Subgrouping patients with low back pain

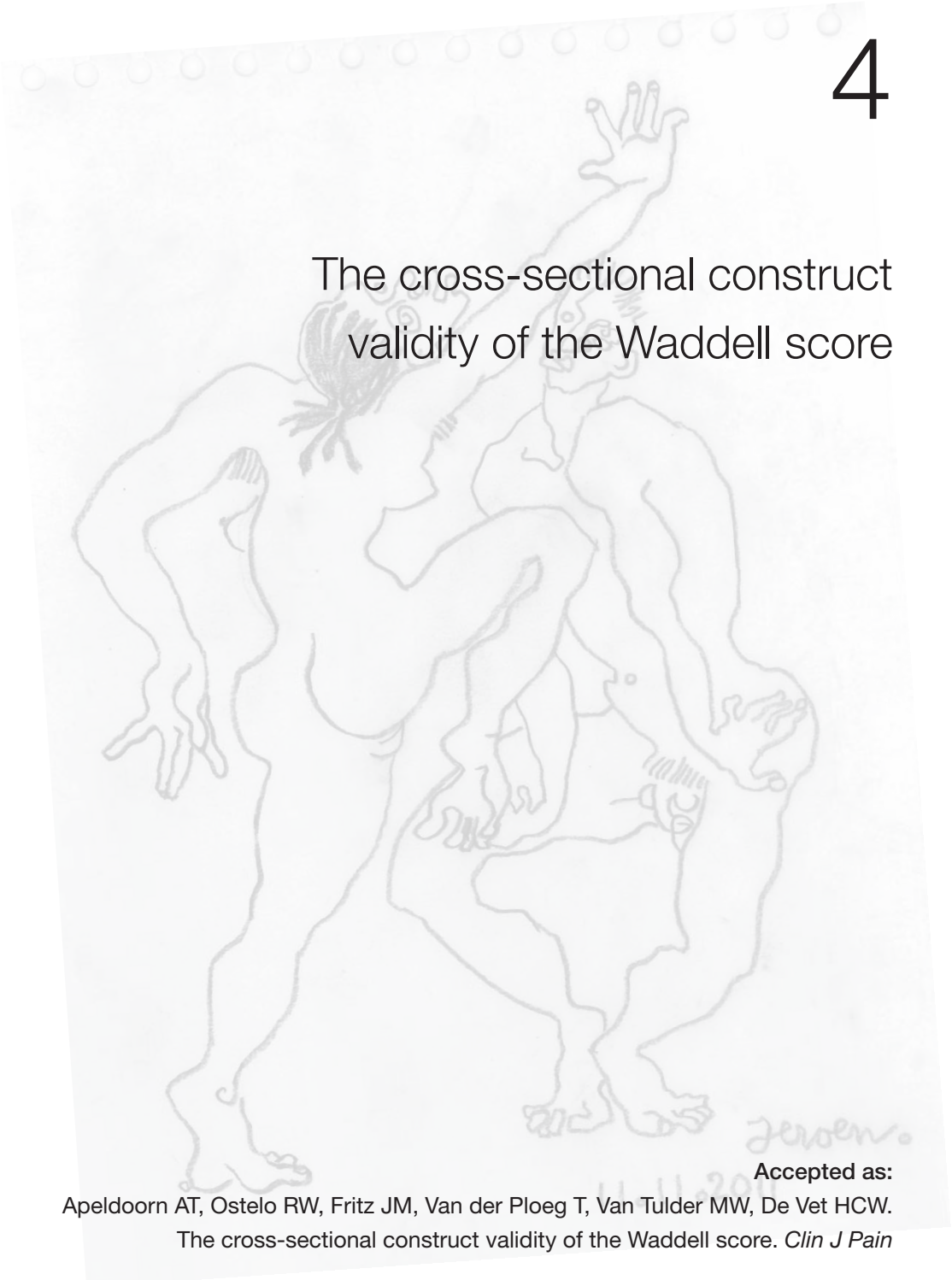
Apeldoorn, A.T.

2012

document version
Publisher's PDF, also known as Version of record

4

The cross-sectional construct validity of the Waddell score



Accepted as:
Apeldoorn AT, Ostelo RW, Fritz JM, Van der Ploeg T, Van Tulder MW, De Vet HCW. The cross-sectional construct validity of the Waddell score. *Clin J Pain*

Abstract

Background: In 1980 the Waddell score, consisting of 8 non-organic or behavioural signs, was developed to measure illness behaviour in patients with low back pain. There is some debate about whether the Waddell score is a valid screening instrument for illness behaviour and psychological distress, or whether it merely reflects elevated pain levels and diminished functional physical capacities.

Objective: The purpose of this study was to examine the construct validity of the Waddell score.

Methods: In this cross-sectional study, a total of 20 hypotheses about the associations between the Waddell score and measures from different domains were formulated a priori, based on a Medline database search (1980-2010). These hypotheses were tested in a sample of 229 patients with chronic low back pain who attended an outpatient rehabilitation centre.

Results: The percentage of hypotheses that were confirmed for the association between the Waddell score and the domain pain was 100%, for the domain physical 80%, for the domain illness behaviour 80% and 50% for the domain psychological. Correlation coefficients and kappa values varied between 0.06 and 0.44 for the measures that were expected to be associated with the Waddell score.

Conclusion: Most of our challenging *a priori* hypotheses were accepted, and the Waddell score was found to have satisfactory cross-sectional construct validity. However, the presence of Waddell signs does not indicate exactly what the specific problems are and must therefore be conceptualized and understood in the total clinical picture of the patient. The association between the Waddell score and measures from different domains is weak. The Waddell score cannot be regarded as a straightforward psychological ‘screener’.

