The accuracy of pain drawing in identifying psychological distress in low back pain—systematic review and meta-analysis of diagnostic studies

Lucia Bertozzi, PT, MSc[1], Anna Rosso, PT[2], Antonio Romeo, PT[2], Jorge Hugo Villafañe, PhD, MSc[3]*, Andrew A. Guccione, PhD, PT, DPT, FAPTA[4], Paolo Pillastrini, PT, MSc[5], Carla Vanti, PT, OMPT, MSc[5]

[1] School of Physical Therapy, Alma Mater Studiorum University of Bologna, Italy
[2] University of Bologna, Italy
[3] IRCCS Don Gnocchi Foundation: Milan, Italy
[4] Department of Rehabilitation Science, College of Health and Human Services, George Mason University, USA
[5] Department of Biomedical and Neuromotor Sciences (DIBINEM), Alma Mater Studiorum, University of Bologna, Italy

Abstract. [Purpose] The aim of this systematic review and meta-analysis was to estimate the accuracy of qualitative pain drawings (PDs) in identifying psychological distress in subacute and chronic low back pain (LBP) patients. [Subjects and Methods] Data were obtained from searches of PubMed, EBSCO, Scopus, PsycINFO and ISI Web of Science from their inception to July 2014. Quality assessments of bias and applicability were conducted using the Quality of Diagnostic Accuracy Studies-2 (QUADAS-2). [Results] The summary estimates were: sensitivity=0.45 (95% CI 0.34, 0.61), specificity=0.66 (95% CI 0.53, 0.82), positive likelihood ratio=1.23 (95% CI 0.93, 1.62), negative likelihood ratio=0.84 (95% CI 0.70, 1.01), and diagnostic odds ratio=1.46 (95% CI 0.79, 2.68). The area under the curve was 78% (CI, 57 to 99%). [Conclusion] The results of this systematic review do not show broad and unqualified support for the accuracy of PDs in detecting psychological distress in subacute and chronic LBP.
Key words: Pain drawing, Stress psychological, Low back pain

INTRODUCTION

A pain drawing (PD) is a simple line drawing of the human body (front and back)[1] on which patients can indicate their pain for both clinical and research purposes[2]. PDs may differ from each other depending on whether the image is simple[3] or detailed[4, 5], presents a frontal or sagittal perspective, or depicts a female or male figure[6]. PDs also use different methods for making notations and scoring the drawing[3, 4, 7]. Although PDs are commonly used to identify pain site, magnitude, and regional extension, their use to identify psychological distress of patients has also been proposed[8]. It is well documented that psychological factors can contribute to how a patient first reports low back pain (LBP) and influence the transition towards chronic pain[9–11]. Moreover, psychological distress has a relevant influence on patients' responses to a variety of therapeutic approaches to LBP, including conservative treatment[12, 13], chemonucleolysis and surgery[4].

Following the scoring method devised by Ransford et al.[7], who concluded that PDs could be used to identify patients with elevated hysteria and hypochondriasis scores on the Minnesota Multiphasic Personality Inventory (MMPI), based on the conformity of the patients' markings of typical or expected pain patterns, other investigators have also proposed using PDs to evaluate an individuals' underlying psychological states, in order to predict outcomes and to avoid unnecessary procedures for LBP patients. Udén et al.[8] introduced an alternative scoring method that classified drawings as "normal" or "organic" versus "abnormal" or "non-organic".

The validity of PD scoring systems for measuring psychological status was systematically reviewed by Carnes et al.[14] who reported on the sensitivity, specificity and positive and negative predictive values of abnormal drawings. They concluded that the available data did not support the assumption that unusual pain drawings or extensive markings indicated a disturbed psychological state, and as a consequence, recommended against using PDs as psychological assessment tools. Given that Carnes' systematic review[14] included dysfunctions in different body structures, a precise conclusion about PDs cannot be drawn with respect

*Corresponding author. Jorge Hugo Villafañe (E-mail: mail@villafane.it)

to LBP. Furthermore, Carnes' analysis differentiated scoring methods and included reference tests covering a large spectrum of psychological disturbances. No meta-analysis was performed and only English-language articles were included, raising the risk of publication bias. Finally, it is important to emphasize that the database search in Carnes' review was performed in the first three months of 2003 and subsequently, different authors have expressed contrasting opinion about on the influence of psychosocial factors on PD in LBP[15–18]. In particular, Abbott[16] and Turk[18] have argued that for some patients, the drawing might be influenced by psychosomatic disorders, and could be used to alert medical and health professionals.

