University of Southampton

CLASSIFICATION OF MUSCULOSKELETAL DISORDERS OF THE NECK AND UPPER LIMB: A POPULATION STUDY

Isabel Claire Reading

Submitted for PhD

Medical Research Council

Environmental Epidemiology Unit

July 2003

UNIVERSITY OF SOUTHAMPTON
ABSTRACT
FACULTY OF MEDICINE
MRC ENVIRONMENTAL EPIDEMIOLOGY UNIT
Doctor of Philosophy

CLASSIFICATION OF MUSCULOSKELETAL DISORDERS OF THE NECK AND UPPER LIMB: A POPULATION STUDY

by Isabel Claire Reading

Soft-tissue musculoskeletal disorders of the neck and upper limb are common. Epidemiological research in this field has been impeded by the lack of an agreed classification system and diagnostic criteria. In 1996 the UK Health and Safety Executive (HSE) proposed diagnostic criteria for eight of these conditions, and a physical examination schedule designed to detect these disorders according to the criteria was subsequently devised. The objectives of this thesis were: 1) to examine the way that symptoms and signs coexist within individuals using cluster analysis, and to compare the observed patterns of clustering with the diagnostic categories proposed by the HSE; and 2) to investigate associated disability, healthcare use, and putative risk factors in the data-driven and the medically based classifications.

A population-based study has been conducted in Southampton, surveying working-aged adults regarding neck and upper limb symptoms. Of 6038 respondents, all symptomatic subjects (N=3152) and a sample of the asymptomatic subjects were invited to attend the physical examination, and 2145 examinations were performed. The findings were analysed: 1) using the medically based HSE criteria, and 2) using cluster analysis to group subjects according to their symptom-sign profiles.

At the neck four symptom-sign clusters were identified: no signs or symptoms; pain only; limited range of movement only; and pain plus a limited range of movement. The data-driven and medically based classification systems were in broad agreement and displayed similar associations with reported disability, healthcare use, mechanical activities and psychological factors.

At the shoulder four profiles were identified; these were characterised by increasing severity of disease involvement and did not distinguish between the five medically based diagnoses. However, comparison of the two classification systems confirmed that there were important differences between the categories identified in each.

Seven robust symptom-sign profiles were identified at the elbow, including two that tallied well with medial and lateral epicondylitis. Associated disability, healthcare use and exposures to mechanical occupational activity were prevalent in these clusters.

At the wrist/hand a total of fourteen symptom-sign profiles were yielded which were differentiated by location and the nature of symptoms and signs. Several of these corresponded well with clinical diagnoses of osteoarthritis and tenosynovitis. A variety of different profiles of sensorineural disturbance was seen, although none of these was clearly consistent with classical carpal tunnel syndrome.

Cluster analysis of symptom-sign profiles provides a unique approach to the classification and diagnosis of musculoskeletal disorders of the neck and upper limb. The information provided by the physical examination in addition to symptom report alone allows important distinctions in disease profile to be made.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This thesis represents one of a number of research avenues pursued as the result of a highly successful community study of musculoskeletal disorders of the neck and upper limbs performed by the Medical Research Council Environmental Epidemiology Unit. I am indebted to members of the capable research team who co-ordinated the study and carried out the impressive number of physical examinations I have analysed. They are: Cathy Linaker, who set up the initial stages of the study and piloted the physical examination schedule; Trish Byng, who undertook the lion's share of the examinations, supported by Claire Ryall, Angela Shipp and Karen Collins; and Vanessa Cox and Ken Cox, who transformed the raw data into workable STATA files. Keith Palmer was a grant holder on the project, and devised the examination schedule alongside Karen Walker-Bone. I am particularly grateful to Karen for the generous amounts of time, support, expertise and friendship that I have enjoyed from her over the last four years.

Cyrus Cooper has been an inspiring supervisor with the right amount of encouragement, cajoling and vision at each meeting. Clive Osmond carefully considered the statistical aspects of my work and provided advice and direction when needed. David Coggon shared his exceptional combination of mathematical and medical insight.

I have received financial support from the Colt Foundation, and have been impressed by their sustained interest in my work.

Holly Syddall continued to muster up interest in my thesis progress as well as providing welcome distractions, as all worthy office-mates should. Kirst Olufsen took on the unenviable task of proofreading my work, and I'm truly grateful for her eagle eyes.

The colour images in Section 1.1.2 were reproduced with permission from http://eduserve.hscer.washington.edu/hubio553/atlas/content.html, last accessed on 6th June 2003.

Finally, heartfelt thanks go out to Graeme and Emily.

Author's Contribution

The author assisted in the design of the study; in securing financial support; in construction of the examination proforma; and in evaluation of the reproducibility of nurse examinations. The author carried out all the data checking, and she agreed any data amendments with either the study research assistant (CHL) or the study rheumatologist (KWB). Five research nurses performed the fieldwork (PB, KC, CHL, CR, AS) and the study was co-ordinated by the research assistant (CHL) and the study rheumatologist (KWB). The author performed all the data analysis presented in this thesis.

Abbreviations

AC joint	acromioclavicular joint
AF	antecubital fossa
BMI	body mass index
BT	bicipital tenoditis
CI	confidence interval
CTD	cumulative trauma disorder
CTS	carpal tunnel syndrome
DIP	distal interphalangeal joint
DIP OA	distal interphalangeal joint osteoarthritis
dQ	de Quervain's tenosynovitis
GP	general practitioner
HSE	UK health and safety executive
ICD-9CM	international classification of disease, 9th revision, clinical modification
κ	Cohen's kappa statistic
LE	lateral epicondylitis
MCP	metacarpal
ME	medial epicondylitis
MRI	magnetic resonance imaging
NSAID	non-steroidal anti-inflammatory drug
NT	numbness and/or tingling
OA	osteoarthritis
OR	odds ratio
PCA	principal components analysis
RA	rheumatoid arthritis
ROM	range of movement
RSI	repetitive strain injury
RT	rotator cuff tendonitis
SAB	subacromial bursitis
SC	shoulder capsulitis
SEP	Southampton examination proforma
SF-36	36-item short-form health survey
SOC90	standard occupational classification 1990

CHAPTER 1: INTRODUCTION

1 Introduction

1.1 Musculoskeletal pain in the upper limb and neck

1.1.1 Overview

Upper limb and neck pain is common in the UK. Population studies suggest a lifetime prevalence of self-reported pain at levels of at least 25% at the neck, 6.7% at the shoulder and upper arm and 14.1% at the elbow and lower arm [1]. Table 1 highlights the results of a selection of prevalence studies exploring self-reported pain in different sites of the upper limb or neck.

Table 1: Prevalence of pain in the neck and upper limb

Study	Country	Gender	Study size	Age range	Point prevalence (%)	Lifetime prevalence (%)
Neck pain:						
Lawrence [2] 1969	UK	M	1803	15 – 75	9.0	27.8
		F	1572		12.0	33.6
Cunningham [3] 1984	US	M,F	6913	25 – 74	10.0	-
Brattberg [4] 1989	Sweden	M,F	1009	18 – 84	-	26.0
Makela [5] 1993	Finland	M,F	8000	>30	10.0	71.0
Cote [6] 1998	Canada	M,F	2184	20 – 69	22.2	66.7
Shoulder pain:						
Lawrence [2] 1969	UK	M,F	3375	15 – 75	-	16.0
Allander [7] 1974	Sweden	M,F	15268	40 – 74	20.0	-
Cunningham [3] 1984	US	M,F	6913	25 – 74	-	6.7
Chard [8] 1987	UK	M	318	>70	17.0	-
		F	326		25.0	-
Makela [5] 1993	Finland	M,F	8000	>30	2.0	-
Pope [9] 1997	UK	M,F	312	18 – 75	20.0	-
Elbow and lower arm:						
Cunningham [3] 1984	US	M,F	6913	25 – 74	-	14.1
Forearm pain:						
Macfarlane [10] 2000	UK	M,F	1953	18 – 65	8.3	-

Studies further indicate that 9% of adults in the UK will consult a general practitioner (GP) at least once during the course of a year with upper limb, lower limb or neck pain [11], although it appears that a high proportion of subjects with musculoskeletal pain do not consult their GP, even for persistent pain [12,13].

Primary and secondary care consultation rates give a useful indication of the impact of musculoskeletal disorders. In primary care, those patients who are likely to consult represent a group whose pain poses particular difficulties, such as impaired work or home activities, or exacerbation of symptoms by daily activities. Secondary referral is generally reserved for patients whose disorders remain persistent, unresponsive to initial treatment and are sufficiently distressing for the patient to return to their GP. Data on referral to secondary care are difficult to come by, and this may be due to the wide range of services available: physiotherapy, sports injuries clinics, rheumatology, orthopaedic and A&E specialist clinics. The care a patient receives is dependent both on local availability and the methods favoured by the GP if secondary care is deemed appropriate.

While the disorders that give rise to upper limb and neck pain are rarely fatal, they are associated with considerable disability: some 25% of subjects with neck pain will continue to have moderate symptoms 10 years later, whilst 7% will become severely disabled [14]. Approximately half of subjects reporting shoulder disorders to their GP still report complaints one year later [15]. Data on the outcome of pain located at the elbow, forearm, wrist and hand are sparse, but again, there is an indication that in a significant minority of cases, symptoms are persistent [16].

In terms of economic impact, neck and shoulder pain form the second most frequent musculoskeletal cause of sickness absenteeism from work after low back pain, and in some industries absenteeism from work due to neck pain alone is comparable to that of back pain [17].

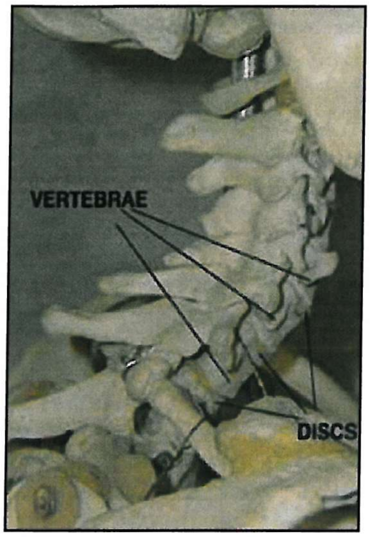
1.1.2 Disorders leading to upper limb and neck pain

Pain in the upper limb and neck can arise from a variety of underlying pathologies. This thesis will consider only soft tissue disorders (i.e. those affecting muscles, capsules, ligaments, tendons, menisci, disks and cartilaginous surfaces), and will describe conditions according to their location, giving indications of differential diagnosis, treatment, outcome and relative frequency.