Introduction

Low back pain (LBP) is a complex phenomenon, and includes psychological and social factors interacting with somatic problems and demographic characteristics. The relative contribution of these factors varies in patients with LBP, and various researchers have attempted to identify factors to target intervention and to optimize outcomes.[1] Several researchers have recommended measuring the degree of illness behaviour.[2-6] Addressing illness behaviour could help to separate patients with a low risk from patients with a high risk for persistent LBP, and enable assessment and treatment tailored to the specific needs of the patient.[6] Illness behaviour can be defined as 'observable and potentially measurable actions and conduct which express and communicate the individual's own perception of disturbed health'.[7]

A well-known measurement instrument to quantify illness behaviour in patients with LBP is the Waddell score.[6] It consists of 8 standardized physical 'non-organic' manoeuvres, and is often used in clinical and research settings.[8] The existence of multiple positive Waddell signs is an indication of psychological distress and indicative of the need for a more careful assessment of the psychological and social factors that may coexist with the physical factors.[6] Although various researchers have examined its psychometric properties, the exact relationship between the Waddell score and psychosocial factors is still unclear, and there is a continuing discussion about its construct validity.[6,9] Therefore, careful examination of the construct validity of the Waddell score is needed.

Construct validity refers to the question whether the relationship between the Waddell score and scores for other factors are in agreement with theoretical expectations. The primary aim of the present study was to examine the cross-sectional construct validity of the Waddell score by testing specific a priori hypotheses about the relationships between the Waddell score and factors measuring demographics, pain intensity, illness behaviour and physical and psychological status. Expected associations were predefined, based on available research findings and theoretical considerations, because construct validity should preferably be assessed by verifying or challenging *a priori* hypotheses.[10] Subsequently, these hypotheses were either accepted or refuted.

Materials and methods

Methodological approach

To formulate the hypotheses, one reviewer (AA) searched the literature in the electronic database Medline from March 1980, the year of the original publication by Waddell et al,[8] until October 2010. The search strategy was restricted to English, French, Dutch and German-language papers and used the following key words; Waddell('s) signs, non(-)organic (physical) signs and non(-)physiological signs. The literature search focused on studies that provided information about the cross-sectional construct validity of the Waddell score in patients with LBP, excluding medico-legal issues. Studies that used modified Waddell signs were excluded. The reference lists of each article found were checked to identify additional studies. The entire search yielded 106 publications that described, discussed or examined the Waddell signs. The reviewer identified 43 publications that fulfilled our criteria.[7,8,11-51]

Recruitment of patients

We recruited consecutive patients in the outpatient rehabilitation department of the Medical Centre in Alkmaar in the Netherlands. Patients were eligible if they were 18 years or older and had suffered from LBP for more than three months (with or without leg symptoms). Exclusion criteria were: unwillingness to participate, inadequate command of the Dutch language, and previous treatment for chronic low back pain (CLBP) in a multidisciplinary rehabilitation setting. Patients with confirmed or suspected specific causes of CLBP (e.g. malignancy, fractures, spinal stenosis, severe cases of spondylolisthesis) were excluded. Patients with osteoarthritis and degenerated or protruded discs were not excluded. The Regional Medical Ethics Committee approved the study protocol. Before entering the study all patients gave written informed consent.

Design

All patients underwent a standardized, comprehensive, multidisciplinary assessment, conducted over a 3-week period, in order to customize interventions to their needs. During this observation period the patients underwent a psychological assessment, which included the completion of a comprehensive set of psychological questionnaires, and an examination by a physical therapist, including questionnaires regarding demographics, physical status, illness behaviour, and pain intensity. During the observation period the psychologist and the physical therapist were blinded to each other's findings, until these were discussed in a multidisciplinary team meeting

to formulate a treatment strategy. The first author (AA) monitored adherence to the study protocol, and collected the reports from the psychologist and physical therapist before the multidisciplinary team meetings and before any treatment started.

Participating physical therapists

All patients were assessed by three participating physical therapists with more than ten years' experience in the clinical assessment and treatment of CLBP in a rehabilitation setting. All the physical therapists had attended four postgraduate Mechanical Diagnosis and Therapy (MDT or McKenzie method) courses (A-D), and one physical therapist had MDT credentials. Two of the three physical therapists had followed a postgraduate course in spinal manipulation and mobilization techniques, and were certified manual therapists.

Waddell score

In 1980 Waddell et al[8] developed a checklist of eight non-organic or behavioural signs, divided into five categories. The non-organic signs became known as Waddell signs, and the checklist of five categories is referred to as the Waddell score. The five categories are tenderness (superficial and non-anatomical tenderness), simulation (axial loading and rotation), distraction, regional disturbances (weakness and sensory) and over-reaction (Table 1). Waddell et al. considered a score to be positive when three or more categories were scored positive. In the present study the scores for the eight individual Waddell signs were summed, according to the recommendations made by Apeldoorn et al,[52] resulting in a minimum score of zero (no signs of illness behaviour) and a maximum score of eight (maximum signs of illness behaviour). At the first day of the patient's observation period, the physical therapist started with the assessment of the Waddell tests, i.e. before history taking and the actual physical examination itself. To prevent a patient's awareness of the evaluation of Waddell signs, the tests were preceded by three active tests measuring lumbar mobility. For part of the patients who participated in the present study (n = 126) the inter-rater and intra-rater reliability of Waddell sign-testing was examined between two of the three participating physical therapists. For these two physical therapists, who examined most of the patients in the present study (98%), inter-rater and intra-rater reliability of the Waddell signs have been found to be moderate (kappa value 0.48, 95% CI 0.30 to 0.65) and good (kappa value 0.65, 95% CI 0.50 to 0.80) respectively.[52]

Table 1 The five categories and eight non-organic signs of the Waddell score

Categories	Non-organic tests	Non-organic signs
I. Tenderness	*Superficial*	Widespread tenderness to light pinches over a wide area of lumbar skin.
	Non-anatomic	Deep tenderness felt over a wide area, not localized to one structure, and often extended to the thoracic spine, sacrum or pelvis.
II. Simulation	*Axial loading*	Low back pain reported when light pressure is applied to the patient's head while standing.
	Rotation	Low back pain reported when the shoulders and pelvis are passively rotated in the same plane, when the patient is standing with the feet together.
III. Distraction	*Distraction*	Inconsistent limitation of straight leg raising in supine and seated positions.
IV. Regional disturbance	*Weakness*	Partial cogwheel 'giving way' in many muscle groups.
	Sensory	Sensory disturbances include diminished sensation to light touch, pinprick, and sometimes other modalities fitting a 'stocking' rather than a dermatomal pattern.
V. Over-reaction	*Over-reaction*	Disproportionate verbalization, facial expression, muscle tension and tremor, collapsing or sweating during examination.

