Contents

Abstract

This study investigated the factors affecting recovery from low back injury. The study population was designed to be representative of patients presenting with back pain following soft tissue injury. A minimum follow up rate of 90 percent was achieved with a length of elapsed time from the injury of up to five years.

Three established methods of assessment of disability and functional capacity (The Oswestry disability scale, The Waddell disability scale and The Waddell physical impairment rating) were compared to each other and a new scale (The Low-Back Outcome Scale) for the measurement of a patient's performance in employment, social activities and activities of daily living. The relationships of these scales to employment, psychological disturbance and other factors were defined. The Outcome Score designed for use in this study provided a comprehensive and discriminating assessment of patient function.

Compensation (particularly lump sum claims), psychological disturbance at review, time off work, and age at injury were important factors in recovery; the diagnosis, type and severity of injury, migrant status and neurological deficits were not.

Eight psychometric instruments were examined, and the combination of The Zung Depression Scale and The Modified Somatic Perception Questionnaire was found to be the most accurate in determining the presence of psychological disturbance in patients with low back pain.

Correspondence

Charles G. Greenough M.D., F.R.C.S.,
Consultant Orthopaedic Surgeon,
Middlesbrough General Hospital,
Ayresome Green Lane,
Middlesbrough,
Cleveland TS5 5AZ, U.K.
Tel +44–642 850222 ext 5811
Fax –942 850710

Introduction

In the United Kingdom (UK) half of a randomised sample of the entire population had at one time or another suffered from back pain, and one fifth of those reporting pain had had time off work in the preceding year (Fairbank 1986). Amongst 2,000 cases of compensated back injuries in Canada the average time off work was 74 days, and 10 percent were off work for more than 6 months (Rossignol et al. 1988). In 1987 in the USA, 5 million people were disabled due to low back pain and a further 12 million had impairment; compensation and medical costs due to back pain in 1984 were $7 billion (Frymoyer and Cats-Baril 1987).

The mechanism of injury is variable. In Denmark the most common injury declared in compensation cases in men was object handling (Josefsen et al. 1987), but in the United Kingdom falls were the most frequent (Troup et al. 1981). Certain types of injury such as jerking or slipping injuries increase the risk of injury and the risk of disability (Magora 1973, Molumphy et al. 1985, Kelsey et al. 1984, Manning and Shannon 1981).

Many previous investigators have drawn patients from pain clinics or from current attenders in orthopaedic or spinal clinics; as a consequence their results were heavily weighted to patients with chronic pain and often failed to include any patients who had recovered from a low back injury. In an attempt to study a cohort of patients taken from general practice, Piterman and Dunt (1987) achieved a follow-up rate of only 63 percent.

In patients with back pain, the apparent results of treatment are influenced by the outcome measure used (Howe and Frymoyer 1985). Subjective measures gave a higher proportion of successful results than objective measures. Further, disability measured by employment capacity, sporting ability, or social function may not directly equate with bodily impairment. A grossly impaired paraplegic may function well in his work and social life, where a "back injured" patient with far less impairment may function poorly.

Two disability scores for the assessment of patients with back pain have been published (Fairbank 1980, Waddell and Main 1984), but not compared with one another. Both systems are self report scores and are thus affected by the patient's perception of pain and his desire to undertake the activities indicated. Neither include employment or more complex social activities. A physical impairment measure was devised by Waddell and Main (1984) in an attempt to measure bodily impairment directly. It may, however, be influenced by the co-operation of the patient during the physical examination, and the reliance on examination findings also introduces the possibility of observer error.

A number of factors may influence the morbidity of injuries to the back, but there is no general agreement on their relative importance. The severity of the injury and the initial treatment would seem to be of importance, but in most studies this has not been assessed. Social grouping and immigrant status exert an influence on recovery from low back injuries, although the association of immigrants with prolonged recovery periods may be due to the large numbers of manual workers in this group (Hewson et al. 1987), but this has not been confirmed. Compensation has been regarded as an important factor in the recovery from low back pain since Rigler coined the term "compensation neurosis" in 1879 (Parker 1977). Compensation influences the results of treatment programs (Krusen and Ford 1958). The effect of settlement on recovery is controversial. Miller (1961) claimed that many of the compensation patients will lose their psychological symptoms and return to work shortly after settlement. Other authors have disputed this view (Mendelson 1982, Tarsh and Royston 1985).

Psychological disturbance has been noted in many patients with back pain (Dzioba and Doxey 1984, Mendelson 1984, Waddell et al. 1984) and may impede recovery. A relationship between psychological disturbance and litigation has been proposed, although the incidence varies with the framework of the observer (Miller 1961, Parker 1977, Repko and Cooper 1983).

Many of the psychometric instruments in current use were not devised in back pain populations (Zung 1965, Zigmond and Snaith 1983, Pilowsky and Spence 1983) and although many have been tested in such populations most of these validation studies were performed in patients with chronic back pain (McCreary et al. 1979, Pheasent et al. 1979, Ransford et al. 1976), and no study has included patients who have recovered from their back pain. The assessment of psychometric instruments in patients with pain or illness is also made difficult as physical limitations may be confused with psychological symptoms in some questions (Watson 1982, Zigmond and Snaith 1983).

The purpose of this investigation was to study in a controlled fashion the factors affecting recovery from low back injury. It was hoped to provide information on the clinical assessment of physical and psychological disability after low back injury, and to provide data from a controlled group of patients on the natural history of such injuries. It was anticipated that this data will allow the determination of the role of the various factors in the recovery from low back injury.

Some of the data have been published by Greenough and Fraser (1989, 1991, 1992)

Methods

Study design

The study was a retrospective cohort study which was controlled for the presence of compensation. A cohort of patients was included from each of the years 1983 to 1987. The period of review was from December 1987 to June 1988, thus cohorts with follow up of one, two, three, four, and five years were generated. Each cohort contained an equal number of men and women, each with an equal number of compensation and non-compensation cases. Only patients who definitely related their back pain to a specific incident were enrolled.

The study was based on the practice of a single orthopaedic surgeon in Adelaide, Australia, RD Fraser, with a predominant interest in spinal surgery. The patients were referred from a number of sources. The source of referral showed a different pattern in the compensation patients as compared with the controls, and between men and women (Table 1).

Commencing with the first of February for each of the five years 1983–1987, the names of consecutive patients with a complaint of back pain seen for the first time (the primary consultation) were obtained from the daily consultation records. The notes of these patients were then examined. Plain radiographs of the lumbosacral spine were available for all patients at the primary consultation. Throughout the period from which the patients were selected the treating surgeon had used a proforma history and examination record for patients with back pain. These records contained a questionnaire which had been completed by the patient on his arrival for primary consultation consultation and had subsequently been used by the surgeon to record other facets of the history. The examination record consisted of a second proforma completed for each patient. Only first time attenders who definitely related their back pain to a specific incident and whose pain started within one week of this incident were eligible for entry into the study. Patients were then excluded if the medical records of the primary consultation were incomplete, if they had sustained any fracture or dislocation of the spine, had undergone spinal surgery subsequent to the injury or were not of working age at the time of injury. Working age was defined as 18–60 years for women, and 18–65 for men. In addition patients living more than 100 km from Adelaide at the time of injury were excluded.

Of the primary consultations, 30 percent were compensation patients and the rest were private. In the compensation cases 52 percent of the consultations were eligible for inclusion, but in the noncompensation patients only 14 percent. This difference was wholly explained by the lack of an identifiable incident in 44 percent of non compensation patients but only in 5 percent of compensation patients. The other exclusions comprised patients who had had surgery (15%), missing records (13%), patients who did not have back pain (7%), patients living more than 100 km from Adelaide (6%) and patients not of the correct age (3%). Men comprised 56 percent of the patients but whereas in the noncompensation patients the sexes were evenly divided, in the compensation cases men comprised 65 percent of the consultations. There was no difference between the sexes in the eligibility for inclusion in either the compensation or the noncompensation group. Over the patient selection period there were no changes in eligibility pattern except that missing records were more common in the earlier years.

Eligible patients from each year were classified into one of four categories: male compensation, male non-compensation, female compensation, female non-compensation. The first 15 patients in each category were forwarded into the study. Thus 60 patients were selected from each of 5 years, making 300 patients in all. To achieve this, 2,857 case records were examined.

Table 1. Source of referral of the study population. Number, *percentages*

Referral	Comp.	Noncomp.	Men	Women	All patients	
General practitioner	88	114	100	102	202	*70*
Other orthopaedic surgeon	27	13	24	16	40	*14*
Second opinion	24	7	12	19	31	*11*
Miscellaneous sources	6	8	6	8	14	*5*
Sum	145	142	142	145	287	

Information from the primary consultation

On the proforma history card the distribution of the pain on presentation was recorded by the selection of one of five choices: back, hip/buttock, leg, all of these, and neck. There was a small overrepresentation of initial pain in the back and leg in the compensation group (Table 2).

The type of injury, nature of the pain, aggravating and relieving factors, ability to work, play sport or exercise, and the need for rest each day were recorded. The initial treatment(s) received following the injury and before the primary consultation were also recorded. The most frequent was physiotherapy (90% of patients), followed by chiropractic treatment (40%), surgical corset (20%), hospital bed rest (17%), traction in hospital (15%), acupuncture (12%) and epidural steroids (9%). The median number of therapies received, 2 (0–6), was the same for the compensation and noncompensation patients.

There was a difference between the compensation and noncompensation patients in the provider of the initial treatment of their back injury (Table 3). The most common provider of initial treatment was the general practitioner in both groups.

There was also a difference in the timing of the initial therapy. 66 (53 %) of the compensation patients sought treatment within a day of the accident, compared with 28 (22 %) of the noncompensation patients (Table 4). There were no differences in the timing of initial therapy between the sexes in either group.

The examination findings were recorded on a second proforma. Build and posture were noted, together with tenderness and the range of spinal movement. A full neurological examination was recorded. Neurological deficit was defined as the loss of a reflex or weakness of a specific muscle or group of muscles but not loss of sensation. 59 patients (39 women and 20 men; $P < 0.01$) had neurological deficits at the time of primary consultation without differences in distribution between compensation and noncompensation patients (Table 5).

At the primary consultation two psychometric instruments were employed. The patient completed a pain drawing (Ransford et al. 1976) and the non-organic physical signs of Waddell et al. (1980) were also recorded.

Table 2. Initial pain distribution. Number, *percentages*

Pain	Comp.	Noncomp.	Men	Women	All patients	
Back	35	46	43	38	81	*28*
Hip	30	25	29	26	55	*19*
Leg	36	46	41	41	82	*29*
All	42	25	27	40	67	*23*
Neck	2	0	2	0	2	*1*

$P < 0.05$, Chi-square

Table 3. Provider of initial therapy following injury. Number, *percentages*

Provider	Comp.	Noncomp.	Total	
General practitioner	82	99	181	*70*
Casualty department	22	9	31	*12*
Physiotherapist	2	12	14	*5.5*
Chiropractor	6	7	13	*5*
Works medical officer	9	0	9	*3.5*
Works nurse	8	1	9	*3.5*
Flying doctor	1	0	1	*0.5*

$P < 0.0001$, Chi-square

Table 4. Timing of initial therapy. Number, *percentages and cumulative percentages*

Initial therapy	Comp.			Noncomp.		
Immediate	32	*26*	*26*	17	*13*	*13*
Same day	34	*27*	*53*	11	*9*	*22*
Following day	28	*23*	*76*	21	*17*	*39*
Within first week	21	*17*	*93*	31	*25*	*64*
Within first month	6	*5*	*98*	13	*10*	*74*
After first month	3	*2*	*100*	32	*26*	*100*

$P < 0.0001$, Chi-square

Table 5. Diagnosis and neurological deficit. Number, *percentages*

Diagnosis	n	With neurological deficit	
Disc prolapse	43	19	*44*
Disc. pain	131	22	*17*
Spondylol.	19	4	*21*
Facet arthrosis	43	2	*5*
Soft tissue str.	31	5	*16*
Other	24	7	*29*
All patients	287	59	*21*

Review

Outcome measures. The following standard outcome assessments were completed for each patient: The Waddell Disability Score (Waddell and Main 1984), the Oswestry Disability Score (Fairbank et al. 1980), and the Physical Impairment Rating (Waddell and Main 1984). These three instruments were devised for the measurement of disability and impairment in low back

Table 6. Low-back Outcome Scale®. Maximum 75 points

		score
Factors scoring 9 points		
Current pain (visual analogue)	7–10	0
	5–6	3
	3–4	6
	0–2	9
Employment (Housewives related to previous abilities)	unemployed	0
	part time	3
	full time lighter work	6
	full time original work	9
Domestic chores /"odd jobs"	none	0
	a few but not many	3
	most, or all but more slowly	6
	normally	9
Sport/active social (dancing)	none	0
	some, much less than before	3
	almost as much as before	6
	back to previous level	9
Factors scoring 6 points		
Resting	resting more than half the day	0
	resting about half each day	2
	little rest needed, occasionally	4
	no need to rest	6
Treatment/consultation	more than once a month	0
	about once a month	2
	rarely	4
	never	6
Analgesia	several times each day	0
	almost every day	2
	occasionally	4
	never	6
Sex life	severely affected / impossible	0
	moderately affected / difficult	2
	mildly affected	4
	unaffected	6
Factors scoring 3 points		
Sleeping	severely affected / impossible	0
Walking	moderately affected / difficult	1
Sitting	mildly affected	2
Travelling	unaffected	3
Dressing		

Table 7. Short Outcome Scale. Maximum 26 points

Current pain	continuous	0
	intermittent	6
Employment	not able to work	0
	able to work	6
Resting	resting more than half the day	0
	resting about half each day	1
	little rest needed, occasionally	2
	no need to rest	3
Three factors scoring 2 points each		
Sport	unable to undertake	0
Exercises	able to undertake	2
Having fun		
Five factors scoring 1 point each		
Sitting	makes the pain worse	0
Standing	does not make the pain worse	1
Walking		
Lifting		
Twisting		

pain populations and have been validated in previous studies (Fairbank et al. 1980, Waddell et al. 1982, Waddell and Main 1984, Waddell et al. 1986, Triano and Schultz 1987, Waddell 1987).

In addition a new Outcome Scale was devised and applied to each patient. This scale is scored from 0 to 75 (Table 6). Each factor is scored on a four point scale to avoid an average option. Pain and active activities are weighted more than treatments and rest required, which in turn are weighted more than passive activities. An important feature of the scale is the attempt to relate the present abilities to previous capacity. In sport for example a regional player who is able only to resume play at a social level cannot be said to have recovered. It is important to note that the Outcome Scale has an inverse relationship to the foregoing disability and impairment scales; increasing disability on the Outcome Scale is indicated by a reduction in the score.

In order to make use of the data from the primary consultation a short Outcome Scale was devised (Table 7). Questions on the proforma completed by the patient at the primary consultation were repeated in exactly the same format in the questionnaire used at review.

Examination. The examination was undertaken under standardised conditions using a proforma recording system, including a clinical examination of the neurological system of the lower trunk and legs. Lumbar flexion was measured as the increase in length in centimetres during full flexion of a 15 cm vertical line drawn in the mid line and based 5 cm distal to the dimples of Venus after Rae et al. (1984). The inappropriate signs of Waddell were recorded.

Follow up. Follow-up intervals for each patient at review were calculated from the date of the accident, and where applicable from the date of settlement of the compensation claim. The time of attendance for each patient was on average 80 minutes. The patients were interviewed and examined by the investigator who, as he at no time had been involved in the patients' treatment, constituted an unbiased observer.

The importance of a complete follow up was recognised. 31 patients had changed address since being seen. At the end of the study all but five of these patients had been located. Patients were contacted twice by letter and then by telephone. 192 patients (64%) attended after the first letter (in 5 patients who did attend it was not recorded at which stage they attended). 98 patients (33%) did not respond to the first letter. Of these, 22 responded to the second letter and 55 (18%) attended after being contacted by telephone. 3 patients returned the questionnaire but would not attend, and 10 of the 18 patients who were contacted

but refused to take part were prepared to answer some questions on the telephone. In total 274 patients (91%) attended for review and a further 13 completed written or telephone questionnaires (total 96%). The information available on the latter 13 patients was incomplete, therefore in the tables the number of patients varies slightly according to the parameter being illustrated.