For all these reasons, another systematic review and specific meta-analysis would be a useful contribution to the assessment of the validity of PDs in identifying psychological distress among subacute and chronic LBP patients. This systematic review expands upon previous studies explicitly targeting a particular population of interest, and focuses on the specific characteristics of outcome measures as inclusion criteria.

SUBJECTS AND METHODS

Our literature search aimed to identify all available studies that evaluated the correlation between qualitative or semi-qualitative PDs scored using only the methods described by Ransford or Udén and psychological distress (anxiety, depression or illness behavior). Records were identified by searching multiple literature databases, including PubMed, EBSCO, Scopus, PsycINFO and ISI Web of Science from their inception to July 2014. The search terms used were: 'pain drawing*' OR 'pain diagram*' OR 'pain measurement*' OR 'pain chart*' OR 'body chart*' OR 'body map*' OR 'mannequin*' OR 'manikin*' OR 'mannikin*' OR 'self-assessment*' combined with 'spinal pain' OR 'back pain' OR 'lumbar pain' OR 'lumbo-pelvic pain' OR 'lumbago' and with 'reliability' OR 'reproducibility of findings'. These keywords were identified after preliminary literature searches. Additional records were searched through other sources to complement the database findings. Two reviewers (AR, AR) independently applied the previously determined inclusion and exclusion criteria to select potentially relevant papers which were been initially identified based on title and abstract. Full-text copies of relevant trials were then obtained and independently evaluated by the reviewers. When a disagreement between reviewers occurred, it was resolved by a meeting held in consultation with another author (CV).

Types of studies − Published diagnostic studies without any restrictions on publication date or language were included.

Types of participants − The participants in selected studies had to be symptomatic adults, 18 years of age or older, with a diagnosis of subacute or chronic specific or non-specific LBP. Pain lasting between 1 and 3 months was categorized as "subacute," and as "chronic" when lasting for more than 3 months. In the absence of this explicit description, pain was considered subacute or chronic when the investigators themselves categorized a subject's pain in those terms. Trials were excluded if subjects with spinal pain without any dis-

tinction among cervical, thoracic or lumbar locations were included; if subjects had received surgery or if any of the participants had received a diagnosis such as myelopathy, fracture, infection, dystonia, tumor, inflammatory disease, or osteoporosis.

Index tests − Among all the types of PDs described in the literature, our analysis was limited to only those qualitative or semi-qualitative PDs that used the methods of Ransford or Udén, either exactly as originally described or with slight variation. These PDs have shown good intra- and inter-examiner reliability[3, 8, 19–23]. Studies concerning other kinds of PDs (e.g. quantitative or descriptive) were excluded. When trials used both Ransford's or Udén's methods and different methods (i.e. quantitavive PDs), only data from qualitative or semi-qualitative PDs were taken into consideration. Moreover, if localizing the site or the area of pain or formulating a prognosis were the main clinical purposes of the study, these trials were excluded.

Reference tests − A further criterion of inclusion was the comparison between PDs and measures of psychological distress. To be eligible for inclusion, studies had to assess anxiety or depression or the presence of an illness behavior according to Waddell's signs and symptoms[24]. Only studies using anxiety or depression scales which demonstrated good psychometric properties and Waddell's signs and symptoms were considered for inclusion. When studies used multidimensional scales, only subscales related to anxiety or depression were considered. Other outcome measures like performance tests, disability measures, medico-legal issues, or imaging were excluded.