The neck:

The neck is a highly complex set of joints based around seven cervical vertebrae through which the spinal cord and vertebral arteries pass. It is continually employed in the vital activities of head stability and movement essential to breathe, eat, talk and balance.

Neck disorders are most commonly of a mechanical nature and affect the joints, ligaments and muscles. Where symptoms are accompanied by radiological changes (such as the development of osteophytes and loss of disc height), they are attributed to **cervical spondylosis**, although the clinical presentation is indistinguishable from mechanical neck pain without radiologic change. Conversely, severe spondylosis is not necessarily accompanied by pain.



Mechanical neck disorders are frequently caused by trauma such as whiplash or sports injuries, but may result from natural wear and tear, degeneration or prolonged poor posture.

Strategies available for the management of mechanical neck pain include analgesics, non-steroidal anti-inflammatory drugs (NSAIDs), heat treatment, manipulation, traction, and perhaps most importantly rest. Surgical intervention is rarely used in mechanical neck pain.

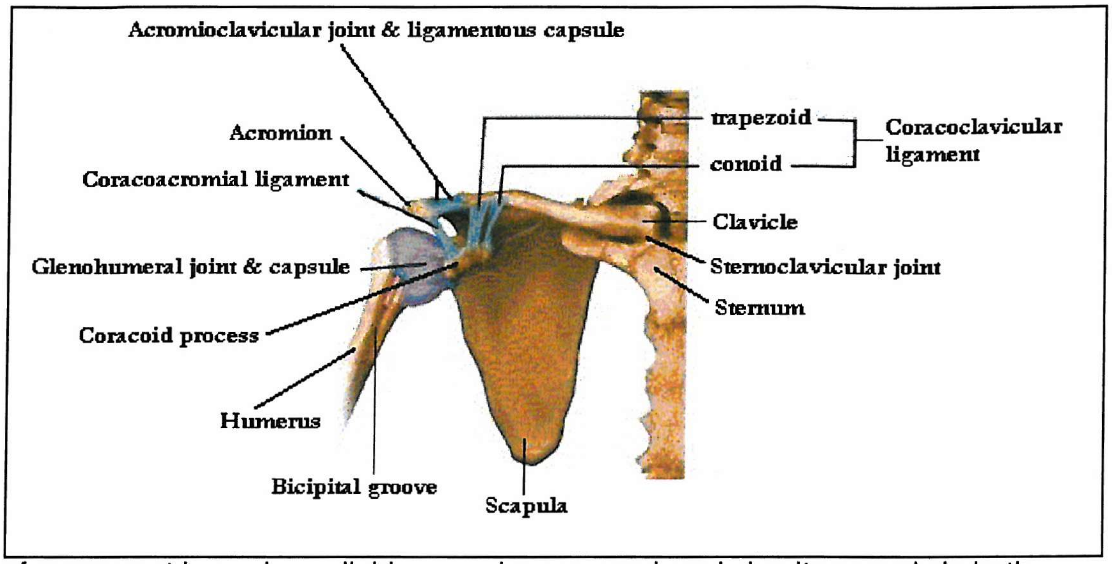
Other causes of neck pain include rheumatoid arthritis, inflammatory disease (in particular ankylosing spondylitis), infection, tumours and thoracic outlet syndrome (discussed later).

Epidemiology

Neck pain is recognised as a common condition, having a lifetime prevalence estimated at around 30% [2,4], although some studies suggest much higher prevalences around 70% [5,6] in adults. The prevalence of neck pain occurring in the previous year has been reported at 30% and of current pain at 10% [3]. Whiplash injuries occur with an incidence of 1 case per 1000 per year [18], and at 12 months, 20% of patients are still symptomatic.

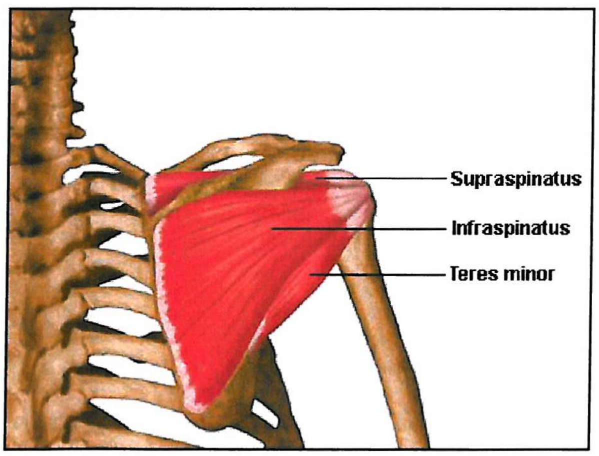
The shoulder:

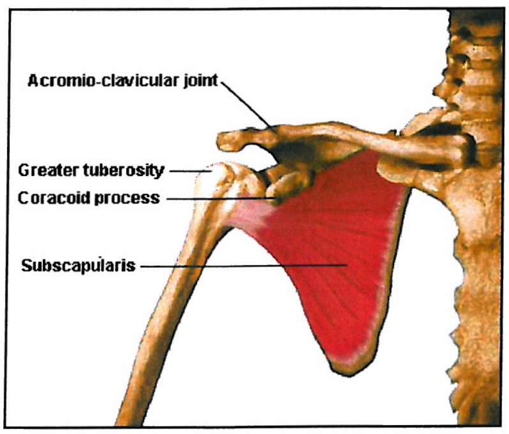
The shoulder is formed by a collection of joints: the glenohumeral, acromioclavicular, sternoclavicular and scapulothoracic, of which the glenohumeral joint causes the greatest amount of pain and disability. This is an incongruous joint (it has a moving axis of rotation and variable joint space with movement) so that whilst a wide range



of movement is made available, muscles are employed simultaneously in both stabilisation and movement. Thus, the normal working order of a variety of muscles and tendons, as well as all four joints, is essential to a single shoulder movement. Differentiation between underlying pathologies at the shoulder presents a challenge since a number of forms of damage lead to a similar outcome in terms of pain and restriction of movement. Special investigations for particular shoulder conditions often lack the sensitivity or specificity required for conclusive diagnosis. Although, therefore, the following descriptions of shoulder disorders appear to imply distinct anatomical and pathological entities, clinical practice might suggest that such exact discrimination is unjustified [19].

Rotator cuff tendonitis (RT). Four muscles make up the rotator cuff: supraspinatus, subscapularis, infraspinatus and teres minor, which are predominantly concerned with maintaining glenohumeral joint stability, along with the biceps tendon. RT





(inflammation of any of the tendons attaching muscles of the rotator cuff to the humerus) therefore produces pain in the deltoid region of the shoulder, and impairs active shoulder movement by allowing glenohumeral joint instability. Normal passive shoulder movement is obtainable because shoulder stability is retained by external means. Likewise a painful arc is characteristic of this condition. (Here shoulder abduction is painful within the middle of the range of movement, but not at the start or end of the movement. This is due to the exertion of the rotator cuff made in order to maintain stability; once the glenohumeral joint has moved to its correct position during shoulder elevation, the pain ceases).

RT can present acutely or with gradual onset: the former is usually seen in younger patients following a trauma, the latter in older patients experiencing chronic tendon changes (which may be due to impingement of the rotator cuff muscles during shoulder movement). Night pain, pain with movement and weakness may all occur in cases of the latter type.

Severe or chronic RT can lead to secondary capsulitis (discussed later), and bicipital tendon or acromioclavicular joint involvement may also present concurrently. Inflammation of the subacromial bursa is closely related to RT (although it will be discussed separately) and is included in RT pathology in clinical textbooks [20].

Management of RT is conservative at the onset, incorporating rest, activity modification and the use of NSAIDs. A local corticosteroid injection may be administered as well, and a strengthening program is used subsequently to restore shoulder function. Surgical intervention is only considered if the shoulder remains unresponsive to treatment one year on. The epidemiology of RT is poorly characterised at present.

Bicipital tendonitis (BT). The biceps tendon, like the rotator cuff, is concerned with stabilising the glenohumeral joint. Thus BT is often secondary to RT or glenohumeral joint instability. The biceps tendon experiences increased stress in attempting to compensate for the primary condition, and tendonitis results. A history of repetitive use or overuse of the tendon such as in carrying is often present. As with RT, young patients often present acutely whilst older ones tend to experience chronic involvement. Pain is usually over the anterior shoulder and may radiate into the biceps muscle. Both passive and active shoulder movements may also elicit pain.

Management of BT requires an assessment of whether the condition is primary or secondary to another underlying pathology. Secondary cases are expected to settle once the primary condition has been addressed. Primary BT is managed with rest and anti-inflammatory drugs. Surgery may be considered in chronic resistant cases. Again a strengthening program of both the biceps tendon and the rotator cuff is employed.

Subacromial bursitis (SAB). The subacromial bursa is situated between the muscles of the rotator cuff and the overlying acromion bone and is attached to both. Its function (along with bursae in general) is to reduce friction between muscle and bone during joint movement. Bursitis (inflammation of the bursa) is often caused by wear and tear or direct trauma, and is also seen as a reaction to RT. (Thus it may be described as part of RT pathology.)

Management takes the form of anti-inflammatory medication and rest.

Adhesive capsulitis. (Used synonymously with shoulder capsulitis and 'frozen shoulder'). The shoulder capsule is a flexible fibrous case, enclosing the glenohumeral joint. Adhesive capsulitis (inflammation of the capsule) is a poorly defined condition of unknown aetiology, characterised by painful global restriction of passive and active glenohumeral movement in all planes. This disorder is thought to have a prevalence of 2 – 3% in the general population (in diabetics this increases to 10 – 20%) although variation in diagnostic criteria makes these estimates difficult to interpret. Subjects presenting with capsulitis are usually aged 40+; a history of preceding minor shoulder injury or strain is common, but may merely reflect the first onset of the condition.

The natural history of shoulder capsulitis is well documented and involves three phases: a painful phase, progressive stiffness with continued pain, and a pain resolution phase leaving profound stiffness. This final stage appears to be self-limiting and recovery is spontaneous and gradual, but may be incomplete.

Management of capsulitis in its painful phase is concentrated on pain reduction and minimisation of joint restriction. The course of, and final recovery from, this condition appears to remain unaffected by treatment, although improvement in range of movement in the final phase of capsulitis may be accelerated following manipulation.

Acromioclavicular (AC) joint dysfunction. Pain located at the AC joint can present both acutely and chronically, the former occurring more frequently. Acute presentation is often due to a direct fall on the point of the shoulder, and the pain remains localised at that point. Range of abduction is often restricted in both active and passive movement.

Management is usually with analgesics and rest for several days or weeks. More serious injury such as complete rupture of the coracoclavicular ligaments may require surgery.

