM specific causes FIQ PCS	Internal consistency - Cronbach's α Test-retest reliability - Pearson's correlations Construct validity: hypotheses testing: Correlation with catastrophizing Pearson's correlations

Albert (2013) **Canada**	Musculoskeletal disorder with absence from work 3m-1y	43	41 (46,5%)	IPQ-R-WD (French) -> with new items TSK PCS PDI-14 PDI SERWS Pain beliefs and perceptions inventory Implicit models of illness questionnaire VAS	Internal consistency - Cronbach's alpha Construct validity: hypotheses testing - multiple regression analyses and Pearson correlation
Chan (2009) **Ireland**	Acute (1) hand injury, surgery required hospital	57	38,2 (21%)	IPQ-R-injury version DASH HISS	Internal consistency - Cronbach's alpha Construct validity: hypotheses testing: Correlation with objective severity and subjective disability - Pearson
Nicholls (2013) **UK**	Knee pain (OA) Hand pain Non-specific LBP	393 2113 1591	63,5 (38%) 65,4 (37%) 43,9 (41%)	IPQ-R	Construct validity: structural validity: CFA (5 domains) - Goodness of fit - Chi², goodness of fit index, Parsimony adjusted GFI, comparative fit index, RMSEA EFA (causes) - PCA with varimax rotation

Glattacker (2009) **Germany**	Orthopaedic 2 rehabilitation clinics	45	45,5 (33,3%)	IPQ-R (German) HADS-D	Test-retest reliability - ICC, Pearson correlation coefficient
Hallegraeff (2013) **The Netherlands**	Acute non-specific LBP < 6 weeks physical therapy providers	84	42 (43%)	Brief IPQ (Dutch) SF36 Health Survey	Internal consistency - Cronbach's alpha Test-Retest reliability - ICC Measurement error - Limits of agreement, Bland Altman Plot Criterion validity: Concurrent validity (Mental Health component of SF-36) - ICC and Pearson correlations

Legend: OA = osteoarthritis, RA = rheumatoid arthritis, FM = fibromyalgia

SIP = sickness impact profile, PANAS = Positive affect and negative affect scale, VAS = visual analogue scale, FIQ = fibromyalgia impact questionnaire, PCS = pain catastrophizing scale, TSK = Tampa scale of kinesiophobia, PDI = pain disability index, PDI-14 = psychological distress index, SERWS = self-efficacy with regard to work capacity, DASH = disabilities of the arm, shoulder and hand, HISS = Hand injury severity score, HADS = Hospital Anxiety and Depression Scale, SF36 Health Survey = Short Form 36 Health Survey

MGM = multiple group method, CFA = confirmatory factor analysis, EFA = exploratory factor analysis, RMSEA = root mean square error of approximation, ICC = intraclass correlation, PCA = principal component analysis

Table 2: Assessment of methodological quality

Author	COSMIN box	Agreement	Clinimetric Outcome	Lowest score
Moss-Morris (2002)	E	5/6	Construct validity: Structural validity: known group method (acute vs chronic) - independent samples t-test	Fair
	B	10/11	Test-retest reliability (RA) - Pearson's correlations	Fair
Van Ittersum	A	9/9	Internal consistency - Cronbach's α	Good
(2009)	E	6/6	Construct validity: structural validity - CFA (MGM)	Good
Van Wilgen (2008)	A	8/9	Internal consistency - Cronbach's α	Fair
	B	9/11	Test-retest reliability - Pearson's correlations	Fair
	F	10/10	Construct validity: Hypotheses testing: Correlation with catastrophizing - Pearson's correlations	Fair
Albert (2013)	A	9/9	Internal consistency - Cronbach's alpha	Fair
	F	10/10	Construct validity: hypotheses testing - Pearson correlation matrix, multiple regression analysis	Fair
Chan (2009)	A	9/9	Internal consistency - Cronbach's alpha	Poor
	F	10/10	Construct validity: Hypotheses testing: Correlation with objective severity and subjective disability - Pearson	Fair
Nicholls (2013)	E	6/6	Construct validity: Structural validity: CFA (5 domains) - Goodness of fit - Chi², goodness of fit index, Parsimony adjusted GFI, comparative fit index, RMSEA EFA (causes) - PCA with varimax rotation	Good

Glattacker (2009)	B	9/11	Test-retest reliability (Orth) - ICC, Pearson correlation coefficient	Fair
Hallegraeff (2013)	A	6/9	Internal consistency - Cronbach's alpha	Poor
	B	5/11	Test-Retest reliability - ICC	Fair
	C	5/11	Measurement error - Limits of agreement, Bland Altman Plot	Fair
	H	3/6	Criterion validity: Concurrent validity (Mental Health component of SF-36) - ICC and Pearson correlations	Poor

Legend: MGM= multigroup method, CFA = Confirmatory factor analysis, EFA = exploratory factor analysis, PCA = principal component analysis, GFI = goodness of fit index, CFI = comparative fit index, RMSEA = root mean square error of approximation, ICC = intraclass correlation, Orth = orthopaedic, RA= rheumatoid arthritis

A= internal reliability, B= reliability, C= measurement error, D= content validity, E=structural validity, F=hypotheses testing, G=cross cultural validity, H=criterion validity