Two authors (AR, AR) independently conducted data extraction. Three other authors (CV, LB, PP) were consulted in the case of persisting disagreement. Reviewers were not blinded to information regarding the authors, journal of origin, or the outcomes of r each paper reviewed. Using a standardized form, data extraction addressed participants, type of PD, reference tests, and the findings that were reported. The methodological quality of studies was assessed by two authors (JHV, CV) using the QUADAS-2 scale, which has been shown to be reliable and valid for rating the quality of diagnostic studies[25]. Trials were not excluded on the basis of quality.

Standard methods recommended for meta-analyses of diagnostic studies were used[26]. ProMeta V.2.0 and Review Manager V.5.2 software were used for the statistical analyses. For each study, computed measures of test accuracy using standard methods were used: Sensitivity, specificity, positive likelihood ratio (LR^+), negative likelihood ratio (LR^-) and diagnostic odds ratio (DOR)[26]. These measures were pooled using a random effects model to forecast the heterogeneity of the comparisons. Since sensitivity and specificity are correlated, their joint distributions were summarized using a summary receiver operating characteristic curve (SROC). SROC plots display the results of individual studies in ROC space with each study plotted as a single sensitivity-specificity point. The size of the points depicted the precision of the estimate, typically scaled according to the inverse of the standard error of the logit (sensitivity) and logit (specificity), or according to their sample sizes. The area under the SROC curve is a global measure of overall

performance. An area under the curve of 1 indicates perfect discriminatory ability[26, 27]. The diagnostic odds ratio (DOR) summarizes the diagnostic accuracy of the index test as a single number that describes how many times higher the odds are of obtaining a positive test result for a person with disease compared to a person without disease. Because the SROC curve and the DOR are not easy to interpret and use in clinical practice, likelihood ratios were also considered, since they are more clinically meaningful[27]. Likelihood ratios greater than 10 or less than 0.1 generate large and often conclusive changes from pre-test to post-test probability; likelihood ratios of 5 to 10 and 0.1 to 0.2 generate moderate shifts in pre-test to post-test probability; likelihood ratios of 2 to 5 and 0.5 to 0.2 generate small (but sometimes important) changes in probability; and likelihood ratios of 1 to 2 and 0.5 to 1 alter probability to a small (and rarely important) degree[28]. Heterogeneity was described using the Q and I-square statistics for pooling sensitivity, specificity, likelihood ratios and diagnostic odds ratio. A significant Q value indicates a lack of homogeneity of findings of studies, but considering the low number of studies for each outcome and knowing that this test is not very indicative in this case, a random-effects model was used. Publication bias was assessed using the Egger's weighted regression method[29] with precision (1/standard error) and log odds ratio plotted.

RESULTS

The database search identified 6,418 studies. Additional eligible studies (n=7) were identified through other sources. After removing duplicates, and then screening titles and abstracts of all the remaining unique articles, 15 full-text articles needed to be assessed to verify their eligibility for the inclusion in the present study. Ultimately, eight of these manuscripts were excluded for various reasons (surgical patients, pain not present at the time of the evaluation, general spinal pain, reference tests not evaluating psychological distress or illness behaviour, missing outcome data, or prognostic study). Thus, seven studies were finally selected for this review (Fig. 1)[30–36].

Overall, the seven papers reported on studies conducted in Europe (Germany, Sweden, UK), Asia (India), and North America (Canada, USA), and had been published between 1983 to 2005, with only 28.6% of them published after 2000. These studies involved a total of 1,622 patients who were enrolled and completed assessments (Table 1). Study samples ranged from 54 to 649 subjects with a mean sample size of 232 participants. The mean age of participants was specified only in four studies, and it was approximately 44.5 years (range: 41.9–47.0). The sex of the participants was indicated in 4 studies, in which a slight majority (56%) of the participants were male.

