M. Grevitt
K. Pande
J. O'Dowd
J. Webb

Do first impressions count?

A comparison of subjective and psychologic assessment of spinal patients

Received: 27 September 1997
Revised: 10 January 1998
Accepted: 20 January 1998

M. Grevitt (✉) · K. Pande · J. Webb
Centre for Spinal Studies and Surgery,
University Hospital,
Queen's Medical Centre,
Nottingham NG7 2UH, UK
Tel. +44 0115 970 9761
Fax +44 0115 970 9991

J. O'Dowd
Frimley Park Hospital NHS Trust,
Portsmouth Road, Frimley Park,
Surrey, UK

Abstract Psychologic factors may have a major influence on the outcome of treatment for back pain. Psychologic disturbance is manifest as emotional distress and may be associated with inappropriate symptoms and signs. Few outcome studies describe the patient population in terms of their psychologic profile. Anecdotal evidence suggested that the routine use of psychologic screening tests in British spine practice was rare. An audit of the prevalent use of psychologic testing amongst a selected group of British spinal surgeons was conducted. This was followed by a prospective, double blind comparison of subjective evaluations of patients with formal psychologic tests. The principal aim was to determine how accurately treating physicians could identify psychologically distressed patients. A postal questionnaire was sent to all consultant members of the British Orthopaedic Spine Society. Details of their current practice and frequency of use of psychologic tests was obtained. In a subsequent study, 125 consecutive new patients attending a back pain clinic were initially evaluated by questionnaires and classified as either psychologically dis-

tressed or non-distressed. These patients were then interviewed and examined by treating physicians, who then allocated them to one of four psychologic categories, using predefined criteria. The two results were compared and sensitivity, specificity and predictive values for the subjective evaluations were calculated.

Sixty-three percent of respondents to the postal survey either never or only occasionally used any form of psychologic testing in assessing back pain patients. The follow-up prospective study demonstrated that experienced spinal surgeons achieved only a 26% sensitivity when trying to identify distressed patients. The specificity for identifying non-distressed patients was 96%. The predictive value of a "distressed" evaluation was 69%. The predictive value for non-distressed patients was 77%. Subjective psychologic assessment of back pain patients has a low sensitivity and predictive value for distressed patients. Formal psychologic screening should be routinely included in the clinical decision making process.

Key words Psychologic tests · Back pain · Patient assessment

Introduction

Chronic low back pain remains a poorly understood condition despite intensive research efforts. There is poor correlation between radiological abnormalities and physical disability; patients may respond very differently to the same pathological condition (e.g. lumbar disc herniation) [2, 3].

The medical outcomes study in the United States has highlighted alterations in general health status after successful treatment of back pain patients. Pain, physical disability and physical role limitation variables of the Short Form-36 questionnaire correlate well with established spinal outcome measures [6]. Mental health and general health perception are less influenced by treatment [10]. Although mental well being and health perception are not substantially altered by treatment of back pain, psychologic factors may nevertheless be a major influence on outcome after treatment. Reviews of failed back surgery emphasise the importance of these psychologic factors and of proper patient selection [14].

Psychologic distress is a disturbance of emotion and mood where both psychologic and physical symptoms may coexist. These symptoms are an exaggeration of the experience of normal individuals to stressful situations. Distress can be clinically manifest as inappropriate symptoms and signs [15]; it can be measured as increased somatic awareness and depression. The former may be assessed by the Modified Somatic Perception Questionnaire (MSPQ) [7]. Depression is a common psychologic finding in chronic back pain and may be quantified using the Zung Depression Index (ZDI) [18].

The Minnesota Multiphasic Personality Inventory (MMPI) has been used as the psychologic assessment tool in many previous studies of low back pain. The MMPI is lengthy and difficult to administer in a general clinical setting. In contrast, the Distress and Risk Assessment Method (DRAM) is easily scored and has been validated on several categories of back pain patient [8]. The DRAM utilises the MSPQ and Zung questionnaires. Patients identified as psychologically distressed using this method are three to four times more likely to have a poor outcome after any form of treatment.

Anecdotal reports suggest that few surgeons in British spinal practice use any formal psychologic assessment. The initial aim of this study was to assess the prevalence of psychologic testing amongst British spinal surgeons. Following this audit, we conducted a prospective study of subjective patient evaluations compared with formal psychologic tests within the setting of a spinal outpatients department. This involved a small group of surgeons from a single specialist spine centre. The main aim of the latter study was to determine how accurately individual doctors identified psychologically distressed patients.