The median follow up from the accident for the male compensation group was 52 (13–192) months and for the male noncompensation group was 56 (12–332) months. For the women it was 47 (16–156) and 53 (9–300) months, respectively.

Interval. The median interval between the incident and the primary consultation for all patients was 10 (0.5–294) months with no differences between the sexes or between compensation groups. The source of referral had no influence on the median interval between the injury and the primary consultation.

Age at injury. The median age at was 34 (18–64) years for men and 38 (18–60) for women with no differences between compensation groups.

Pain. The current pain distribution, aggravating and relieving factors were noted. Pain was recorded using a standard 10 cm visual analogue system with anchors of *no pain at all* and *worst back pain imaginable.*

Type of injury

The type of injury was classified without regard to the degree of violence (Table 8). Although no differences were observed between the sexes, there was a difference in distribution between the compensation and the control groups ($P < 0.001$, Chi-square).

Slip. The patient slipped on a surface but not down a step or drop, without falling to the ground.

Impact. The patient suffered a direct impact to the back. This could be due to direct trauma where the patient was struck by an object, or indirect trauma for example in a rear end motor vehicle collision.

Jerk. The patient received an unexpected jerk or strain during an activity, usually lifting or pulling. Examples included lifting an object when some other support suddenly gave way, or pulling a crow bar which slipped out of position.

Lift. The incident occurred during a lifting procedure but excluding any sudden jerk.

Twist. This included any form of twisting injury but excluded lifting objects from one side (classified under lifting). Examples included reaching for a map on the back seat of a car, twisting to reach into a confined space under a sink and playing certain ball games e.g. squash (without collision).

Fall. This was defined as any unexpected drop from any height. The patient did not have to fall bodily to the ground. This category included stepping off kerbs etc. which had not been observed, jumping down a drop which was larger than expected, falling off ladders, machinery etc. The criterion used was that there should have been a definite physical jolt.

Bend. These incidents occurred during bending or stooping with no more than a trivial load. Examples in this category included bending to pick up small objects (cutlery, clothing) or to tie shoelaces.

Others. Six injuries could not be classified under these headings. These included simple jogging, running for a bus, the lotus position in yoga and sexual intercourse.

Table 8. Types of injury sustained by the study population related to compensation and sex. Number, *percentages*

Injury	Comp.	Noncomp.	Men	Women	All patients	
Slip	8	1	5	4	9	*3*
Impact	30	5	15	20	35	*12*
Jerk	16	7	13	10	23	*8*
Lift	57	36	47	46	93	*32*
Twist	11	24	20	15	35	*12*
Fall	15	35	27	23	50	*18*
Bend	8	28	12	24	38	*13*
Other	0	6	3	3	6	*2*

Diagnosis

From the patient's records, the diagnoses made following the primary consultation were recorded. The diagnoses were based on the following criteria:

Prolapsed intervertebral disc. A prolapsed intervertebral disc was suspected in patients with sciatica predominating over back pain, with evidence of nerve root tension (Lasegue test, crossed straight leg raise test, femoral stretch test) and often with signs of nerve root dysfunction (loss of a reflex, or weakness, or wasting of a specific muscle or muscle group). The diagnosis, however, was only made where these signs were confirmed by a CT scan demonstrating prolapse affecting an appropriate root.

Discogenic pain. This was a syndrome of "mechanical" pain associated with symptoms of instability. Mechanical pain was defined as episodic severe pain caused often by sudden movements or jarring, or unguarded twisting or bending. This was not associated with root compression, and prolapse was absent. It would be accompanied by some symptoms of facet joint pain in many cases. Plain radiography revealed either no abnormality or some slight reduction in disc height with or without minor retrolisthesis or very early spondylophyte formation. Flexion and extension views were not routinely performed. Discography, which was

only performed when surgery was contemplated, demonstrated degenerative changes together with reproduction of the patient's pain.

Spondylolysis and spondylolisthesis. The fracture of the pars interarticularis in spondylolysis was demonstrated on oblique projections. Spondylolisthesis was determined from the lateral projection of the lumbosacral spine. On examination there was evidence of an extension catch. In most of these cases there were no symptoms or signs of root compression. Where root compression was suspected from the physical signs, defined in the same way as for prolapsed intervertebral disc, it was confirmed only if a CT scan demonstrated root encroachment at the affected level.

Facet joint pain. This was a syndrome of back pain which was worse after but not usually during exercise, with stiffening of the back on sitting or first thing in the morning, some radiation to buttock, thigh or occasionally lower leg and with more pain on extension or on straightening from a flexed position (Jackson et al. 1988). There was no evidence of root compression or tension. Radiographic confirmation of sclerosis, loss of joint space, osteophytosis and cyst formation was often noted on plain films or CT scans although not a requirement.

Soft tissue strain. This included those patients in whom the pain appeared to be posturally related. Examination, plain radiographs and CT scans (when performed) were normal, and the patients responded usually to increased fitness and a change in working or habitual posture.

Miscellaneous diagnoses. These comprised nerve root irritation, osteochondritis, pseudarthrosis in a previous fusion and lateral stenosis.

There were no differences in the distribution of the various diagnoses related to gender or compensation. Half the patients had discogenic (or mechanical) pain, 15 percent had facet joint arthropathy and 14 percent were shown to have a disc prolapse, 11 percent were considered to have soft tissue strain, 6 percent spondylolisthesis or spondylolysis, and the remaining 8 percent had a number of miscellaneous diagnoses.

Table 9. Injury Severity Score

Mild	*score 1*
Bending	
Lifting a light object	
Running	
Moderate	*score 2*
Bending or twisting with movement	
Lifting a moderate weight (20–40 lbs.)	
Pushing or pulling a load	
Jolt or drop of less than three feet	
Struck in back by a light object	
Rear end collision with minor car damage	
Moderately severe	*score 3*
Lifting a heavy weight	
Jerked while lifting a moderate weight	
Jerked while pushing or pulling	
Fall four to six feet	
Rear end collision with major car damage	
Severe	*score 4*
Jerked while lifting a heavy weight	
Maximum pull with a jerk	
Fall more than six feet	
Struck in back by heavy object	
High speed vehicular injury (without fracture)	

Injury severity. An injury severity score was assigned to each patient (Table 9). In this table it was recognised that it was not possible to quantify the energy absorbed by the spine in any particular accident. In addition as this was a retrospective study it was not reasonable to attempt grade an incident which might have occurred some years before. The severity score does serve, however, to separate the trivial episode from the major accident. The individual definitions were designed to allow the easy assignment of an injury severity score to the common types of injury.

In the men the average scores were 2.43 ± 0.77 and 2.41 ± 0.80 and in women 2.09 ± 0.50 and 1.91 ± 0.65 for the the compensation and noncompensation patients, respectively. There was some decrease of severity with age in the men, but not in the women (Table 10).

Table 10. Variation of Injury Severity Score with age. Number, mean score and *SD*

Age groups:		18–30			31–45			46–65		
Men	Comp.	20	2.75	*0.85*	30	2.27	*0.65**	15	2.40	*0.83*
Men	Noncomp.	30	2.66	*0.80*	35	2.26	*0.74**	8	2.13	*0.83*
Women	Comp.	21	2.14	*0.48*	29	1.97	*0.50*	20	2.20	*0.52*
Women	Noncomp.	25	1.92	*0.64*	33	1.85	*0.67*	12	2.08	*0.67*

*Difference from 18–30 group, $P < 0.05$

Table 11. Treatment following primary consultation related to diagnosis (some patients had more than one treatment)

Diagnosis	Classes	Excerc.	Facet inj.	Epid.	Surgery (refused)
Disc prolapse	4	25	0	4	13
Discogenic pain	40	97	6	9	3
Spondylol.	6	13	2	1	2
Facet arthropathy	6	21	22	1	0
Soft tissue strain	5	23	1	1	0
Miscellaneous	1	10	0	9	0

Table 12. The distribution of compensation and noncompensation patients between social groups for men and women

Social group	Men		Women	
	Comp.	Noncomp.	Comp.	Noncomp.
1	0	3	0	3
2	2	4	3	2
3	11	31	29	17
4	25	23	15	38
5	18	7	11	9
6	15	2	13	1

Treatment following primary consultation (Table 11). Education classes were run by a physiotherapist and consisted of four tutorials on the anatomy and physiology of the spine, the causes of low back pain, modifications of the activities of daily living, and discussion of the relationship of back pain to stress. All patients attending the classes also undertook the exercise regime. These exercises comprised walking or swimming, at least three times per week for approximately one hour. These exercises were supervised by the patients themselves.

Facet joint injections were performed under image intensifier control using 0.5% Marcaine and 100 mg hydrocortisone. Epidurals were performed by an anaesthetist at a low lumbar interspace using 0.25% Marcaine and a depot preparation of steroid. Disc excision was recommended for a number of patients with disc prolapse and fusion for a few with discogenic pain or spondylolysis. In the patients in this study the proposed surgery had been refused.

A number of patients received more than one treatment modality. There were no differences in the treatments given for each diagnosis between the compensation groups.

Employment. The patient was asked for his current employment, whether this was part time or full time, and whether this job was different from his employment at the time of the injury. If the patient had changed employment following the injury it was assessed with the patient during the interview whether the new job was lighter than the original. The patient was also asked to state how many days on average he would lose off work each month due to back pain. The time off work was defined as the time taken to return to work or to formal termination of employment. The date of formal termination was checked with the patient during the interview.

Social group. Occupation at the time of injury was obtained and this was checked during the interview. The patients were then assigned to socio-economic groups according to the Ann Daniel Scale which was constructed for the Australian population and is a standard instrument for classification in Australia (Daniel 1983). To simplify analysis, Daniel's six main divisions were combined into three groups chosen in order to combine levels of education and physical demand which were approximately similar. Daniel groups 1 to 3 were combined, group 4 remained separate, and groups 5 and 6 were combined. In the multivariate analysis the original Daniel groups were used as the variable for social group.

Of the patients who responded there were 105 in Groups 1–3, 101 in Group 4 and 76 in Groups 5 and 6. There was a relative preponderance of lower socio-economic status in the noncompensation patients (Table 12).

Social group had little influence on the type of injury sustained, although in the noncompensation patients lifting and twisting injuries were more common in Groups 1–3 than in Groups 5 and 6. Social group had no association with the diagnosis made.

Migrant status. Migrants were defined as those patients whose first language was not English.

The incidence of migrants when adjusted for social group tended to be but marginally greater in the compensation patients with perhaps increased unemployment in the migrants notably in social Groups 5 and 6 (Figure 1).

Compensation. Details of compensation claims were explored in the questionnaire. The type of claim was established together with settlement status and the date of settlement where applicable. The majority of the compensation patients were claiming under the South Australian Workman's Compensation Scheme, which was unchanged during the period under study. Under this scheme, following injury the worker would initially receive intermittent payments. Later, liability for any one incident could be commuted to a lump sum payment at the initiation of the worker or indirectly by the insurance company. This payment could be claimed whether or not the worker was able to work. In addition to the Workers' Compensation Scheme, which also covered travel to and from work, in some cases a negligence or third party claim might also be entered.

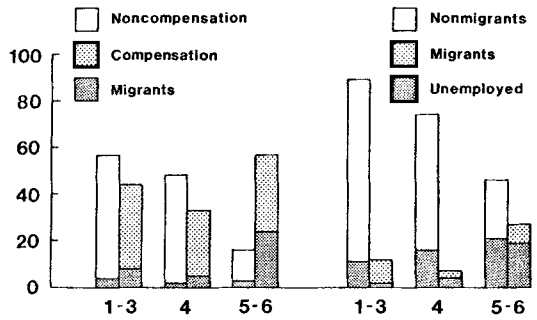

Figure 1. The incidence of migrants in the compensation and noncompensation groups and the incidence of unemployment in nonmigrants and migrants, controlling for social group. [Reproduced with permission, Greenough and Fraser 1989]

Patients not covered under the Workman's Compensation Scheme were either claiming lump sums from common law actions or intermittent payments from private insurance policies.

There were 12 housewives in the noncompensation group and 8 in the compensation group (the latter comprised three road traffic accidents, two accidents in shops and three private sickness policies). In total 20 men and 16 women claimed intermittent payments and 51 men and 52 women claimed lump sums (Table 13).

At follow up, of the patients with lump sum claims 35/51 men had achieved settlement, as had 35/52 women. The median time to settlement was 29 (4–135) months and the median time post settlement was 22 (0.3–120) months.

Psychological disturbance. The psychometric instruments used in this work were:

- Pain *drawing* (Ransford et al. 1976)
- Inappropriate *symptoms* (Waddell et al. 1984)
- Inappropriate *signs* (Waddell et al. 1980)
- Illness Behaviour Questionnaire *(IBQ)* (Pilowsky and Spence 1983)
- Modified Somatic Perception Questionnaire *(MSPQ)* (Main 1983)
- Hospital anxiety scale *(HAD(a))* (Zigmond and Snaith 1983)
- Hospital depression scale *(HAD(d))* (Zigmond and Snaith 1983)
- Zung depression scale *(Zung)* (Zung 1965)

Table 13. Type of claim made by compensation claimants

	Men	Women
Intermittent payments		
Workers' compensation	16	15
Insurance policies	4	1
Lump sums		
Workers' compensation	30	28
Third party	7	18
Both	14	6

These instruments had either been designed for use in a back pain population (signs, symptoms, MSPQ and drawing), assessed in back pain patients (IBQ (Main and Waddell 1987), Zung (Main and Waddell 1984)), or designed to separate physical from psychological symptoms in a population with physical illness (HAD(a) and HAD(d)).

Due to an oversight in preparing the questionnaires the first 15 patients did not complete the IBQ or Zung depression scale. 10 patients refused physical examination preventing scoring of the Signs. Occasional patients missed some questions of some of the tests. For these reasons the numbers of patients completing each test varied slightly.

Each test was scored using cut-off values previously published (Zung 1965, Pilowsky and Spence 1975, Ransford et al 1976, Waddell et al. 1980, Main 1983, Zigmond and Snaith 1983, Waddell et al. 1984). The cut-off values used were as follows :-

• Pain Drawing	3 or more
• Inappropriate symptoms	Men 2 or more
	Women 4 or more
• Inappropriate signs	Men 2 or more
	Women 4 or more
• IBQ Discriminate Function	70 or more
• MSPQ	Men 7 or more
	Women 10 or more
• Hospital Anxiety Scale	11 or more
• Hospital Depression scale	11 or more
• Zung Depression Scale	55 or more

If for an individual patient three or more of these tests were positive then that patient was defined as disturbed. If less than three tests were positive then they were defined as non-disturbed, unless fewer than six of the eight tests were available in which case they were excluded from the analysis (one patient).

Using these criteria a quarter of the patients were defined as psychologically disturbed with equal numbers of men and women and a higher incidence in the compensation and non compensation patients. After controlling for compensation there was no difference in the incidence of psychological disturbance between the social groups (Table 14). Within social groups compensation significantly influenced the incidence of disturbance in Groups 1–3 and 5–6.

Table 14. Distribution of psychological disturbance with social group

Social group	Compensation		Noncompensation		
	Disturb.	Not dist.	Disturb.	Not dist.	*P*- value
1–3	17	28	3	56	<< 0.0001
4	10	30	7	52	ns
5–6	23	34	1	18	< 0.01
P- value	ns		ns		

Comparison of psychometric instruments

Using the combined result of all eight tests obtained as above as the 'Gold Standard', the specificity and sensitivity of each individual test was calculated for a number of cut-off values, to determine the optimum cut-off score. The specificity of the test was the number of non disturbed patients correctly identified (true negatives) expressed as a percentage of the total of non disturbed patients. The sensitivity of the test was the number of disturbed patients correctly identified (true positives) expressed as a percentage of the total of disturbed patients. Specificity and sensitivity vary inversely as the cut-off score is changed. As the non disturbed patients comprised the majority, a high specificity would reduce the overall number of mis-classifications. For this reason cut off scores yielding a specificity greater than 85 percent were chosen.