The quality of the eligible studies was assessed using the revised Quality Assessment for Studies of Diagnostic Accuracy (QUADAS-2) tools[27]. Overall, the quality of the studies was poor to moderate (Table 2). For the patient selection domain, three studies[30, 33, 34] had a high risk of bias because of inappropriate inclusion or exclusion criteria. Four studies[30, 33–35] raised substantial concerns regarding their applicability to our research question. For the index test
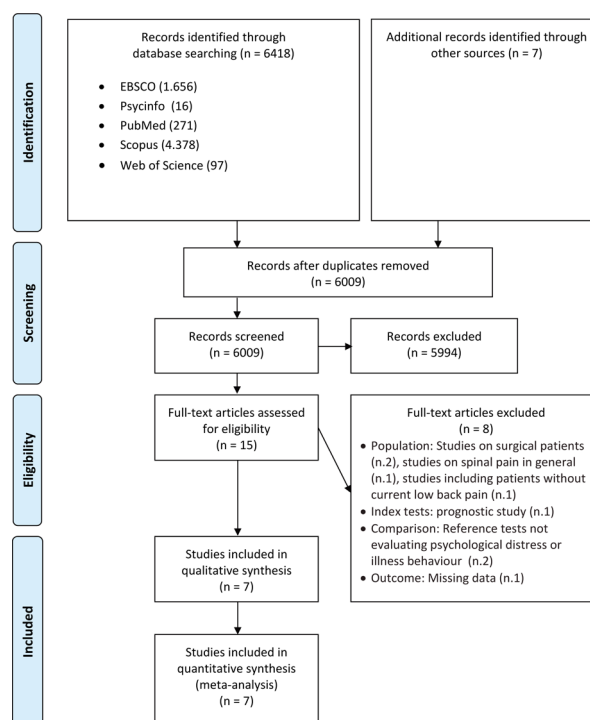
Fig. 1. Flow diagram of studies through the different phases of the review

domain, three studies[30, 34, 35] had a high risk of bias. None of the studies specified the threshold used. For two studies, there were strong concerns regarding the applicability of a study[30, 34] because variations in execution or interpretation of the PD may have affected estimates of its diagnostic accuracy. With respect to the QUADAS-2 criterion of risk of bias relative to the reference standard, the reference standard for all studies was considered to be appropriate. However, one study[31] raised concerns regarding the applicability of the reference standard to the research question. The most common methodological concerns was the failure to report patient flow and timing which occurred in 87% of the studies[30, 31, 33–36].

All seven studies, including 10 pair-wise comparisons, were included in the meta-analysis and their sensitivities and specificities are shown on the forest plot (Fig. 2). The forest plot shows the studies in order of sensitivity. Study-specific estimates of sensitivity and specificity are shown, with their 95% confidence intervals. Figure 2 shows the sensitivity and the specificity of each study with point estimates and confidence intervals. Sensitivity estimates were lower and more variable (range 4–86%) than those of specificity (range 40–100%). The area under the curve was 78% (CI, 57–99%), indicating that PDs demonstrated only a modestly acceptable discriminatory power to identify psychological distress in subacute and chronic LBP patients. Table 3 shows the summary measures of test accuracy with 95% CIs and tests for heterogeneity. Publication bias was not evident in the