Table 2 Hypotheses for the cross-sectional construct validity of the Waddell score

Hypotheses		Expected values	Estimated association for this study	Hypotheses accepted (yes or no)
Waddell scores (0-8) are not associated with:				
Demographic				
1.	higher age	-0.20, 0.20	-0.05	yes
Physical signs and symptoms				
2a.	lower back surgery†	-0.20, 0.20	0.11	yes
2b.	longer duration of CLBP	-0.20, 0.20	-0.01	yes
Waddell scores (0-8) are associated with:				
2c.	a greater fingertip-to-floor distance while bending forward with straight knees	0.21, 0.50	0.35	yes
2d.	a decreased active lumbar extension ROM	0.21, 0.50	0.21	yes
2e.	a decreased active lumbar flexion ROM	0.21, 0.50	0.09	no
Pain				
3a.	elevated current CLBP and/or pain in the leg	0.21, 0.50	0.34	yes
3b.	elevated CLBP in the past week	0.21, 0.50	0.33	yes
3c.	elevated pain in de leg due to the lower back in the past week	0.21, 0.50	0.30	yes
Illness behaviour				
4a.	elevated scores for the pain drawing	0.21, 0.50	0.25	yes
4b.	Waddell symptoms	0.21, 0.50	0.37	yes
4c.	more tablets each day for current CLBP	0.21, 0.50	0.21	yes
4d.	more different types of therapists ever consulted for CLBP	0.21, 0.50	0.06	no
4e.	employed, but currently on sick leave due to CLBP†	0.21, 0.50	0.39	yes

Hypotheses		Expected values	Estimated association for this study	Hypotheses accepted (yes or no)
Psychological variables				
5a.	higher scores for the BDI	0.21, 0.50	0.23	yes
5b.	higher total scores for the SCL-90-R	0.21, 0.50	0.24	yes
5c.	higher scores for somatization (subscale of the SCL-90-R)	0.21, 0.50	0.44	yes
5d.	higher scores for anxiety (subscale of the SCL-90-R)	0.21, 0.50	0.13	no
5e.	higher scores for catastrophizing (subscale of the CSQ)	0.21, 0.50	0.17	no
5f.	higher scores for the TSK	0.21, 0.50	0.13	no
Number of hypotheses that were accepted				15/20

Values are Spearman correlation coefficients, unless otherwise indicated. Kappa values were used for dichotomous variables (†). To calculate kappa values, a cut-off score of ≥2 Waddell signs was used as the threshold defining a positive Waddell score.
CLBP, chronic low back pain; ROM, range of motion; BDI, Beck Depression Inventory; SCL-90-R, Symptom Checklist-90-Revised (modified Dutch version); CSQ, Coping Strategy Questionnaire; TSK, Tampa Scale of Kinesiophobia

Literature research and hypotheses

As outlined in the introduction, the basic approach was to test a priori formulated hypotheses about the Waddell score by evaluating correlations between the Waddell score and a wide range of measures. The measures were categorized in five domains that are commonly assessed in studies addressing the Waddell score; demographics, physical, pain, illness behaviour and psychological. The hypotheses were based on the literature and theoretical considerations and were agreed upon by all authors before they were tested. As in previous observations, we did not expect to find correlation coefficients of more than 0.50. If a relationship was expected, we expected to find correlation coefficients between 0.21 and 0.50. These cut-off values were arbitrarily chosen, but are in line with general recommendations for weak associations.[53,54] The hypotheses will be discussed in detail below, and are presented in Table 2.

Demographic variables

Age. Several studies found no significant associations between the Waddell score and age.[8,15,17,35,37] In the present study we therefore expected to find no associations, or only poor associations between the Waddell score and age (hypothesis 1, Table 2). Evidence regarding the association between the Waddell score and gender is contradictory,[8,12,15,20,27,30,31,37,41,48] so no specific hypotheses were formulated for gender.

Physical signs and symptoms

Lower back surgery. The Waddell score and lower back surgery have not been found to be associated,[15,37,50] and in the present study we expected to find no association, or only a poor association (hypothesis 2a, Table 2).

Duration of CLBP. The literature reported no associations between the Waddell score and the duration of CLBP, and consequently we also expected to find no association between the Waddell score and the duration of CLBP (hypothesis 2b, Table 2).

Mobility. The relationship between the Waddell score and limited lumbar mobility has been widely studied.[13,15,24,32,35,47] Positive associations were found in all studies, but no correlation coefficients were provided. In addition, limited lumbar flexion was found to load on the same factor as the Waddell score in two out of three studies.[8,15,42] In the present study, lumbar mobility was assessed with three active range of motion (ROM) measurements. Flexion and extension ROM of the lumbar spine was measured with the Schober skin distraction method re-modified by Van Adrichem et al,[55] and global flexion mobility was measured with the fingertip-to-floor test.[56] These measurements have been found to be adequately reliable and valid.[55-57] We expected to find a positive

association between high Waddell scores and limited mobility (hypotheses 2c-2e, Table 2).

Pain

Pain. A multitude of studies have documented an association (range r = 0.29 to r = 0.53) between higher Waddell scores and elevated self-reported pain scores.[7,15,16,22,37,38,50] Dickens et al,[15] using factor analysis, found that the Waddell score and pain loaded on the same factor (factor loadings > 0.60), but this was not found in another study.[8] In the present study, pain was measured according to self-reported visual analogue scales (VAS) of 100 mm, ranging from 0 ('no pain') to 100 ('unbearable pain'). The VAS have been found to be user-friendly, valid and reliable.[58-60] We expected that the Waddell score would be positively associated with elevated pain (hypotheses 3a-3c, Table 2).

Illness behaviour variables

Pain drawing. The pain drawing consists of an outline of the body, on which patients are asked to indicate their pain. The Waddell score and the pain drawing were found to be associated in four studies, with correlation coefficients ranging from r = 0.21 to r = 0.27.[8,17,34,42] In one study the degree of association varied, depending on the scoring method for the pain drawing (r = -0.03 to 0.26).[39] In the present study, the pain drawing was scored according to Margolis et al,[61] who divided the pain drawing into 45 anatomical areas. The body areas are assigned weights according to the percentage of body surface they cover. Possible scores vary from 0 (no pain) to 100 (total body pain). Inter-rater reliability was found to be good in 51 CLBP patients.[62] We expected to find a positive association between high Waddell scores and high scores for the pain drawing (hypothesis 4a, Table 2).