In addition to the eight individual instruments, the performance of two combinations was assessed. The Zung depression scale was rescored 0–60 instead of 20–80 and added to the result of the MSPQ, producing a combined MSPQ and Zung scale from 0 to 99. The HAD(a) and the HAD(d) were combined by simple addition.

The IBQ data was also analysed according to the factors suggested by Main and Waddell (1987). Three scales were calculated; Affective and Hypochondriacal Disturbance (AHD), Life Disruption (LD) and Social Inhibition (SI).

Further assessment of the ability of each test to distinguish the disturbed and non disturbed groups was provided by comparing the scores of disturbed and non disturbed patients. In a multiple regression analysis each individual test was used in turn as the dependent variable and regressed against the other seven to determine which of the instruments had most of its variance explained by the remaining instruments. The two combinations were regressed against the remaining six instruments.

To assess the influence of low back pain, social group, migrant status or compensation status a number of pairs of sub-groups were compiled:

- current pain < 5 versus pain ≥ 5 (Analogue pain scale)
- social groups 1–3 versus groups 4–6
- compensation claimants versus noncompensation cl.
- migrants versus non-migrants
- employed versus unemployed

For each member of the pair optimum cut-off scores were determined and compared both with each other and with the score previously determined for the whole population.

Reporting of results and statistics

Compensation was found to have a profound influence on the recovery from low back injury. The results for many factors are therefore reported for the compensation patients and the noncompensation patients separately, to avoid masking differences between these groups.

Statistical analysis was undertaken using a number of methods. Most group comparisons were made using the Wilcoxon Rank Sum Test. Comparison of the injury severity scores was made using the *t*-test, and the results are reported as mean ± standard deviation. Four fold table analysis was carried out using Fisher's Exact Probability test and the Chi-square test.

Multivariate analysis was performed using the Minitab statistical package (release 7) running on a Gould machine. The total variance explained (r^2) is reported, together with regression equations and the significance of individual regression coefficients.

In the multivariate analysis the combined score of the MSPQ and the Zung depression scale was used as the variable for psychological disturbance. Compensation was coded as 0) noncompensation, 1) intermittent payments or 2) lump sum payments. Employment was coded as 0) unemployed, 1) part time work, 2) full time lighter work or 3) full time original work. Men were coded as 1, women as 0.

Factors such as diagnosis containing options which were not arranged in a hierarchy were incorporated using a separate field for each diagnosis, with each diagnosis being coded as 1 if present or 0 if absent. Thus there are 6 fields for diagnosis and 8 for type of injury.

Results

Comparison of disability and outcome measurements

The new Outcome scale provided a satisfactory spread of results. Comparing the compensation patients and the noncompensation patients the range of scores observed was wide (Figure 2), although in the non-compensation group the least disabled fifth of the patients were contained within a range of the 4 points 72–75. This, however, compares favourably with the Oswestry Disability score and the disability score of Waddell where the least disabled one fifth were contained in a range of the 3 points 0–2 and 1 point (zero), respectively.

To allow direct comparison of the coefficients in the regression analysis the four outcome and disability scales were each transformed to scales of 1–100. Correlation coefficients of the Outcome scale with the Oswestry disability score, the disability score of Waddell, and the physical impairment rating of Waddell were significant in all cases but only exceeded 0.8 for the Oswestry disability score (Table 15). The impairment rating accounted for only 40 percent of the variation of the Outcome scale. Each score was used as the dependent variable in regression analysis against the other three (Table 16).

The effect of psychological disturbance at review on the four scales was assessed along with five other independent variables (Table 17). Physical impairment was least affected by these variables, but was still influenced by psychological disturbance. The influence of

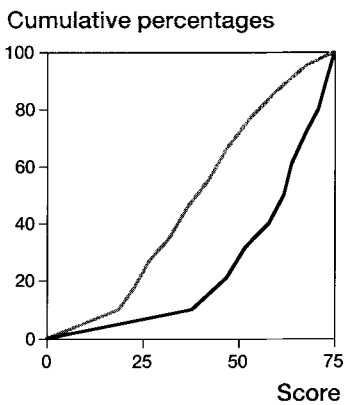

Figure 2. Score of the Outcome scale in the study population. Grey line compensation and black line noncompensation cases.

Table 15. Correlation coefficients for the Outcome Scale with Disability (Oswestry), Disability (Waddell) and Physical impairment. All scales transformed to 0–100

Regression equation	Corr. coeff.	*P*-value
Outcome = 95.0 – 1.20 • Disability (Oswestry)	–0.87	< 0.001
Outcome = 85.9 – 0.68 • Disability (Waddell)	–0.74	< 0.001
Outcome = 85.3 – 1.06 • Physical impairment	–0.63	< 0.001

Table 16. Multiple regression of Outcome, Disability and Impairment measurements. All scales transformed to 0–100. Coefficient, significance

	Outcome	Oswestry	Waddell	Impairment
Total variance	79	79	59	41
Independent variable				
Constant	97	49	41	26
Outcome	...	–0.48***	–0.37**	–0.20**
Oswestry	–0.91***	...	0.62**	0.21*
Waddell	–0.16**	0.15**	...	0.02
Impairment	–0.22 *	0.13 *	0.05	...

*$P < 0.05$, **$P < 0.001$, ***$P < 0.0001$

Table 17. Multiple regression of the four outcome and disability scales against six independent variables. All scales transformed to 0–100. Coefficient, significance

	Outcome	Oswestry	Waddell	Impairment
Total variance	65	57	48	26
Independent variable				
Constant	85	–6.1	2.4	1.3
Sex	7.1[b]	–5.4[b]	–7.6[a]	–0.63
Age	–0.19[a]	0.23[a]	0.23[a]	0.13
Social group	–1.7	2.0[a]	0.79	1.5[a]
Compensation	–4.8[b]	5.0[c]	5.1[a]	1.7
Employment	7.4[c]	–1.9[a]	–3.9[a]	–1.5
Psychiatric disturbance	–0.66[c]	0.57[c]	0.91[c]	0.31[c]

[a]$P < 0.05$, [b]$P < 0.001$, [c]$P < 0.0001$

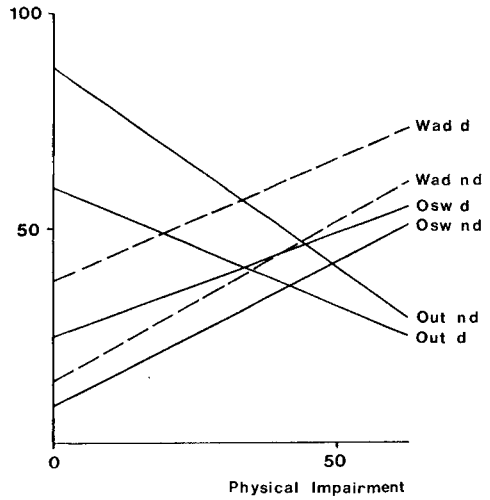

Figure 3. Regression Plots of Low-back Outcome Score and Oswestry and Waddell disability scores against physical impairment for disturbed and not disturbed patients. All scales transformed to 0–100.
Wad d = Waddell disability, disturbed patients,
Wad nd = Waddell disability, non disturbed patients,
Osw d = Oswestry disability, disturbed patients,
Osw nd = Oswestry disability, non disturbed patients,
Out d = Low-back Outcome Score, disturbed patients,
Out nd = Low-back Outcome Score, non disturbed patients.
[Reproduced with permission, Greenough and Fraser 1991]

Table 18. Outcome scores for different lengths of follow up. Maximum Outcome score = 75. All scales transformed to 0–100. Number, median (range)

Age at injury	Compensation			Noncompensation			*P*-value C vs N
18–30	37	41	(17–72)	49	64	(22–75)	< 0.001
31–45	57	42	(6–75)	66	63	(18–75)	< 0.001
46–65	32	32	(16–74)	18	52	(27–72)	< 0.001
P-value		a			b		

[a]46–65 group versus 18–30, $P < 0.05$.
[b]46–65 group versus 18–30, $P < 0.01$, and versus 31–45, $P < 0.02$.

Table 19. Influence of age at injury on Outcome score at follow up. Number, median (range)

Follow up (years)	Compensation			Noncompensation			*P*-value C. vs Nonc
0–1	no patients			6	72	(48–75)	
1–2	16	44	(17–73)	14	59	(27–75)	< 0.05
2–3	19	52	(6–58)	23	63	(18–75)	< 0.02
3–4	29	35	(16–69)	18	70	(32–75)	< 0.0001
4–5	25	39	(20–72)	29	59	(19–75)	< 0.001
5–6	20	50	(28–62)	18	63	(25–75)	< 0.01
6–10	14	40	(9–75)	13	64	(36–75)	< 0.02
> 10	7	32	(21–46)	15	63	(27–72)	< 0.01

psychological disturbance on the disability scales appeared to be reduced at increasing levels of physical impairment (Figure 3).

Referral

There was no difference in the Outcome score between the various referral sources when controlled for compensation.

Length of follow up

There were no changes in the Outcome scores with increasing length of follow up from the injury (Table 18). Compensation patients fared worse than noncompensation patients at all stages of follow up.

Interval between incident and primary consultation

There was an association between the Outcome score at review and the length of interval between the incident and the primary consultation, but this was observed in the compensation patients only; those seen less than six months from injury had a better outcome than those seen at more than six months (48 (6–73) versus 35 (9–75), $P < 0.01$). Corresponding figures for the noncompensation patients were 63 (18–75) and 63 (19–75).

No differences were seen in the Short Outcome scores at primary consultation between those seen less or more than six months from the injury.

Sex

There was no difference in the new Outcome score between the sexes in the compensation patients but there was in the noncompensation patients (men (n 69) 66 (23–75), women (n 64) 54 (18–75); $P < 0.001$). This difference between the sexes in the noncompensation patients was observed in all the contributing factors counting for six or more points except employment and sex life.

No differences between men and women were observed in the Short Outcome score at primary consultation in either compensation group.

Age at accident

The Outcome score at follow up showed a reduction with increasing age at accident. The association with compensation, however, was more marked (Table 19).

Multivariate analysis of follow up interval, sex, and age

Analysis was performed using these four variables as independent variables together with compensation, social class and psychological disturbance at review. The lack of influence of length of follow up on the Outcome score was confirmed (Table 20). The total variance explained by these seven variables was 57 percent.

Table 20. Multiple regression analysis of Outcome scale against seven independent variables

Independent variable	Points on scale	Coeff.	t	*P*-value
Constant	...	88		
Sex	2	4.9	3.2	< 0.002
Age	48	–0.23	–3.0	< 0.005
Social group	6	–2.4	–3.4	< 0.002
Compensation	3	–6.8	–7.1	< 0.0001
Psychiatric disturbance	100	–0.64	–8.0	< 0.0001
Interval	294	–0.087	–2.0	< 0.05
Follow up	300	–0.064	–1.6	ns

Type of accident

Outcome at follow up was calculated for each accident group. There were no differences between the groups within the noncompensation patients, but there were in the compensation patients. These did not achieve significance individually but taking the two accident types with the best outcome (fall and bend) and comparing them with the three with the worst outcome (twist, slip and jerk) there was a difference (30 (22 - 59) versus 50 (16 - 66); $P < 0.05$).

Multivariate analysis of Outcome score against accident type demonstrated associations between slipping and jerking injuries and poor Outcome scores ($P < 0.02$ and $P < 0.05$, respectively). The regression equation is given below, the total variance explained was 9 percent.

Outcome = 61 – 24.2 Slip – 17.6 Impact
– 18.6 Jerk – 12.2 Lift – 10.1 Twist
– 6.8 Fall – 5.9 Bend + 2.8 Other

Table 21. Relative risks of particular diagnoses with respect to type of injury, ratios of 2 or more only

Diagnosis	Injury producing increased incidence	with respect to	Ratio	*P*-value
Disc prolapse	Twist	Fall	3.8	< 0.05
		Impact	3.8	
		Jerk	2.6	
	Bend	Impact	3.0	
		Fall	3.0	
		Jerk	2.0	
Discogenic pain	Lift	Jerk	2.0	< 0.05
	Twist	Jerk	2.0	
Facet arthropathy	Jerk	Impact	3.3	
		Twist	3.3	
	Fall	Impact	2.3	
		Twist	2.3	
Spondylol.	Fall	Impact	4.3	

Diagnosis

After controlling for compensation, there were no significant in the Outcome scores between the diagnostic groups. Multivariate analysis confirmed the lack of significant association between the Outcome score and diagnosis. The regression equation is given below, the total variance explained was only 3 percent.

Outcome = 75 – 24.8 Prolapse – 27.6 Discog. pain
– 19.8 Spondylolysis/olisthesis
– 22.6 Facet arthropathy
– 25 Soft tissue strain – 27.3 Other

The coefficients were similar and none achieved significance. Only one diagnosis was made in each patient.

Diagnosis related to accident type

Impact was 3.4 times more likely to precede soft tissue strain than all other types of injury ($P < 0.001$). Bending preceded the group of miscellaneous diagnoses 2.8 times more often ($P < 0.05$). Soft tissue strain occurred less often after twisting injuries ($P < 0.05$).

No other combination of accident and diagnosis differed from the group as a whole. However, comparing one mode of accident with another there were associations with different diagnoses in two instances (Table 21). Thus disc prolapse was associated with twisting and bending more often than with impact, falls or jerks. Discogenic pain resulted more often from lifting or twisting injuries then from jerks. Jerking strains were not associated with prolapse or discogenic pain, but did produce an excess of facet joint pain. Spondylolisthesis and spondylolysis were seen more often after falls than after impact.

Multivariate analysis was performed for each diagnosis in turn against accident type. Little of the variance

Table 22. Regression of diagnoses against injury type, significant associations only

Diagnosis	Variance explained	Injury type	Coeff.	*P*-value
Disc prolapse	5.6	Other	0.50	< 0.02
Discogenic pain	2.7	None		
Spondylol.	3.4	None		
Facet arthropathy	4.0	None		
Soft tissue strain	11	Impact	0.37	< 0.02
Other	4.4	None		

Table 24. Outcome scores at follow up related to initial pain distribution. Number, median (range)

Pain	Compensation		Noncompensation		*P*-value
Back	30	38 (16–75)	46	66 (25–75)	< 0.0001
Hip	25	43 (23–69)	25	63 (38–75)	< 0.001
Leg	35	37 (9–74)	43	63 (18–75)	< 0.0001
All above	37	41 (6–73)	23	52 (19–75)	ns

Table 23. Outcome scores at follow up related to injury severity and age. Number, median (range)

Age	Severity 1 & 2		Severity 3 & 4		*P*-value
Comp.					
18–30	29	44 (17–72)	12	40 (18–62)	ns
31–45	48	39 (6–75)	11	48 (19–55)	ns
46–65	22	29 (16–74)	8	35 (24–66)	ns
Noncomp.					
18–30	33	64 (27–75)	19	64 (32–75)	ns
31–45	54	63 (18–75)	12	63 (19–75)	ns
46–65	15	48 (27–71)	3	59 (38–72)	ns

Table 25. Neurological deficit and initial pain distribution.

Pain distribution	Compensation Neurological deficit		Noncompensation Neurological deficit	
	Yes	No	Yes	No
Back	4	30	4	42
Hip	3	25	2	23
Leg	14	22	12	34
All	13	29	7	20
P-value	< 0.05		ns	
Chi-square				

of the diagnoses was explained by accident type, in only two instances was there an association of diagnosis with accident (Table 22). The regression equation used was

Diagnosis = 0.0 + n (coefficient x accident type)

This analysis confirmed the influence of impact on soft tissue strain and also demonstrated a relationship between "other" injuries and disc prolapse, which had not been sought by simple analysis owing to the small numbers.

Accident severity score

There was some variation of the Outcome score with severity particularly within the compensation group, but this showed no consistent trend. Analysis by age and compensation status revealed no effect of severity on outcome (Table 23). Correlation calculations of severity score with Outcome score revealed no correlation between the two. For compensation patients the correlation coefficient (r) was –0.02 and for the noncompensation patients it was –0.03.