Table 1. Characteristics of included studies

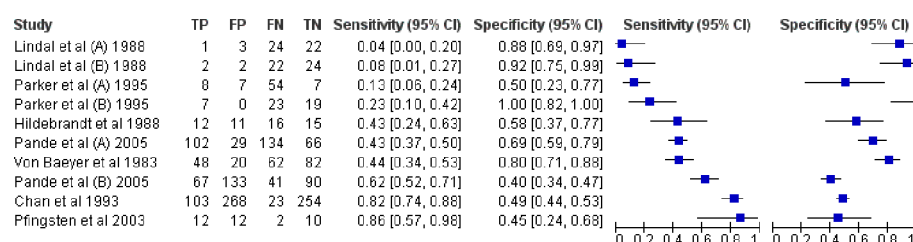| Author	Participants	Outcome measures	Reported results
Von Baeyer CL. (1983)[35]	212 chronic LBP patients. 2 samples: California (111), and Saskatchewan (101) heterogeneous as to diagnosis and chronicity of back pain.	1) Pain drawing by Ransford 2) MMPI (Minnesota Multiphasic Pain Inventory)	Not dichotomous data: higher correlation coefficient = 0.28 for the Hypochondriasis (Hs) scale (p<0.01) Dichotomous data: regarding Hs, 61% agreement of the cases overall, but only 44% of patients with elevated Hs score were correctly identified by their drawings scores.
Hildebrandt J. (1988)[31]	54 subjects (32 men, 22 women) randomly chosen among all chronic LBP patient visited in the authors' pain clinic within 1 month. Mean age: 46.4 years (men) and 44,5 years (women).	1) Pain drawing slightly changed from Ransford scoring method. 2) Erweiterte Revidierte Mehrdimensionale Schmerzskala (ERMSS)	No correlation between PDs scores (> or <2) and affective score of ERMSS.
Lindal E. (1988)[32]	Group 1: 54 subjects with LBP (22 men, 32 women) randomly contacted without any prior knowledge of their medical history. Mean age: 46.8 years. Group 2: 13 subjects with LBP scheduled for back surgery. Mean age: 51.1 years.	1) Pain drawing by Udén. 2) Rod-and-Frame Test (RFT) 3) Metacontrast Technique (MCT) 4) Eysenck Personality Inventory	PDs classified as non-organic and possibly non-organic more common in group 2 (p<0.01) No significant correlation between the abnormality of PD and depression (MCT).
Chan CW. (1993)[30]	49 subjects with chronic LBP (387 men, 264 women) classified as worker's compensation LBP patients (484) and as medicolegal LBP patients (167).	1) Pain drawing by Udén with patients sorted by individually pooled pain categories 2) Waddell score	51.3% of patients with Waddell score <2 had nonorganic PD; 94.6% of patients with Waddell scores = 5 had nonorganic PD; 81.7% of patients with high Waddell scores had nonorganic PD.
Parker H. (1995)[23]	Cohort 1: 100 subjects with LBP aged up to 65 years, without any major concurrent psychiatric illness, rheumatologic or neoplastic disease, or severe spinal structural abnormality. Cohort 2: 100 subjects with chronic LBP aged up to 65 years, without any major concurrent psychiatric illness, rheumatologic or neoplastic disease, or severe spinal structural abnormality.	1) Pain Drawing by Ransford, Pain Drawing by Margolis and Body Map Scoring System 2) Modified Somatic Perception Questionnaire (MSPQ) 3) Modified Zung (ZUNG) 4) Behavioral Signs (INSIGN) 5) Behavioral Symptoms (INSYMPT) 6) Distress and Risk Assessment Method (DRAM)	The Modified Ransford scores did not show consistent correlations with the measures of psychological distress.
Pfingsten M. (2003)[34]	109 chronic LBP patients (62 women 57 men), median age 41.9 years, mean duration of pain 49.9 months. Subsample composed of 36 patients: group 1 (clearly without psychological distress) = 22 patients; group 2 (clear psychological distress) = 14 patients.	1) Pain drawing by Ransford 2) Depression score (ADS) 3) Affective score (SES/A) 4) Other scales used: Numerical Rating Scale, Pain Disability Index	Not dichotomous data: mild correlation (of a score >2) with ADS (r= 0.15) and SES (r=0.27). Dichotomous data: Ransford's method showed sensitivity = 86% and specificity = 46%.
Pande KC (2005)[36]	331 patients (123 women, 208 men), with non-specific LBP, degenerative disc disease or lumbar stenosis. Mean age: 47.8± 15.1 (men) and 46.1± 13.4 (women)	1) Pain drawing by Ransford 2) Hospital Anxiety and Depression Scale (HADS)	1) Performance of Pain Drawing respect in identifying patients with any degree of depression: sensitivity = 62%, specificity = 40%, positive predictive value = 69 % 1) Performance of Pain Drawing respect in identifying patients with any degree of anxiety: sensitivity =43%, specificity = 70%, positive predictive value =78%

Table 2. Methodological quality of included studies. Tabular presentation of QUADAS-2 results

STUDY	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Chan, 1993	☹	☹	☺	☹	☹	☹	☺
Hildebrandt, 1987	☺	☺	☺	☹	☺	☺	☹
Lindal, 1998	☺	☺	☺	☺	☺	☺	☺
Pande, 2005	☺	☺	☺	☹	☺	☺	☺
Parker, 1995	☹	☺	☺	☹	☹	☺	☺
Pfingsten, 2003	☹	☹	☺	☹	☹	☹	☺
Van Baeyer, 1983	?	☹	☺	☹	☹	?	?