Table 1 Postal questionnaire employed in the audit of use of psychologic tests in UK spinal practice

Q1. How many fusion / stabilisation operations for low back pain do you perform per year?

- ☐ None
- ☐ 10–20
- ☐ 21–50
- ☐ > 50

Q2. Do you use any form of psychometric or psychologic tests in your practice?

- ☐ Never
- ☐ Only occasionally in cases of clinical doubt
- ☐ Only in those patients considered surgical candidates
- ☐ In all patients

Q3. What instruments do you use?

- ☐ Pain diagram
- ☐ Beck Depression Inventory
- ☐ Zung Depression Index
- ☐ Modified Somatic Perception Questionnaire
- ☐ Waddell's inappropriate signs
- ☐ Minnesota Multiphasic Personality Inventory
- ☐ Others (please state below)

Methods

Audit of British practice

All consultant members of the British Orthopaedic Spine Society were sent a postal questionnaire that sought details of their current spinal practice. The society represents the majority of orthopaedic surgeons that have a declared spinal subspecialty interest. Most have a general orthopaedic practice but 14 members (23%) perform complex or major spinal surgery. The average age of this group is 50 years (95% CI 48–52 years); all members are male. Neurosurgeons are not members of this society.

Besides the number of spinal fusions performed for low back pain each year, each surgeon was asked how often or in what circumstances he employed any form of psychologic assessment. If appropriate, each respondent was then asked to indicate those measurement instruments or tests that were used (Table 1).

Comparison of subjective with psychological assessment

One hundred and twenty-five consecutive patients were entered into this study. All were new patients referred to either acute or chronic back pain clinics of the two senior authors (JO'D and JKW). At the initial visit and immediately prior to their consultation, each patient completed an MSPQ [6] and ZDI [18]. Patients were seen by either a consultant or a spinal fellow. The consultants were practising spinal surgeons; each fellow had a minimum of 6 months' specialist spinal experience. Each assessing physician was unaware of the results of the above questionnaires. Eight doctors participated in this study; all were males and the average age was 40 years (range 35–55 years). This was an unselected group of physicians that were normally employed in the clinics from which the study patients were drawn. The patients were randomly allocated to participating doctors. Each made an approximately equal number of evaluations and no patient was assessed twice.

Table 2 Psychologic categories: criteria for subjective assessment

1. Normal (abbreviation N)
 - No evidence of distress or abnormal illness behaviour
 - Optimistic about outcome
 - Readily accepts responsibility for improvement in back disability
2. At risk (R)
 - More distress, no illness behaviour
 - Reasonably optimistic and accepts advice proffered
 - May feel “down” occasionally but not clinically depressed
3. Distressed – depressive (DD)
 - Very distressed about back pain
 - Feels “gloomy” about outcome of treatment
 - Negative attitude to treatment plan proposed
 - Feels helpless to influence outcome – “doctor, do something to make me feel better”
 - Elements of clinical depression
4. Distressed – somatic (DS)
 - Very distressed
 - May demonstrate inappropriate symptoms/signs
 - Very descriptive language to describe pain
 - Other somatic symptoms
 - Pessimistic about outcome, feelings of helplessness

Derived from Main et al. [8]

Following a history and examination each doctor allocated the patient to one of four DRAM psychologic categories (normal, at risk, distressed–depressive, distressed–somatic). The criteria for each category were defined prior to the study; they were derived from Main et al. [8] (Table 2).

Patients below 18 years of age and those whose native language was not English were excluded. Patients with known psychiatric illness, tumours, infections or previous spinal surgery were not included in the study.

At the end of the consultations the completed questionnaires were scored by an independent assessor (KP). The definitive DRAM category for each patient was then derived on the basis of the MSPQ and ZDI scores.

The data were collated at the end of the study and the doctors' assessments were compared with the DRAM categories. To facilitate calculation of sensitivities and specificities the normal and at-risk categories were amalgamated as a “non-distressed” group, and the distressed–depressive and distressed–somatic categories combined as a “distressed” group. A true-positive result was defined as when the surgeon's assessment of the patient as being distressed was later matched by the definitive DRAM category. A true-negative result was when a subjective assessment as a non-distressed patient coincided with the final DRAM category. The sensitivity, specificity and predictive values were calculated. In this context, sensitivity is a measure of the doctors' ability to distinguish as positive those patients that are psychologically distressed. Specificity measures the ability to correctly identify those patients who are non-distressed. In contrast, the predictive value of an abnormal assessment (i.e. distressed patient) relates to the proportion of patients with these DRAM categories who are truly distressed.