The distribution of diagnoses made following particular accident types was not influenced by the severity of the accident.

Initial pain distribution

Outcome scores at review related to the initial pain distribution demonstrated little variation in the compensation patients between the different pain distributions. In the noncompensation group those patients who indicated pain in all three areas had a reduced outcome compared with those who indicated back pain alone ($P < 0.01$) and with those who indicated leg pain ($P < 0.05$). There were differences between the compensation and noncompensation groups for most of the initial pain groups (Table 24).

Initial neurological deficit

Naturally, patients reporting leg pain or pain in all three areas were more likely to have a neurological deficit on examination (Table 25). There were no differences between the compensation patients and the noncompensation patients in the distribution of neurological deficit between the various pain distributions, but men complaining of pain in all three areas had less deficits than women (2 out of 26 versus 17 out of 40; $P < 0.01$).

The presence of a neurological deficit at the primary consultation did not alter the outcome (Table 26). Multivariate analysis confirmed the lack of influence of neurological deficit on the Outcome score (see below).

Table 26. Outcome scores at follow up related to neurological deficit. Number, median (range)

	Neurol. deficit		No neurol. deficit		*P*-value
Men					
Comp.	12	38 (18–74)	54	39 (17–75)	ns
Noncomp.	8	61 (32–75)	60	66 (23–75)	ns
Women					
Comp.	18	37 (6–66)	44	42 (9–72)	ns
Noncomp.	15	51 (19–75)	52	59 (18–75)	ns

Table 27. Short Outcome scores at primary consultation related to initial treatment. Median (range)

Treatment	n	Score	*P*-value[a]
All Treatments	287	11 (0–25)	...
Physiotherapy[b]	159	10 (0–24)	ns
Chiropractic[c]	19	12 (0–25)	ns
Surgical belt	46	8 (0–25)	< 0.001
Bed rest	49	11 (0–22)	ns
Traction	43	12 (0–22)	ns
Epidural	25	8 (0–21)	< 0.01
Acupuncture	35	11 (0–25)	ns

[a]Selected treatment versus All treatments.
[b]Patients who also had chiropractic excluded.
[c]Patients who also had physiotherapy excluded.

Initial treatment and Short Outcome score

Analysis of the Short Outcome score recorded at the time of the primary consultation in relation to the type of treatment received following the accident showed that surgical corsets and epidurals were associated with a lower score (Table 27).

Timing of initial treatment

In the noncompensation patients there was no difference in the median Outcome score at review between the 49 patients who sought initial treatment early and the 75 who sought treatment later. In the compensation patients, however, the Outcome score at review in the 90 patients treated within 2 days was worse than in the 32 patients treated later. This reduction of Outcome score with earlier treatment in the compensation patients was not associated with any difference in average accident severity score.

Treatment following primary consultation

The results of treatment are expressed in the difference in the Short Outcome scores between the primary consultation and review. Overall the compensation and noncompensation patients showed an equal improvement. The compensation patients, however, had lower Short Outcome scores both at the primary consultation and at review (Table 28).

Table 28. Outcome Scores at primary consultation and at review. Number, median (range)

	Compensation		Noncompensation		*P*-value
Prim. cons.	138	9 (0–23)	141	15 (1–25)	< 0.0001
Review	135	13 (0–25)	135	21 (2–25)	< 0.0001

Table 29. Short Outcome Score differences related to treatment. Number, median (range)

	Exc. plus classes		Exc. alone		*P*-value
Comp.					
All	27	7 (–8–20)	61	3 (–9–14)	ns
Disco.	21	7 (–8–20)	32	1 (–9–13)	0.05
Noncomp.					
All	23	6 (–8–20)	64	5 (–3–21)	ns
Disco.	12	9 (–8–20)	25	6 (–3–20)	ns

The addition of educational classes seemed to improve the response to the exercise regime, although this did not achieve significance. Controlling for compensation demonstrated that the difference was more marked in the compensation patients (Table 29).

A number of patients (mainly with prolapse) were offered surgery but declined.

There was a marked difference in the response of patients with facet arthropathy to facet joint injections. The noncompensation patients showed an increase in their Short Outcome scores from the time of primary consultation to review but the compensation patients were worse than before.

Multivariate analysis of the difference in the Short Outcome score between the primary consultation and review against treatment given at the primary consultation confirmed the value of educational classes. The total variance explained was 5 percent and the regression equation was:

$$\begin{aligned}\text{Short Outcome difference} &= 5.19 + 2.33\ \text{Classes} + 0.58\ \text{Exercises} \\ &\quad + 2.01\ \text{surgery (offered)} \\ &\quad + 1.71\ \text{facet injection} - 0.51\ \text{Epidural}\end{aligned}$$

Of the coefficients only that of the educational classes was significant ($P < 0.005$).

Table 30. Months off work, 20 housewives excluded. Median (range)

	n	Months off work	P-value C. vs Nonc.
Men			
All compensation	70	12 (0.25–84)	< 0.001
Lump sum cases	50	22 (0.5–84)	< 0.0001
Intermittent payments	20	3 (0.25–30)	< 0.01
Noncompensation	70	0.25 (0–180)	
Women			
All compensation	61	15 (0–132)	< 0.001
Lump sum cases	45	24 (0–132)	< 0.0001
Intermittent payments	16	3.5 (0–57)	< 0.05
Noncompensation	54	0.5 (0–22)	

Table 31. Months off work related to social group, housewives excluded. Median (n)

Social group	Comp.	Noncomp.	P-value
Men			
1–3	3 (13)	0.25 (37)	< 0.001
4	13 (24)	0.25 (21)	<< 0.0001
5–6	22 (32)	3 (9)	< 0.01
Women			
1–3	4 (30)	1 (20)	< 0.05
4	33 (8)	0.25 (27)	< 0.005
5–6	24 (23)	0.25 (7)	< 0.001

Table 32. Outcome scores at follow up related to time off work, housewives excluded. Number, median (range)

	< 6 months		≥ 6 months		P-value
Men					
Comp.	23	55 (21–75)	43	36 (17–61)	< 0.01
Noncomp.	61	66 (35–75)	8	57 (23–75)	< 0.02
Women					
Comp.	31	44 (22–72)	31	30 (6–61)	< 0.001
Noncomp.	48	62 (19–75)	5	32 (18–53)	< 0.01

Table 33. Influence of social group on Outcome score at follow up. Number, median (range)

Social group	Compensation		Noncompensation		P-value C. vs Nonc
1–3	41	45 (19–74)	60	67 (23–75)	<< 0.0001
4	38	39 (6–75)	56	54 (18–75)	< 0.001
5–6	50	33 (9–62)	18	62 (36–75)	<< 0.0001
P-value	a		b		

[a] 1–3 versus 5–6, $P < 0.001$.
[b] 1–3 versus 4, $P < 0.001$.

Time off work

The median time off work (housewives excluded) was quite different between the compensation group and the controls (Table 30). The difference was almost entirely due to the lump sum claimants. After controlling for social group, the differences in time off work were still marked (Table 31).

Multiple regression analysis of time off work was performed against sex, age, social group, and compensation. The total variance explained was 32 percent and the regression equation was

$$\text{Time off work} = -15.5 - 0.33\ \text{Sex} + 0.20\ \text{Age} + 3.09\ \text{Social group} + 9.54\ \text{Compensation}$$

Coefficients were significant for age ($P < 0.05$), social group ($P < 0.002$) and compensation ($P < 0.0001$).

In the compensation patients, there was a strong association between prolonged time off work and employment outcome. Of 25 men off work for less than six months 5 were unemployed at review whereas of 45 off work for more than six months 30 were unemployed. Corresponding figures for the women were 1 unemployed out of 25 against 25 unemployed out of 34. There was no such trend in the noncompensation patients. Taking both sexes together, of patients off work for less than six months 5 were unemployed out of 98 as compared with 4 unemployed out of 13 patients off work for six months or more.

The Outcome scores demonstrated an decreased outcome in both compensation patients and noncompensation patients with increasing time off work (Table 32).

Social group

In the compensation patients subsequent employment was associated with socio-economic group. Two out of 13 men in Groups 1–3 were unemployed at review versus 24/32 in Groups 5–6. Similarly, while 9/31 women in Groups 1–3 were unemployed, 15/24 in Groups 5–6 were out of work. There was no association between social group and unemployment in the noncompensation patients.

The Outcome score was related to social group with Groups 1–3 having a better outcome both in the compensation and noncompensation groups. There were also differences in the Outcome scores between the compensation and noncompensation patients (Table 33).

Table 34. Multiple regression analysis of employment against six independent variables

Independent variable	Points on scale	Coeff.	t	*P*-value
Constant	...	3.9		
Sex	2	–0.05	–0.43	ns
Age	48	–0.01	–1.6	ns
Social group	6	–0.14	–2.5	< 0.02
Compensation	3	–0.34	–4.4	< 0.0001
Time off work	180	–0.02	–6.1	< 0.0001
Psychiatric disturbance	100	–0.02	–3.6	< 0.0001

Table 36. Employment outcome at follow up, 20 housewives excluded. Number

	Empl.	Unempl.	*P*-value C. vs Nonc
Men			
All compensation	35	37	3.0×10^{-7}
Lump sum cases	18	34	3.0×10^{-10}
Intermittent payments	17	3	ns
Noncompensation	63	8	
Women			
All compensation	34	29	3.0×10^{-6}
Lump sum cases	22	26	2.4×10^{-7}
Intermittent payments	12	3	ns
Noncompensation	51	4	

Table 35. Outcome scores at follow up related to migrant status. Number, median (range)

Social group	Migrants		English speakers		*P*-value
Comp.					
1–3	8	51 (20–74)	33	44 (19–73)	ns
4	8	28 (16–48)	30	43 (6–75)	< 0.05
5–6	19	27 (9–62)	31	40 (16–61)	< 0.02
Noncomp.					
1–3	6	68 (54–72)	55	67 (23–75)	ns
4	3	41 (32–42)	51	55 (18–75)	< 0.05
5–6	3	69 (49–75)	15	60 (36–72)	ns

Table 37. Effect on Outcome score, reported pain, Oswestry Disability Score and Physical Impairment (Waddell) of compensation status. Number, median (range)

	Compensation		Noncompensation		*P*-value
Men					
Outcome	68	40 (17–75)	70	66 (23–75)	<< 0.0001
Pain	70	5 (0–9)	70	2 (0–10)	<< 0.0001
Oswestry	65	32 (2–68)	69	9 (0–47)	< 0.001
Impairm.	68	14 (–2–30)	67	5 (–4–20)	< 0.001
Women					
Outcome	62	39 (6–72)	67	54 (18–75)	<< 0.0001
Pain	64	5 (0–10)	65	4 (0–9)	< 0.01
Oswestry	60	32 (7–74)	65	18 (0–60)	< 0.01
Impairm.	65	12 (–2–27)	64	10 (–4–25)	ns

Multivariate analysis of employment against social group, compensation, time off and three other variables confirmed the influence of time off work and social group on subsequent employment (Table 34). The total variance explained was 53 percent.

Migrant status

After controlling for social group there was no association of migrant status with unemployment at review. Some differences in the Outcome score at review between the migrants and the native English speakers were observed even after controlling for social group and compensation (Table 35). Multivariate analysis of the Outcome score against migrant status and additional variables, however, indicated that migrant status had no independent effect on the Outcome score (Table 42).

In the compensation patients the crude incidence of psychological disturbance was increased in the immigrants. 21/39 immigrants were disturbed compared with 30/104 native English speakers. In the noncompensation patients no such increase was seen. Multiple regression analysis of psychological disturbance against migrant status and some other variables, however, indicated that migrant status made no significant contribution (Table 40).

Compensation

Employment at follow up differed markedly between the compensation groups and the controls (Table 36). This was more prominent in the lump sum claimants. Comparison of employment outcome was made after controlling for social grouping (Figure 2). Differences existed between the compensation patients and the controls in nearly all groups. Multiple regression analysis of employment against compensation and a number of other variables confirmed the effect of compensation (Table 43).

The Outcome scores at review were reduced in the compensation patients compared with the noncompensation patients. Compensation also exerted an influence on reported pain, the Oswestry disability score, and the physical impairment rating of Waddell (Table 37).

The type of lump sum claim (worker's compensation or third party) had no effect on employment outcome.

Table 38. Psychological disturbance at follow up. Number

	Disturbed	Not disturbed	*P*-value C. vs Nonc.
Men			
All compensation	25	47	3.8×10^{-4}
Lump sum cases	21	31	9.7×10^{-5}
Intermittent payments	4	16	ns
Noncompensation	7	67	
Women			
All compensation	26	45	1.8×10^{-4}
Lump sum cases	20	33	3.1×10^{-4}
Intermittent payments	6	12	ns
Noncompensation	6	61	

There were 14 men and 6 women making both types of claim. Excluding these dual claimants, for men, 19/30 worker's compensation patients were unemployed at review as compared with 4/7 third party claimants. For women the figures were 16/28 and 12/18, respectively.

Psychological disturbance at follow up was associated with compensation (Table 38), almost entirely due to the lump sum claimants. Multivariate analysis of the influence of compensation and other variables on psychological disturbance confirmed the role of compensation (Table 40).

Settlement of the claim did not influence either employment status or the incidence of psychological disturbance at follow up. Furthermore, time from settlement did not appear to influence either of these variables. The proportion of employed to unemployed, and of disturbed to not disturbed remained unchanged up to 5 years post settlement (Figure 4). Even in the compensation patients employed after settlement there was no reduction of disturbance; 2/13 men and 5/14 women were disturbed post settlement.

Multiple regression analysis confirmed that settlement did not contribute to psychological disturbance or employment status at review. The regression equations were

Psychological disturbance =
20.7 – 0.22 Sex + 0.20 Age + 0.95 Social group
+ 0.047 Settlement

Total variance explained was 7 percent. No coefficient achieved significance.

Employment =
4.35 – 0.48 Sex – 0.019 Age – 0.32 Social group
– 0.029 Psychological disturbance
+ 0.00044 Settlement

Total variance explained was 26 percent. Coefficients achieved significance for social group and psychological disturbance.

Following settlement there was some improvement of the Outcome score in the men but not in women.

Multiple regression analysis of Outcome score against settlement and other variables demonstrated no effect on settlement on outcome. The total variance explained was 27 percent and the regression equation was

Outcome score =
70.6 + 2.97 Sex – 0.098 Age – 3.22 Social group
– 0.54 Psychological disturbance
+ 0.095 Settlement

The coefficients achieved significance for social group and psychological disturbance.

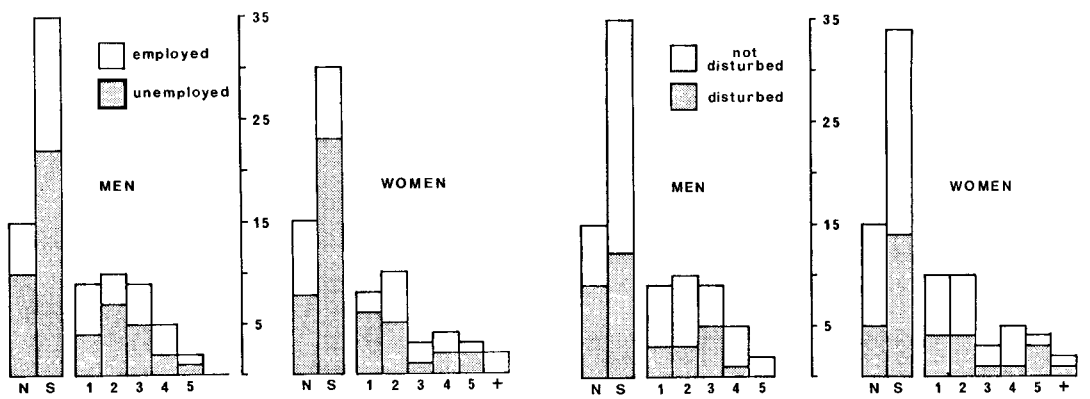

Figure 4. Effects of settlement of a compensation claim on employment and psychological disturbance.N = not settled, S = settled (all cases), 1 = settled 0 –1 years, 2 = settled 1–2 years, 3 = settled 2–3 years, 4 = settled 3–4 years, 5 = settled 4–5 years, and + = settled 5 or more years.