Waddell symptoms. Apart from the non-organic signs, which are the focus of the present study, Waddell et al.[46] developed a checklist of seven behavioural or non-organic symptoms. This checklist is a self-reported measure that focuses on symptoms that are vague, and are not usually found in clinical practice in patients with CLBP, such as pain in the tip of the tailbone and pain in the entire leg. One point is scored for every symptom present, and all the points are summed up to calculate a total score, with a minimum of zero (no symptoms of illness behaviour) and a maximum of seven (maximum symptoms of illness behaviour). The test-retest reliability was found to be adequate in 30 patients with LBP,[46] but there is conflicting evidence about the internal reliability of the symptoms.[46,57] The Waddell score has been found to correlate (r = 0.33 and 0.50) with the Waddell symptom checklist.[8,15] We therefore expected to

find a positive association between the Waddell score and the Waddell symptoms (hypothesis 4b, Table 2).

Pain medication. Pain medication has been found to be related to illness behaviour.[4] For many patients with CLBP pain medication must be considered dysfunctional pain behaviour.[63,64] In the present study, the patients were asked how many tablets a day they took for their current CLBP. They were not asked about the exact type of pain medication. We expected that high Waddell scores would be positively associated with more pain medication (hypothesis 4c, Table 2).

Health care consumption. Pain-related use of health care has been found to be a symptom of illness behaviour,[4] and related to higher Waddell scores.[45] In the present study, the patients were asked how many different types of conservative therapists (e.g. osteopath, chiropractor, physical therapist and acupuncturist) they had ever consulted for their LBP. Each type of therapist was given a score of one. They were not asked about the total number of treatment sessions. To help the patients to remember these therapists, they were given an extensive list of different types of health care providers. We expected to find a positive association between a high Waddell score and a greater number of different types of therapists ever consulted for LBP (hypothesis 4d, Table 2).

Employed, but currently on sick leave due to CLBP. Sick leave is greatly influenced by patient efforts and other considerations concerning the environment,[65] and has been found to be a strong predictor for poor outcome in patients with LBP in primary care.[66] We are not aware of any studies that have explored the relationship between current sick leave due to CLBP and high Waddell scores, but based on theoretical considerations we expected to find a positive association (hypothesis 4e, Table 2).

Psychological variables

Depression. Several studies have concluded that there is a link between high Waddell scores and depression,[15,17,31,37,50] but one other study found no association.[33] In the present study, depression was measured with the Dutch translation of the Beck Depression Inventory (BDI).[67,68] The BDI consists of 21 graded items, ranging in severity from 0-3. It has good psychometric properties for the measurement of depression, but in patients with CLBP a confounding effect has been found for three items measuring somatic symptoms.[69] In the present study we expected to find a positive association between high Waddell scores and elevated BDI scores (hypothesis 5a, Table 2).

Psychological distress and psychopathology. Several studies have found a relationship between high Waddell scores and the subscales for hypochondriasis and hysteria on the Minnesota Multiphasic Personality Inventory (MMPI) ($r = 0.10$ to 0.46).[8,17,26,27,33,37]

Using other reference standards than the MMPI, associations have been found with anxiety,[15,31,50] somatization,[31] and psychological distress.[7,29] In the present study, psychological distress and psychopathology were measured with the Dutch version of the Symptom Checklist-90-Revised (SCL-90-R).[70,71] This version contains the following eight dimensions; agoraphobia, anxiety, depression, somatization, cognitive-performance deficits, interpersonal sensitivity-mistrust, acting-out hostility, and sleep difficulties. It has been found to be internally consistent and generalizable to different patient samples, including pain patients,[72] and subscales of the SCL-90-R have been used successfully in previous back pain studies.[73,74] We expected to find associations between high Waddell scores and the total SCL-90-R score, and its subscales somatization and anxiety (hypotheses 5b-5d, Table 2).

Catastrophizing and fear avoidance. These two cognitive-behavioural factors are closely related to each other and to depression.[75] Unfortunately, no data are available on the relationship between these psychological factors and the Waddell score. In the present study, catastrophizing was measured with the Coping Strategy Questionnaire (CSQ) subscale catastrophizing.[76] The CSQ is a 44-item questionnaire that assesses seven dimensions. It is a widely used measurement instrument, with satisfactory internal consistency and validity.[77-79] The Dutch translation has been found to have adequate validity.[80] Fear avoidance was measured with the Dutch version of the Tampa Scale of Kinesiophobia, which consists of 17 questions.[75,81] The reliability and validity of the Dutch version has been found to be adequate.[75,82] Based on theoretical considerations, we expected to find positive associations between high Waddell scores and elevated scores for the CSQ subscale catastrophizing and the Tampa Scale of Kinesiophobia (hypotheses 5e and 5f, Table 2).

Statistical analysis

Appropriate statistical methods were used to describe the patients' characteristics. The a priori formulated hypotheses were tested by calculating Spearman's correlation coefficients between the Waddell score and several measures that fall into the 5 aforementioned domains. To estimate the association between the Waddell score and dichotomous factors, the Waddell score was dichotomized and kappa values were calculated. Based on previous findings a positive Waddell score was defined at the level of two or more positive Waddell signs.[57]

To obtain information about the validity of the different measures representing different domains, we performed exploratory factor analysis (principal components analysis using Varimax [oblique] rotation). For this analysis only continuous measures were

used, and no data were included from subscales of questionnaires with a total score (e.g. anxiety, subscale of the SCL-90-R). The Bartlett test for sphericity (with $p < 0.05$) and a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy of ≥ 0.6 were used in this factor analysis. The number of factors was set at four, because we wanted to find out whether the selected measures correlated well with the four dimensions. A factor was considered to be important if its eigenvalue exceeded 1.0.[83]

Factor analysis is based on complete cases, and missing data can reduce the precision of calculated statistics. Therefore, prior to factor analysis, missing variables were imputed. We used the expectation maximization (EM) method, which estimates missing values in an iterative process. These results were compared with the results from another imputation method (maximum likelihood estimation method) and with the model that emerged without using imputation methods. Provided a meaningful structure was found, correlation coefficients were subsequently calculated for the association between the four different factors and the Waddell score, because the latter was not included in the model. All analyses were performed with SPSS version 15.0.