☺ Low Risk   ☹ High Risk   ? Unclear Risk

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Lindal et al (A) 1988	1	3	24	22	0.04 [0.00, 0.20]	0.88 [0.69, 0.97]		
Lindal et al (B) 1988	2	2	22	24	0.08 [0.01, 0.27]	0.92 [0.75, 0.99]		
Parker et al (A) 1995	8	7	54	7	0.13 [0.06, 0.24]	0.50 [0.23, 0.77]		
Parker et al (B) 1995	7	0	23	19	0.23 [0.10, 0.42]	1.00 [0.82, 1.00]		
Hildebrandt et al 1988	12	11	16	15	0.43 [0.24, 0.63]	0.58 [0.37, 0.77]		
Pande et al (A) 2005	102	29	134	66	0.43 [0.37, 0.50]	0.69 [0.59, 0.79]		
Von Baeyer et al 1983	48	20	62	82	0.44 [0.34, 0.53]	0.80 [0.71, 0.88]		
Pande et al (B) 2005	67	133	41	90	0.62 [0.52, 0.71]	0.40 [0.34, 0.47]		
Chan et al 1993	103	268	23	254	0.82 [0.74, 0.88]	0.49 [0.44, 0.53]		
Pfingsten et al 2003	12	12	2	10	0.86 [0.57, 0.98]	0.45 [0.24, 0.68]		

0 0.2 0.4 0.6 0.8 1   0 0.2 0.4 0.6 0.8 1

Fig. 2. Forest plot of the estimates of sensitivity and specificity of PDs in identifying psychological distress in subacute and chronic LBP

Each solid square indicates the point estimate of sensitivity and specificity calculated for each study with 95% confidence intervals. Studies are presented in descending order of sensitivity

Table 3. Summary measures of test accuracy for all studies and tests of heterogeneity

All comparisons (n=10)	Summary measure of test accuracy* (95% CI)	Test for heterogeneity[†] (95% CI) p value
Sensitivity	0.45 (0.34, 0.61)	*
Specificity	0.66 (0.53, 0.82)	*
LR⁺	1.23 (0.93, 1.62)	*
LR⁻	0.84 (0.70, 1.01)	*
DOR	1.46 (0.79, 2.68)	*

*Random effects model; CI: confidence interval; [†] Q test for heterogeneity;
LR⁺: positive likelihood ratio; LR⁻: negative likelihood ratio;
DOR: diagnostic odds ratio

meta-analysis of the seven studies (t = −1.09; p=0.31).

DISCUSSION

This study aimed to search and review the scientific evidence for the validity of qualitative PDs in identifying psychological distress in subjects with subacute or chronic LBP. Seven diagnostic studies with large sample sizes, mostly involving adults with non-specific LBP, were selected and subjected to meta-analysis. The summaries of sensitivity and specificity, the LR⁺ and LR⁻ values, as well as the large range of the results (sensitivity range = 4–86%; specificity range = 40–100%) do not allow us to confirm the supposed discriminative power of PDs without reservation. Moreover, the area under the ROC curve of 78% suggests that there is only modest validity in using PDs to detect anxiety, depression or illness behaviour in LBP patients. Furthermore, on the basis of the QUADAS-2 results, all of the data were drawn from poor to moderate quality studies.

Although the index tests the selected studies were homogeneous (i.e., only qualitative PDs evaluated according to Ransford or Udén were selected), very different reference tests were employed to detect psychological disturbances. In fact, the Waddell's score, the Minnesota Multiphasic Pain Inventory, the Erweiterte Revidierte Mehrdimensionale Schmerzskala, the Metacontrast Technique, the Modified Somatic Perception Questionnaire, the Modified Zung Questionnaire, the Behavioral Signs, the Behavioral Symptoms, the Distress and Risk Assessment Method, the Depression score, and the Affective score were all used in the selected studies. These different reference standards document anxiety, depression, or illness behaviour across a wide conceptual swath.