Other shoulder disorders. These include calcific tendonitis, glenohumeral instability and hand-shoulder syndrome, and will not be discussed further in any detail.

Shoulder pain associated with general medical conditions such as osteoarthritis, stroke, multiple sclerosis, inflammatory arthritis and diabetes mellitus accounts for a minor proportion of all shoulder pain [8].

Epidemiology

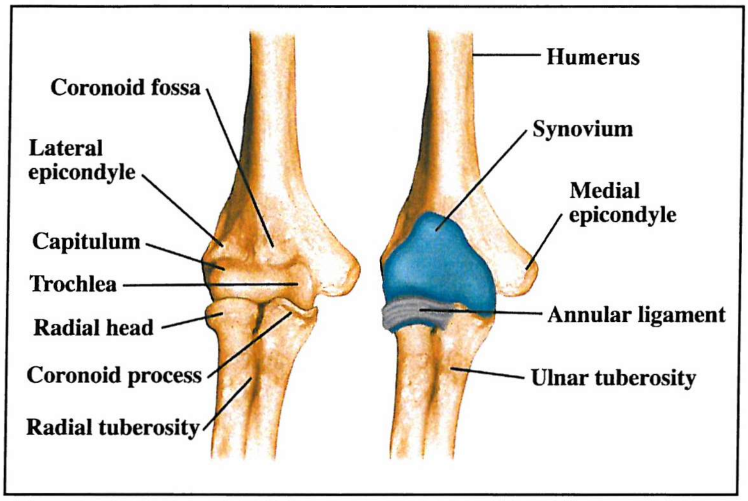
Epidemiological studies of shoulder pain have shown a wide disparity in prevalence rates, particularly in population surveys. This is largely due to the ambiguity of where the shoulder region starts and ends, and variation in the exact wording of the enquiry into pain presence. Pope *et al* [9] demonstrated this when they surveyed a population registered with a Stockport general practice, using four approaches to ask about shoulder pain in the previous month. Prevalences ranged from 31% – 48% across the four definitions, with a direct question 'During the last month, have you experienced pain in your shoulder(s) lasting more than 24 hours?' unaccompanied by a body diagram yielding the lowest prevalence.

In other studies, the prevalence of current shoulder pain in adults has been reported at around 20% [7,9] and at 2% [5], with a one-month prevalence at levels of 6.7% through to 48% [3,9,21-23]. Lawrence (1969) has reported a lifetime prevalence of shoulder pain of 16% [2].

Neck, shoulder and arm pathology overlap frequently, as demonstrated in a UK national survey of occupational exposure to hand-transmitted vibration [24]. Of those reporting shoulder pain 54% (n=1283) also reported neck pain and 27% also reported both neck and arm pain.

The elbow:

The elbow joint is a hinge joint that also allows a degree of rotation. The distal end of the humerus hinges to both the radius and ulna bones of the forearm, which are also linked to each other. At the elbow joint the humerus has a number of boney projections, two of which are the lateral and medial epicondyles. Ligaments for elbow joint stability and muscles for elbow movement are attached to these projections, and from these two main soft tissue disorders arise. The other common elbow condition occurs at the vulnerable elbow point, the olecranon. Referred pain, particularly from the neck or shoulders also occurs.



Lateral epicondylitis. (Tennis elbow). Inflammation of the common extensor tendons at the lateral epicondyle occurs in 1 – 3% of the population [20], usually between the ages of 40 and 60 years. The cause is unclear, and onset is generally gradual and spontaneous. Pain at the lateral epicondyle may spread up and down the arm, but often remains localised. Pain is elicited with resisted wrist extension and grip may be impaired.

Management involves anti-inflammatory treatment, rest and possibly splinting in mild cases, although the efficacy of these regimens is disputed [20]; a local corticosteroid injection is more effective in established lateral epicondylitis (around 90% of subjects showing improvement). Surgical intervention may be considered in resistant cases. Lateral epicondylitis has exhibited varying outcomes when under study, with relapse rates at 6 months ranging between 18% – 50% and reports of continued minor pain for up to 5 years in some subjects. Generally, however, lateral epicondylitis is considered to be a self-limiting condition, which improves within a year regardless of treatment strategy.

Medial Epicondylitis. (Golfer's elbow). This is analogous to lateral epicondylitis, but involves the common flexor tendons at the medial epicondyle. Medial epicondylitis is around 15 times less common than lateral epicondylitis, and is often milder and more localised. Pain and tenderness at the medial epicondyle with pain elicited with resisted wrist flexion are the distinguishing features.

Management and prognosis are the same as those for lateral epicondylitis.

Olecranon bursitis. Bursitis located at the elbow has been documented at various sites, but is primarily seen at the olecranon, where the bursa lies between it and the skin. Due to its superficial position, injury to the olecranon bursa is often due to external trauma, and swelling occurs readily and visibly. Pain is elicited on pressure to the olecranon (such as when leaning on the elbow).

Management is usually with a local steroid injection, although aspiration of the bursa may be required first to reduce the swelling and rule out infection as the primary cause. Olecranon bursitis may also be associated with rheumatoid arthritis.

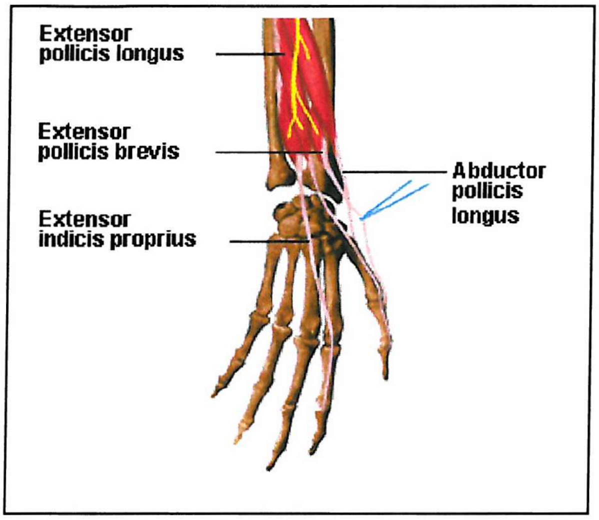
Epidemiology

Few population-based studies have investigated the prevalence of elbow pain. However, a recent study found a one-month prevalence of elbow pain of 6% in a UK population aged over 16 years [21]. One-month prevalence in a Finnish survey doubled to 14% in 50 – 64 year olds from a prevalence of 7% in 40 – 49 year olds. Lifetime prevalence of elbow pain was estimated as 4% (self-report) and 1% (physician identified) in a US population survey [3].

Lateral epicondylitis has been reported with a point prevalence of 2.5% in a Swedish population study. Subjects reporting pain were examined by a physician and were diagnosed with lateral epicondylitis on the basis of pain lasting for at least one month exacerbated by carrying, together with distinct tenderness over the lateral epicondyle and pain on resisted pronation.

The wrist and hand:

The wrist and hand comprise 30 bones and numerous tendons, and are served by three major nerves: the median, radial and ulnar. Soft tissue wrist and hand conditions related to nerve compression will be discussed separately.



Tenosynovitis. Tendons at the wrist are enclosed in slippery smooth membranes, synovia, which can become inflamed (tenosynovitis) and painful. This may be seen in association with rheumatoid arthritis or direct trauma. **De Quervain's syndrome** is a particular tenosynovitis affecting the abductor pollicis longus and extensor pollicis brevis tendons (running through the wrist and thumb). Pain in the radial aspect of the wrist and thumb is evident during pinching, gripping and other thumb/wrist movements. Finkelstein's test may be positive (here the subject makes a fist with the thumb in the palm enclosed by the fingers, and passive ulnar deviation elicits a pain response). A history of repetitive activity involving pinching along with wrist movement may be present, and the condition is most frequently seen in women between the ages of 30 and 50 years.

Management includes NSAIDS, splinting of the wrist and thumb, and activity modification. Whilst this scheme is usually effective, persistent cases may require local corticosteroid injection(s) or surgical intervention.

Trigger finger or thumb (Stenosing digital tenosynovitis) is another well-defined form of tenosynovitis in which a flexor tendon sheath of a finger or thumb becomes inflamed and thickens. A nodule forms on the tendon within the thickened synovium, and this causes an obstruction as the tendon passes through a sheath during movement. The digit is prevented from completing the movement and becomes locked. Since a sharp pull will overcome the obstruction, a triggering digit motion is seen. Again, a history of repetitive gripping motion is often present, and management is the same as for de Quervain's tenosynovitis.

Dupuytren's contracture. Flexion of one or more fingers due to nodular thickening of the palmar fascia (the fibrous lining of the palm) is a relatively common condition seen particularly in men and with older age. With no progression, this disorder is painless and causes no impaired function. However, severe deformity, pain and impaired hand function can occur. Surgical intervention may be necessary in this extreme case. The aetiology of this condition is poorly understood.

Other soft tissue wrist/hand disorders. Vascular impairment in the fingers such as Raynaud's phenomenon, and hyperextension of finger joints such as seen in swan-neck, Boutonniere and Z deformities are further causes of hand pain. These will not be discussed in any detail.

Epidemiology

Epidemiological studies of wrist/hand pain have generally focused on specific disorders rather than undifferentiated wrist/hand pain. However, Urwin and colleagues[21] recently reported that 12% of a general population sample had reported wrist/hand pain lasting more than a week over the past month.

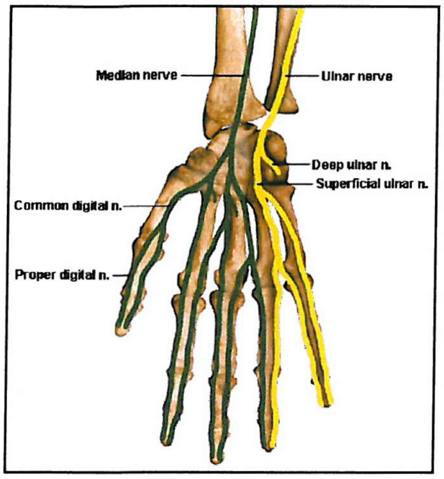
There are no reported studies on tenosynovitis that include a physical examination. In a US population study, 2% of subjects reported a physician diagnosis of tendonitis [3].

Dupuytren's contracture has been reported with a prevalence of 9% – 35% in over 75 year olds, with a lower prevalence in middle age, (0.5% – 16.3% in 45 – 54 year olds [25-28]) and no occurrence at younger ages. Variation in reported prevalences probably arises from the difficulty in diagnosing mild or early stage cases.

