J Rehabil Med 2013; 45: 177–185

ORIGINAL REPORT

SUB-GROUPING PATIENTS WITH NON-SPECIFIC LOW BACK PAIN BASED ON CLUSTER ANALYSIS OF DISCRIMINATORY CLINICAL ITEMS

Evdokia Billis, PhD[1,2], Christopher J. McCarthy, PhD[3], Chris Roberts, PhD[2], John Gliatis, MD[4], Maria Papandreou, PhD[5], George Gioftsos, PhD[6] and Jacqueline A. Oldham, PhD[2]

From the [1]Department of Physiotherapy, Technological Educational Institute (TEI) of Patras, Greece, [2]University of Manchester, Manchester, [3]Imperial College Healthcare NHS Trust, London, UK, [4]Orthopaedic Department, University Hospital of Patras, [5]Department of Physiotherapy, Technological Educational Institute (TEI) of Athens and [6]Department of Physiotherapy, Technological Educational Institute (TEI) of Lamia, Greece

Objective: **To identify potential subgroups amongst patients with non-specific low back pain based on a consensus list of potentially discriminatory examination items.**
Design: **Exploratory study.**
Participants: **A convenience sample of 106 patients with non-specific low back pain (43 males, 63 females, mean age 36 years, standard deviation 15.9 years) and 7 physiotherapists.**
Methods: **Based on 3 focus groups and a two-round Delphi involving 23 health professionals and a random stratified sample of 150 physiotherapists, respectively, a comprehensive examination list comprising the most "discriminatory" items was compiled. Following reliability analysis, the most reliable clinical items were assessed with a sample of patients with non-specific low back pain. *K*-means cluster analysis was conducted for 2-, 3- and 4-cluster options to explore for meaningful homogenous subgroups.**
Results: **The most clinically meaningful cluster was a two-subgroup option, comprising a small group ($n=24$) with more severe clinical presentation (i.e. more widespread pain, functional and sleeping problems, other symptoms, increased investigations undertaken, more severe clinical signs, etc.) and a larger less dysfunctional group ($n=80$).**
Conclusion: **A number of potentially discriminatory clinical items were identified by health professionals and sub-classified, based on a sample of patients with non-specific low back pain, into two subgroups. However, further work is needed to validate this classification process.**

Key words: non-specific low back pain; classification; sub-grouping; Greece; cluster analysis.

J Rehabil Med 2013; 45: 177–185

Guarantor's address: Evdokia Billis, Lecturer in Physiotherapy, Department of Physiotherapy, Branch Department of Aigion, Technological Educational Institute (TEI) of Patras, Psaron 6, Aigion, GR-25100, Greece. E-mail: ebillis@teipat.gr

Submitted May 11, 2011; accepted September 17, 2012

INTRODUCTION

The classification of patients with non-specific low back pain (NSLBP) into subgroups has a number of advantages. First, this approach is thought to be superior in guiding treatment compared with other approaches (1–4). Secondly, classifying patients based on their signs and symptoms has demonstrated diagnostic success in studies investigating specific low back pain (LBP) (5) and is a feasible, practical, cost-effective and clinically applicable process. Finally, clinical reasoning is facilitated by the utilization of diagnostic/prognostic classification systems, especially in cases where a specific cause of disease is indefinable (e.g. in NSLBP) (6).

However, a primary shortcoming of classification systems is their subjective nature. Most classification systems have been developed using a judgmental rather than a statistical approach (basing their development on small samples and relying on clinicians' personal experience and expertise); therefore this is likely to bias the process, and reduce the generalizability of the results (7, 8). Furthermore, little work has been undertaken to address issues of reliability and to ensure agreement on the selection of discriminatory criteria for the classifying criteria. As such, many classification systems lack validity and reliability, and do not lead to confidence in the derived subgroup profiling. Thus, although there is emerging evidence that classification systems are the best method for diagnosing and/or guiding treatment for NSLBP, basic steps towards their development require further elaboration.

In view of the high incidence of LBP in the general population and the lack of an established NSLBP taxonomy adopted by clinicians, it would seem appropriate to implement basic steps in subgroup development by trying to reach consensus on reliable clinical items that are believed to be discriminatory for identifying different NSLBP subsets. Consensus amongst experienced clinicians has important benefits. First, if a high level of consensus is reached, it is more likely that these agreed items will be useful and worth further consideration and study. It is also possible that these items will suggest a degree of discriminative validity in identifying LBP subsets. Secondly, the outcomes may facilitate cross-cultural comparisons, since cultural trends in the diagnosis of LBP are suggested across different settings (7, 9).

Thus, the aims of this study, which forms the second part of a two-part series of studies, were to explore consensus on potentially discriminatory clinical items in a NSLBP assessment list, and to identify whether homogenous sub-groups are

 doi: 10.2340/16501977-1100

 ISSN 1650-1977

being developed. The initial study (10) reported the inter-tester reliability of this discriminatory list, enabling the identification of reliable items, which then formed the examination procedure for investigating sub-group developments.