The incidence of recent treatment (in the month before review) for low back pain was not affected by settlement. Of those patients whose claims had been settled, patients who had recently sought treatment were more disturbed than those who had not.

Half of the of patients (62 percent of the men and 38 percent of the women) claiming lump sum compensation stated that they would not go through the claim procedure again under similar circumstances.

The reasons given were:
the process had been too stressful ;
the procedure was too slow;
it had caused too much family trauma;
it appeared to reduce the treatment that they were given;
they had become depressed;
they felt it reduced their ability to find a job.

Psychological disturbance at review

Psychological disturbance was associated with unemployment; 24/181 employed patients were disturbed as compared with 34/77 unemployed patients. Psychological disturbance was also closely associated with compensation; analysis after controlling for unemployment and compensation status demonstrated a close correlation between the two (Table 39). In the unemployed patients there was no significant effect of compensation on the incidence of distress, similarly in the compensation patients there was no significant effect of employment, suggesting both contribute in the aetiology of disturbance.

Multivariate analysis of psychological disturbance against compensation and employment status revealed a strong association with each. The regression equation was

Psychological disturbance =
29.4 + 1.75 Compensation – 2.57 Employment

The total variance explained was 26 percent and the value of t for compensation was 3.7, and for employment was 5.0. As the employment scale used had four points to the compensation scale's three, employment accounted for approximately twice as much variation as compensation when no other factors were considered.

A more detailed analysis of psychological disturbance against seven variables was also performed (Table 40). The total variance explained was 41 percent.

Adjusting for the number of points on the scales, the maximum contributions were 17 for current pain, –6.5 for employment, and 4.5 for compensation.

Table 39. Psychological disturbance, unemployment and compensation. Number

	Compensation		Noncomp.		*P*-value[a]
	Disturb.	Not dist.	Disturb.	Not dist.	
Employed	19	50	5	107	<< 0.0001
Unemployed	29	37	5	6	ns
P-value[b]	ns		< 0.001		

[a]Compensation versus noncompensation.
[b]Employed versus unemployed.

Table 40. Multiple regression analysis of psychological disturbance against seven independent variables

Independent variable	Points on scale	Coeff.	t	*P*-value
Constant	...	22	5.6	
Sex	2	–1.2	–1.1	ns
Age	48	0.085	1.6	ns
Social group	6	0.049	0.09	ns
Compensation	3	1.7	2.2	< 0.05
Employment status	4	–1.8	–3.1	< 0.005
Migrant status	2	0.83	0.54	ns
Pain score	11	1.7	6.7	< 0.0001

Table 41. Effect on Outcome score of psychological disturbance and employment status. Number, median (range)

Employed	Disturbed		Not disturbed		*P*-value[a]
Comp.					
Yes	17	44 (9–74)	49	52 (24–75)	ns
No	30	27 (6–50)	25	29 (18–55)	ns
P-value[b]	< 0.01		<< 0.0001		
Noncomp.					
Yes	5	42 (36–61)	106	64 (18–75)	< 0.01
No	5	25 (19–47)	4	50 (35–63)	ns
P-value[b]	ns		< 0.05		

[a]Disturbed versus Not disturbed.
[b]Employed versus Not employed.

The Outcome score at review was reduced in patients with psychological disturbance even after controlling for employment status, but this was not always significant (Table 41).

Multivariate analysis

Multivariate analysis of the Outcome score was performed against thirteen variables (Table 42). The total variance explained was 59 percent.

Table 42. Multiple regression analysis of Outcome score against thirteen independent variable

Independent variable	Points on scale	Coeff.	t	*P*-value
Constant	...	82		
Sex	2	4.8	2.8	< 0.01
Age	48	−0.16	−1.9	ns
Follow up	300	0.066	1.6	ns
Interval	294	−0.084	1.8	n.s.
Injury severity	4	−0.48	−0.39	ns
Neuro. deficit	2	0.67	−0.30	ns
Time off	180	−0.1	−2.4	< 0.02
Social group	6	−1.8	−2.4	< 0.02
Compensation	3	−6.2	−5.5	< 0.0001
Migrant status	2	−2.7	−1.2	ns
Psychological disturbance	100	0.58	−6.5	< 0.0001
Smoker	2	1.5	0.81	ns
Marital status	2	−2.1	−1.0	ns

Table 43. Multiple regression analysis of employment against thirteen independent variables

Independent variable	Points on scale	Coeff.	t	*P*-value
Constant	...	82		
Sex	2	−0.095	−0.74	ns
Age	48	−0.009	−1.4	ns
Follow up	300	0.006	2.0	ns
Interval	294	−0.007	−2.0	ns
Injury severity	4	0.11	1.2	ns
Neuro. deficit	2	0.054	0.33	ns
Time off	180	−0.023	−5.8	< 0.0001
Social group	6	−0.13	−2.2	< 0.05
Compensation	3	−0.33	−4.0	< 0.0001
Migrant status	2	−0.22	−1.4	ns
Psychological disturbance	100	−0.020	−3.4	< 0.002
Smoker	2	0.045	0.34	ns
Marital status	2	−0.048	−0.31	ns

Table 44. Optimum cut-off scores, *percentages*

Test	Patients	Cut-off Score	Specificity	Sensitivity
Pain Drawing	All	2 or more	*91*	*42*
	Men	not discrimin.		
	Women	2 or more	*86*	*61*
Inapp. Symptoms	Men	3 or more	*93*	*44*
	Women	3 or more	*87*	*62*
Inapp. Signs	Men	1 or more	*86*	*44*
	Women	1 or more	*84*	*48*
IBQ (Discr. Func.)	All	76 or more	*87*	*53*
IBQ (Aff. & Hypo.)	All	10 or more	*89*	*62*
IBQ (Life disrupt.)	All	9 or more	*94*	*40*
IBQ (Soc. inhibit.)	All	not discrimin.		
MSPQ	Men	7 or more	*91*	*65*
	Women	10 or more	*93*	*69*
HAD anxiety	All	11 or more	*92*	*67*
HAD depression	All	9 or more	*90*	*68*
Zung depression	All	56 or more	*92*	*73*
MSPQ + Zung	Men	29 or more	*91*	*84*
	Women	33 or more	*96*	*85*
HAD(a) + HAD(d)	All	18 or more	*88*	*78*

Table 45. Scores for disturbed and not disturbed patients. Number, median (range)

Test	Patients	Not disturbed			Disturb.			*P*-value
Drawing	All	206	0	(0– 5)	62	1	(0– 5)	b
	M	104	0	(0– 3)	34	0	(0– 3)	ns
	W	102	0	(0– 5)	28	2	(0– 5)	b
Symptoms	M	106	0	(0– 4)	34	2	(0– 5)	b
	W	104	0	(0– 5)	29	3	(0– 6)	b
Signs	M	103	0	(0– 3)	32	0	(0– 3)	a
	W	99	0	(0– 4)	29	0	(0– 3)	a
IBQ (D.F.)	All	196	54	(21–99)	59	77	(45–99)	b
IBQ (AH)	All	194	4	(0–14)	60	11	(0–18)	b
IBQ (LD)	All	193	4	(0–10)	60	8	(3–10)	b
IBQ (SI)	All	194	2	(0– 5)	61	2	(0– 5)	ns
MSPQ	M	106	3	(0–15)	34	8	(0–39)	b
	W	103	3	(0–10)	29	12	(4–25)	b
HAD (a)	All	208	6	(0–17)	63	11	(2–20)	b
HAD (d)	All	208	3	(0–22)	63	10	(2–19)	b
Zung	All	196	40	(25–75)	59	61	(41–89)	b
MSPQ/Zung	M	100	15	(0–43)	32	35	(25–80)	b
	W	96	18	(0–37)	27	39	(24–58)	b
HAD Comb.	All	208	10	(0–26)	63	22	(4–39)	b

[a] $P < 0.01$, [b] $P < 0.0001$

Multivariate analysis of employment status was performed against thirteen variables (Table 43). The total variance explained was 54 percent.

The measurement of psychological disturbance

Four tests, the MSPQ, the Hospital Anxiety Scale (HAD(a)), the Hospital Depression Scale (HAD(d)) and the Zung Depression Scale had specificities and sensitivities above 85 percent and 64 percent, respectively. The specificities and sensitivities of each test at a number of cut-off scores plotted in Figure 5 clearly demonstrated the differences in discriminating ability of the various tests. The selected optimum cut-off values are given, together with specificities and sensitivities, in Table 44, and data for each test are shown in Table 45.

Of the four tests found to be less discriminating, the Pain Drawing showed poor sensitivity in the women and was not discriminating in the men. The Inappropriate Symptoms were discriminating in the women but insensitive in the men. The Inappropriate

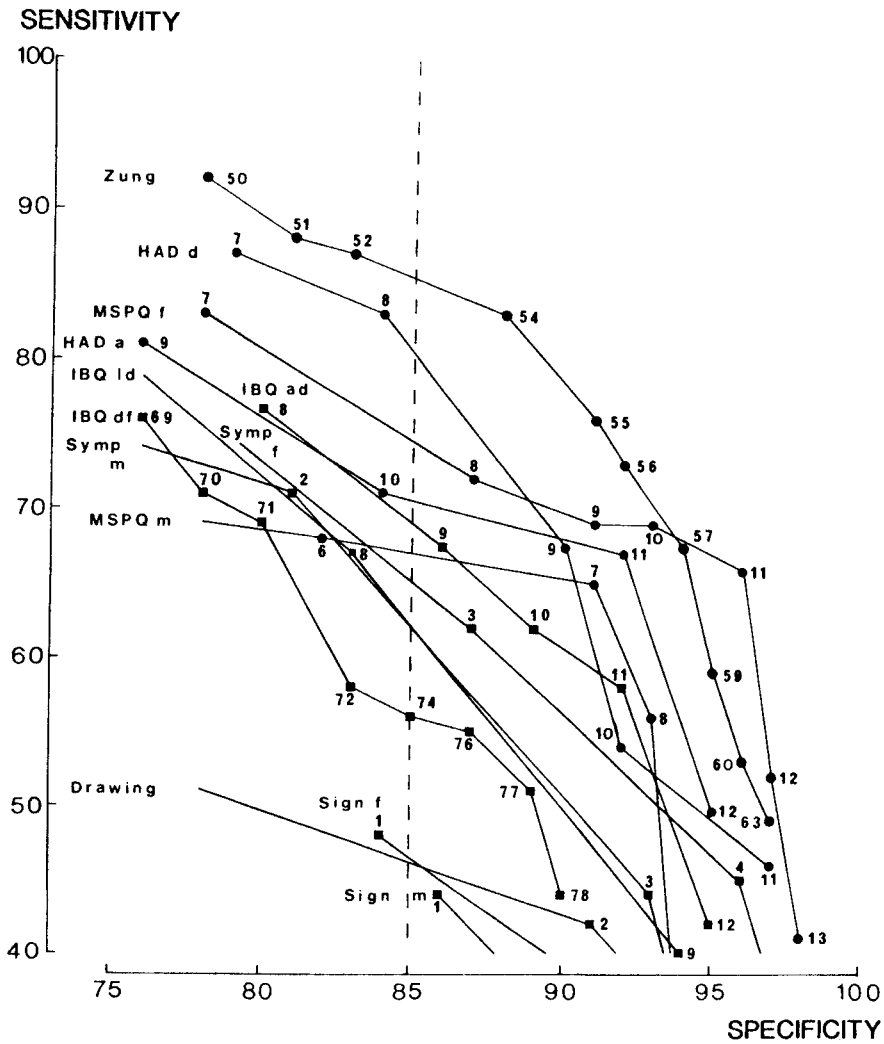

Figure 5. Sensitivity and specificity of the psychometric instruments plotted for a number of cut-off scores. The four most discriminating tests are represented by circles, the others by squares. [Reproduced with permission, Greenough and Fraser 1992]

Drawing	Pain Drawing
Symp m	Inappropriate Symptoms, men
Symp f	Inappropriate Symptoms, women
Sign m	Inappropriate Signs, men
Sign f	Inappropriate Signs, women
MSPQ m	Modified Somatic Perception Questionnaire, men
MSPQ f	Modified Somatic Perception Questionnaire, women
IBQ df	Illness Behaviour Questionnaire, discriminant function
HAD a	Hospital Anxiety Scale
HAD d	Hospital Depression Scale
Zung	Zung Depression Scale
IBQ ad	Illness BehaviourQuestionnaire, affectiveand hypochondriacal disturbance
IBQ ld	Illness Behaviour Questionnaire, life disruption

Signs were insensitive owing to a relative infrequency of occurrence (86 percent of normals and 54 percent of disturbed patients scored zero). There was no difference between the sexes for the IBQ, but the test was insensitive.

Two of the new IBQ scales distinguished between the two groups (Affective and Hypochondrical Disturbance and Life Disruption). Both scales were relatively insensitive. The Social Inhibition scale showed no difference between the populations.

Of the four most discriminating tests, the MSPQ needed cut-off scores to be different for men and women, but proved both specific and sensitive. The Hospital Anxiety Scale and the Hospital Depression Scale showed no sex differences, and had high specificities and sensitivities. The Zung Depression Scale again showed no sex differences but good specificity and sensitivity.

Table 46. Total variance explained and individual tests having significant coefficients of variation

Instrument used as dependent variable	Total variance explained	Instruments having significant coefficients of variation							
		Drawing	Symp.	Signs	IBQ	MSPQ	HAD (a)	HAD (d)	Zung
Pain Drawing	27.7		< 0.005			< 0.005			
Symptoms	47.7	< 0.005		< 0.05	< 0.0001	< 0.0001			
Signs	20.0		< 0.05						
IBQ (D.F.)	33.4		< 0.0001				< 0.005	< 0.0001	< 0.005
MSPQ	48.1	< 0.005	< 0.0001				< 0.05	< 0.05	
HAD (a)	51.1				< 0.005	< 0.05		< 0.0001	< 0.0001
HAD (d)	63.6				< 0.0001	< 0.05	< 0.0001		< 0.0001
HAD Comb.	59.6					< 0.0001			< 0.0001
Zung	59.6				< 0.005		< 0.0001	< 0.0001	
MSPQ/Zung	68.4	< 0.005	< 0.005	< 0.05			< 0.0001	< 0.0001	

Table 47. Effect of pain on optimum cut-off values for instruments where cut-off scores change. Cut off value (Specificity, Sensitivity)

Test	All	Pain < 5	Pain ≥ 5
Pain Drawing	2 (91, 42)	ns	3 (95, 21)
Symptoms Women	3 (87, 62)	a	2 (95, 33)
Signs Women	1 (84, 48)	a	2 (95, 33)
IBQ (D.F.)	76 (87, 53)	ns	88 (88, 29)
HAD(a) + HAD(d)	18 (89, 76)	17 (89, 87)	19 (86, 76)

[a]Only one patient in one group

Table 48. Effect of social group on optimum cut-off values for instruments where cut-off scores change. Cut off value (Specificity, Sensitivity)

Test	All	Group 1–3	Group 4–6
Symptoms M	3 (93, 44)	2 (93, 33)	3 (90, 50)
Symptoms W	3 (87, 62)	4 (94, 56)	3 (95, 55)
Signs M	1 (96, 44)	ns	2 (95, 40)
Signs W	1 (84, 48)	1 (95, 45)	2 (93, 33)
IBQ (D.F.)	76 (87, 53)	72 (92, 68)	84 (95, 35)
HAD(a) + HAD(d)	18 (89, 74)	17 (96, 85)	19 (87, 67)

The combined tests, the MSPQ + Zung and the HAD(a) + HAD(d) performed better than any of the single tests, with high specificities and sensitivities. Considerable differences between the tests were observed in the total variance explained on the multivariate analysis (Table 46).

The four instruments with high specificities and sensitivities also had a large part of their variance explained by the other instruments. The two combinations chosen also scored highly in this manner, and the combination of MSPQ and Zung had substantial contributions from five of the six remaining instruments.