Nerve entrapment:

There are a number of recognised sites and nerves that can be involved in nerve entrapment in the arm. The most common, median and ulnar nerve compression, are discussed in detail.

Median nerve compression. The median nerve can undergo entrapment at a number of sites, the most common being at the carpal tunnel in the wrist (**carpal tunnel syndrome (CTS)**). Here compression occurs due to pressure exerted on the carpal tunnel externally (such as from swelling of the surrounding tissues) or from an



increase in the volume being contained in the tunnel (such as from tenosynovitis). Symptoms include sensory loss in the palm, thumb, index, middle and half of the ring fingers, and a dull aching pain may accompany this. Pain may also radiate to the antecubital area of the elbow and the lateral shoulder. Sleep disturbance due to abnormal sensation is common and thenar muscle wasting may occur. Whilst clinical examination, imaging and nerve conduction testing all contribute to the diagnosis CTS, none is conclusive.

Management depends on the cause of the compression, and includes splinting, local corticosteroid injection, NSAIDS or surgery. Keyboard use and other activities involving wrist flexion may be associated with CTS [29].

Median nerve compression in the forearm also occurs, but is less frequent and may be mistaken for CTS.

Epidemiology

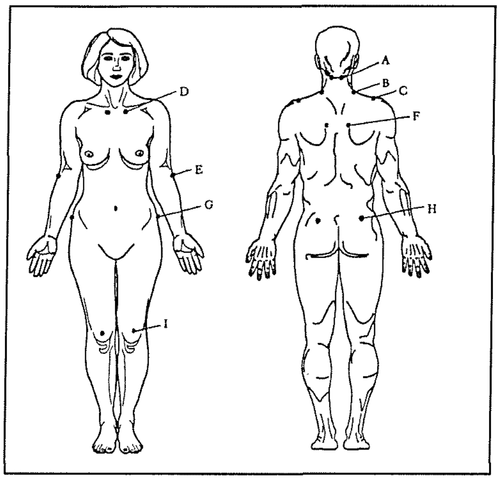
The investigation of carpal tunnel syndrome presents a particular epidemiological challenge: how to identify and classify an entity which, when seen in the general population, apparently presents at times with symptoms, signs or positive nerve conduction tests, but not often with all three, or even two of these [30]. Additionally CTS should be viewed as a continuing process rather than a state of which the classical (diagnosable) case represents only a part.

Using the definition of symptoms of pain, tingling, or numbness in the thumb, index or middle fingers occurring twice a week or more with abnormal median nerve conduction at the wrist (indicated by a difference of greater than 0.4msec in distal sensory latency to the ring finger between the median and ulnar nerve fibres, the median latency being the greater), the point prevalence of CTS has been estimated to be 9.2% in women and 0.6% in men in a Dutch population [31].

Ulnar nerve compression. The ulnar nerve may become trapped at the elbow (**cubital tunnel syndrome**) or at the wrist (**ulnar tunnel syndrome**). At either site compression may be due to trauma, and prolonged extreme elbow flexion may cause cubital tunnel syndrome. Sensory loss is experienced in the ulnar nerve distribution (that is the little and ring fingers), and hypothenar muscle wasting may be seen. **Thoracic outlet syndrome** describes ulnar compression at the thoracic outlet (on the neural pathway between the neck and shoulder) but is a very rare condition in the UK. The treatment strategy for ulnar nerve compression is as for CTS.

Fibromyalgia:

Fibromyalgia is a condition characterised by chronic musculoskeletal pain and tenderness at defined discrete locations as marked on the figure (A – I tender points). Neither evidence of muscle or tendon inflammation, nor abnormal laboratory or radiological investigations are present. Other symptoms including fatigue, sleep disturbance, headaches, irritable bowel syndrome, paraesthesiae, Raynaud's-like symptoms, depression and anxiety may also be experienced. The condition is usually seen between the ages of 30 – 50 years and 80% – 90% of patients are women. This condition has been the topic of much debate in terms of its existence, definition and relationship to other pain syndromes. In particular, rheumatoid arthritis, Sjőgren's syndrome, systemic lupus erythematosis, hyperthyroidism, myositis and neuropathies should be considered in the differential diagnosis of fibromyalgia.



Management is aimed at enhancing peripheral and central analgesia, diminishing mood disturbances, improving sleep and increasing blood flow to muscle and superficial tissues. Outcome is varied. Some patients respond with attention to physical fitness and ergonomic factors only, coping well with their fibromyalgia as a nuisance condition, while others report no significant change in their moderate or severe pain after three years [20].

Consensus criteria for the diagnosis of fibromyalgia were developed in 1990 [32], prior to which a variety of other diagnostic labels were in circulation (fibrositis, psychogenic rheumatism, myogelosis, muscle pain syndrome) each capturing different aspects of the wide symptom profile seen with fibromyalgia. Thus, epidemiological investigation pre-1990 cannot be compared across studies and is difficult to interpret. Post-1990 population studies in the US and Sweden suggest a point prevalence of 2% [33,34].

Non-specific arm pain:

Non-specific upper limb disorders are characterised by recurring or persistent pain, muscular fatigue, numbness, aching, and stiffness in part or all of the arm. Physical signs are not generally present, although tenderness may be seen. Lack of evidence of underlying pathology accounting for the pain is the key feature. Research into non-specific upper limb pain has particularly centred on that related (or believed to be related) to work. Syndromes defined by their occupational cause such as cumulative trauma disorder (CTD), occupational overuse syndrome (OOS), repetitive motion disorder and perhaps the most widely known repetitive strain injury (RSI) all come under the umbrella term of **work-related non-specific upper limb disorder** and are overlapping terms. It is noted that these are all anatomically non-specific as well as lacking in pathological specificity. Nevertheless, such syndromes can include specific musculoskeletal disorders within their definition (for example, if a case of tenosynovitis is due to wrist overuse and repetitive movements, it would qualify also as a case of OOS or RSI [29,35]). It can be seen that the exact nature of non-specific upper limb disorders and thus their diagnosis is far from clear.

Some clinicians and researchers view non-specific pain only as a diagnosis of exclusion, where symptoms are reported but have no accompanying pathology, thus ruling out specific conditions. Others take a more extreme opinion and refute the relevance of such disorders to the medical profession, suggesting that the phenomenon of work-related non-specific pain is less a medical condition than it is a socio-political phenomenon [36]. Such opinion holders cite the Australian epidemic of RSI in the mid-1980's as a prime example: reports of RSI and workers' compensation claims soared following the change in government policy allowing keyboard workers to be compensated for RSI. A national crisis developed. Later, the rules concerning compensation returned to more stringent guidelines, whereon the level of RSI reporting as well as compensation claims returned to their original low state.

There is little doubt that non-specific work-related upper limb disorders do exist and can cause considerable disability and distress. It is also undisputed that societal awareness, acceptability and belief systems strongly influence the reporting of these conditions, which are in turn influenced by economic factors (compensation and time lost from work) as well as the media. Non-specific work-related upper limb disorders are generally believed to involve a strong psychological component, and are viewed by some clinicians as a chronic pain syndrome, operating from a similar basis as other chronic pain syndromes such as fibromyalgia or chronic low back pain.

What has yet to be agreed is the medical approach to this entity. Research into the area would be particularly welcome considering that reports of non-specific work-related upper limb and neck disorders have increased in most countries during the 1990's [29], and that current treatment is often ineffective [36].

1.1.3 Risk factors for musculoskeletal pain in the upper limb and neck

Age and sex

The occurrence of self-reported pain increases with age at all locations [1]. For particular disorders, however, such as lateral epicondylitis, tenosynovitis, de Quervain's syndrome, carpal tunnel syndrome and fibromyalgia, prevalence tails off in the older age groups. This may be a cohort phenomenon, or reflect the fact that older people are less likely to report their symptoms in a questionnaire [1]. Alternatively, having retired, subjects may no longer carry out the work activities that exacerbated symptoms or required that symptoms be resolved.

Women are found to experience, or report, pain more frequently than men; the most extreme difference being for fibromyalgia, where 80% – 90% of patients are women. Dupuytren's contracture flouts this trend and is seen predominantly in men. Lateral epicondylitis is seen equally in both sexes.

Social class

Higher consulting rates to GPs for bursitis and tendonitis (all sites) have been seen in manual social class groups compared to non-manual groups in the UK [1]. Since the elevation of rates could not be explained by a general tendency for the manual groups to consult more readily, the physical load of manual work may be the explanatory factor.

Psychological factors

Self-reported pain, particularly at the neck and shoulder and with fibromyalgia, has been shown to correlate with measures of depression and stress [1]. However, whether psychological factors are a cause or an effect of musculoskeletal pain is less clear [37]. A study performed in the late 1980's went some way to address this issue.

The prospective study of 902 men and women based in Finland suggested that psychological symptoms predicted musculoskeletal symptoms 10 years on, after allowing for baseline pain.

Body mass index

Body mass index (BMI) has been cited as a risk factor for CTS [31], with an 8% increase in risk for each unit increase in BMI (kg/m^2). Evidence for BMI as a risk factor for neck and shoulder pain is inconsistent [1].

Occupation

Numerous epidemiological studies have attempted to assess occupational factors in the occurrence of musculoskeletal disorders. However, they should be interpreted with the methodological challenges of such investigation in mind. Studies frequently compare the prevalence of disorders in a screened group of workers with that in another set of workers. However, if symptoms are associated with an occupation or activities involved in that occupation, workers may deliberately avoid or leave such work and favour other occupations. Thus, any comparison between groups will underestimate any relationship between that occupation and condition (the healthy worker effect). Studies focused on particular work activities as putative risk factors for specified conditions may overlook other complaints or exposures and hence miss a confounding risk factor or more generalised associations. The effects of age and length of time in the job also need consideration.

Studies have investigated butchers, packers, garment workers, assembly workers, scissors-makers, heavy machine operators and others. They are generally compared to workers in less physically strenuous occupations - office workers, workers in the same industry but carrying out a different task, or the general working population. Specific mechanical occupational activities such as using vibratory tools, prolonged neck flexion, prolonged elevation of the shoulder, repetitive tasks using the hands and wrists, bending or twisting the arm or hand have all been associated with individual disorders [29]. Psychosocial factors at work such as work monotony, job control, job satisfaction, social support and job demands also appear to play a part in musculoskeletal disorders [29].

1.2 Diagnostic criteria

Diagnostic labels are the physician's shorthand for a wealth of information on a presenting condition: useful diagnoses encapsulate the nature of a condition and an understanding of the processes leading up to and beyond the current state. Making an accurate diagnosis can therefore be useful in directing effective management and treatment, and in facilitating disease prevention via research or dissemination of information.