Results

A total of 229 patients were included in the study (63% females and 37% males). Their mean age was 43.9 years (SD 11.0), and the median duration of their CLBP was 84 months (interquartile range 24-180). Table 3 shows the frequency distributions and descriptive statistics for the main characteristics of the participating patients.

Some data were missing for the Waddell score (1.7%), and for the variables that were compared to the Waddell score (5.8%). The Waddell score had a positively skewed scale-distribution, and transformation to normality with standard procedures was not possible. We therefore calculated Spearman's correlation coefficients (non-parametric method) to examine the associations between the Waddell score and the different continuous measurements. To calculate correlation coefficients and kappa values missing data were not imputed, because the percentage missing was low. As can be seen in Table 2, a total of 75% of the hypotheses were accepted. The percentage of hypotheses that were confirmed for the domain pain was 100%, for the domain physical 80%, for the domain illness behaviour 80% and 50% for the domain psychological. The Spearman's correlation coefficient with age was -0.05, and varied for physical signs and symptoms between -0.01 and 0.35, for pain intensity between 0.30 and 0.34, for illness behaviour between 0.06 and 0.37, and for psychological factors between 0.13 and 0.44.

To carry out factor analysis with complete data, we imputed missing data (in total 6.7%), using the EM method. Next, a 4-factor solution was generated. The KMO measure of sampling adequacy was 0.68, and the Bartlett test for sphericity was $p < 0.001$, which confirmed that the data were appropriate for the factor model. Inspection of the 4-factor model revealed that factor 1 represented the dimension psychological, factor 2 represented pain, factor 3 represented physical and factor 4 represented illness behaviour (Table 4). All four factors had eigenvalues of > 1.0 and a cumulative explanatory variance of 57%. The 4-factor structure that emerged with the EM method was comparable to the model for which the MLE method or no imputation method was used (data not shown). Additionally, Spearman's correlation coefficients were computed between the four factors and the Waddell score. As shown in Table 4, the correlations varied between 0.15 (factor 1, labelled psychological) and 0.35 (factor 2, labelled pain).

Table 3 Main characteristics of the participants (n=229)

Female (%)	63.3
Duration of CLBP in months†	84 (24-180)
Age in years	43.9 (11.0)
Lower back surgery (%)	19.6
Marital status	
Married/living with a partner (%)	87
Single/divorced (%)	13
Employed (%)	84.4
Employed and currently working (%)	57.4
Currently with modified work (%)	27.4
Currently working less hours (%)	50.5
Employed, but currently on sick leave due to CLBP (%)	42.6
Not employed/retired (%)	15.6
Dutch nationality (%)	96.1
Fingertip-to-floor distance (0 – 90 cm)†	18 (0-31)
Active lumbar extension ROM‡	13.9 (0.7)
Active lumbar flexion ROM╣	20.3 (1.3)
Current CLBP/leg pain (0-100)	53.2 (23.2)
CLBP in past week (0-100)	60.0 (20.5)
Pain in the leg in past week (0-100)	39.3 (29.7)
Pain drawing (0-100)	19.2 (12.5)
Waddell symptoms (0-7)	2.6 (1.5)
Waddell signs (0-8)†	1 (0-3)
Two or more Waddell signs (%)	47.6
Number of tablets a day for CLBP†	0 (0-2)
Number of different types of therapists ever consulted for CLBP†	3 (2-5)
Depression (BDI, 0-63)	10.6 (7.2)
Psychological distress and psychopathology (SCL-90-R, 90-450)	144.0 (39.6)
Somatization (subscale SCL-90-R, 0-60)	23.8 (6.7)
Anxiety (subscale SCL-90-R, 0-50)†	13 (11-16)
Fear avoidance (TSK, 17-68)	38.4 (7.1)
Catastrophizing (subscale CSQ, 0-60)	19.9 (12.6)

Values are the mean and the standard deviation, unless otherwise indicated. CLBP, chronic low back pain; ROM, range of motion; BDI, Beck Depression Index; SCL-90-R, Symptom Checklist-90-Revised (modified Dutch version); TSK, Tampa Scale of Kinesiophobia; CSQ, Coping Strategy Questionnaire,
† Median and interquartile range; ‡ Range 15 cm (no mobility) to <10 cm (hypermobility);
╣ Range 15 cm (no mobility) to >25 cm (hypermobility)

Table 4 Varimax rotated factor-loading matrix for 229 CLBP patients

	Factor 1	Factor 2	Factor 3	Factor 4
Impairment				
Fingertip-to-floor distance			.762	
Active lumbar extension ROM‡			.719	
Active lumbar flexion ROM╣			- .768	
Pain				
Current CLBP and/or pain in the leg		.900		
CLBP in past week		.887		
Pain in the leg due to the lower back in past week		.648		
Illness behaviour				
Behavioural symptoms				.526
Daily use of pills for current CLBP			.326	.340
Pain drawing				.702
Different types of therapists ever consulted for CLBP				.647
Psychological signs				
BDI	.866			
SCL-90-R	.840			
Catastrophizing (subscale of the CSQ)	.781			
Tampa Scale of Kinesiophobia	.569			
Total variance explained (%)	23.8	13.2	11.9	8.5
Correlations between the Waddell score and factors 1-4†	.15	.35	.25	.16

Only factor scores of ≥0.30 are reported.
ROM, range of motion; CLBP, chronic low back pain; BDI, Beck Depression Inventory; SCL-90-R, Symptom Checklist-90-Revised (modified Dutch version); CSQ, Coping Strategy Questionnaire
‡ Range 15 cm (no mobility) to <10 cm (hypermobility); ╣ Range 15 cm (no mobility) to >25 cm (hypermobility); † Spearman's correlation coefficients