Psychometric instruments and confounding factors

119 patients scored five or more on the visual analogue pain scale and were allocated to the pain plus group, with the others termed pain minus. In the pain plus patients there was a variation of the optimum cut-off score from that of the whole sample population in the Pain Drawing, the Inappropriate Symptoms (women), the Inappropriate Signs (women) and the IBQ. Both the combination scores demonstrated changes in the optimum cut-off score in the pain plus group; the HAD(a) + HAD(d) combination showed a variation in the pain minus group (Table 47). Some less discriminating tests failed to distinguish the disturbed patients in the subgroups.

102 patients in social Groups 1–3 were compared with the remaining patients in Groups 4–6. There were variations of the optimum cut-off scores for the Inappropriate Symptoms and the MSPQ + Zung in Groups 1–3, the Inappropriate Signs and the HAD(a) + HAD(d) in Groups 4–6 and the IBQ in all social groups (Table 48). The Inappropriate Signs (men) failed to distinguish the disturbed patients in Groups 1–3.

Compensation for an injury was being claimed by 137 patients. Optimum cut-off scores were altered in the MSPQ (men) for the compensation patients and in the HAD(d), the MSPQ + Zung (men) and the HAD(a) + HAD(d) for both groups (Table 49). Again some of the less discriminating tests failed to distinguish the disturbed patients in the sub-groups.

The 49 immigrants were compared with native English speakers. Optimum cut-off scores were unchanged between the groups except for the Inappropriate Symptoms in the women in the immigrants and for the IBQ and the MSPQ + Zung (women) in both groups (Table 50). The Inappropriate Signs failed to distinguish the disturbed patients in the immigrants.

78 patients were unemployed at review. Optimum cut-off scores varied between the employed and the

Table 49. Effect of compensation on optimum cut-off values for instruments where cut-off scores change. Cut off value (Specificity, Sensitivity)

Test	All	Noncomp.	Comp.
Symptoms M	3 (93, 44)	2 (92, 50)	3 (86, 46)
Symptoms W	3 (87, 62)	3 (92, 75)	4 (93, 44)
Signs Women	1 (84, 48)	ns	2 (92, 36)
IBQ (D.F.)	76 (87, 53)	76 (88, 67)	78 (89, 42)
MSPQ Men	7 (91, 65)	7 (95, 50)	8 (86, 62)
HAD (d)	9 (90, 68)	8 (91, 83)	10 (87, 57)
MSPQ/Zung M	29 (91, 84)	29 (97, 87)	31 (90, 71)
MSPQ/Zung W	33 (96, 85)	33 (100, 75)	34 (91, 83)
HAD(a) + HAD(d)	18 (89, 76)	16 (86, 75)	19 (87, 76)

Table 50. Effect of migrant status on optimum cut-off values for instruments where cut-off scores change. Cut off value (Specificity, Sensitivity)

Test	All	English	Migrants
Symptoms W	3 (87, 62)	3 (88, 60)	4 (89, 67)
IBQ (D.F.)	76 (87, 53)	78 (88, 59)	79 (89, 45)
MSPQ/Zung M	29 (91, 84)	29 (94, 89)	31 (88, 62)
HAD(a) + HAD(d)	18 (89, 76)	18 (88, 91)	18 (88, 70)

Table 51. Effect of employment on optimum cut-off values for instruments where cut-off scores change. Cut off value (Specificity, Sensitivity)

Test	All	Employed	Unemployed
Symptoms M	3 (93, 44)	2 (86, 54)	4 (90, 24)
Symptoms W	3 (87, 62)	3 (88, 56)	4 (79, 47)
Signs Women	1 (84, 48)	1 (89, 33)	3 (92, 20)
IBQ (D.F.)	76 (87, 53)	73 (89, 48)	80 (92, 45)
MSPQ Men	7 (91, 65)	6 (87, 62)	8 (85, 67)
MSPQ Women	10 (93, 69)	8 (88, 78)	11 (81, 38)
HAD (a)	11 (92, 67)	10 (85, 73)	11 (91, 64)
HAD (d)	9 (90, 68)	7 (86, 99)	12 (85, 36)
Zung	56 (92, 73)	54 (92, 90)	62 (85, 53)
MSPQ/Zung M	29 (91, 84)	28 (94, 92)	33 (89, 63)
MSPQ/Zung W	33 (96, 85)	33 (100, 99)	36 (89, 73)
HAD(a) + HAD(d)	18 (89, 76)	18 (93, 91)	21 (85, 64)

unemployed patients in almost all instruments. The largest differences were elevations of the cut-off scores in the unemployed patients (Table 51).

Each of the eight instruments and the two combinations was subjected to multivariate analysis against a number of independent variables (Table 52). In this table, the scales of each psychometric instrument have been transformed to 0–100 to allow comparison of the coefficients of the independent variables. Additionally, the coefficients have been multiplied by the number of points on the scale of their respective independent variable to analyze the degree of influence of the maximum change of each independent variable. Only independent variables with significant coefficients are reported.

In the case of the MSPQ/Zung combination, use of the original cut-off score in compensation patients produced a mis-classification of 6 men and 2 women out of 122 scored patients (5%). Similarly in the immigrants, 4 men were mis-classified out of 44 patients (9%) and in the unemployed patients 8 men and 4 women out of 61 scored patients were mis-classified (20%). In the unemployed patients, use of the original cut-off score produced mis-classification of 13 percent of the patients with the HAD(a/d) and between 5 and 16 percent for the other instruments.

Table 52. Multivariate analysis of each psychometric instrument against sex, age, current pain, social class, compensation, migrant status and employment

Test	Total variance	Coefficient and *P*-value			
		Pain	Empl.	Comp.	Sex
Drawing	21	26^{c}			
Symptoms	32	21^{c}	1.6^{a}		0.28^{a}
Signs	19	13^{b}			
IBQ (D.F.)	32	32^{c}			
MSPQ	23	18^{c}	1.7^{a}		
HAD(a)	18	24^{c}	2.3^{a}		
HAD(d)	39	33^{c}	3.4^{b}	6.8^{a}	
Zung	41	21^{c}	2.6^{b}		
MSPQ/Zung	40	19^{c}	1.9^{b}	4.0^{a}	
HAD Comb.	33	30^{c}	3.4^{b}		

$^{a}P < 0.05$, $^{b}P < 0.005$ and, $^{c}P < 0.0001$

Discussion

Patient selection and pattern of referral

The population sample used in this study was chosen to be representative for patients referred to an orthopaedic surgeon with an interest in spinal problems. The use of consecutive patients from the day sheets was designed to reduce sample bias, and only the selection criteria detailed in the methods were used to further define the population.

In Australian practice the general practitioner would invariably refer back pain patients with Worker's Compensation claims to a specialist if symptoms failed to settle rapidly enough to allow an early return to work. Thus the patients studied were likely to have a broad spectrum of severity of back pain and also include those with relatively minor problems.

Referrals for second opinions might be expected to be different in their behaviour from primary G.P. referrals. However, these patients did not prove more chronic in their presentation as the interval from the accident to the first consultation was not increased and indeed no difference in their Outcome scores was demonstrated.

Exclusion of patients who underwent subsequent surgery removed a more disabled group and thus may improve the resemblance of the study population to that which might present to a general orthopaedic specialist. The population sample used in this study contained patients who, although having suffered from low back pain in the past, were not all having pain when seen. They may be considered more representative of the population presenting to orthopaedic specialists than either patients with chronic pain or "normal" patients with no history of back pain.

A further advantage of the sampling frame was the amount of complete and detailed information collected in a standardised manner by a single observer, which is unusual for a retrospective study. Identical questions could be asked at review, allowing comparison of the patients' condition at review with that at presentation. It was hoped that the Inorganic signs of Waddell et al. (1980) and the pain drawing (Ransford 1976) would provide a measure of the psychological status of the patient at presentation. Unfortunately I found that these two psychometric tests were least discriminating for psychological disturbance.

The achievement of a follow up rate of 91 percent adds weight to the findings of the study. Using the criteria of Nachemson and LaRocca (1987) a worst case analysis is unnecessary.

Performance of the Outcome score and Disability measurements

The Outcome score devised for the assessment of the patients appeared to perform satisfactorily. There was good correlation with the Oswestry Disability score (Fairbank et al. 1980) and with the disability rating of Waddell and Main (1984). The new Outcome score, however, provided an increased spread of scores among the patients with the least disablement, particularly in the women. This makes it more useful in the assessment of back function in the least disabled groups. The improved discriminating ability of the new Outcome score between employed and unemployed patients, compared with the two disability scores, is in part due to the inclusion of employment itself in the scoring system, but this only represented 12 percent of the score, and other factors such as the inclusion of high demand activities may also be important.

Unlike the disability and Outcome scores which are completed by the patient, the physical impairment rating of Waddell is open to observer bias. Additionally the "non test" straight leg raise which forms part of the impairment rating has poor reproducibility (McCombe 1989) and considerable diurnal variation in patients with prolapsed discs (Porter 1990). The rating proved to be unsatisfactory as a measure of general outcome, as it did not reflect factors such as employment or activities of daily living. It did, however, make an independent contribution to the Outcome scale in multiple regression analysis, although this contribution was small.

The Outcome score, the Oswestry disability score, and the disability score of Waddell were all influenced by psychological disturbance. The physical impairment rating of Waddell was also influenced in this way, but to a lesser extent. It has been shown previously that the range of flexion as measured for the impairment rating is influenced by psychological factors (Rae et al. 1984). The presence of psychological disturbance caused separation of the regression lines of disability and Outcome scores against the physical impairment rating. Although this is in general agreement with the findings of Waddell and Main (1984), in the present study the difference made by psychological disturbance was significant only with lower disability and impairment scores. As the influence of psychological disturbance on the physical impairment rating was not greater with increased impairment, this suggests that psychological disturbance plays an important role in the expression of disability in minimally impaired patients, but not in

severely impaired patients. As the impairment rating itself is increased by psychological disturbance, the finding of disability in a psychologically disturbed patient with a normal impairment rating may well be suggestive of somatisation.

Length of follow up

It is of interest that the outcome results did not change with the length of follow up. This suggests that a prognosis based on the situation at one year is of considerable value and that further developments of the patient's condition are unlikely.

Interval between accident and primary consultation

In the compensation patients there was an association between increasing delay between injury and primary consultation and poor Outcome scores. This may result from either an increasing resistance to treatment with increasing duration of the complaint or selection out over time of the more serious injuries which are eventually referred. The lack of change of the Outcome score with increasing interval in the noncompensation patients suggests that such selection might not occur. It may be that over the initial months of the compensation claim some other factor operates to worsen the prognosis. This may be a prolonged time off work, in part instigated by the need to demonstrate the severity and the time taken to initiate claim procedures.

Sex

The difference in Outcome score between the sexes in the noncompensation patients was not due to any one or two factors in the Outcome scale as differences were observed in all factors except employment and sex life, neither was it due to differences at presentation. The difference may lie in the ability to resume full activities in all spheres and not just a few, as there was no difference between the sexes in the compensation patients who had lower overall scores.

Age at accident

Reduced Outcome scores at review were associated with increasing age. The accident severity, however, was greatest in the younger patients, implying that their ability to recover is greater. This observation is in agreement with the findings of Bigos et al. (1986).

Type of accident

The frequency with which an accident was claimed as the precipitating cause of back pain was quite different in the compensation patients and the noncompensation patients. This may merely indicate that claiming compensation depends to a great deal on the demonstration of causation. Further work is needed in this area in a work-place or primary health care setting. The major modes of injury were lifting and falling. Data reported from the USA also contained a large proportion of lifting injuries (48%), with falls less common (4%) (Klein et al. 1984). The incidence of a jerk during lifting was less common than the 24 percent reported in a study of physiotherapists (Molumphy et al. 1985), but that represented the working pattern of a specialised group. Contrary to the findings of Josefsen et al. (1987), no differences in the types of accident were observed between the sexes in this study.

In the compensation patients twisting, slipping and jerking injuries produced more disability than other injuries, a finding in agreement with other studies (Manning and Shannon 1981, Magora 1973, Molumphy et al. 1985, Kelsey et al. 1984). In contrast to other studies (Troup et al. 1981), falls had a better outcome than most other accident types. These differences, however, were not observed in the noncompensation patients. This suggests that it may not be the mechanism of the accident but some other factor which was associated with the reduction in the outcome score. Slipping and jerking accidents, for example slipping on an oil spillage on the floor, may be more demonstrably the fault of a third party.

Diagnosis

There appeared to be little relationship between the diagnosis and Outcome scores. The diagnoses of prolapsed disc and spondylolysis and spondylolisthesis were supported by radiographic findings consistent with the clinical picture. It is possible that in some of the cases of spondylolysis and spondylolisthesis the symptoms may have been arising elsewhere in the spine as both of these radiographic features may be found in symptom-free spines (Frederickson et al. 1984). The diagnosis of discogenic pain was basically a clinical one and as such is open to more error. Facet-joint arthropathy was not always confirmed radiographically as CT scans were not available for every patient. However, the diagnostic criteria were applied consistently by a single treating surgeon throughout the period of patient recruitment, and the groups of patients therefore had a measure of internal consistency.

Multivariate analysis demonstrated that the six diagnoses together contributed only 3 percent to the variance of the Outcome score. Treatments currently direct-

ed at specific pathological entities are thus addressing only a small fraction of the etiology of the disability. This has important implications for the management of patients with a low back pain, suggesting more attention needs to be given to psychological and socio-economic factors.

Diagnosis related to injury type

The relationship of injury type to diagnosis has not been previously explored in a well defined group using consistent diagnostic criteria. The association of impact with soft tissue strain is hard to explain on pathological grounds, but this association was predominantly found in the compensation patients as only one noncompensation patient was included in the group. The patients in this group also had a higher incidence of psychological disturbance than the population sample as a whole and a possible explanation is that this type of postural soft tissue strain was a result of increased muscle tone secondary to stress. Bending injuries were associated with the group of other diagnoses, and this is difficult to explain.

Kelsey et al. (1984) found that combined lifting and twisting were associated with an increased incidence of disc prolapse. In the present study twisting injuries were associated with both disc prolapse and discogenic pain, and this supports the concept of an important role for torsion both in the generation of disc prolapse (Farfan et al. 1970) and possibly in the formation of outer annular tears (Vernon-Roberts and Pirie 1977, Hilton and Ball 1984). Bending injuries were also associated with disc prolapse lending some support to the concept of prolapse being a hyperflexion injury (Adams and Hutton 1982), although these injuries were not sustained under any load. Lifting injuries were associated with discogenic pain and could possibly be mediated through the production of annular tears with subsequent degeneration (Osti et al. 1990). The apparent connection between facet arthropathy and jerking or falling injuries may represent secondary changes following a joint injury but the median interval between accident and diagnosis was only 10 months and this might not be sufficient for such changes to occur. An alternative explanation would be the further injury of a previously degenerate joint e.g. fracture of an osteophyte precipitating symptoms. No causal relationship could be ascertained for the association of falls and spondylolisthesis or spondylolysis. The accident severity score of falls in patients with these diagnoses was no greater than that in other falls.

Accident severity

The arbitrary nature of the accident severity score is open to criticism. The score used in this study, however, was used to distinguish the trivial from the severe and at least was applied uniformly throughout. Accident severity varied with age but appeared to play little part in outcome or in determining the diagnosis.

Initial pain and neurological deficit

The pain pattern recorded at the primary consultation varied between the compensation patients and the controls, with more of the compensation patients indicating pain in all three areas. This was not associated with increased evidence of neurological deficit in the these patients and may reflect the increased incidence of psychological disturbance in the same manner as the pain drawing attempted to do. It is interesting that the group indicating pain in all areas fared worse at review in the noncompensation patients but were no different from the rest in the compensation claimants.