Diagnostic labels have been built up over time as patterns of disease (including natural history and possible treatments as well as presentation) have been identified and documented. For example, in the rheumatology field many disease patterns have been recognised in the last 50 – 60 years, including ankylosing spondylitis, psoriatic arthropathy and polymyalgia rheumatica, which would have previously been compressed into one of three prevailing diagnoses: gout, rheumatism or arthritis [38]. As a result, more effective treatment and management for these different conditions have been developed. Similarly, pre-1930, no distinction was made between angina pectoris and myocardial infarction. The invention of the cardiograph led to the recognition of myocardial infarction, and whilst no new disease had been discovered, the new diagnostic label allowed physicians to predict prognosis with greater accuracy [39].

The diagnostic label of 'low back pain' is a prime example of a diagnosis which possibly covers a variety of underlying pathologies, but which is 'good enough' in terms of directing treatment and predicting outcome. It represents an active shift away from the tendency to diagnose on the basis of pathology, and may, for this reason, be controversial. The use of this diagnosis also illustrates the point that a diagnosis should be viewed as a prediction, not as an absolute state, and that the diagnoser should be continually following-up their patient and re-evaluating their diagnosis [39]. In the low back pain context, this involves looking out for characteristics of serious underlying pathology such as infection or systemic inflammatory disease (known as 'red flag' diagnoses), whilst remaining aware that the majority of low back pain in the general population or in primary care is simply that, low back pain [40].

Making a diagnosis thus depends on the recognition of a previously identified pattern, often chosen from two or three original diagnostic hypotheses after seeking further discriminatory evidence. This is a complex task and requires knowledge not only of a patient's complaints, but also of background information on the patient (previous medical history, age, sex, environmental exposures) and the natural history of disease. The latter may include the age and sex distribution, prevalence, geographical distribution and risk factors, the last of these being known as 'yellow flags' in the low back pain diagnostic process. Diagnosing a complaint is therefore context specific, as has already been alluded to: a general practitioner does not expect to refer all patients for possible cerebral tumour when they present with a headache, although the possibility of this rare condition is borne in mind.

Diagnosis and classification of musculoskeletal disorders of the neck and upper limb

Considering the aspect of diagnosis concerned with the physical presentation of a complaint, reaching a definite diagnosis in the area of musculoskeletal upper limb and neck disorders may not be straightforward, as has been previously indicated. Shoulder disorders present similarly and may coexist. Compression of the median nerve at the carpal tunnel has no definitive clinical features and abnormal nerve conduction is seen both with and without symptoms in the general population. Fibromyalgia and thoracic outlet syndrome are controversial as clinical entities, and cervical spondylosis is, in practice, often used as a catch-all label for neck pain not otherwise specified. It has also been suggested that both neck pain and shoulder pain presenting in primary care fit similar models of diagnostic process to that for low back pain [40].

Variability amongst physicians in their diagnostic practices, even when they agree on the physical findings, has been documented [41]. This is perhaps unsurprising, since until recently, no widely recognised and clearly defined diagnostic criteria existed for the majority of musculoskeletal upper limb and neck disorders [42-44]. The range of classification schemes including those based purely on descriptive designations (e.g. painful wrist), those based on pathology (e.g. tenosynovitis) and those which attribute a cause for the disorder (e.g. RSI and work-related upper limb disorder) has further muddied the waters. Even if a classification system can be agreed and diagnostic criteria defined, making a correct diagnosis relies on the accuracy of the signs, symptoms, diagnostic tests and investigations used as diagnostic criteria both in terms of validity and repeatability.

Hence, research into musculoskeletal disorders in the upper limb and neck has been severely hampered by a lack of reliable case definition, and thus incomparability of study reports. Epidemiological studies also make further requirements of diagnostic criteria beyond those for clinical practice: that they are cheap, simple and reliable enough to be used on a large (and hence heterogeneous) population sample, such that case definition and ascertainment are ensured. Diagnostic criteria used in epidemiological research must also be unambiguous, even if, as a consequence, they are somewhat arbitrary.

In the last 10 – 20 years a number of researchers have proposed classification schemes and diagnostic criteria for upper limb and neck disorders (particularly those pertaining to work-relatedness). Buchbinder *et al* [42] reviewed four of these, paying particular attention to their appropriateness for purpose, validity, repeatability, feasibility and generalisability (Table 2). This methodological framework for appraisal was based on the premise that classification schemes should be tested for and meet standard measurement principles before their utility is accepted. The classification systems reviewed were chosen because they were concerned with one or more soft tissue disorders in the upper limb or neck, and with the frequency, aetiology, diagnosis, treatment or prognosis of such disorders. Classification systems with a narrow focus (such as shoulder disorders solely) were not considered.

Waris *et al* [45] devised a scheme to detect neck and upper limb disorders in epidemiologic surveys. Diagnostic criteria were based on expert opinion and a detailed literature review and were clearly presented. The classification system was simple, relied on a clinical examination alone, and took about one hour to perform. However, the examiner required special training and cases with examination findings that were discrepant with the evident clinical diagnosis were referred to an expert. No attempt was made to test the either the validity of the classification scheme, or the validity or repeatability of the diagnostic criteria. In particular the researchers felt that the system had a considerable weakness in being non-comprehensive; nerve entrapments and non-specific entities were likely to remain unclassified in this scheme. The authors used the classification scheme in a variety of occupational settings [46-49] and in these patients were frequently placed into two or more categories. Whether this reflects true coexistence of disease, or a trait of the classification scheme is unclear.

Table 2: Description of four classification systems for musculoskeletal disorders of the neck and upper limb

Author	**Waris *et al*** [45]	**Viikari-Juntura** [50]
Country	Finland	Finland
Year	1979	1983
Purpose	Case finding/ screening to determine occurrence in occupational health surveys	As for Waris *et al*
Domain	Upper limb and neck disorders, known or anticipated relation to work	Upper limb and neck disorders, known or anticipated relation to work
Specific exclusions	Inflammatory diseases, chronic arthrosis	Nil
Categories	Tension neck syndrome Cervical syndrome Thoracic outlet syndrome Humeral tendinitis Frozen shoulder syndrome Acromioclavicular syndrome Lateral and medial epicondylitis Peritendinitis and tenosynovitis Pronator teres syndrome Carpal tunnel syndrome	As per Waris *et al* Plus: Infraspinous tendinitis Olecranon bursitis Carpal ganglion Painful 1st carpometacarpal joint Osteoarthritis of finger joints Posterior interosseous nerve entrapment Ulnar nerve entrapment at the elbow Ulnar nerve entrapment at Guyon's tunnel
Additional axes	Nil	Nil

Author	**Silverstein** [51]	**McCormack *et al*** [52]
Country	US	US, Canada
Year	1985	1990
Purpose	Case finding in industry, to determine association with biomechanical risk factors	Case finding to determine occurrence in a manufacturing workforce
Domain	Upper extremity 'cumulative trauma disorders' (CTD)*	All neck, upper limb disorders, particularly tendinitis and related disorders
Specific exclusions	Localised osteoarthrosis of interphalangeal joints. Also exclusions listed below in definition of CTDs.	Nil
Categories	End point categories: Ulnar nerve compression (Guyon tunnel) Carpal tunnel syndrome Trigger finger Tendinitis, tenosynovitis, de Quervain's disease Non-specific pattern of pain, numbness or tingling Degenerative joint disease (LOA) Lateral and medial epicondylitis Olecranon bursitis Radial nerve compression (radial tunnel syndrome) Median nerve compression (pronator teres syndrome) Ulnar nerve compression (cubital tunnel syndrome) Bicipital tendinitis Rotator cuff tendinitis 'Frozen shoulder' Degenerative joint disease Tension neck syndrome Scapulocostal syndrome	Group A: Carpal tunnel syndrome Epicondylitis Tendinitis Shoulder: miscellaneous Ganglion Neck: miscellaneous Group B: Myalgia Arthralgia Other groups Miscellaneous
Additional axes	Nil	Severity: mild, moderate, severe

*Meets both interview and physical examination (PE) criteria, positive PE and interview; only meets interview criteria, positive interview; otherwise negative CTD. On interview criteria: one of above end points; symptoms lasting more than 1 week + /or occurring 20+ times in the previous year: no evidence of acute traumatic onset; no related systemic disease; onset since current job. Physical examination: characteristic signs of end points; rule out other conditions with referred symptoms.

Viikari-Juntura [50] used the Waris *et al* classification scheme with some modifications to case definition, and added further categories. Whilst these covered forms of nerve entrapment omitted from Waris *et al*'s system, still no provision was made for non-specific entities. No specific exclusions from the system were made, and again the examination relied on a trained examiner with specialist skills, and took in excess of one hour to perform. No investigation of validity or repeatability was made on the categories or the diagnostic criteria.

Silverstein's classification system [51] was confined to the investigation of 'cumulative trauma disorders', a label that implies causation and that was felt to be unproven for many of the categories included in the scheme. Case definitions were clearly stated, although the examination required special training to perform. No assessment of validity or repeatability was made on these diagnostic criteria, and it was unclear whether the scheme had been used in any other settings.

The fourth classification system, presented by McCormack *et al* [52] used the International Classification of Disease, 9th revision, Clinical Modification (ICD-9CM) to record the physician's diagnosis, which they then classified according to their own scheme. Diagnostic criteria were given, although many were based on the judgement of the authors and physicians involved with the study. How the ICD-9CM code was classified was not described. No assessment of validity or repeatability was made on the diagnostic criteria and the system had not been used in other settings.

All four of these classification schemes used diagnostic labels that implied an underlying pathology that could only be inferred but not ascertained by a clinical examination (e.g. degenerative joint inflammation). Buchbinder *et al* concluded that these classification schemes were unsatisfactory according to a number of their methodological criteria. In particular generalisability, validity and repeatability were poorly investigated, and feasibility for use in large studies was limited.

Other authors have proposed different bases for the classification of neck and upper limb disorders. Kuorinka and Viikari-Juntura [53] have suggested a hierarchical classification system with five main categories of neck or upper limb disorder into which discrete diagnoses should fit (Table 3). Their system was devised with the working population alone in mind. Again, these categories were based on the authors' experience, and any work to validate the system has not been published.

Nørregaard *et al*[54] argue that since clinical criteria are often poorly defined and lack objective appraisal, a simple phenomenological classification of pain according to duration and area of distribution could form a more 'honest' basic scheme, with further tests to confirm or subdivide these categories (Table 4). Most of their suggested tests are the 'gold standard' tests, which may exclude their use in large studies due to expense and availability.