METHODS

Sample

A convenience sample of Greek patients with NSLBP was invited to participate in the study. Patients were diagnosed and referred for physiotherapy by specialist doctors (mostly orthopaedic specialists) working predominantly within the private sector. The evaluations took place in 3 different testing sites (physiotherapy clinics) within Greece (Athens, Patras and Lamia) and volunteers were recruited consecutively as they visited each clinic over a 10-month period. Patients were excluded if: (*i*) their native language was not Greek, (*ii*) their LBP was due to specific spinal pathology, (*iii*) they had undergone lumbar surgery, (*iv*) they were pregnant, or (*v*) they had severe neurological problems (that influenced their cognitive or motor performance). A total of 106 patients with LBP consented to participate. In addition, 7 physiotherapists (5 men, 2 women), clinically experienced in treating LBP, agreed to perform the assessments (2 based in Athens, 2 in Patras and 3 in Lamia). Ethical approval was obtained from the ethics committees of the Technological Educational Institute (TEI) of Lamia, Greece and the University of Manchester, UK.

Testing procedure

The first study accompanying this publication (10) reports the inter-tester reliability of a suggested discriminatory list, derived from the consensus of a large number of health professionals (11, 12). This reliability study enabled the identification of items that presented with a kappa value of 0.41 or more, corresponding to moderate and substantial/excellent agreement (13). These items were included in a final examination list. Items with kappa value of less than 0.41 were excluded from the examination, as they were considered to be of low reliability. In addition, items with almost perfect percentage agreement (for which kappas could not be computed) were not included; the reason for their exclusion was the non-existence of positive findings on these items, which indicated that they would not have any discriminative ability (14), and were thus discarded from the examination. Based on these criteria, the final examination list comprised 82 items (51 oral questions and 31 clinical tests). Thus, each physiotherapist (PT) conducted a full clinical examination (Appendix I), addressing the items that were deemed reliable, as reported in the accompanying study (10).

In the first section of the examination, which involved history-taking, the PT read each question from the examination booklet and recorded the answers. The second section involved a detailed physical examination. All PTs were familiarized with each testing procedure, following a training session directed by the principal investigator. Training was accompanied by a booklet summarizing the key examination points. Each PT assessed 15 patients, except for one PT who assessed 16. In addition, self-administered outcome measures for pain, disability and psychosocial status, which had undergone a cross-cultural validated adaptation procedure into the Greek language, and were highly recommended for use in LBP studies (15, 16) were also administered to the sample. For each patient the whole evaluation procedure lasted approximately 40 min (15 min for the questionnaires and 25 min for the examination).

Data analysis

To investigate whether there were distinct clinical subgroups within the sample, a cluster analysis approach was utilized. Cluster analysis is an exploratory procedure that aims to classify data into subgroups (17). In this study, a *K*-means non-hierarchical cluster analysis approach was utilized; this method is considered appropriate for producing exactly *K* different cluster solutions of maximum possible distinction. However, as this procedure cannot indicate how many clusters exist in the given data, in order to obtain an indication of the possible clustered options, hierarchical clustered modelling (14) was applied initially, and the dendrogram (formed by the hierarchical clustering) suggested further investigation of 2–6 cluster solutions. In addition, from the cluster analysis of 4 memberships, two of the groups contained two patients each, whose characteristics did not appear to be clinically different from the other two subgroup options. Thus, as this 4 membership option was not able to provide clinically meaningful results, a decision was made to stop the statistical cluster analysis at the 4 membership option. Thus, a *K*-means cluster analysis was carried out from 2- up to 4-group membership. *K*-means clustering generates an analysis of variance (ANOVA) table, whereby the mean on each item (for each cluster option) was explored to assess how "distinct" the within-cluster

Table I. *Characteristics of patients with non-specific low back pain (*n = *106)*

Characteristics	% (*n*)	Outcome measures	Mean (SD)
Male	40.6 (43)	VAS – Present pain intensity	3.31 (2.63)
Married/living with partner	47.2 (50)	VAS – Average pain intensity	4.11 (2.33)
Urban place of stay	73.6 (78)	VAS – Pain at best	1.25 (1.44)
Work		VAS – Pain at worst	7.52 (2.05)
Public sector	26.4 (28)		
Private sector	30.1 (32)	RMDQ	6.48 (5.17)
Sedentary (e.g. secretary)	25.5 (27)		
Active/manual occupation	25.4 (27)	ODI	21.06 (15.28)
Housewife/pensioner/student	49 (52)		
Health professionals seen		FABQ – Work	16.28 (11.29)
Specialist (orthopaedic, neurosurgeon etc.)	39.4 (42)	FABQ – Physical Activity	14.11 (6.11)
Physiotherapist	50.1 (54)		
Bed rest, days		HAD – Anxiety subscale	7.92 (4.35)
1–3	5.7 (6)	HAD – Depression subscale	4.37 (3.11)
3–7	7.5 (8)		
>7	18.7 (20)	PCS – Rumination	7.86 (4.67)
Sick leave	3.8 (4)	PCS – Magnification	4.01 (3.03)
Claiming compensation	0 (0)	PCS – Helplessness	7.48 (5.62)

SD: standard deviation; VAS: visual analogue scale (0–10), RMDQ: Roland-Morris Disability Questionnaire (0–24), ODI: Oswestry Disability Index (0–100), FABQ: Fear-Avoidance Beliefs Questionnaire (FABQ Work: 0–42, FABQ Physical Activity: 0–24), HAD: Hospital Anxiety and Depression Scale (HAD subscales: 0–21), PCS: Pain Catastrophizing Scale (PCS – Rumination: 0–16; PCS – Magnification: 0–12; PCS – Helplessness: 0–24).

differences were in relation to the *p*-value and the magnitude of the F values. Thus, both F and *p*-values were used for descriptive purposes. Statistically significant *p*-values set at 95% level ($p < 0.05$) were also used to distinguish between items. Cross-tabulations on each item within each cluster solution were also utilized as a means to determine whether observed differences in magnitude across the clusters were clinically "important" in size; and from the resulting χ^2 values the strength of association of each variable assigned to each group was tested. Finally, the distribution of each item across the subgroups was used. All data were analysed in SPSS (version 15.0).