In both compensation and noncompensation patients there were associations between leg pain (either alone or combined with the other two areas) and a neurological deficit on primary examination. As might be expected a neurological deficit was more common in the patients with disc prolapse. A surprising finding, however, was the lack of a difference in outcome between the patients with an initial neurological deficit and those without. This is agreement with the findings of Currey et al. (1979) but is contrary to those of Berquist-Ullman and Larsson (1977) and Troup et al. (1981). The lack of influence of a neurological deficit on outcome implies considerable redundancy of capability of the lower limb in most activities, however, none of these patients had a drop foot.

Initial treatment

The initial treatment received following the accident was not very successful. No treatment modality was rated as successful by more than one third of the patients as a whole. There was also little difference in the assessment ratings between the compensation and noncompensation patients (except for physiotherapy). The current work suffers, however, from being retrospective, and in addition the patients were inclined to aggregate the initial treatment with subsequent therapy over the course of time. During the interviews the investigator gained a strong impression that no treatment provided more than temporary relief, either in the compensation or the non compensation groups. As expected the compensation patients had a higher proportion of works personnel or casualty departments

providing the first therapy, but in both groups the G.P. provided initial care most often.

There was no correlation between the initial therapies prescribed following the injury and the diagnoses made at the primary consultation with the specialist. This lack of agreement in other practitioners as to which therapy is suitable for which clinical presentation further supports the finding that none of the treatments was of any specific value in any of the conditions. Two therapies (surgical belt and epidural injection) were associated with worse Short Outcome scores at the primary consultation. In this retrospective study the cause of this association cannot be determined; it may be that these treatments were less effective than the others or that they were prescribed in more disabled patients.

There was a difference in the pattern of time elapsed from the accident to the first consultation with a doctor or nurse. More than half of the compensation patients were seen the same day, whereas only one fifth of the noncompensation patients were seen this early. It is most interesting to note that the compensation patients seen within the first 2 days fared worse at follow up than those seen later. As the accident severity scores were the same in each of the two groups it suggests as one possibility that the personality of the patient was important in determining the urgency of treatment and in part the subsequent outcome. This is supported by the observation that the Short Outcome score at the primary consultation showed no differences between the two groups, with the implication that the initial disability was similar.

Treatment following primary consultation

Treatment following the primary consultation was similar in the compensation patients and the noncompensation patients. The success of this treatment cannot be quantified as there is no control group which received no treatment. It is however clear that the educational classes contributed to the improvement in the Short Outcome scores between the primary consultation and follow up, a finding also noted by Berquist-Ullman and Larsson (1977) and Sikorski (1985). This supports my impression, gained from the interviews, that many patients found these classes helpful. The patients treated by exercises alone had a similar increase in Short Outcome score to patients treated with other modalities, in contrast to the findings of Manniche et al. (1988) who reported improved results from an exercise regime. Their regime was, however, supervised and intensive, in contrast to the self administered treatment given to patients in this study. With the more specific therapies, of which the best example is injections for facet arthropathy, the improved response of the noncompensation patients over the compensation patients was marked and agrees with the findings of Krusen and Ford (1958). These authors studied the response to a rehabilitation programme of 272 compensation claimants and 237 noncompensation patients. They found 89 percent of the noncompensation group were rated as improved compared with 56 percent of the compensation group. Compensation patients also had more treatments than noncompensation patients.

Time off work

The present work demonstrated a strong link between increasing time off work following an injury and subsequent unemployment. This association could only be seen in the compensation patients as there were insufficient noncompensation patients unemployed at follow up for analysis. There was also an inverse correlation between the Outcome score at review and increasing time off work, both in the compensation and the noncompensation groups. This association is a possible factor in the poor performance of the compensation patients, who were encouraged by their advisors not to return to work early. The loss of regular exercise at work and in daily travel may lead to further spinal decompensation and increased symptoms.

Social group

In compensation patients there was an association of social group with employment status at follow up and this observation agrees with the findings of Tunturi (1979) who was examining the results of spinal fusion. No such association, however, was found in the noncompensation patients, which questions whether back injured workers in heavy jobs are necessarily more at risk of loss of their working capacity. Outcome measurements at review were also influenced by social group, but again more consistently in the compensation patients than in the noncompensation patients.

Within the compensation patients no variation occurred in diagnosis or in type of accident with social group. In the noncompensation cases, lifting injuries were less common and falls more common in Groups 5–6. It is possible that patients would be tempted to ascribe their injury to a lift performed at work rather than one at home, which would be more difficult with a fall, but there was no evidence to support this suggestion.

Migrant status

The influence of migrant status on subsequent employment was small and after the patients were controlled for social group there was no difference between immigrants and native English speakers. This agrees with the findings of Hewson et al. (1987) who found that although there was an increase in the numbers of immigrants overall, this was not significant after controlling for social group.

Although more of the migrants were unemployed at review this was secondary to differences in the distribution of social groups between the immigrants and the native English speakers. Similarly, although immigrants had more psychological disturbance, when other factors were included in multivariate analysis migrant status made no significant contribution. Differences in Outcome scores in both the compensation and noncompensation patients between the immigrants and the native English speakers were not reflected by a contribution of migrant status in the multi variate analysis. Thus there was no evidence to support any suggestion that immigrants are overrepresented in compensation claims or that they have a less satisfactory recovery.

Patients classified as immigrants to Australia in this study population were predominantly from Greece and Italy. They may differ from immigrants of refugee or illegal immigrant status to Europe or the U.S.A., but are not likely to differ from other immigrant groups to these countries where routine immigration is controlled in a similar manner to Australia. A proportion of the first generation immigrants spoke poor English which may have contributed to the popular prejudice against them as a group claiming more than their fair share of state benefits.

Compensation

Although there was an increased proportion of second referrals and consultant referrals amongst the compensation patients compared with the noncompensation patients analysis of the Outcome scores revealed no differences between these groups.

The results of this study showed that compensation, in particular lump sum compensation, is associated with an increased rate of unemployment, and a prolonged time off work. The differences are still very marked and highly significant after controlling for social class. The unemployment rate of 49 percent in the compensation patients was somewhat higher than the rate of 33 percent reported by White (1969) for conservative therapy but in that study the follow up was only 81 percent.

In contrast to the findings of Leavitt et al. (1982), the compensation patients reported more pain and disability than matched controls, which was further accentuated in those who were unemployed or had a psychological disturbance. To my knowledge this is the first controlled study of this type on the effects of compensation on the recovery from low back pain.

The compensation patients in this study were more disturbed than controls. The analysis suggested that compensation and unemployment have independent approximately equal effects on the incidence of disturbance.

Patients claiming a lump sum appeared much worse off than those on intermittent payments only. However, in the Workers' Compensation Scheme some patients became involved in a lump sum claim only when a dispute arose concerning their accident, when in some cases the insurance company ceased payments unilaterally. Thus in the lump sum group there was an increased incidence of disputed claims, or of patients who had been off work for some time previously on intermittent payments. However, patients claiming lump sums in a negligence or third party claim had an equal disability at review to those claiming lump sums in the Workers' Compensation system. Thus the higher awards available to third party claimants did not improve the outcome.

Neither unemployment nor psychological disturbance was affected by settlement, either immediately or after several years. Following settlement, even in the employed patients psychological disturbance continued in many, even five years post settlement. From this it could be inferred that these patients are more easily stressed than the controls, or that there is a certain personality type more likely to be involved in a claim. Although no information concerning the pre-injury status of these patients was available, if either of the postulates are correct then the present system has allowed them to take a path which has very poor outcome. Parker (1977) found that approximately one fifth of his compensation patients had some evidence of pre-existing neurotic symptoms, but this was assessed retrospectively. Murphy and Cornish (1984) reported that some scales of the MMPI together with the pain drawing could predict which patients with acute injuries would become chronic. There is a need for a suitably controlled prospective study in the work place to examine the pre-injury status more closely.

Settlement of the claim was associated with a small increase in the Outcome score in the men, but there was no change in reported pain, Oswestry disability score or Waddell's physical impairment rating. Talo et al. (1989) reported similar findings.

Following settlement the patients seeking treatment were more likely to have psychological disturbance. As the response to therapy in compensation patients

depends in part on the extent of psychological disturbance (Dzioba and Doxey 1984), this has important implications if operation is contemplated. In addition, compensation has been found to reduce the response of psychological disturbance itself to treatment (Trief and Stein 1985).

The results of this study demonstrated that although the compensation patients do badly, the noncompensation patients do very well. It may be argued that the noncompensation patients have a significant incentive to return to work as soon as possible to restore their income. This hypothesis is supported by our findings in a number of self-employed patients in the compensation group who had a private sickness benefits policy. In order to minimise premiums these were often arranged to pay only a subsistence income, significantly below the patients normal income. These patients as a group behaved similarly to the noncompensation patients. A comparison of patients in New Zealand on minimal payments with similar patients in the USA also supports this hypothesis (Carron et al. 1985).

Many explanations might be advanced for the poor results achieved by the compensation system. In this study the three factors which appeared most relevant were the delay in return to work, the adversarial nature of the system, and the effect of the claim history on prospective employers. Return to work as part of a treatment programme increases the success rate (Catchlove and Cohen 1982). This was confirmed in this study; the patients who returned to work in less than six months had a greatly increased chance of employment at follow up than those who were out of work longer. Under the South Australian Workers' compensation system the employer was liable for any injury or disability arising at work, regardless of pre-existing conditions. This produced a strong incentive for the employer, under pressure sometimes from the insurance companies, to terminate the employment of an injured worker. Thus part time work or light duties which might otherwise have been valuable in a graduated return to work were often unavailable.

Secondly, the adversarial system is associated with a poor outcome, the lump sum claimants contributing most to the poor results of the compensation group. The patients claiming intermittent payments did not differ in many of the outcome parameters from the controls. Under the adversarial system patients were required to recount their symptoms and disabilities on numerous occasions to lawyers, doctors and the court. I gained a strong impression that the adjustment of the patient to his back pain was the crucial factor. This adjustment appeared reduced under the stressful conditions of litigation, for example some patients were followed by agents of the insurance companies using video cameras gathering evidence to produce in court. This may reduce compliance with rehabilitation exercise regimes. The adversarial system was also associated with delays; the median time to settlement in this series was nearly two and a half years.

The system of medical reports appeared unsatisfactory. Haddad (1987) found that in 1,818 compensation claims the claimants' specialists awarded no disability in 0.5 percent of cases whereas the insurance companies' specialists awarded no disability in 75 percent. A similar dichotomy became apparent to the investigator during this review. Such differing medical opinions may produce feelings of stress. The lack of quantifiable parameters makes the value of medical reports in the context of back pain questionable. In this regard the Physical Impairment Rating of Waddell was unhelpful, as it was affected by psychological disturbance and also contributed only a small amount to outcome.

Thirdly, those patients who had settled their cases and who were actively looking for work found that their history of a claim for back injury was a significant handicap. This difficulty with employment was one of the most common reasons cited by those patients who had indicated they would not go through the claim procedure again.

It thus appears that the compensation system, particularly the lump sum system, acts directly and powerfully against the long term interests of the patient. Further, this adverse effect continues after settlement, for at least 5 years in this study. I would recommended that the lump sum system of compensation for low back injury should be abolished, that legal involvement should not occur, and that a continuous payment system should be considered for these cases.

Psychological disturbance at review

Psychological disturbance was present in 36 percent of the compensation patients and 9 percent of the noncompensation patients. Main and Waddell (1984) found an incidence of 15–20 percent in chronic pain patients and 5 percent in the normal population. The incidence of psychological disturbance was not related to social group, and differences existed between the compensation patients and the noncompensation patients in all group bands although they were not significant for social Group 4. Psychological disturbance had a significant effect on the Outcome score at follow up in all social groups.

Multivariate analysis showed that both compensation and unemployment were associated with an increased incidence of psychological disturbance. In this study the effect of unemployment was more marked than that

of compensation. Current pain as recorded by the visual analogue pain scale was even more strongly associated with psychological disturbance, but whether the association was causal could not be determined.

Each of the outcome and disability scales was influenced by both compensation and employment status. The Outcome scale was more strongly affected by employment, but employment is part of the scale itself. In the compensation patients, the effect of psychological disturbance on the various self-report ratings when controlled for employment was considerably reduced and significant only in two instances, whereas if the patients were controlled for psychological disturbance unemployment still had an effect on these scales. In the compensation patients as a whole, therefore, the profound effect of psychological disturbance on the outcome measures used in research and clinical practice may be due mainly to the strong association of disturbance with unemployment.

In the noncompensation patients the contrary situation pertained. Although the numbers in the unemployed and disturbed groups were small, psychological disturbance had a marked effect on the self-report scales, particularly in the employed patients where there was a large group. No effect was seen of unemployment, except in the Outcome score in which unemployment itself is included.

This difference between the compensation patients and the noncompensation patients is interesting. It suggests the possibility of a learned behaviour pattern in the compensation patients of regarding themselves as disabled which is particularly pronounced in the unemployed patients. These patients have all had to justify their disability on a number of occasions during the claim procedure. The higher scores seen in the compensation patients may be masking the effect of psychological disturbance in this group. Unemployment in the noncompensation patients (which at 8 percent was average for Adelaide) was not linked to the effects of the previous back accident and the patients did not regard themselves as disabled.

Comparison of psychometric methods

The comparison of the psychometric methods demonstrated that many of those used in patients with low-back pain fail to accurately distinguish psychological disturbance. The ability of this study to discriminate between the tests may have been improved by the broad population base.

The assessment of psychometry in patients with pain or illness is also made difficult as physical limitations may be confused with psychological symptoms in some questions (Zigmond and Snaith 1983, Watson 1982). Using the Minnesota Multiphasic Personality Inventory (MMPI) for example, elevation of the Hs and Hy scores in pain patients has been shown to be due to their endorsement of physical symptoms relevant to their pain rather than their being hypochondriacal (Watson 1982). In the tests which performed well in this study (the Hospital Anxiety and Depression Scales, the Modified Somatic Perception Questionnaire, and the Zung Depression Scale) there was very little possibility of such confusion; indeed the first three were designed with this objective in view.

In this study psychological disturbance was defined using a combination of all the test scores. To improve comparability with other published work, the use of the MMPI was considered. In the MMPI, however there is overlap of physical with psychological symptoms, and the MMPI has also been shown to distinguish poorly between patients with functional diagnoses and those with organic diagnoses (Schwartz and Krupp 1971) and has a considerable overlap between such patients (McCreary et al. 1977). The cumbersome nature of the MMPI (it contains 566 items, takes 90 minutes or more to complete. and requires a reading ability of high school equivalent) was also disadvantageous in a study of this nature where patients attended at the investigators' request and not as part of a treatment programme. The educational requirements also excluded a number of the population sample, particularly in the immigrant group.

The resemblance of the shape of some of the curves suggests they may be measuring the same features. The HAD(a) and MSPQ curves closely resembled each other but still differed markedly in form from the two depression scales, the Zung and the HAD(d), which in turn are closely related to each other. This observation lends support to the use of these methods in pairs of dissimilar tests when they may complement each other by measuring different aspects of disturbance.

Pain, social group, compensation, migrant status and employment all affected the cut-off scores of some of the tests. For any method to be universally applicable these effects should be minimal, or test results would have to be analysed for each separate subgroup. Migrant status affected the tests least, supporting the concept that immigrants do not differ in their psychological responses. For the compensation, pain plus and social groups 4–6 subgroups, the difference in cut-off scores may be due to the observed increase in median scores in the nondisturbed patients. This suggestion is supported by the generally higher specificities and sensitivities in the noncompensation, pain minus and social groups 1–3 sub-groups (Tables 3–5).

The Pain Drawing was found to be of little value because of its relative insensitivity. This poor correla-

tion agrees with that found by Van Baeyer et al. (1983), who reported a specificity of 80 percent, with a sensitivity of 44 percent.

The distribution of Inappropriate Symptoms in this study agreed with the distribution for all back patients reported by Main and Waddell (1984). This test discriminated between the two groups quite well, but had lower specificities and sensitivities than the best of the tests used in this study and showed an increased variability with pain, social group, and compensation. Items are also obtained during the history taking, allowing the possibility of observer bias or accidental omission.