Discussion

Although the psychometric properties of the Waddell score have been studied extensively in the past 30 years, there is still a heated debate about the construct validity of the Waddell score. Most noticeable are the inconsistent conclusions in reviews published by Waddell,[6] compared to Fishbain et al.[9] Contrary to Waddell,[6] Fishbain et al[9] concluded that the Waddell score is not an indicator of abnormal illness behaviour or a disturbed psychological state of the patient, but a reflector of elevated pain and diminished functional physical capacities. These inconsistent findings can be partly explained by the fact that the latter included studies that did not use the original Waddell signs, but the poor methodological quality of this study may also have played a role. Our results are more in line with those of Waddell,[6] who concluded that the Waddell score measures physical factors, pain, illness behaviour and psychological factors. However, although most of our a priori hypotheses were accepted, it is difficult to offer a clear and precise description of the construct the Waddell score measures. A wide range of factors can contribute to the expression of Waddell signs, such as illness worry, sick role, unemployment, hypochondriasis, depression, anxiety to examination, central sensitization, learned behaviour and the disease state. For example, recent research shows that in patients with CLBP brain structure and function can alter[84] and the patient can display Waddell signs as a response to these changes.[85] Given the substantial inter-individual variability of how people respond to health problems, the differences in social norms, cultural models and health care systems and the probably limited usefulness of the description of the Waddell score in traditional diagnostic domains, it is not very surprising that we found only weak relationships between the Waddell score and the domains measured. For clinical practice, the presence of Waddell signs does not indicate exactly what the specific problems are, and they must therefore be conceptualized and understood in the total clinical picture of the patient. It is evident that the Waddell score cannot be regarded as a straightforward psychological 'screener'.[86]

One of our challenges was to formulate hypotheses. Our a priori hypotheses were based on the available evidence in the literature, but if the relevant studies did not provide the required data, we formulated hypotheses based on theoretical grounds. We chose correlation coefficients between -0.20 and 0.20 if we did not expect to find an association, and correlations between 0.21 and 0.50 if we expected to find a weak association. Most of the observed correlations between the Waddell score and the different domains were weak. The clinical relevance of these weak correlations is questionable, although there is preliminary evidence that the Waddell score may contribute to more cost-effective

rehabilitation aimed at return to work,[87] and to the identification of patients with CLBP who might benefit from additional psychological assessment.[57]

In the present study, the cross-sectional construct validity of the Waddell score was assessed using a wide range of measures with overall accepted and satisfactory reliability and validity, because no 'gold standard' was available. The measures we used were clustered into one of the following five domains; demographics, physical factors, pain, illness behaviour and psychological factors. The results of factor analysis confirmed that the measures did form groups based on the domains as expected, thereby giving credence to the used grouping.

Some limitations in the present study need to be taken into consideration. A first point of concern is the risk of bias due to diagnostic suspicion in the physical therapist assessments of the signs and symptoms. This form of bias occurs when knowledge or suspicion of a relationship leads to preferential findings. We tried to minimize this bias by testing the Waddell signs and the mobility measurements at the beginning of the examination procedure, i.e. before questionnaire intake, history taking, and the actual physical examination itself. However, we believe that any presence of such a bias would have caused an over-estimation of the associations that were found. Secondly, our data were collected in an outpatient rehabilitation centre. The patients were predominantly from a Caucasian, middle-aged working population with CLBP. However, it cannot be ruled out that the construct validity of the Waddell score might differ for other populations. Therefore, the results of this study should be interpreted in the light of our specific study population. Finally, in the present study the literature was screened for associative patterns between the Waddell score and different factors, without using specific quality criteria and without a systematic review approach. Most of our hypotheses were confirmed, and we therefore estimate that our approach was appropriate.

In summary, a better understanding of illness behaviour in patients with CLBP is important to increase our knowledge of the aetiology and the maintenance of the syndrome, and to optimize treatment strategies.[5,6] In the present study we examined the cross-sectional construct validity of the Waddell score, a commonly used illness behaviour-screening measurement instrument for patients with LBP. We found evidence that it measures a combination of pain intensity, illness behaviour, physical dysfunction and psychological functioning. However, the associations with the different domains were, in general, weak, which underscores the complexity of illness behaviour as measured with the Waddell score.

References

1. Billis EV, McCarthy CJ, Oldham JA. Subclassification of low back pain: a cross-country comparison. *Eur Spine J* 2007;16:865-79.
2. Hutten MM, Hermens HJ, Zilvold G. Differences in treatment outcome between subgroups of patients with chronic low back pain using lumbar dynamometry and psychological aspects. *Clin Rehabil* 2001;15:479-88.
3. Keefe FJ, Bradley LA, Crisson JE. Behavioral assessment of low back pain: identification of pain behavior subgroups. *Pain* 1990;40:153-60.
4. Rief W, Ihle D, Pilger F. A new approach to assess illness behaviour. *J Psychosom Res* 2003;54:405-14.
5. Turk DC, Rudy TE. Persistent pain and the injured worker: Integrating biomedical, psychosocial, and behavioral factors in assessment. *J Occup Rehabil* 1991;1:159-79.
6. Waddell G. *The back pain revolution.* 2nd ed. Edinburgh: Churchill Livingstone, 2004.
7. Waddell G, Pilowsky I, Bond MR. Clinical assessment and interpretation of abnormal illness behaviour in low back pain. *Pain* 1989;39:41-53.
8. Waddell G, McCulloch JA, Kummel E, et al. Nonorganic physical signs in low-back pain. *Spine* 1980;5:117-25.
9. Fishbain DA, Cole B, Cutler RB, et al. A structured evidence-based review on the meaning of nonorganic physical signs: Waddell signs. *Pain Med* 2003;4:141-81.
10. 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.
11. Blom A, Taylor A, Whitehouse S, et al. A new sign of inappropriate lower back pain. *Ann R Coll Surg Engl* 2002;84:342-3.
12. Chan CW, Goldman S, Ilstrup DM, et al. The pain drawing and Waddell's nonorganic physical signs in chronic low-back pain. *Spine* 1993;18:1717-22.
13. Cooke C, Menard MR, Beach GN, et al. Serial lumbar dynamometry in low back pain. *Spine* 1992;17:653-62.
14. Cooke C, Dusik LA, Menard MR, et al. Relationship of performance on the ERGOS work simulator to illness behavior in a workers' compensation population with low back versus limb injury. *J Occup Med* 1994;36:757-62.
15. Dickens C, Jayson M, Creed F. Psychological correlates of pain behavior in patients with chronic low back pain. *Psychosomatics* 2002;43:42-8.
16. Dirks HF, Wunder J, Reynolds J, et al. A scale for predicting nonphysiological contributions to pain. *Psychother Psychosom* 1996;65:153-7.
17. Doxey NC, Dzioba RB, Mitson GL, et al. Predictors of outcome in back surgery candidates. *J Clin Psychol* 1988;44:611-22.