RESULTS

A total of 106 patients (43 males, 63 females) participated in the study, mean age 36 years (standard deviation (SD) 15.9 years, age range 18–70 years). Fifty-eight patients (54.7%) had LBP for more than 12 weeks and, for most of them (87.7%), the pain was of a recurrent nature. The sample's profile and main examination findings are shown in Tables I and II, respectively. The examinations were conducted by 7 PTs with mean clinical experience with LBP of 11.8 years (range 7–19 years).

K-means cluster analysis generated an ANOVA table, whereby each item's mean explored how "distinct" the clusters were.

For each cluster option 104 patients were finally computed (2 were missed due to missing data).

Two-cluster option

The 2-subgroup option consisted of a small group of 24 patients with NSLBP and a larger group of 80 patients with NSLBP. A total of 29 items (21 from the patient's history and 8 from physical examination) yielded statistically significant values. The small group (Group 1) had greater "severity" in their presentation and outcome measures' scores compared with Group 2. More than half of this subgroup's patients had pins and needles (58.3%), neck pain (75%), restricted lumbar movements (62.5–66.7%) and had their lumbar spines radiographed (83.3%), as opposed to less than half from the larger group (23.7–42.5%) with these

Table II. *Characteristics of the sample's symptoms and clinical presentation (*n = *106)*

Clinical presentation characteristics	% (*n*)	Clinical examination findings	% (*n*)
Pain location		Posture	
Mainly in the back	83.0 (88)	Normal	29.2 (31)
Mainly in the leg	10.4 (11)	Lordotic	34.0 (36)
L-sided back pain (body chart)	71.7 (76)	Active movements (lumbar)	
R-sided back pain (body chart)	68.9 (73)	Pain in flexion	33.0 (35)
L buttock pain (body chart)	38.7 (41)	Pain in extension	42.5 (45)
R buttock pain (body chart)	34.9 (37)	Pain in right-side flexion	17.9 (19)
L foot (sole) pain (body chart)	7.5 (8)	Centralization in flexion	38.7 (41)
L foot pain (dorsum) (body chart)	2.8 (3)	Combined movements	
Type of pain		Restricted extension with R SF	38.7 (41)
Dull	38.7 (41)	Restricted extension with L SF	49.1 (52)
Deep	69.8 (74)	Pain in flexion with R SF	19.8 (21)
Sharp	45.3 (48)	Pain in extension with R SF	40.6 (43)
Diffuse	34 (36)	Pain in extension with L SF	47.2 (50)
Pain and activity		Restricted posterior pelvic tilt	37.7 (40)
Mainly at rest	60.4 (64)	Accessory (P–A) movements	
Mainly in motion	55.7 (59)	L1 pain reproduction	13.2 (14)
Relieving positions – Lying	62.3 66)	S1 pain reproduction	22.6 (24)
Aggravating position – Sitting	34.0 (36)	Palpation (trigger points, etc.)	
Aggravating position – Lying	15.1 (16)	Upper lumbar area	22.6 (24)
Chronicity of episode		Sacroiliac area	24.5 (26)
Chronic (over 12 weeks)	54.7 (58)	Neurological examination	
Recurrent episode	87.7 (93)	L2 dermatome – abnormality	4.7 (5)
LBP getting better	50.0 (53)	L3 dermatome – abnormality	4.7 (5)
Diurnal pattern		L4 dermatome – abnormality	12.3 (13)
Pain waking at night	18.9 (20)	SLR – Pain reproduction (positive test)	14.2 (15)
Pain worse in the morning	42.5 (45)	SLR – Positive neurodynamic	18.9 (20)
Other symptoms		Therapists' clinical impression	
Stiffness	44.3 (47)	Closing pattern	45.3 (48)
Pins and needles	32.1 (34)	Impairment dysfunction	34.0 (36)
Investigations		Good prognosis for recovery	91.5 (97)
Radiographs (X-ray)	52.8 (56)	Biomedical domain of influence	91.5 (97)
MRI	25.5 (27)	Psychological/social domain of influence	6.5 (7)
Medical history and other problems			
Neck ache	46.2 (49)		
Other musculoskeletal (deformity, leg length)	28.3 (30)		
Work and function			
Hobbies severely affected by LBP	40.6 (43)		
Daily activities severely affected by LBP	30.2 (32)		
Psychosocial – exaggerated pain behaviour	39.6 (42)		

LBP: low back pain; L: left; R: right; SF: side flexion; MRI: magnetic resonance imaging; P-A: posteroanterior (glide); SLR: straight leg raise.

symptoms. In addition, the smaller group comprised more patients with deep pain, predominant leg pain, aggravation in lying and night pain. Two pain location items (left foot and anterior leg pain) were present in larger percentages (20.8% and 12.5%) in the small group compared with the larger one (3.7% and 0%), indicating greater peripheralization of symptoms. A larger proportion of the small group presented with stiffness (66.7%) and dragging feet (25%) and had undertaken MRI diagnostic tests (40%) compared with the larger group (36.2%, 2.5% and 18.7%, respectively). In addition, straight leg raise (SLR) presented with pain and positive responses in 33.3% and 37.5%, respectively, in the small group, compared with the large one (8.7% and 13.8%, respectively). Table III summarizes the ANOVA outputs, χ^2 scores and clinical item distributions for the two membership option (due to the large volume of data, 82 items being included, only statistically significant items are presented).