The Wilcoxon rank sum test was used to compare groups; analysis of categorical data was by Fisher's exact probability test. The level of significance was taken as $p < 0.05$. All statistical calculations were performed using a software package (Minitab, v9.2 for Windows, State College, PA).

Fig. 1 Distribution of the number of fusion/stabilisation operations performed by each surgeon per year **Fig. 2** Frequency of use of psychologic tests in each surgeon's practice

Results

Audit of British practice

Sixty questionnaires were sent and 47 returned (78% response rate).

Figure 1 illustrates the distribution of the number of fusion operations performed by each surgeon. Figure 2 shows the frequency of use of psychometric testing in each consultant's practice; 29 (63%) never or only occasionally used any formal tests. The remainder used some form of psychologic test; the most commonly used were a pain drawing and inappropriate signs test (Table 3). Seven respondents (15%) referred patients to a clinical psychologist for formal evaluation.

Table 3 Psychologic tests in use by surgeons

	No. of surgeons (%)
Pain drawing	21 (58)
Beck Depression Index	1 (3)
Zung Depression Index	13 (36)
Modified Somatic Perception Questionnaire	11 (31)
Inappropriate Signs test	20 (56)
Minnesota Multiphasic Personality Inventory	4 (11)
Others ^a	5 (14)

^aHospital Anxiety Depression Scale, SM Scale, Pain Self Efficacy Questionnaire, General Hospital Questionnaire

Table 4 Results: summary of DRAM categories (*N* Normal, *R* at risk, *DD* distressed-depressive, *DS* distressed-isomantic)

Subjective assessment	Definitive DRAM category			
	N	R	DD	DS
N	24	31	10	3
R	12	19	9	4
DD	0	3	3	0
DS	0	1	3	3
Total (%)	36 (29)	54 (43)	25 (20)	10 (8)

There was no significant association between the number of fusions performed per year and the use of psychometric testing.

Comparison of subjective with psychologic assessment

There were 59 males and 66 females in the study; the mean age was 45 years (95% CI 43–48 years). Patients attending the acute back clinic had an average duration of symptoms of 4 months (range 2–6 months). The mean Oswestry Disability Score [4] was 45% (95% CI 35%–55%). Those attending the chronic back clinic had a mean symptom duration of 67 months (range 24–96 months). This latter group had a mean Oswestry Disability Score of 57% (95% CI 53%–62%). The patients attending the acute and chronic back pain clinics were compared; they did not differ in age and sex ratio and there was no significant difference in the prevalence of definitive DRAM categories. The distribution of subjective DRAM categories was similar amongst the participating surgeons.

The results from both clinics were pooled and the distribution of DRAM categories is summarised in Table 4. There were 9 true-positive, 86 true-negative, 4 false-positive and 26 false-negative results. This gave a sensitivity of 26% and specificity of 96%. For patients assessed as “distressed” the predictive value was 69%. The predictive value for non-distressed patients was 77%. Depression

was more likely to be missed than anxiety (overall accuracy 12% compared with 30% respectively).

Discussion

There is a high rate of mental distress associated with chronic back pain. One cross-sectional study reported that 77% of patients met lifetime diagnostic criteria of at least one psychiatric diagnosis, with 59% demonstrating current symptoms at the time of their initial consultation [11]. The latter figures should be contrasted with the fact that 63% of the respondents to our survey never, or only occasionally, used any psychologic testing.

The reasons for this infrequent usage of formal psychologic assessment were not explored in the questionnaire. Several respondents commented that widespread adoption of psychologic testing was hindered by the logistic problems of scoring questionnaires in a busy outpatients clinic. Intuitive interpretation of these scores can be difficult for the clinician, which limits their usefulness in the context of a consultation. Current psychologic tests used in the assessment of back pain patients vary in the complexity of the scoring methods. The pain drawing was the most frequently used test in this study; one reason may be that it is simple and quick to administer. As well as documenting the patient's symptoms the psychologic status may be inferred without formal scoring. The initial report of this test demonstrated a good correlation between the hypochondriasis (Hy) and hysteria (Hs) scales of the MMPI [12]. Other authors have not confirmed this relationship [1]. Using a new scoring method for the pain drawing Parker et al. [9] found the test to have poor sensitivity and specificity in identifying distressed patients or distinguishing organic versus non-organic pain patterns. Although many surgeons use the pain drawing it is a crude indicator of psychologic distress. These later studies do not support the usefulness of the pain drawing as the sole method of psychologic screening of back pain patients.