18. Gaines WG Jr, Hegmann KT. Effectiveness of Waddell's nonorganic signs in predicting a delayed return to regular work in patients experiencing acute occupational low back pain. *Spine* 1999;24:396-400.
19. Greenough CG, Fraser RD. Comparison of eight psychometric instruments in unselected patients with back pain. *Spine* 1991;16:1068-74.
20. Hadjistavropoulos HD, Craig KD. Acute and chronic low back pain: cognitive, affective, and behavioral dimensions. *J Consult Clin Psychol* 1994;62:341-9.
21. Hadjistavropoulos HD, LaChapelle DL. Extent and nature of anxiety experienced during physical examination of chronic low back pain. *Behav Res Ther* 2000;38:13-29.
22. Hasenbring M, Marienfeld G, Kuhlendahl D, et al. Risk factors of chronicity in lumbar disc patients. A prospective investigation of biologic, psychologic, and social predictors of therapy outcome. *Spine* 1994;19:2759-65.
23. Hayes B, Solyom CA, Wing PC, et al. Use of psychometric measures and nonorganic signs testing in detecting nomogenic disorders in low back pain patients. *Spine* 1993;18:1254-9.
24. Hirsch G, Beach G, Cooke C, et al. Relationship between performance on lumbar dynamometry and Waddell score in a population with low-back pain. *Spine* 1991;16:1039-43.
25. Karas R, McIntosh G, Hall H, et al. The relationship between nonorganic signs and centralization of symptoms in the prediction of return to work for patients with low back pain. *Phys Ther* 1997;77:354-60.
26. Lacroix JM, Powell J, Lloyd GJ, et al. Low-back pain. Factors of value in predicting outcome. *Spine* 1990;15:495-9.
27. Lehmann TR, Russell DW, Spratt KF. The impact of patients with nonorganic physical findings on a controlled trial of transcutaneous electrical nerve stimulation and electroacupuncture. *Spine* 1983;8:625-34.
28. Main CJ, Waddell G. Psychometric construction and validity of the Pilowsky Illness Behaviour Questionnaire in British patients with chronic low back pain. *Pain* 1987;28:13-25.
29. Main CJ, Wood PL, Hollis S, et al. The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine* 1992;17:42-52.
30. Manchikanti L, Pampati V, Fellows B, et al. Characteristics of chronic low back pain in patients in an interventional pain management setting: a prospective evaluation. *Pain Physician* 2001;4:131-42.
31. Manchikanti L, Fellows B, Singh V, et al. Correlates of non-physiological behavior in patients with chronic low back pain. *Pain Physician* 2003;6:159-66.
32. Marras WS, Parnianpour M, Ferguson SA, et al. The classification of anatomic- and symptom-based low back disorders using motion measure models. *Spine* 1995;20:2531-46.
33. Maruta T, Goldman S, Chan CW, et al. Waddell's nonorganic signs and Minnesota Multiphasic Personality Inventory profiles in patients with chronic low back pain. *Spine* 1997;22:72-5.

34. Masferrer R, Prendergast V, Hagell P. Colored pain drawings: preliminary observations in a neurosurgical practice. *Eur J Pain* 2003;7:213-7.
35. Menard MR, Cooke C, Locke SR, et al. Pattern of performance in workers with low back pain during a comprehensive motor performance evaluation. *Spine* 1994;19:1359-66.
36. Newton M, Thow M, Somerville D, et al. Trunk strength testing with iso-machines. Part 2: Experimental evaluation of the Cybex II Back Testing System in normal subjects and patients with chronic low back pain. *Spine* 1993;18:812-24.
37. Novy DM, Collins HS, Nelson DV, et al. Waddell signs: distributional properties and correlates. *Arch Phys Med Rehabil* 1998;79:820-2.
38. Ohlund C, Lindstrom I, Areskoug B, et al. Pain behavior in industrial subacute low back pain. Part I. Reliability: concurrent and predictive validity of pain behavior assessments. *Pain* 1994;58:201-9.
39. Parker H, Wood PL, Main CJ. The use of the pain drawing as a screening measure to predict psychological distress in chronic low back pain. *Spine* 1995;20:236-43.
40. Prkachin KM, Schultz I, Berkowitz J, et al. Assessing pain behaviour of low-back pain patients in real time: concurrent validity and examiner sensitivity. *Behav Res Ther* 2002;40:595-607.
41. Reesor KA, Craig KD. Medically incongruent chronic back pain: physical limitations, suffering, and ineffective coping. *Pain* 1988;32:35-45.
42. Rose MJ, Reilly JP, Slade PD, et al. A comparative analysis of psychological and physical models of low back pain experience. *Physiotherapy* 1995;81:710-6.
43. Rucker KS, Metzler HM, Kregel J. Standardization of chronic pain assessment: a multiperspective approach. *Clin J Pain* 1996;12:94-110.
44. Vendrig AA, De Mey HR, Derksen JJ, et al. Assessment of chronic back pain patient characteristics using factor analysis of the MMPI-2: which dimensions are actually assessed? *Pain* 1998;76:179-88.
45. Waddell G, Bircher M, Finlayson D, et al. Symptoms and signs: physical disease or illness behaviour? *Br Med J* 1984;289:739-41.
46. Waddell G, Main CJ, Morris EW, et al. Chronic low-back pain, psychologic distress, and illness behavior. *Spine* 1984;9:209-13.
47. Waddell G, Somerville D, Henderson I, et al. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine* 1992;17:617-28.
48. Waddell G, Richardson J. Observation of overt pain behaviour by physicians during routine clinical examination of patients with low back pain. *J Psychosom Res* 1992;36:77-87.
49. Watson PJ, Poulter ME. The development of a functional task-oriented measure of pain behaviour in chronic low back pain patients. *J Back Musc Rehab* 1997;9:57-9.
50. Weaver CS, Kvaal SA, McCracken L. Waddell signs as behavioral indicators of depression and anxiety in chronic pain. *J Back Musc Rehab* 2003;17:21-6.
51. Werneke MW, Hart DL. Centralization: association between repeated end-range pain responses and

behavioral signs in patients with acute non-specific low back pain. *J Rehabil Med* 2005;37:286-90.