Three-cluster option

The 3-cluster option comprised a large group of patients ($n = 73$) and 2 smaller ones ($n = 9$ and $n = 22$ patients, respectively). Twenty-two history and 8 physical examination items achieved statistical significance. From the distribution of outcomes and clinical features, one of the groups ($n = 22$) appeared more "severe" compared with the other two groups in terms of referred and widespread (i.e. neck) pain, other symptoms, recurrent episodes, investigations and restricted movement. However, the other groups did not show any consistent differences (indicative of a particular pattern) in their characteristics.

Table III. *Cluster analysis output for 2 subgroups (presenting the items with* $p < 0.05$*)*

	F-value	χ^2-test	Group 1 ($n = 24$)	Group 2 ($n = 80$)
Clinical items, n *(%)*				
Left foot (sole) pain (body chart)	8.028	0.006	5 (20.8)	3 (3.7)
Left anterior leg pain (body chart)	11.209	0.001	3 (12.5)	0 (0)
Left foot pain in dorsum (body chart)	11.209	0.001	3 (12.5)	0 (0)
Deep pain	7.783	0.007	22 (91.7)	50 (62.5)
Mainly in back (pain)	24.044	<0.001	13 (54.7)	74 (92.5)
Mainly in leg (pain)	24.580	<0.001	8 (33.3)	2 (2.5)
Relieving position – Lying	6.535	0.012	20 (83.3)	44 (55)
Aggravating position – Lying	9.709	0.003	8 (33.3)	7 (8.7)
Pain getting worse	4.669	0.033	7 (29.2)	9 (11.2)
Diurnal – Pain waking at night	12.605	0.001	10 (40)	9 (11.2)
Diurnal – Pain preventing from sleeping	25.198	<0.001	9 (37.5)	3 (3.7)
Other symptoms – Stiffness	7.314	0.008	16 (66.7)	29 (36.2)
Other symptoms – Pins and needles	11.083	0.001	14 (58.3)	19 (23.7)
Other symptoms – Dragging feet	14.780	<0.001	6 (25)	2 (2.5)
Investigations – X-ray performed	13.721	<0.001	20 (83.3)	34 (42.5)
Investigations – MRI performed	5.488	0.021	10 (40)	15 (18.7)
Investigations – Other investigations performed	21.103	<0.001	9 (37.5)	4 (5)
Type of work – Sedentary type	8.556	0.005	5 (20.8)	43 (53.7)
Type of work – Involving carrying weights	5.738	0.019	6 (25)	6 (7.5)
Daily physical activities severely affected	12.222	0.001	14 (58.3)	18 (22.5)
Musculoskeletal problems – Neck ache	11.390	0.001	18 (75)	30 (37.5)
Combined movements – Restricted extension with R SF	7.313	0.008	16 (66.7)	23 (28.7)
Combined movements – Restricted extension with L SF	3.952	0.049	15 (62.5)	26 (32.5)
Combined movements – Painful flexion with R SF	6.024	0.016	16 (66.7)	35 (43.7)
Ba Posterior pelvic tilt-restricted	12.464	0.001	9 (37.5)	12 (15)
SLR – Pain reproduction	9.709	0.003	8 (33.3)	7 (8.7)
SLR – Positive neurodynamic test	7.028	0.01	9 (37.5)	11 (13.8)
P–A glides – Pain in L1	4.567	0.035	6 (25)	7 (8.7)
Prognosis	5.050	0.027	20 (83.3)	77 (96.2)
Outcome measures, mean (SD)				
VAS – Present pain intensity			4.86 (2.74)	2.96 (2.45)
VAS – Average pain intensity			6.83 (1.87)	3.55 (2.19)
VAS – Pain at best			2.46 (1.8)	1.08 (1.4)
VAS – Pain at worst			8.59 (1)	7.15 (2.25)
Roland-Morris Disability Questionnaire			10.15 (7.3)	5.98 (4.9)
Oswestry Disability Index (ODI)			31.69 (18.9)	19.65 (14.13)
FABQ – Work			21.38 (8.39)	15.76 (12)
FABQ – Physical Activity			16.77 (5.61)	13.94 (6.06)
HADS – Anxiety			8.46 (4.84)	7.27 (4.03)
HADS – Depression			5.92 (3.63)	3.67 (2.62)
PCS (total)			26.54 (13.06)	19.73 (10.76)

MRI: magnetic resonance imaging; R: right; L: left; SF: side flexion; SLR: straight leg raise; P–A: posteroanterior (glide); SD: standard deviation; VAS: visual analogue scale; FABQ: Fear-Avoidance Beliefs Questionnaire; HADS: Hospital Anxiety Depression Scale; PCS: Pain Catastrophizing Scale.

Four-cluster option

The 4-subgroup option entailed a large group ($n = 79$), a smaller one ($n = 21$) and two 2-patient groups. Eighteen out of 36 items achieved statistical significance; however, in view of the limited number of patients in the two subgroups, consistent distribution patterns or distinctive characteristics across groups were not obtained.