We tested an individual's ability to identify distressed patients in the second part of this study. The setting of the study was chosen to replicate as closely as possible the normal working practices of the unit. Each new patient completes an Oswestry Disability questionnaire and visual analogue pain scale prior to the consultation. The addition of the MSPQ and Zung questionnaires was not a great additional burden to the patients. These are well validated measures of anxiety and depression in back pain patients. The patients assessed in this study had a similar distribution of DRAM categories to a group attending a general back pain clinic [8].

All doctors were briefed about the study and the criteria for the DRAM categories. This introduced an element of observer bias to the study. In this respect each physician may have been more aware of psychologic distress in

their patient assessments. Our results should therefore represent the "best effort" in subjective evaluations. The results demonstrate that surgeons are generally good at identifying non-distressed patients in a given back pain group (specificity 96%). We did not perform reliability tests of our subjective evaluations. A test-retest method would have been logistically difficult and would have eliminated the "gut instinct" element of this study. In most instances of false-negative or false-positive results the disparity between the subjective and definitive DRAM categories was so great that we believe it unlikely that variation in one category in either direction would substantially alter these results.

Many of the patients referred to the chronic back pain clinic had significant disability and had failed to respond to a period of conservative treatment. These patients were potential candidates for spinal fusion. Spinal fusion for low back pain remains controversial and the results are unpredictable. In one study of workers' compensation cases there was only 50% pain reduction in less than half the patients; 30% returned to work, and the remainder received disability benefit after an average of two repeat operations [5]. There are many factors other than the technical competency of the surgeon that influence the outcome in spine surgery. A prospective evaluation of lumbar spine fusion showed strong correlation of outcome with aspects of the MMPI [13].

It is clear from the above that identification of distressed patients is as important as quantifying disability in assessment prior to spinal fusion. Given the low usage of psychologic tests in British spinal practice, many surgeons must rely on subjective evaluations prior to making management decisions. This study demonstrates that a group of experienced spine surgeons did not accurately identify distressed patients in a typical back pain population (sensitivity 26%). The predictive value for non-distressed patients was 77%. This implies that in a quarter of instances where a surgeon assesses a patient as psychologically "normal" he may be wrong. The latter figure is similar to the success rate quoted in the majority of the literature relating to outcome after spinal fusion [15]. We speculate whether failures in reported surgical series rep-

resent inadequate screening of distressed patients. There are few studies examining physicians' ability to identify psychologic distressed patients. Wiltse and Rocchio [16] published a double-blind study indicating that outcome of treatment for disc disease correlated best with the Hy and Hs scales of the MMPI. In this study the pre-operative ratings by the surgeon of the degree to which the patient's symptoms were psychogenic proved to have considerable prognostic value ($R = -0.51$). These findings conflict with our results but the psychologic grading system used was comparatively crude and had not been validated. Psychologically distressed patients have a worse outcome after treatment [8]. These individuals may be more appropriately treated by psychologic counselling and pain management programmes. We argue that surgery should be avoided in these individuals given the uncertain outcome and potential risks and complications of surgery. This implies that psychologic testing should be performed on all potential surgical candidates.

The prospective study involved a small number of physicians in a single spine practice and the generalisability of the findings may be questioned. The study was designed to reflect normal office practice; the results suggest the need for replication in other orthopaedic settings. Further studies may then provide answers as to how and by whom these patients are best evaluated.

Conclusions

An audit of British spinal practice revealed a low usage rate of psychologic testing when assessing back pain patients. A double-blind comparison of subjective assessment with psychologic testing showed a low sensitivity in identifying distressed patients. We believe that formal psychologic screening should be routinely included in the clinical decision-making process.

Acknowledgements We wish to thank the following colleagues who also participated in this study: Norbert Boos, Rabi Khazim, Damien McDermott and Wojcik Niezwywinski.