Table 3: Proposed basis of classification of neck and upper limb symptoms in working populations

	Symptoms and characteristics		Examples
1	Temporary symptoms of overuse	Fatigue, stiffness, soreness following strenuous exercise	Athletes muscle soreness Pain in neck muscles after a strenuous work spell in typewriters
2	Conditions with long-lasting pain with pathological changes and functional loss	Chronic pain (often at tendons, etc.) possible after prolonged work stress.	Tenosynovitis and periteninitis of wrist and forearm, humeral epicondylitis
3	Primary fibromyalgia	Symptom complex with general aches and pains, prominent stiffness, general fatigue, poor sleep, anxiety, chronic headache, irritable bowel syndrome	'Cervical/scapula fibrositis'
4	General diseases with musculoskeletal symptoms	Various symptoms	Rheumatoid arthritis Degenerative arthritis
5	Psychogenic manifestations	Symptoms in which psychic features dominate Symptoms of psychological origin	Marked musculoskeletal pains in a depressive person

Classification according to Kuorinka and Viikari-Juntura[53]

Table 4: Proposed classification of neck and upper limb pain

Duration	Area of distribution	Diagnostic terms	Subdivision or confirmation by
Acute	Localised tendon	Acute peritendinitis, tendon tears	MRI, ultrasonography, arthroscopy, biopsy
	Localised muscle	Acute muscle strain	
	Regional	Delayed onset muscle soreness	
	Regional nerve related	Entrapments	Neurophysiology, MRI
	Regional joint related	Vertebral dysfunction, distortion	Clinical tests
	Generalised	Acute somatic disease, overstraining	Blood tests
Persisting	Localised tendon	Tendinosis and /or tendinitis	
	Localised muscle	Trigger points	
	Regional	Regional fibromyalgia/ myalgia	
	Regional nerve related	Entrapments	Neurophysiology, MRI
	Regional joint related	Vertebral dysfunction, joint disease	
	Generalised	Somatic disease, fibromyalgia	Blood tests, biopsies

Classification according to Nørregaard *et al*[54]

The diagnosis of shoulder disorders has received much attention, and experts are in broad agreement (based on clinical experience) that most cases can be satisfactorily diagnosed with a careful history and physical examination [55-58]. More challenging cases may benefit from magnetic resonance imaging (MRI) or diagnostic injections [56-59]. Broad classifications of shoulder disorders were in agreement with Booth and Marvel [60], who proposed seven categories of shoulder disorder based on anatomical location and systemic process:

1) musculoskeletal disorders
2) trauma
3) systemic disease
4) neoplasms
5) frozen shoulder
6) neurovascular disorders
7) referred pain.

In contrast to the above system, the physical examination findings and medical history of shoulders presenting in general practice in the Netherlands were classified by data driven methods [61]. Only three distinct shoulder profiles were seen: one characterised by long duration of pain but no limitation of movement; another by long duration of pain and some limitation of movement; and a third smaller group by recent pain and moderate to severe limitation of movement. The authors concluded that the suitability of the more detailed shoulder classifications advocated in the medical literature was doubtful in a general practice setting, since their findings did not indicate that more categories of shoulder disorder were meaningful in terms of prognosis or shoulder function. In a second paper by these authors [62], the conclusion was again that a more detailed classification of shoulder disorders was not needed to determine a successful therapeutic strategy, although the three shoulder classifications suggested there were of a slightly different nature to those in the previous paper.

Classification of carpal tunnel syndrome has undergone much study in recent years. Early 1990's studies used neurophysiological testing as the gold standard for CTS diagnosis, and all concluded that medical history and physical examination were of limited use as diagnostic tools because of their poor agreement with the gold standard [63-65]. Later literature has shifted position from viewing neurophysiological testing as the gold standard [30,66]. Consensus criteria for CTS in epidemiological studies agreed in 1998 by 12 medical researchers [66] suggested a combination of symptoms, physical examination and electrodiagnostics in the ideal situation. In the absence of neurometry testing, the use of hand symptom diagrams leading to a classification of classical, probable, possible or unlikely CTS alongside physical examination and recording of sleep disturbance was recommended. It is of interest to note that these researchers could not reach a consensus opinion on the classification of subjects with classical or probable symptoms but normal electrodiagnostics. No attempt was made to validate these criteria in a population.

1.2.1 The HSE criteria set

In response to the growing awareness of the difficulties surrounding case definition for the epidemiological study of musculoskeletal neck and upper limb disorders, in 1997 the UK HSE (Health and Safety Executive) sponsored a Delphi exercise to establish case definitions for several clinical conditions of the upper limb associated with work [67]. The Delphi technique is a method of collecting and combining the experience and judgement of experts. In this exercise, a core group of 29 experts from the UK participated, representing the fields of rheumatology (6), surgery (3), occupational health (8), epidemiology (3), general practice (1), physiotherapy (2), ergonomics (3), psychiatry and psychology (2) and pain physiology (1). All were involved in clinical management of patients or epidemiological investigation of work related upper limb disorders. A neurologist also agreed to participate.

At the first stage of the exercise, all participants were asked to complete proformas and recruit two colleagues to do likewise on:

1) carpal tunnel syndrome
2) tenosynovitis of the wrist
3) pain syndrome of the forearm or hand
4) lateral epicondylitis
5) frozen shoulder

and optionally on:

6) de Quervain's tenosynovitis
7) shoulder tendonitis
8) shoulder capsulitis
9) thoracic outlet syndrome.

Each participant's proforma included a definition, major and minor diagnostic criteria, relevant comments and their professional affiliation. Participants were also asked if there were other conditions that they felt should be discussed. Aspects of work-relatedness were not to be included at this stage.

The results of the first stage were reviewed at a workshop (Birmingham, UK, February 1997) by the core participants. An agreed case definition and diagnostic criteria were recorded at the end of each discussion. The third stage of the process was to send out these agreed case definitions to the core participants after the meeting for further review. A total of 430 individual responses was generated: from 26 for shoulder capsulitis to 45 for CTS. Shoulder capsulitis and frozen shoulder were mostly considered to be manifestations of the same condition and were amalgamated for discussion at the workshop.

Tables 5 to 9 list the disorders, consensus definitions, surveillance criteria and additional features agreed by the Delphi exercise.

Table 5: Consensus definition and criteria for shoulder disorders

Disorder	Definition	Surveillance criteria	Additional features
Shoulder capsulitis	A condition characterised by current or past pain in the upper arm, with global restriction of glenohumeral movement in a capsular pattern.	History of unilateral pain in the deltoid area and equal restriction of active and passive glenohumeral movement in a capsular pattern (external rotation > abduction > internal rotation).	It was noted that as well as surveillance criteria the development of an agreed staging system and measures of severity would be useful investigative and clinical tools.
Shoulder tendonitis	Symptomatic painful inflammation or degeneration of the tendons of the rotator cuff or biceps.	Rotator cuff: history of pain in the deltoid region and pain on one or more resisted active movements (abduction of the supraspinatus external rotation of the infraspinatus, teres minor; internal rotation of the subscapularis). Biceps: history of anterior shoulder pain and pain on resisted active flexion of elbow or supination of forearm.	

Table 6: Consensus definition and criteria for elbow disorders

Disorder	Definition	Surveillance criteria
Lateral epicondylitis	A lesion at the common extensor origin of the lateral epicondyle of the humerus causing the effects noted in the criteria.	Lateral epicondylar pain and epicondylar tenderness and pain on resisted extension of the wrist.
Medial epicondylitis	A lesion at the common flexor origin of the medial epicondyle of the humerus causing the effects noted in the criteria.	Medial epicondylar pain and epicondylar tenderness and pain on resisted flexion of the wrist.

Table 7: Consensus definition and criteria for wrist and hand disorders

Disorder	Definition	Surveillance criteria	Additional features
Tenosynovitis of the wrist	Inflammation of the extensor or flexor tendon sheaths of the wrist	Pain on movement localised to the affected tendon sheaths in the wrist and reproduction of pain by resisted active movement of the affected tendons with the forearm stabilised	History of crepitus, tenderness or swelling over the affected tendon sheaths.
de Quervain's tenosynovitis	Painful swelling of the first extensor compartment containing extensor pollicis brevis and adductor pollicis longus.	Pain which is centred over the radial styloid and tender swelling of first extensor compartment and either pain reproduced by resisted thumb extension or positive Finkelstein's test.	

Table 8: Consensus definition and criteria for nerve entrapment disorders

Disorder	Definition	Surveillance criteria	Additional features
Carpal tunnel syndrome	A clinical syndrome caused by compression of the median nerve as it passes through the carpal tunnel	Pain, or paraesthesia, or sensory loss in the median nerve distribution and one of: Tinel's test positive, Phalen's test positive, nocturnal exacerbation of symptoms, motor loss with wasting of abductor pollicis brevis, and abnormal nerve conduction time.	No signs or symptoms in the little finger and on the dorsum of the hand, no other cause apparent, history of successful steroid injection or surgery.
Thoracic outlet syndrome	A constellation of symptoms and signs in the arm or hand caused by compression of the neurovascular bundle at the thoracic outlet.	None formulated.	This was considered to be a very rare condition in UK practice.

Table 9: Consensus definition and criteria for non-specific pain disorders

Disorder	Definition	Surveillance criteria	Additional features
Non-specific diffuse forearm pain	Pain in the forearm in the absence of a specific diagnosis or pathology.	Pain in the forearm and failure to meet the diagnostic criteria for other specific diagnoses and diseases.	Loss of function, weakness, cramp, muscle tenderness, allodynia, and slowing of fine movements.

A common view during the Delphi exercise was that non-specific diffuse forearm pain was a diagnosis made by exclusion. It was agreed that further research was needed to establish whether a distinct forearm pain syndrome exists, or whether this is a term for cases that cannot be otherwise classified (as suggested by the criteria proposed).

No consistent evidence suggested that different professional groups responded systematically differently in their case definitions or diagnostic criteria, and considerable agreement was reached by the core participants.

Some shortcomings of the surveillance criteria and the process were acknowledged:

1) The extent and distribution of symptoms were not clearly defined in all the conditions considered (for example, restriction of movement as a criterion for shoulder capsulitis was not defined).
2) A means of excluding systemic disorders when considering work related biomechanical disorders is necessary.
3) It was noted that gold standard diagnostic tests do not exist for most of these conditions, hence validation of the criteria (or particular components of the criteria) is problematic.
4) Duration of condition was not used in these criteria because it was not believed to add to the diagnosis. The exception to this would be shoulder capsulitis, which has a well-documented pattern of progress.
5) Severity of conditions was not used in these criteria because it has poorly defined measures.
6) The HSE criteria set formed at this Delphi exercise has some obvious and important omissions. Pain referred from the neck was not discussed, but was recognised as being a major consideration in any overall classification system.