DISCUSSION

This study aimed to explore a list of reliable and potentially discriminatory items for NSLBP, on their ability to distinguish amongst different patient subsets. The sample utilized was predominantly recruited by medical referrals from the private sector, which is well developed in Greece (18), and consisted of a mix of patients with acute and chronic LBP, who were moderately disabled by LBP. The sample's profile had comparable features to that of some previous classification studies (19, 20), and a marginally less severe clinical profile compared with some others (21).

Following cluster analysis within the 2-cluster option, the small group had a more severe clinical profile compared with the larger group. Based on the distribution of the clinical items on the 3- and 4-cluster options, it could be argued that they both included a small and a larger group that possessed similarities to the 2-subgroup option. However, in the 3-cluster option one of the smaller groups was not distinctively different to the other two, whereas for the 4-group option, two of the groups were extremely small and their distribution patterns did not subsequently indicate any clinically meaningful solutions. Thus, given the above information, the 2-subgroup option provided the most comprehensible and clinically meaningful presentation profiles (compared with the other 2-cluster options). Furthermore, despite outcome measures not being included in the cluster model, the scores for the small group on the outcomes of pain intensity, disability, anxiety, etc. were much higher, further verifying the presence of more severe cases in this group.

Although direct comparisons between this study's two subgroups and previous classification reports are difficult to perform with accuracy due to methodological differences, some similarities are evident. Pain location, aggravating factors, 24-h pain patterns, pain response to movement, and symptom duration are criteria included in several European (19, 22, 23), Canadian (24), New Zealand (25) and US (26) studies. Similarly, a wide range of physical features have been utilized, most common of which are SLR (22–24, 27) and lumbar mobility pain provocation tests (28, 29). These similarities give weight to the inclusion of these factors in such classification approaches.

From the studies limited to 2–3 subgroups, comparisons are still difficult due to variability in their designs and lack of reported detail on item distribution across the generated groups. Whilst Langworthy & Green (30) identified 3 criteria with similarity to this study (pain getting worse, night pain and exacerbation in lying), the subgroups' distribution on these items was not detailed. Similarly, despite the interesting results in the study by Hill et al. (21) comparing 2 validated questionnaire instruments in their ability to identify subgroups requiring early intervention, limited information was given on their subgroups' distribution characteristics. McCarthy et al. (19) were the only group that adopted a similar approach to this study and generated similar results. Their cluster analysis pointed towards two subgroups; a smaller more dysfunctional one (called the "hypervigilant" group), and a larger (less dysfunctional) one. Distribution of age, gender and outcome measures across their subgroups was comparable with this study. In addition, distribution on several clinical items (i.e. SLR test, upper lumbar palpation, pain provocation tests) pointed towards more positive (severe) responses for the hypervigilant group. The presence of a UK study (19) with similarities to this Greek one provides some confidence in our results, and their possible generalizability in wider population samples.

Some issues were of clinical importance. First, a large set of signs, symptoms and clinical measures have been utilized. Secondly, most previous classification studies lack the standards required to be deemed methodologically "rigorous" (i.e. lack of statistically developing clusters, utilization of empirical methods, lack of item reliability, etc.). Thirdly, apart from utilizing a statistical procedure for generating subgroups, which is considered more sophisticated than utilizing other methods (i.e. observational or judgmental approaches) (8), the present study selected items suggested to be "discriminatory" by a consensus procedure involving a large and representative sample of health professionals dealing with NSLBP within Greece (11). In addition, only the most reliable of these selected items were utilized (10) in order to improve confidence and objectivity in outcomes. In addition, this study aimed to provide a meaningful, practical and useful taxonomy within the Greek healthcare setting. In view of existing cultural variations (7), it has not been assumed that classification systems developed in a given cultural setting can be adopted and utilized in the Greek setting. Therefore, this exploratory study tried to develop clinical subgroups based on practical, clinically applicable, reliable and generalizable classifying criteria. However, whether this approach is clinically useful as an assessment-based process for targeting treatment or as a prognostic guiding path is currently untested and needs to be further investigated.

In terms of the methodology utilized, the *K*-means cluster approach was considered more appropriate than a hierarchical one because it is easy to use, reliable (no "multiplicity" effects, repeatable cluster generation, etc.), has the ability to produce distinct non-overlapping clusters (14, 17) and has been used in LBP exploratory studies (23, 30, 31). However, in view of its limitation in indicating *a priori* the exact number of existing clusters, a second clustering approach, the hierarchical method, a commonly used adjunct approach (17, 32), was utilized to verify the generated clusters and increase the validity of the findings (14). In addition, interpretation of the resulting partitions by descriptive means (as shown in Table III) was conducted, as recommended, for evaluating their clusters' clinical utility (14). However, whilst cluster analysis has merit as a data-driven analysis procedure (as indicated above), it has pointed towards a relatively simple model of analysis,

which can provide only basic discrimination of data. Thus, given the complexity of the NSLBP problem, as well as the fact that there is currently no optimal classification scheme, it appears that further testing is required in order to ensure reproducibility and enhance credibility of the study's subgroups. Perhaps utilizing more complex analyses (such as neural networks or data-mining) or even combining approaches, to allow subgroups' comparison and observation of similarities and differences between approaches would be desirable and beneficial. Interestingly, in Kent et al.'s review (33) on the research methods utilized for sub-grouping LBP, they proposed a method framework comprising 6 stages, including hypothesis setting and testing studies, validation studies, as well as impact analyses studies. They highlighted the need for sub-grouping research to proceed through all phases of study, in order for the developed subgroups to gain credibility, generalizability and applicability within clinical practice.