52. Apeldoorn AT, Bosselaar H, Blom-Luberti T, et al. The reliability of nonorganic sign-testing and the Waddell score in patients with chronic low back pain. *Spine* 2008;33:821-6.
53. Portney LG, Watkins MP. *Foundations of clinical research. Applications to practice.* Upper Saddle River, NJ: Prentice Hall Health, 2000.
54. Munro BH, ed. *Statistical methods for health care research.* 3rd ed. Philadelphia: Lippincott-Raven, 1997.
55. Van Adrichem JA, Van der Korst JK. Assessment of the flexibility of the lumbar spine. A pilot study in children and adolescents. *Scand J Rheumatol* 1973;2:87-91.
56. Perret C, Poiraudeau S, Fermanian J, et al. Validity, reliability, and responsiveness of the fingertip-to-floor test. *Arch Phys Med Rehabil* 2001;82:1566-70.
57. Apeldoorn AT, Bosselaar H, Ostelo RW, et al. Identification of patients with chronic low back pain who might benefit from additional psychological assessment. *Clin J Pain* 2012;28:23-31.
58. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983;16:87-101.
59. Revill SI, Robinson JO, Rosen M, et al. The reliability of a linear analogue for evaluating pain. *Anaesthesia* 1976;31:1191-8.
60. Sriwatanakul K, Kelvie W, Lasagna L, et al. Studies with different types of visual analog scales for measurement of pain. *Clin Pharmacol Ther* 1983;34:234-9.
61. Margolis RB, Tait RC, Krause SJ. A rating system for use with patient pain drawings. *Pain* 1986;24:57-65.
62. Margolis RB, Chibnall JT, Tait RC. Test-retest reliability of the pain drawing instrument. *Pain* 1988;33:49-51.
63. Crisostomo RA, Schmidt JE, Hooten WM, et al. Withdrawal of analgesic medication for chronic low-back pain patients: improvement in outcomes of multidisciplinary rehabilitation regardless of surgical history. *Am J Phys Med Rehabil* 2008;87:527-36.
64. Dersh J, Mayer TG, Gatchel RJ, et al. Prescription opioid dependence is associated with poorer outcomes in disabling spinal disorders. *Spine* 2008;33:2219-27.
65. Fordyce WE. On the nature of illness and disability: an editorial. *Clin Orthop Relat Res* 1997;336:47-51.
66. Dunn KM, Jordan KP, Croft PR. Contributions of prognostic factors for poor outcome in primary care low back pain patients. *Eur J Pain* 2010;15:313-9.
67. Bouman TK, Luteijn F, Albersnagel FA, et al. Enige ervaringen met de Beck depression inventory (BDI) (Some experiences with the Beck depression inventory). *Gedrag* 1985;13:13-24.
68. Beck AT, Steer RA, Brown GK. *Beck Depression Inventory.* San Antonio,TX: Psychological Corporation, 1996.

69. Wesley AL, Gatchel RJ, Garofalo JP, et al. Toward more accurate use of the Beck Depression Inventory with chronic back pain patients. *Clin J Pain* 1999;15:117-21.
70. Derogatis LR. *SCL-90: Administration, scoring and procedures manual-I for the R(evised) version and other instruments of the psychopathology rating scale series.* Baltimore, MD: Clinical Psychometrics Research Unit, Johns Hopkins University School of Medicine, 1977.
71. Arrindell WA, Ettema JHM. *Symptom checklist. Handleiding bij een multidimensionele psychopathologie-indicator (SCL-90; Manual for a multidimensional psychopathology-indicator).* Lisse: Swets & Zeitlinger bv, 1986.
72. Arrindell WA, Barelds DP, Janssen IC, et al. Invariance of SCL-90-R dimensions of symptom distress in patients with peri partum pelvic pain (PPPP) syndrome. *Br J Clin Psychol* 2006;45:377-91.
73. Dionne CE, Bourbonnais R, Fremont P, et al. Determinants of 'return to work in good health' among workers with back pain who consult in primary care settings: a 2-year prospective study. *Eur Spine J* 2007;16:641-55.
74. Nickel R, Egle UT, Eysel P, et al. Health-related quality of life and somatization in patients with long-term low back pain: a prospective study with 109 patients. *Spine* 2001;26:2271-7.
75. Vlaeyen JW, Kole-Snijders AM, Boeren RG, et al. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995;62:363-72.
76. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. *Pain* 1983;17:33-44.
77. Main CJ, Waddell G. A comparison of cognitive measures in low back pain: statistical structure and clinical validity at initial assessment. *Pain* 1991;46:287-98.
78. Jensen MP, Turner JA, Romano JM, et al. Coping with chronic pain: a critical review of the literature. *Pain* 1991;47:249-83.
79. Lawson K, Reesor KA, Keefe FJ, et al. Dimensions of pain-related cognitive coping: cross-validation of the factor structure of the Coping Strategy Questionnaire. *Pain* 1990;43:195-204.
80. Spinhoven P, Ter Kuile MM, Linssen AC, et al. Pain coping strategies in a Dutch population of chronic low back pain patients. *Pain* 1989;37:77-83.
81. Kori SH, Miller RP, Todd DD. Kinisophobia: a new view of chronic pain behavior. *J Pain Manag* 1990;3:35-43.
82. Goubert L, Crombez G, Van Damme S, et al. Confirmatory factor analysis of the Tampa Scale for Kinesiophobia: invariant two-factor model across low back pain patients and fibromyalgia patients. *Clin J Pain* 2004;20:103-10.
83. Bowling A. *Research methods in health: investigating and health services.* 2nd ed. Philadelphia, PA: Open University Press, 2002.
84. Wand BM, Parkitny L, O'Connell NE, et al. Cortical changes in chronic low back pain: current state of the art and implications for clinical practice. *Man Ther* 2011;16:15-20.

85. Ranney D. A Proposed Neuroanatomical Basis of Waddell's Nonorganic Signs. *Am J Phys Med Rehabil* 2010;89:1036-42.
86. Main CJ, Waddell G. Behavioral responses to examination. A reappraisal of the interpretation of 'nonorganic signs'. *Spine* 1998;23:2367-71.
87. Kool JP, Oesch PR, De Bie RA. Predictive tests for non-return to work in patients with chronic low back pain. *Eur Spine J* 2002;11:258-66.