Table 3: Adaptations of the IPQ-R in the included studies

Author (Year) **Questionnaire**	**"My illness"** **was changed into…**	**Illness identity**	**Beliefs domain**	**Causes**	**Total**
Moss-Morris (2002) **IPQ-R (English)**	/	14	50° 38°°	18	**70**
Van Ittersum (2009) **IPQ-R-FM (Dutch)**	My fibromyalgia	14	37	18	**69**
Van Wilgen (2008) **IPQ-R-FM (Dutch)**	My fibromyalgia	14	37	26	**77**
Albert (2013) **IPQ-R-WD (French)**	My current health condition	16*	52**	20***	**88**
Chan (2009) **IPQ-R-injury version**	My injury	14	38	18	**70**
Nicholls (2013) **IPQ-R (English)**	My hand/knee/back pain or problem	/	/	/	**/**
Glattacker 2009 **IPQ-R (German)**	/	14	32	18	**64**
Hallegraeff (2013) **Brief IPQ (Dutch)**	My low back pain	/	/	/	**8**

Legend: ° items in the first principle components analysis, °° remaining items

* 5 items removed, 7 added ** 26 new items added ***3 items removed, 5 added

Table 4: Internal consistency and test-retest reliability of the IPQ-R

			Internal consistency Cronbach's alpha				Test-retest reliability			
							Pearson correlations			ICC
			IPQ-R			IPQ-R-WD adapted	IPQ-R			
			Van Ittersum et al. 2009	Van Wilgen et al. 2008	Albert et al. 2013		Moss-Morris et al. 2002	Van Wilgen et al. 2008	Glattacker et al. 2009	
			FM, n=196	FM, n=51	work disability due to MSD, n=43		6 months, RA	3 weeks, FM	4 days, Orth	
Illness identity	Identity	Identity	/	/	/	/	. 57***	.24	.66	.66
Beliefs domain	Timeline	Timeline cyclical	0.75	0.77	0.58	0.58	.55**	.69**	.87	.87
		Timeline acute/chronic	0.80	0.80	0.81	0.81	.35**	.77**	.66	.65
	Consequences	Consequences	0.77	0.64	0.59	0.77	.74***	.75**	.72	.71
	Control/cure	Personal control	0.77	0.83	0.59	0.68	.57***	.57**	.71	.69
		Treatment control	0.79	0.67	0.73	0.77	.50***	.72**	/	/
	Emotional representations	Emotional representations	0.81	0.86	0.81	0.87	.81***	.72**	.78	.78
	Illness coherence	Illness coherence	0.79	0.51	0.80	0.83	.53***	.55**	.56	.55
Causes domain	Causes	Psychological attribution	0.82	0.90			0.82***	.85**		
		Risk factors	0.55	0.48			0.72***	.69**		
		Immunity	0.62	0.47			0.58***	.73**		

		Accident or chance	0.14	0.00-0.61		0.53***	.62**	

Legend: IPQ-R-WD = Illness Perception Questionnaire Revised Work Disability, FM = fibromyalgia, MSD = Musculoskeletal disorder, RA = rheumatoid arthritis, Orth = orthopaedics. ICC = intaclass coefficient. **p<0.01, ***p<0.001

Table 5: Results of hypothesis testing for construct validity of the IPQ-R

Article and population	Questionnaires	Relationship with	Results
Van Wilgen (2008) FM	IPQ-R-FM (Dutch) FIQ PCS	catastrophizing (Pearson's correlations)	- **Catastrophizing** related to a low understanding of the symptoms and positively related to the more cyclical nature of FM and an emotional representation - **Anxiety** was related to experiencing more consequences of FM, to an emotional representation of FM, and to more psychological attributions and more FM-

			specific attributions. - **Feeling depressed** was related to a low score for illness coherence, an emotional representation and more psychological attributions
Chan (2009) Acute hand injury, surgery required	IPQ-R-injury version DASH HISS	objective severity and subjective disability (Pearson Product Moment Correlations)	No significant correlation between DASH/HISS scores and all the components of IPQ-R
Albert (2013) musculoskeletal disorder with absence from work 3m-1y	IPQ-R-WD (French) -> with new items TSK PCS PDI-14 PDI SERWS PBPI IMIQ VAS	Convergent validity (multiple regression analyses and Pearson correlation)	Adjusted r^2 between .33 and .70 ($p \leq .001$) Moderate to strong correlations for each dimension with six theoretically-related variables: TSK, PCS, PDI, PDI-14, PBPI, IMIQ No significant relation with VAS or SERWS

SERWS = self-efficacy with regard to work capacity, VAS = visual analogue scale, DASH = disabilities of the arm, shoulder and hand, HISS = Hand injury severity score, TSK = Tampa scale for kinesiophobia, PCS = pain catastrophizing scale, PDI = pain disability index, IMIQ = Implicit models of illness questionnaire, PDI-14 = psychological distress index, PBPI = pain beliefs and perceptions inventory

Table 6: Results of factor analysis for construct validity of the IPQ-R

Article Patient population	n	Method		Dimension (number of items)	Result
Van Ittersum (2009) **FM**	196	CFA	MGM	Beliefs domain (38) Causal (18)	7 factor-model: -> 55% of the variance 4 factor- model: -> 50% of the variance
Nicholls (2013) **knee pain (OA)** **hand problem** **acute non-specific LBP**	 330 1621 1319	CFA	Goodness of fit - Chi^2, GFI, Parsimony adjusted GFI, CFI, RMSEA	Beliefs domain (38)	7 factor-model: goodness-of-fit statistics were below the criteria

	EFA	PCA varimax rotation	Causal (18)	Knee: 5 factors -> 62% of the variance Hand: 4 factors -> 56% of the variance LBP: 3 factors -> 51% of the variance

Legend: CFA = Confirmatory factor analysis, EFA = exploratory factor analysis, MGM = multigroup method, PCA = principal component analysis, GFI = goodness of fit index, CFI = comparative fit index, RMSEA = root mean square error of approximation, OA = osteoarthritis, LBP = low back pain

Figure 1: Flowchart of the selection process