The exercise was viewed very much as a starting point for further research rather than an end in itself [43,68].

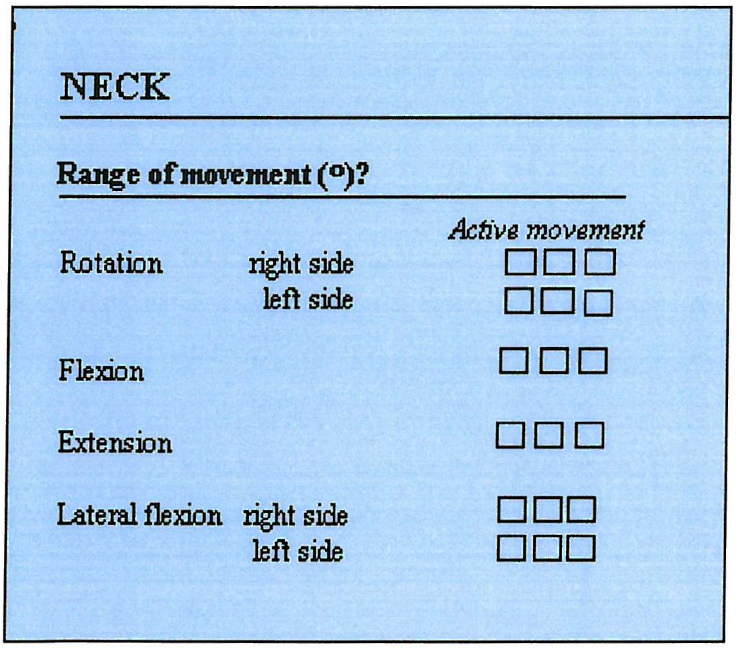
1.2.2 The Southampton Examination Proforma

Palmer *et al* in 1998 [69] suggested that a valid, repeatable diagnostic physical examination for the neck and upper limb area would present substantial progress beyond the HSE criteria set.

These authors have since devised a structured physical examination schedule, 'The Southampton Examination Proforma' (SEP). It has been designed to provide information on the neck and arm to enable diagnosis both of the conditions defined in the HSE Delphi exercise, plus other common musculoskeletal conditions including cervical spondylosis, acromioclavicular joint disorder, subacromial bursitis, olecranon bursitis and fibromyalgia tender spots. The examination has been designed to be performed by trained research nurses, and is accompanied by instructions outlining the anatomical locations of the neck, shoulder, elbow and wrist, and details of how to perform the examination (Appendix I).

Figure 1 shows the neck section of the Southampton physical examination. A question regarding neck pain experienced for a day or longer in the last 7 days is also included in a nurse interview, which is performed before the physical examination takes place (Appendix II). All neck and shoulder movements are measured using a goniometer (neck rotation) or plurimeter (all other movements).

Figure 1: Physical examination schedule for the neck


NECK

Range of movement (°)?

		Active movement
Rotation	right side	☐☐☐
	left side	☐☐☐
Flexion		☐☐☐
Extension		☐☐☐
Lateral flexion	right side	☐☐☐
	left side	☐☐☐

Cervical spondylosis is diagnosed on the basis of current neck pain plus restricted range of movement in any direction. Restricted range of movement was classed as: <80° for left and right rotation, < 60° for flexion, < 70° for extension and <45° for right and left lateral flexion.

Figure 2 shows the shoulder section of the Southampton physical examination. Two extra shoulder diagnoses were added to those discussed at the Delphi exercise - acromioclavicular (AC) joint disorder, and subacromial bursitis.

Figure 2: Physical examination schedule for the shoulder

SHOULDERS

Left Side

1. History: Where is the pain located?

	Yes
No pain	☐
Deltoid area	☐
Anterior shoulder	☐
Acromioclavicular joint	☐
Subacromial bursa	☐
Diffuse	☐
Elsewhere?	☐

(describe) __________

2 Palpation: Where is it maximally tender?

	Yes
No tenderness	☐
	☐
	☐
	☐
	☐
	☐
	☐

(describe) __________

3 Pain on resisted movement?

	No	Yes
a) Elbow flexion	☐	☐
b) Forearm supination	☐	☐
c) External rotation	☐	☐
d) Internal rotation	☐	☐
e) Abduction	☐	☐

	No	Yes	Range of movement (°)?	
Painful arc?	☐	☐	☐☐☐ (started)	☐☐☐ (stopped)

4 Stress test, acromioclavicular joint

	No	Yes
Acromioclavicular joint pain on adduction?	☐	☐

5 Range of movement (°)?

	Active Movement	*Passive Movement*
a) Abduction	☐☐☐°	☐☐☐°
b) Forward flexion	☐☐☐°	☐☐☐°
c) Extension	☐☐☐°	☐☐☐°
d) External rotation	☐☐☐°	☐☐☐°
e) Internal rotation	☐☐☐°	☐☐☐°

Diagnostic criteria for AC joint disorder were based on Waris *et al*'s review of diagnostic criteria [45] and those for subacromial bursitis were based on clinical experience, since no published or agreed diagnostic criteria were available for this disorder (Table 10). The Southampton schedule has deviated from the criteria agreed during the Delphi exercise in that shoulder capsulitis requires current pain (as opposed to current or past pain) that is not necessarily unilateral, and includes a painful arc as a physical sign indicating rotator cuff tendonitis. The restriction of movement has been quantified for the three directions stated in the Delphi criteria.

Table 10: Criteria for shoulder disorders according to the SEP

Condition	Criteria
Shoulder capsulitis	1) Pain anywhere in the shoulder; *plus* 2) Deficit in active *and* passive movements for *either* abduction *or* external rotation *or* internal rotation. Deficits are defined as: <140° for abduction, <70° for external rotation, <90° for internal rotation
Rotator cuff tendinitis	1) Pain anywhere in the shoulder; *plus* 2) Induced shoulder pain on *either* resisted external rotation, internal rotation, abduction *or* a painful arc
Bicipital tendinitis	1) Pain in the anterior shoulder region; *plus* 2) Induced pain on resisted elbow flexion *or* forearm supination
Acromioclavicular Joint disorder	1) Pain over the AC joint; *plus* 2) Tenderness over the AC joint; *plus* 3) A positive AC joint stress test
Subacromial bursitis	1) Pain over the subacromial bursa; *plus* 2) Tenderness over the subacromial bursa

Figure 3 shows the elbow section of the Southampton physical examination. Examination of the posterior elbow allows olecranon bursitis to be diagnosed, using criteria employed by Viikari-Juntura [50] and Silverstein [51]. Lateral and medial epicondylitis criteria were left unchanged from those suggested by the Delphi exercise (Table 11).

Figure 3: Physical examination schedule for the elbow

ELBOWS

Right Side

1 History: Where is the pain located?

	Yes
No pain	☐
Lateral elbow	☐
Medial elbow	☐
Posterior elbow	☐
Other	☐
(describe) ________	

2 Palpation: Where is it maximally tender?

	Yes
No tenderness	☐
	☐
	☐
	☐
	☐
(describe) ________	

Other observations/procedures:

	No	Yes	Crepitus? Yes
Pain lateral elbow on resisted wrist extension?	☐	☐	☐
Pain medial elbow on resisted wrist flexion?	☐	☐	☐
Swelling over posterior elbow joint?	☐	☐	

Table 11: Criteria for elbow disorders according to the SEP

Condition	Criteria
Lateral epicondylitis	1) Pain over the lateral elbow; *plus* 2) Tenderness over the lateral elbow; *plus* 3) Induced pain over the lateral elbow on resisted wrist extension
Medial epicondylitis	1) Pain over the medial elbow; *plus* 2) Tenderness over the medial elbow; *plus* 3) Induced pain over the medial elbow on resisted wrist flexion
Olecranon bursitis	1) Pain over the posterior elbow; *plus* 2) Tenderness over the posterior elbow; *plus* 3) Fluid-filled swelling over the posterior elbow

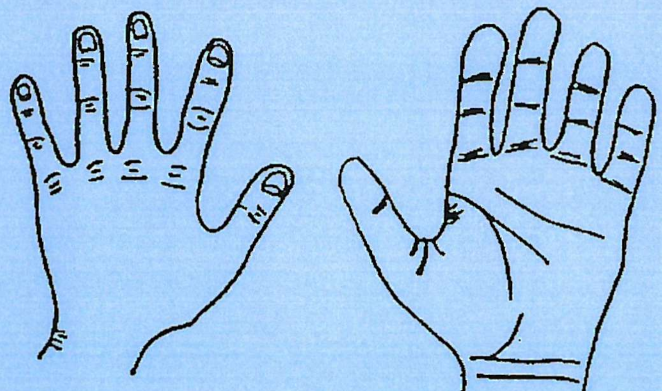
Figure 4 shows the wrist and hand section of the physical examination. Components of the hand examination allow for the diagnosis of Dupuytren's contracture, and an indication of osteoarthritis can be detected by noting the presence of Heberden's nodes. Diagnostic criteria for tenosynovitis, de Quervain's syndrome and carpal tunnel syndrome remain unchanged from those proposed by the Delphi exercise (Table 12), although no nerve conduction tests are included in the Southampton examination. The Katz hand diagram needs to be shaded on at least two of the thumb, index and middle fingers, but not on the dorsum or palm of the hand in order to be graded 'classical', as described by Katz [70].

Figure 4: Physical examination schedule for the wrist and hand

KATZ HAND DIAGRAM

Right Side

If the subject has indicated tingling or numbness in the hand(s)/arm(s) in the past 7 days (question 30), indicate where it (they) occurred by shading the affected parts on the diagram below.