A limitation of the present study is that most discriminatory items utilized were biomedically based despite the profound role of psychosocial factors in NSLBP (34, 35). Psychosocial factors were excluded on the basis of their poor reliability. Nevertheless, certain social factors (work, hobbies, physical activities, etc.) and the psychosocial measures' scores (Hospital Anxiety and Depression Scale, Fear-Avoidance Beliefs Questionnaire, Pain Catastrophizing Scale, presented in Table I) differed across the two groups, indicating greater psychosocial overlay for the small group compared with the large one. Interestingly, only 4 studies incorporated biopsychosocial elements in their classification (36–39); yet, they were utilized in a very different and, consequently, non-comparable way. Another limitation concerns the sample utilized, which was limited in terms of representing the more "disabled" patients; the presenting sample consisted of LBP patients with, in general, low levels of disability. This could have precluded the potential development of another subgroup with more "dysfunctional" features. However, this exploratory study constitutes only the first step towards developing a classification system within a particular cultural context.

In conclusion, this cross-sectional exploratory study identified the existence of two distinct subgroups by utilizing a cluster analysis approach; a small group with more "severe" and widespread clinical signs and symptoms, and a large group with low severity, dysfunction and symptom presentation. This preliminary study forms the first step in developing a classification system within Greece based on discriminatory and reliable clinical criteria. However, despite its advantages, cluster analysis provides a simplistic method of subgroup research, and further work should thus explore each subgroup's clinical and diagnostic utility, in larger samples.

ACKNOWLEDGEMENTS

The authors would like to thank physiotherapists Ioannis Stathopoulos, Panagiotis Trigkas, Savvas Spanos, George Krekoukias and Nikolaos Strimpakos, as well as all the patients who participated in this study. This research was funded by the Manipulation Association of Chartered Physiotherapist's Elsevier Science and Doctoral Awards for Research in Manipulative Physiotherapy.

REFERENCES

1. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. Spine 2003; 28: 1363–1371.
2. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying subgroups of patients with acute/subacute "non-specific" low back pain: results of a randomized clinical trial. Spine 2006; 31: 623–631.
3. Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. Ann Intern Med 2004; 141: 920–928.
4. Kent P, Mjosund HL, Petersen DHD. Does targeting manual therapy and/or exercse improve patient outcomes in nonspecific low back pain? A systematic review. BMC Medicine 2010; 18: 22.
5. Vroomen PC, de Krom MC, Knottnerus JA. Diagnostic value of history and physical examination in patients suspected of sciatica due to disc herniation: a systematic review. J Neurol 1999; 246: 899–906.
6. Kent P, Keating JL. Classification in nonspecific low back pain: what methods do primary care clinicians currently use? Spine 2005; 30: 1433–1440.
7. Billis EV, McCarthy CJ, Oldham JA. Subclassification of low back pain: a cross-country comparison. Eur Spine J 2007; 16: 865–879.
8. Riddle DL. Classification and low back pain: a review of the literature and critical analysis of selected systems. Phys Ther 1998; 78: 708–737.
9. Aceves-Avila FJ, Ferrari R, Ramos-Remus C. New insights into culture driven disorders. Best Pract Res Clin Rheumatol 2004; 18: 155–171.
10. Billis E, McCarthy CJ, Gliatis J, Gittins M, Papandreou M, Oldham JA. Inter-tester reliability of discriminatory examination items for sub-classifying non-specific low back pain. J Rehabil Med 2012; 44: 851–857.
11. Billis E, McCarthy CJ, Gliatis J, Stathopoulos I, Papandreou M, Oldham JA. Which are the most important discriminatory items for subclassifying non-specific low back pain? A Delphi study among Greek health professionals. J Eval Clin Pract 2010; 16: 542–549.
12. Billis EV, McCarthy CJ, Stathopoulos I, Kapreli E, Pantzou P, Oldham JA. The clinical and cultural factors in classifying low back pain patients within Greece: a qualitative exploration of Greek health professionals. J Eval Clin Pract 2007; 13: 337–345.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174.
14. Milligan GW, Cooper MC. Methodology review: clustering methods. Appl Psychol Meas 1987; 11: 329–354.
15. Pincus T, Santos R, Breen A, Burton AK, Underwood M. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. Arthritis Rheum 2008; 59: 14–24.
16. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. Spine 2000; 25: 3100–3103.
17. Everitt BS, Landau S, Leese M. Cluster analysis. London: Arnold; 2001.
18. Siskou O, Kaitelidou D, Papakonstantinou V, Liaropoulos L. Private health expenditure in the Greek health care system: where truth ends and the myth begins. Health Policy 2008; 88: 282–293.
19. McCarthy CJ, Roberts C, Gittins M, Oldham JA. A process of subgroup identification in non-specific low back pain using a standard clinical examination and cluster analysis. Physiother Res Int 2012; 2: 92–100.
20. George SZ, Delitto A. Clinical examination variables discriminate among treatment-based classification groups: a study of construct validity in patients with acute low back pain. Phys Ther 2005; 85: 306–314.
21. Hill JC, Dunn KM, Main CJ, Hay EM. Subgrouping low back pain:

a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. Eur J Pain 2010; 14: 83–89.
22. Coste J, Paolaggi JB, Spira A. Classification of nonspecific low back pain. II. Clinical diversity of organic forms. Spine 1992; 17: 1038–1042.
23. Coste J, Paolaggi JB, Spira A. Classification of nonspecific low back pain. I. Psychological involvement in low back pain. A clinical, descriptive approach. Spine 1992; 17: 1028–1037.
24. Spitzer W, LeBlanc F, Depuis M. Scientific approach to the assessment and management of activity related spinal disorders: report of the Quebec Task Force on Spinal Disorders. Spine 1987; 12: S16–S21.
25. McKenzie R. The lumbar spine: mechanical diagnosis and therapy. Waikanae: Spinal Publications Ltd; 1981.
26. Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. J Orthop Sports Phys Ther 2007; 37: 290–302.
27. Bertilson BC, Bring J, Sjoblom A, Sundell K, Strender LE. Interexaminer reliability in the assessment of low back pain (LBP) using the Kirkaldy-Willis classification (KWC). Eur Spine J 2006; 15: 1695–1703.
28. McKenzie R. Subgroups in non specific spinal disorders. In: McKenzie R, editor. The lumbar spine: mechanical diagnosis and therapy. Waikanae: Spinal Publications Ltd; 1981, p. 34–37.
29. Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, McDonnell MK, Bloom NJ. Movement system impairment-based categories for low back pain: stage 1 validation. J Orthop Sports Phys Ther 2003; 33: 126–142.
30. Langworthy JM, Breen AC. Rationalizing back pain: the development of a classification system through cluster analysis. J Manipulative Physiol Ther 1997; 20: 303–310.
31. Bergstrom G, Bodin L, Jensen IB, Linton SJ, Nygren AL. Long-term, non-specific spinal pain: reliable and valid subgroups of patients. Behav Res Ther 2001; 39: 75–87.
32. Gironda RJ, Clark ME. Cluster analysis of the pain outcomes questionnaire. Pain Med 2008; 9: 813–823.
33. Kent P, Keating J, Leboeuf-Yde C. Research methods for subgrouping low back pain. BMC Med Res Methodol 2010; 10: 62.
34. Clays E, De BD, Leynen F, Kornitzer M, Kittel F, De BG. The impact of psychosocial factors on low back pain: longitudinal results from the Belstress study. Spine 2007; 32: 262–268.
35. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002; 27: E109–E120.
36. Huyse FJ, Lyons JS, Stiefel FC, Slaets JP, de Jonge P, Fink P, et al. "INTERMED": a method to assess health service needs. I. Development and reliability. Gen Hosp Psychiatry 1999; 21: 39–48.
37. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum 2008; 59: 632–641.
38. O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. Man Ther 2005; 10: 242–255.
39. Harper AC, Harper DA, Lambert LJ, Andrews HB, Lo SK, Ross FM, et al. Symptoms of impairment, disability and handicap in low back pain: a taxonomy. Pain 1992; 50: 189–195.

APPENDIX I. Examination list

A.	HISTORY		
1.	PRESENT SYMPTOMS		
•	BODY CHART. Please locate areas of pain, referred pain, etc. (A body chart diagram divided into 20 consecutively numbered body areas, for description of pain was provided)		
•	QUALITY OF PAIN. How would you describe your pain?		
	Dull ache	☐ YES	☐ NO
	Intense pain	☐ YES	☐ NO
	Superficial	☐ YES	☐ NO
	Deep ache	☐ YES	☐ NO
	Sharp	☐ YES	☐ NO
	Diffuse	☐ YES	☐ NO
	Localized	☐ YES	☐ NO
	Mainly in the leg	☐ YES	☐ NO
	Mainly in the back	☐ YES	☐ NO
	PAIN BEHAVIOUR & POSTURES/ACTIVITIES		
•	When do you get your pain?		
	At rest	☐ YES	☐ NO
	When moving/during movement	☐ YES	☐ NO
•	Describe your most relieving position/activity		
	Bending	☐ YES	☐ NO
	Sitting	☐ YES	☐ NO
	Standing	☐ YES	☐ NO
	Lying (describe position)	☐ YES	☐ NO
•	Describe your most aggravating position/activity		
	Bending	☐ YES	☐ NO
	Sitting	☐ YES	☐ NO
	Lying (describe position)	☐ YES	☐ NO
	Other:	……	
•	PAIN STATUS		
	Is the pain getting better?	☐ YES	☐ NO
	Is the pain staying the same?	☐ YES	☐ NO
	Is the pain getting worse?	☐ YES	☐ NO
•	24-HOUR PAIN BEHAVIOUR.		
	When do you mostly get your primary pain?		
	Waking them at night	☐ YES	☐ NO
	Preventing them from getting back or getting to sleep	☐ YES	☐ NO
	Worse in the morning	☐ YES	☐ NO
	Worse in the evening	☐ YES	☐ NO
•	CHARACTERISTICS OF OTHER SYMPTOMS		
	Any symptoms other than pain? (Mark areas in body chart)		
	Stiffness	☐ YES	☐ NO
	Pins and needles	☐ YES	☐ NO
	Dragging feet	☐ YES	☐ NO
2	HISTORY OF CONDITION/SYMPTOMS		
•	When did the symptoms start?	…… months	
•	INVESTIGATIONS		
	X-ray	☐ YES	☐ NO
	Blood tests	☐ YES	☐ NO
	MRI	☐ YES	☐ NO
	Other	☐ YES	☐ NO
•	Is this the first pain episode? If YES go to section 3	☐ YES	☐ NO
•	PREVIOUS EPISODES		
	Was previous episode(s) of the same type?	☐ YES	☐ NO
3	FUNCTION		
•	TYPE OF WORK. Describe the type of work		
	Primarily sedentary	☐ YES	☐ NO
	Primarily repetitive movements	☐ YES	☐ NO
	Primarily carrying weights	☐ YES	☐ NO
•	HOBBIES & DAILY ACTIVITIES		
	Does the back problem severely affect the patient's hobbies? Describe how:……	☐ YES	☐ NO
	Does the back problem severely affect the patient's level of daily physical activity? Describe how:……	☐ YES	☐ NO
4	MEDICAL HISTORY		