One potential limitation of our study concerns our decision to analyze non-homogeneous outcome measures together. This option was preferred by our team because the specific relevance of any single one of these mood or behavioral disturbances in inducing a shift from acute to chronic LBP, or in obstructing recovery, has been elusive[12]. Furthermore, two selected studies[32, 36] analyzed both anxiety and depression within their samples, and failed to find substantial differences in the accuracy of PDs in identifying anxious or depressed subjects. This evidence strongly suggests that the

diagnostic accuracy of PDs relative to psychological distress does not change whether the types of distress are assessed separately or together.

Our conclusion dawn from the few selected studies of PDs is similar to that of the previous systematic review conducted by Carnes et al.[14], who also considered the quality of the studies that provided the data for analysis. In the light of our results, the current literature does not support the broad and unqualified use of qualitative PDs to identify psychological distress in subacute or chronic LBP, despite the value PDs may have in assessing symptoms, pain patterns or other clinical phenomena. However, future meta-analyses using higher quality studies may lead to different conclusions.

ACKNOWLEDGEMENT

REFERENCES

1) Malliou P, Gioftsidou A, Beneka A, et al.: Measurements and evaluations in low back pain patients. Scand J Med Sci Sports, 2006, 16: 219–230. [Medline] [CrossRef]

2) Ferrari S, Manni T, Bonetti F, et al.: A literature review of clinical tests for lumbar instability in low back pain: validity and applicability in clinical practice. Chiropr Man Therap, 2015, 23: 14. [Medline] [CrossRef]

3) Hägg O, Fritzell P, Hedlund R, et al. Swedish Lumbar Spine Study: Pain-drawing does not predict the outcome of fusion surgery for chronic low-back pain: a report from the Swedish Lumbar Spine Study. Eur Spine J, 2003, 12: 2–11. [Medline]

4) Block AR, Ben-Porath YS, Marek RJ: Psychological risk factors for poor outcome of spine surgery and spinal cord stimulator implant: a review of the literature and their assessment with the MMPI-2-RF. Clin Neuropsychol, 2013, 27: 81–107. [Medline] [CrossRef]

5) Pillastrini P, Ferrari S, Rattin S, et al.: Exercise and tropism of the multifidus muscle in low back pain: a short review. J Phys Ther Sci, 2015, 27: 943–945. [Medline] [CrossRef]

6) Lee JS, Kim TH, Kim DY, et al.: Effects of selective exercise for the deep abdominal muscles and lumbar stabilization exercise on the thickness of the transversus abdominis and postural maintenance. J Phys Ther Sci, 2015, 27: 367–370. [Medline] [CrossRef]

7) Ransford A, Cairns D, Mooney V: The pain drawing as an aid to the psychological evaluation of patients with low back pain. Spine, 1976, 1: 1–127. [CrossRef]

8) Udén A, Aström M, Bergenudh H: Pain drawings in chronic back pain. Spine, 1988, 13: 389–392. [Medline] [CrossRef]

9) McWilliams LA, Goodwin RD, Cox BJ: Depression and anxiety associated with three pain conditions: results from a nationally representative sample. Pain, 2004, 111: 77–83. [Medline] [CrossRef]

10) Lee JC, Lee SK, Kim K: Comparison of abdominal muscle activity in relation to knee angles during abdominal drawing-in exercises using pressure biofeedback. J Phys Ther Sci, 2013, 25: 1255–1257. [Medline] [CrossRef]

11) Isgro M, Buraschi R, Barbieri C, et al.: Conservative management of degenerative disorders of the spine. J Neurosurg Sci, 2014, 58: 73–76. [Medline]

12) Foster NE, Thomas E, Bishop A, et al.: Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. Pain, 2010, 148: 398–406. [Medline] [CrossRef]

13) Park SD: Reliability of ultrasound imaging of the transversus deep abdominial, internal oblique and external oblique muscles of patients with low back pain performing the drawing-in maneuver. J Phys Ther Sci, 2013, 25: 845–847. [Medline] [CrossRef]