References

1. Baeyer CL von, Bergstrom KJ, Boodwin MG, Boodwin SK (1983) Invalid use of pain drawings in psychological screening of back pain patients. *Pain* 16: 103-107
2. Boos N, Reider R, Schade V, Spratt V, Semmer N, Aebi M (1995) The diagnostic accuracy of magnetic resonance imaging, work perception and psychosocial factors in identifying symptomatic disc herniations. *Spine* 20: 2613-2615
3. Deyo RA (1994) Magnetic resonance imaging of the lumbar spine. (Editorial) *N Engl J Med* 331: 115-116
4. Fairbank JCT, Cooper J, Davies JB, O'Brien JP (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66: 271-273
5. Franklin GM, Haug J, Heyer NJ, McKeeffrey SP, Picciano JF (1994) Outcome of lumbar fusion in Washington State Workers' Compensation. *Spine* 19: 1897-1904

6. Grevitt M, Khazim R, Webb J, Mulholland R, Shepperd J (1997) The Short Form-36 health survey questionnaire in spine surgery. *J Bone Joint Surg [Br]* 79B: 48–52
7. Main CJ (1983) The modified somatic perception questionnaire. *J Psychosom Res* 27: 503–514
8. Main CJ, Wood PL, Hollis S, Spanswick CS, Waddell G (1992) The distress and risk assessment method – a simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine* 17: 41–52
9. Parker H, Woods PLR, Main CJ (1995) The use of the pain drawing as a screening measure to predict psychological distress in chronic low back pain. *Spine* 20: 236–243
10. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB (1995) Assessing health related quality of life in patients with sciatica. *Spine* 20: 1899–1909
11. Polatin PB, Kirney RK, Gatchel RJ, Lillo E, Mayer TG (1985) Psychiatric illness and chronic low back pain. The mind and the spine – which goes first? *Spine* 10: 932–936
12. Ransford AO, Cairns D, Mooney V (1976) The pain drawing as an aid to the psychologic evaluation of patients with low back pain. *Spine* 1: 127–134
13. Turner RS, Lerding WC (1985) Correlation of the MMPI with lumbosacral spine fusion results: a prospective study. *Spine* 10: 932–936
14. Waddell G, Kummell EG, Lotto WN et al (1979) Failed lumbar disc surgery and repeat surgery following industrial injuries. *J Bone Joint Surg [Am]* 61A: 201–207
15. Waddell G, Main CJ, Morris EW, Di Paola M, Gray ICM (1984) Chronic low back pain, psychological distress and illness behaviour. *Spine* 9: 209–213
16. Wiltse LL, Rocchio PD (1975) Preoperative psychological tests as predictors of success of chemonucleolysis in the treatment of the low back pain syndrome. *J Bone Joint Surg [Am]* 57A: 478–483
17. Zdeblick TA (1995) The treatment of degenerative lumbar disorders – a critical review of the literature. *Spine* 20: 126–137 S
18. Zung WWK (1965) A self-rating depression scale. *Arch Gen Psychiatry* 32: 63–70

REVIEWER'S COMMENT

The authors present results from a specialist back clinic in a secondary care setting, finding 35 patients (28%) to be psychologically distressed. This figure is similar to the 23% found by Greenough and Fraser [1] and the 29% found by Main et al. [2]. The experienced clinicians in the study by Grevitt et al., using a subjective evaluation, diagnosed distress in only 13 patients (10%) and were significantly more likely to miss depression than somatisation ($P < 0.05$).

This important study tells clinicians that relying on subjective assessment alone will lead to an under-diagno-

sis of distress (an extremely important prognostic factor in the treatment of low back pain) and particularly an under-diagnosis of depression, which is a treatable condition. Failure to diagnose depression may be an aetiological factor in failed back surgery. All clinicians involved in the treatment of low back pain should use an instrument to formally assess psychological distress in their patients.

References

1. Greenough CG, Fraser RF (1991) Comparison of eight psychometric instruments in unselected patients with back pain. *Spine* 16: 1068–1074
2. Main CJ, Wood PLR, Hollis S, Spanswick CC, Waddell G (1992) The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate risk of poor outcome. *Spine* 17: 42–52

C. G. Greenough
Consultant Orthopaedic Surgeon
Middlesbrough General Hospital
Ayresome Green Lane
Middlesbrough, Cleveland TS5 5AZ, UK