•	DRUG HISTORY					
	Does any drug affect his back problem?	☐ YES	☐ NO			
	Doses of NSAIDs					
	If patient takes NSAIDs is the dose high?	☐ YES	☐ NO			
	Dose:……					
•	OTHER MUSCULOSKELETAL PROBLEMS					
	Neck pain	☐ YES	☐ NO			
	Leg length inequality	☐ YES	☐ NO			
•	PREVIOUS SURGERY	☐ YES	☐ NO			
	Describe:……					
•	GYNAECOLOGICAL HISTORY					
	Does the patient have any menstrual or hormonal problems linked to the LBP?	☐ YES	☐ NO			
•	POST-NATAL BACKACHE					
	Is the current linked with post-natal backache?	☐ YES	☐ NO			
5	PSYCHOSOCIAL HISTORY					
	Is the patient's behaviour affected by the following:	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
•	PAIN BEHAVIOUR OF THE PATIENT (i.e. fear of pain, expectation of pain to increase with activity/work, belief that pain is uncontrollable/harmful/hypervigilance, etc.)	☐	☐	☐	☐	☐
•	PSYCHOLOGICAL & EMOTIONAL STATUS					
	Impact of patient's problem towards psychological status					
	Does patient believe he has a pathological condition?	☐	☐	☐	☐	☐
	Is patient clear of what things make him better/ worse?	☐	☐	☐	☐	☐

B. CLINICAL EXAMINATION

Standing

6 OBSERVATION OF POSTURE. What posture best describes the patient? Normal ☐ YES ☐ NO Lordotic ☐ YES ☐ NO

7 ACTIVE MOVEMENTS

	HYPERMOBILE	NORMAL	RESTRICTED	PAIN REPRODUCTION
• LUMBAR RANGE				
Flexion	☐	☐	☐	☐ YES ☐ NO
Extension	☐	☐	☐	☐ YES ☐ NO
Right-side flexion	☐	☐	☐	☐ YES ☐ NO

	Peripheralizing/ increasing symptoms	No change	Centralizing/reducing or alleviating symptoms
• REPEATED MOVEMENTS			
Lumbar flexion	☐	☐	☐

	RESTRICTION		PAIN REPRODUCTION	
• COMBINED MOVEMENTS				
Flexion with right-side flexion	☐ YES	☐ NO	☐ YES	☐ NO
Extension with right-side flexion	☐ YES	☐ NO	☐ YES	☐ NO
Extension with left-side flexion	☐ YES	☐ NO	☐ YES	☐ NO

• Based on the above movements the patient presents with

Closing pattern ☐ YES ☐ NO Impairment dysfunction ☐ YES ☐ NO

Supine

7 ACTIVE MOVEMENTS (continued)

	Hypermobile	Normal	Restricted	Reproducing pain
• pelvic ROM				
Posterior pelvic tilt	☐	☐	☐	☐ YES ☐ NO

8 NEUROLOGICAL EXAMINATION. Response of symptomatic side Symptomatic side: LEFT ☐ RIGHT ☐

	ABSENT	REDUCED	NORMAL	HYPERSENSITIVE
• SENSATION				
L2	☐	☐	☐	☐
L3	☐	☐	☐	☐
L4	☐	☐	☐	☐

	VERY LIMITED	LIMITED	NORMAL	Reproducing pain	Positive response
• NEURODYNAMIC TESTS					
SLR	☐	☐	☐	☐ YES ☐ NO	☐ YES ☐ NO

Prone

9 PASSIVE JOINT ASSESSMENT & PALPATION

	Hypermobile	Normal	Restricted	Increases symptoms / reproducing pain
• ACCESSORY MOVEMENTS (P-A glides)				
L1	☐	☐	☐	☐ YES ☐ NO
S1	☐	☐	☐	☐ YES ☐ NO

• PALPATION (checking for tenderness, trigger points, etc.) Upper lumbar paraspinal areas ☐ YES ☐ NO Sacro-iliac area ☐ YES ☐ NO

10 CLINICAL REASONING

• Domain with strongest influence of patient's symptoms? ☐ Biomedical ☐ Psychological ☐ Social

• Prognosis? ☐ Good ☐ Poor

P–A: posteroanterior (glide); ROM: range of motion; LBP: low back pain; MRI: magnetic resonance imaging; NSAIDs: non-steroidal anti-inflammatory drugs.