14) Carnes D, Ashby D, Underwood M: A systematic review of pain drawing literature: should pain drawings be used for psychologic screening? Clin J Pain, 2006, 22: 449–457. [Medline] [CrossRef]

15) Mannion AF, Balagué F, Pellisé F, et al.: Pain measurement in patients with low back pain. Nat Clin Pract Rheumatol, 2007, 3: 610–618. [Medline] [CrossRef]

16) Abbott JH, Foster M, Hamilton L, et al.: Validity of pain drawings for predicting psychological status outcome in patients with recurrent or chronic low back pain. J Man Manip Ther, 2015, 23: 12–19.

17) Hildebrandt J: Prediction of psychosocial factors by pain drawing in patients with chronic back pain. Pain Med, 2014, 15: 1067–1069. [Medline] [CrossRef]

18) Turk D, Melzack R: Handobook of Pain Assessment, 3rd ed. Guilford, 2010.

19) Andersen T, Christensen FB, Høy KW, et al.: The predictive value of pain drawings in lumbar spinal fusion surgery. Spine J, 2010, 10: 372–379. [Medline] [CrossRef]

20) Ohnmeiss DD: Repeatability of pain drawings in a low back pain population. Spine, 2000, 25: 980–988. [Medline] [CrossRef]

21) Ohnmeiss DD, Vanharanta H, Ekholm J: Relationship of pain drawings to invasive tests assessing intervertebral disc pathology. Eur Spine J, 1999, 8: 126–131. [Medline] [CrossRef]

22) Ohnmeiss DD, Vanharanta H, Guyer RD: The association between pain drawings and computed tomographic/discographic pain responses. Spine, 1995, 20: 729–733. [Medline] [CrossRef]

23) Regio S, Notaristefano T, Palminiello A: [Death conference: hindsight in heart surgery]. G Ital Cardiol, 1998, 28: 392–396 (Rome). [Medline]

24) Waddell G, McCulloch JA, Kummel E, et al.: Nonorganic physical signs in low-back pain. Spine, 1980, 5: 117–125. [Medline] [CrossRef]

25) Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2 Group: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med, 2011, 155: 529–536. [Medline] [CrossRef]

26) Irwig L, Macaskill P, Glasziou P, et al.: Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol, 1995, 48: 119–130, discussion 131–132. [Medline] [CrossRef]

27) Deeks JJ: Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. BMJ, 2001, 323: 157–162. [Medline] [CrossRef]

28) Jaeschke R, Guyatt G, Sackett DL: Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA, 1994, 271: 389–391. [Medline] [CrossRef]

29) Egger M, Davey Smith G, Schneider M, et al.: Bias in meta-analysis detected by a simple, graphical test. BMJ, 1997, 315: 629–634. [Medline] [CrossRef]

30) 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: 1717–1722. [Medline] [CrossRef]

31) Hildebrandt J, Franz CE, Choroba-Mehnen B, et al.: The use of pain drawings in screening for psychological involvement in complaints of low-back pain. Spine, 1988, 13: 681–685. [Medline] [CrossRef]

32) Líndal E, Udén A: Cognitive differentiation, back pain, and psychogenic pain drawings. Percept Mot Skills, 1988, 67: 835–845. [Medline] [CrossRef]

33) 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: 236–243. [Medline] [CrossRef]

34) Pfingsten M, Baller M, Liebeck H, et al.: [Psychometric properties of the pain drawing and the Ransford technique in patients with chronic low back pain]. Schmerz, 2003, 17: 332–340. [Medline] [CrossRef]

35) Von Baeyer CL, Bergstrom KJ, Brodwin MG, et al.: Invalid use of pain drawings in psychological screening of back pain patients. Pain, 1983, 16: 103–107. [Medline] [CrossRef]

36) Pande KC, Tripathi S, Kanoi R: Limited clinical utility of pain drawing in assessing patients with low back pain. J Spinal Disord Tech, 2005, 18: 160–162. [Medline] [CrossRef]