The Inappropriate Signs were also found to be insensitive, with nearly 80 percent of all patients scoring zero. Even in the disturbed group, only 30 percent scored more than one. The discrepancy between these results and those of Waddell et al. (1980) may be due to differences in the populations studied, or to inter-observer variation. Waddell's subjects were Canadians and Britons, both normal subjects and referrals to ordinary orthopaedic clinics and problem back clinics. It has not been shown that Australians differ from Britons or Canadians, and the results from the other tests in this study suggest that they do not. It remains probable that the difference observed is due mainly to the inter-observer variability reported by McCombe et al. (1989). This reduces the value of the Inappropriate Signs in the assessment of these patients.

The results of this study confirm the reservations of Main and Waddell (1987) concerning the use of the Illness Behaviour Questionnaire (discriminant function) in patients with back pain, the more so as the present sample was drawn from the population in which this test was developed. The present results, however, do not confirm the discriminating power of the three factors proposed by these authors; indeed the factor they designated social inhibition failed to distinguish at all between the two groups. The Hospital Anxiety and Depression Scales, the Modified Somatic Perception Questionnaire, and the Zung Depression Scale were all simple to administer and score, and were all acceptable to the patients whose compliance with these tests was good. It was chosen to present the results in terms of a single cut-off value, in order to identify tests which may be used quickly and easily in clinical practice.

The combination of the MSPQ and the Zung depression scale provides an instrument which is least affected by other factors such as pain, social group, compensation and migrant status. Using the cut-off score given here, it provides a tool for the assessment of the psychological status of patients with low back pain which is simple, convenient and accurate.

Summary and future research

The Outcome Scale designed for use in this study provided a comprehensive and discriminating assessment of patient function. This scale should now be tested in further work in particular to assess the effect of conservative or surgical therapy. It is important that such studies should be prospective and that the scale be administered at presentation. This will allow comparison of the patients' condition before treatment and at follow up. The physical impairment rating makes only a small contribution to the Outcome scale but has some validity as an independent factor and may be worth including in some studies. As all the scales were influenced by psychological disturbance it is important that reports of future work should characterise the psychological profile of the study population.

The length of follow up did not affect the outcome, suggesting an accurate prognosis could be given at one year. Patients who were older at the time of the accident fared worse, but accident severity did not influence the outcome. Neurological deficit was not associated with a poorer prognosis.

The type of accident, although associated with diagnosis in some cases, does not appear to have a direct relationship to recovery. The diagnosis in patients with back pain following an accident plays little part in determining the recovery, neither do specific therapies directed to particular pathologies. Education classes, however, are associated with improved outcome. Time off work is closely associated with outcome, and treatment programmes should therefore aim at early return. Immigrants are not more likely to be unemployed or psychologically disturbed and regression analysis indicates that migrant status is not a factor in the outcome. Previous findings may be explained by the variation in social groups and other factors.

Compensation status plays a major role in the recovery from low back injury. Lump sum claims are more deleterious than intermittent payments. The type of lump sum claim is not a factor in outcome. Settlement makes no difference to the patient's recovery. Reform of the compensation system is indicated.

Psychological disturbance is also a major determinant of outcome. This work does not allow determination of the psychological status of the patient at the time of accident and further work in an industrial setting is required to clarify whether disturbance is the cause or the result of poor treatment.

A psychometric method has been identified which has a high specificity and sensitivity and is easy to apply in clinical practice.

The finding that compensation and psychological disturbance are prognostic factors of greater magnitude

than diagnosis or treatment is clearly relevant for the design of future studies of low back pain. These factors must be rigidly controlled or be included in the analysis, and multifactorial analysis may be required.

Acknowledgements

During the period of the study the author was Spinal Research Fellow in the Spinal Unit, Royal Adelaide Hospital, Adelaide, South Australia. Mr Robert D Fraser MD FRACS is the Head of the Unit and the study was made on patients treated by Dr Fraser.

At the time of the study the author was the holder of the Thackray Orthopaedic Travelling Fellowship and also received support from the Adelaide Bone and Joint Research Foundation.

This study formed the thesis for the degree of MD at Cambridge University, awarded in 1991. Adelaide Bone and Joint Research Foundation, Aesculap Limited, Ethicon Limited, South Tees Acute Hospital Trust and Stratec Medical supported publication of this thesis.

References

Adams M A, Hutton W C. Prolapsed Intervertebral Disc. A Hyperflexion Injury. Spine 1981; 7: 184-191.

Bergquist-Ullman M, Larsson U. Acute Low Back Pain in Industry. Acta. Orth. Scand. 1977; Supp. 170: 1-117.

Bigos S J, Spengler D M, Martin N A, Zeh J, Fisher L, Nachemson A. Back Injuries in Industry: a Retrospective Study. 3. Employee Related Factors. Spine 1986; 11: 252-256.

Carron H, DeGood D E, Tait R. A Comparison of Low Back Pain Patients in the United States and New Zealand: Psychosocial and Economic Factors Affecting Severity of Disability. Pain 1985; 21: 77-89.

Catchlove R, Cohen K. Effects of a Directive Return to Work Approach in the Treatment of Workmen's Compensation Patients with Chronic Pain. Pain 1982; 14: 181-191.

Currey H L F, Greenwood R M, Lloyd G G, Murray R S. A Prospective Study of Low Back Pain. Rheumatol. Rehabil. 1979; 18: 94-104.

Daniel A. Power, Privilege and Prestige. Longman Cheshire Pty Limited, Melbourne 1983: 197-206.

Dzioba R B, Doxey N C. A Prospective Investigation into the Orthopaedic and Psychological Predictors of Outcome of First Lumbar Surgery Following Industrial Injury. Spine 1984; 9: 614-623.

Fairbank J C T. The Incidence of Back Pain in Britain. In Back Pain. Methods For Clinical Investigation and Assessment (Eds. Huskins DWL, Mulholland RC). Manchester University Press 1986: 1-12.

Fairbank J C T, Couper J, Davies J B, O'Brien J P. The Oswestry Low Back Pain Disability Questionnaire. Physiother. 1980; 66: 271-273.

Farfan H F, Cossette J W, Robertson G H, Wells R V, Kraus H. The Effects of Torsion on the Lumbar Intervertebral Joints: the Role of Torsion in the Production of Disc Degeneration. J. Bone Joint Surg. [Am] 1970; 52-A: 468-497.

Fredrickson B E, Baker D, McHolick W J, Yuan H A, Lubicky J P. The Natural History of Spondylolysis and Spondylolisthesis. J. Bone Joint Surg. [Am] 1984; 66-A: 699-707.

Frymoyer J W, Cats-Baril W. Predictors of Low Back Pain Disability. Clin. Orthop. 1987; 221: 89-98.

Greenough C G, Fraser R D. The Effects of Compensation on Recovery from Low-Back Injury. Spine 1989; 14: 947-955.

Greenough C G, Fraser R D. Comparison of Eight Psychometric Instruments in Unselected Patients with Back Pain. Spine 1991; 16: 1068-1074.

Greenough C G, Fraser R D. Assessment of Outcome in Patients with Low-Back Pain. Spine 1992; 17: 36-41.

Haddad G H. Analysis of 2932 Workers' Compensation Back Injury Cases. The Impact on the Cost to the System. Spine 1987; 12: 765- 769.

Hewson D, Halcrow J, Brown C S. Compensable Back Pain and Migrants. Med. J. Austral. 1987; 147: 280-284.

Hilton R C, Ball J. Vertebral Rim Lesions in the Dorsolumbar Spine. Ann. Rheum. Dis. 1984; 43: 302-307.

Howe J, Frymoyer J W. The Effects of Questionnaire Design on the Determination of End Results in Lumbar Spine Surgery. Spine 1985; 10: 804-805.

Jackson R P, Jacobs R R, Montesano P X. Facet Joint Injection in Low-Back Pain: a Prospective Statistical Study. Spine 1988; 12: 966-971.

Josefsen K, Boss A H, Andersen V, Biering-Sorensen F. One Year's Notifications of Back Injuries to the National Social Security Office, Especially with Respect to Nature of Cause, Medical Diagnoses and Consequences. In: Musculoskeletal Disorders at Work (Ed. Buckle P). Taylor and Francis, London 1987: 68-73.

Kelsey J L, Githens P B, White III A A, Holford T R, Walter S D, O'Connor T, Ostfeld A M, Weil U, Southwick W O, Calogero J A. An Epidemiologic Study of Lifting and Twisting on the Job and Risk for Acute Prolapsed Lumbar Intervertebral Disc. J. Orthop. Res. 1984; 2: 61-66.

Klein P B, Jensen R C, Sanderson L M. Assessment of Workers' Compensation Claims for Back Strains/sprains. J. Occup. Med. 1984; 26: 443-448.

Krusen E M, Ford D E. Compensation Factor in Low Back Injuries. J. Amer. Med. Ass. 1958; 166: 1128-1133.

Magora A. Investigation of the Relation between Low Back Pain and Occupation 4. Physical Requirements: Bending, Rotation, Reaching and Sudden Maximal Effort. Scand. J. Rehab. Med. 1973; 5: 186-190.

Main C J. The Modified Somatic Perception Questionnaire (MSPQ). J. Psychosom. Res. 1983; 27: 503-514.

Main C J, Waddell G. The Identification of Psychological Abnormality in Patients with Chronic Low Back Pain. Available from C J Main, Principal Clinical Psycholgist, Hope Hospital, Eccles Old Road, Salford M6 8HD, UK, 1984.

Main C J, 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.

Manniche C, Hesselsoe G, Bentzen L, Christensen I, Lundberg E. Clinical Trial of Intensive Muscle Training for Chronic Low Back Pain. Lancet 1988: 1473-1476.

Manning D P, Shannon H S. Slipping Accidents Causing Low-back Pain in a Gear Box Factory. Spine 1981; 6: 70-72.

McCombe P F, Fairbank J C T, Cockersole B C, Pynsent P B. Reproducibility of Physical Signs in Low-back Pain. Spine 1989; 14: 908-918.

McCreary C, Turner J, Dawson E. The Mmpi as a Predictor of Response to Conservative Treatment for Low Back Pain. J. Clin. Psychol. 1979; 35: 278-284.

Mendelson G. Not "cured by a Verdict". Effect of Legal Settlement on Compensation Claiments. Med. J. Austral. 1982; 2: 132-134.

Mendelson G. Compensation, Pain Complaints and Psychological Disturbance. Pain 1984; 20: 169-177.

Miller H. Accident Neurosis 1. Br. Med. J. 1961: 919-925.

Miller H. Accident Neurosis 2. Br. Med. J. 1961: 992-998.

Molumphy M, Unger B, Jensen G M, Lopopolo R B. Incidence of Work Related Low-back Pain in Physical Therapists. Physical Therapy 1985; 65: 482-486.

Murphy K A, Cornish R D. Prediction of Chronicity in Acute Low Back Pain. Arch. Phys. Med. Rehabil. 1984; 65: 334-337.

Nachemson A L, Larocca H. Editorial: Spine 1987. Spine 1987; 12: 427-429.

Osti O L, Vernon-Roberts B, Fraser R D. A Sheep Model of Disc Degeneration Initiated by a Peripheral Annulus Tear. Read at Int. Soc. Study Lumbar Spine, Boston 1990.

Parker N. Accident Litigants with Neurotic Symptoms. Med. J. Austral. 1977; 2: 318-322.

Pheasent H C, Gilbert D, Goldfarb J, Herron L. The MMPI as a Predictor of Outcome in Low-back Surgery. Spine 1979; 4: 78-.

Pilowsky I, Spence N D. Manual for the Illness Behaviour Questionnaire (IBQ). Avaiable from Professor I. Pilowsky Dept. Psychiatry, University of Adelaide, Adelaide, South Australia 5000 1983.

Piterman L, Dunt D. Occupational Lower-back Injuries in a Primary Medical Care Setting: a Five Year Follow-up Study. Med. J. Austral. 1987; 147: 176-279.

Porter R W, Trailescu I F. Diurnal Changes in Straight Leg Raising. Spine 1990; 15: 103-106.

Rae P S, Waddell G, Venner R M. A Simple Technique for Measuring Lumbar Spinal Flexion. J. Roy. Coll. Surg. Edin. 1984; 29: 281-284.

Ransford A O, Cairns D, Mooney V. The Pain Drawing as an Aid to the Psychologic Evaluation of Patients with Low-back Pain. Spine 1976; 1: 127-134.

Repko G R, Cooper R. A Study of the Average Workers' Compensation Case. J. Clin. Psychol. 1983; 39: 287-295.

Rossignol M, Suissa S, Abenhaim L. Working Disability Due to Occupational Back Pain: Three-year Follow-up of 2,300 Compensated Workers in Quebec. J. Occup. Med. 1988; 30: 502-505.

Schwartz M S, Krupp N E. The MMPI "conversion V" Among 50,000 Medical Patients: a Study of Incidence, Criteria and Profile Elevations. J. Clin. Psychol. 1971; 27: 89-95.

Sikorski J M. A Rationalized Approach to Physiotherapy for Low-Back Pain. Spine 1985; 10: 571-579.

Talo S, Hendler N, Brodie J. Effects of Active and Completed Litigation on Treatment Results: Workers' Compensation Patients Compared with Other Litigation Patients. J. Occup. Med. 1989; 31: 265-269.

Tarsh M J, Royston C. A Follow-up Study of Accident Neurosis. Br. J. Psychiat. 1985; 146: 18-25.

Triano J J, Schultz A B. Correlation of Objective Measure of Trunk Motion and Muscle Fuction with Low-back Disability Ratings. Spine 1987; 12: 561-565.

Trief P, Stein N. Pending Litigation and Rehabilitation Outcome of Chronic Back Pain. Arch. Phys. Med. Rehabil. 1985; 66: 95-99.

Troup J D G, Martin J W, Lloyd D C E F. Back Pain in Industry. A Prospective Survey. Spine 1981; 6: 61-69.

Tunturi T, Kataja M, Keski-Nisula L, Lapinsuo M, Lepisto P, Paakala T, Patiala H, Rokkanen P. Posterior Fusion of the Lumosacral Spine. Acta Orthop. Scand. 1979; 50: 415-425.

Vernon-Roberts B, Pirie C J. Degenerative Changes in the Intervertebral Discs of the Lumbar Spine and their Sequelae. Rheumatol. Rehabil. 1977; 16: 13-21.

Von baeyer C L, Bergstrom K J, Brodwin M G, Brodwin S K. Invalid Use of Pain Drawings in Psychological Screening of Back Pain Patients. Pain 1983; 16: 103-107.

Waddell G. Clinical Assessment of Lumbar Impairment. Clin. Orthop. 1987; 221: 110-120.

Waddell G, Main C J. Assessment of Severity in Low-back Disorders. Spine 1984; 9: 204-208.

Waddell G, Main C J, Morris E W, Di Paola M, Gray I C M. Chronic Low-back Pain, Psychologic Distress, and Illness Behaviour. Spine 1984; 9: 209-213.

Waddell G, Main C J, Morris E W, Venner R M, Rae P S, Sharmy S H, Galloway H. Normality and Reliability in the Clinical Assessment of Backache. Br. Med. J. 1982; 284: 1519-1523.

Waddell G, McCulloch J A, Kummel E, Venner R M. Nonorganic Physical Signs in Low-back Pain. Spine 1980; 5: 117-125.

Waddell G, Morris E W, Di Paola M P, Bircher M, Finlayson D. A Concept of Illness Tested as an Improved Basis for Surgical Decisions in Low-back Disorders. Spine 1986; 11: 712-719.

Watson D. Neurotic Tendencies Among Chronic Pain Patients: an MMPI Item Analysis. Pain 1982; 14: 365-385.

White A W M. Low Back Pain in Men Receiving Workmen's Compensation: a Follow-up Study. Can. Med. Ass. J. 1969; 101: 61-67.

Zigmond A S, Snaith R P. The Hospital Anxiety and Depression Scale. Acta psychiatr. scand. 1983; 67: 361-370.

Zung W W K. A Self-rating Depression Scale. Arch. Gen. Psych. 1965; 12: 63-70.

Zung W W K. From Art to Science. the Diagnosis and Treatment of Depression. Arch. Gen. Psych. 1973; 29: 328-337.