Diagnosis: classical ☐ probable ☐ possible ☐ unlikely ☐

FOREARMS AND HANDS

Left Side

1 History: location of pain (on movement) Palpation: maximum tenderness? Swelling?

	History: location of pain (on movement) Yes	Palpation: maximum tenderness? Yes	Swelling? Yes
dorsal forearm	☐	☐	☐
palmar foram	☐	☐	☐
dorsal wrist	☐	☐	☐
palmar wrist	☐	☐	☐
radial wrist	☐	☐	☐
medial wrist	☐	☐	☐
other	☐	☐	☐

(describe) ______ *(describe)* ______ *(describe)* ______

2 Pain on resisted movement Crepitus?

	No	Yes	Crepitus? Yes
radial wrist	☐	☐	☐
medial wrist	☐	☐	☐
finger extension	☐	☐	☐
finger flexion	☐	☐	☐

3 Hand examination

		No	Yes		No	Yes
Muscle wasting	thenar eminence	☐	☐	hypothenar eminence	☐	☐
Dupuytren's contracture		☐	☐			
Heberden's nodes		☐	☐			

Light touch:	*normal*	*abnormal*		Thumb base:	No	Yes
thumb	☐	☐		Pain?	☐	☐
index finger	☐	☐		Tenderness?	☐	☐
little finger	☐	☐				

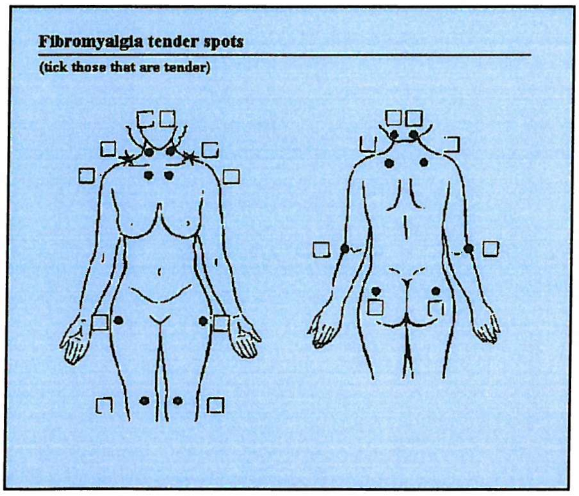
	No	Yes		No	Yes
Positive Phalen's test?	☐	☐			
Positive Tinel's test?	☐	☐			
Weakness of thumb abduction	☐	☐	thumb opposition	☐	☐
Pain on resisted left thumb extension?	☐	☐			
Positive Finkelstien test?	☐	☐			

Table 12: Criteria for wrist and hand disorders according to the SEP

Condition	Criteria
Tenosynovitis	1) Pain over the dorsal wrist *plus* induced pain over the dorsal wrist on resisted finger extension; *or* 2) Pain over the palmar wrist *plus* induced pain over the palmar wrist on resisted finger flexion; *or* 3) Pain over the ulnar aspect of the wrist *plus* induced pain over the ulnar aspect of the wrist on resisted ulnar flexion of the wrist; *or* 4) Pain over the radial aspect of the wrist *plus* induced pain over the radial aspect of the wrist on resisted radial flexion of the wrist, *in the absence of de Quervain's disease of the wrist*
de Quervain's disease of the wrist	1) Pain over the radial aspect of the wrist; *plus* 2) Tenderness over the radial aspect of the wrist; *plus* 3) Pain on resisted thumb extension *or* a positive Finkelstein's test
Carpal tunnel syndrome	1) Katz diagram positive (classical) *or* sensory loss (impairment of light touch in thumb and index but not little finger); *plus* 2) Positive Tinel's sign *or* positive Phalen's test *or* motor loss (wasted thenar eminence *or* weakness of thumb abduction *or* weakness of thumb opposition) *or* disturbed sleep in the last 7 days due to numbness or tingling in the arms or hands

Figure 5 shows the fibromyalgia tender spots that the Southampton schedule examines. The diagnostic criteria are an amended version of the 1990 American College of Rheumatology classification criteria, and involve only tender spots (11 of the 18 are required for a positive diagnosis), but no investigation into widespread pain.

Figure 5: Physical examination schedule for fibromyalgia tender spots



No mention of non-specific pain is explicit in the diagnostic algorithm used with the Southampton physical examination, but since pain is recorded at all sites throughout the neck and upper limb, a diagnosis of exclusion such as that agreed by the Delphi participants can be made.

Repeatability and validity of the Southampton Examination Proforma

The repeatability and validity of the SEP have been reported in a hospital setting [71]. Repeatability was investigated between two observers in 43 subjects attending a soft tissue rheumatism clinic at Southampton General Hospital. All subjects had been referred between November 1997 and May 1998 because of neck or upper limb symptoms, and all agreed to participate in the study. The two physical examinations were spaced a few minutes apart for each subject during their clinic visit.

The between-observer repeatability of physical signs included in the examination proforma on both left and right limbs was assessed using Cohen's kappa statistic (κ) (Table 13). Kappa (κ) indicates the measure of agreement observed above that expected by chance. Thus a κ of 0.00 indicates no agreement other than that expected by chance, a κ of 0.4 – 0.75 indicates good agreement, a κ of above 0.75 indicates excellent agreement and a κ below zero indicates worse agreement than that expected by chance. Kappa (κ) is also scaled by the quantity of disagreement that could possibly have been seen: in the measurement of a rare observation, there is far less chance of disagreement simply because so few positive findings will be reported that could be opposed by another observer.

Shoulder signs and most elbow signs were observed with good or excellent agreement, except for medial elbow tenderness. The hand examination showed poor agreement for pain elicited on resisted finger movement, thenar muscle wasting, Tinel's test and thumb weakness. For all physical signs, however, the small numbers of positive findings will partially account for the extreme κ values near 0.00. (For example, if one observer records no positive findings and the other records even one positive finding, κ by definition, will be 0.00).

A second analysis was performed by the author of this thesis, which took into account the fact that the 86 limbs could not be assumed to be independent observations. κ^a in Table 13 shows the kappa coefficient when left and right limb measurements were considered together for each subject. Thus the two observers could agree in both limbs, disagree in both limbs, or agree in one limb but not the other. Agreement in one limb but not the other was given a weight of 0.2 (Table 14). Hence, the partial disagreement was considered to be more important than the partial agreement, and the possibility that measuring two limbs within one person led to the two observers being more likely to have concordant findings within that individual was compensated for.

Table 13: The between-observer repeatability of physical signs recorded in the SEP

Signs	N	Observer 1 / Observer 2				$\kappa^{\dagger}$	κ^{a}
Shoulder		-/-	-/+	+/-	+/+		
Any shoulder tenderness	86	68	2	3	13	0.80	0.77
Shoulder pain on resisted elbow flexion	86	70	3	1	12	0.83	0.80
Shoulder pain on resisted forearm supination	86	73	3	3	7	0.66	0.67
Shoulder pain on resisted external rotation	86	66	2	1	17	0.90	0.88
Shoulder pain on resisted internal rotation	86	74	4	3	5	0.54	0.54
Shoulder pain on resisted abduction	86	67	0	5	14	0.81	0.79
Painful arc	86	78	1	0	7	0.93	0.92
Positive AC joint stress test	86	76	3	0	7	0.80	0.81
Elbow							
Lateral Elbow tenderness	86	72	5	0	9	0.75	0.76
Medial Elbow tenderness	86	81	4	1	0	-0.02	-0.03
Posterior Elbow tenderness	86	85	0	0	1	1.00	1.00
Other Elbow tenderness	86	86	0	0	0	1.00	1.00
Lateral elbow pain on resisted wrist extension	86	78	2	1	5	0.75	0.74
Medial elbow pain on resisted wrist flexion	86	83	1	0	2	0.79	0.79
Swelling posterior elbow	86	85	0	0	1	1.00	1.00
Wrist and Hand							
Radial wrist tenderness	86	84	1	0	1	0.66	0.66
Medial wrist tenderness	86	84	0	1	1	0.66	0.55
Forearm, dorsal or palmar wrist tenderness	86	82	1	1	2	0.65	0.54
Any swelling in the forearm/wrist	86	86	0	0	0	1.00	1.00
Pain on resisted radial wrist movement	86	84	0	1	1	0.66	0.66
Pain on resisted medial wrist movement	86	84	0	1	1	0.66	0.55
Pain on resisted finger extension*	86	82	0	4	0	0	0
Pain on resisted finger flexion*	86	84	0	2	0	0	0
Muscle wasting (thenar)*	86	85	1	0	0	0	0
Muscle wasting (hypothenar)	86	86	0	0	0	1.00	1.00
Dupuytren's contracture	86	78	1	1	6	0.84	0.83
Heberden's nodes	86	68	1	0	17	0.96	0.95
Abnormal light touch:							
thumb	86	82	3	0	1	0.39	0.27
index finger	86	82	0	2	2	0.66	0.66
little finger	86	82	0	0	4	1.00	1.00
Positive Phalen's test	86	85	0	0	1	1.00	1.00
Positive Tinel's test*	86	85	1	0	0	0	0
Weakness of thumb abduction*	86	85	1	0	0	0	0
Weakness of thumb opposition	86	82	2	2	0	-0.02	-0.03
Pain, resisted thumb extension	86	81	1	2	2	0.55	0.54
Positive Finkelstein's test	86	83	1	0	2	0.79	0.79
Katz hand diagram classical	84	83	0	0	1	1.00	1.00

* these items will have a kappa of zero by definition

† the standard error of κ ranged from 0.00 to 0.11 for all measurements, and from 0.00 to 0.15 for κ^{a}

κ^{a} was very similar to κ for all physical signs (Table 13). The largest changes were for forearm, dorsal or palmar wrist tenderness (κ=0.65, κ^{a} =0.54), pain on resisted medial wrist movement (κ=0.66, κ^{a} =0.55) and abnormal thumb sensation (κ=0.39, κ^{a}=0.27).

Table 14: Weights for κ^a: Kappa adjusted for two limbs examined within each individual

Weights for κ^a		Observer 1: right limb / left limb			
		-/-	-/+	+/-	+/+
Observer 2: right limb / left limb	-/-	1	0.2	0.2	0
	-/+	0.2	1	0	0.2
	+/-	0.2	0	1	0.2
	+/+	0	0.2	0.2	1

Table 15 summarises the repeatability of recorded shoulder and neck movements in this study. A substantial proportion (over 25%) of observations were at least 20° apart between the two observers for active and passive forward flexion, and all movements had a worst disagreement of 20° – 60°, suggesting that only severe limitation of movement can be reliably detected using pleurimetry and goniometry. The results for passive shoulder abduction and external rotation were similar to those reported by Croft *et al.* [72].

Table 15: Repeatability of measurements of neck and shoulder movement in the Southampton Examination Proforma

Based on 86 shoulders and 43 necks	Median range of movement: observer 1	Median range of movement: observer 2	Median difference (obs. 1 – obs. 2)	% of differences that are: ≤ 10°	% of differences that are: ≤ 20°	Worst disagreement (°)
Shoulder:						
Range of active movement (°):						
- abduction	160.0	160.0	7.5	64	87	40
- forward flexion	150.0	140.0	10.0	56	70	60
- extension	62.5	60.0	0.0	72	88	50
- external rotation	57.5	60.0	0.0	73	87	50
- internal rotation	110.0	100.0	10.0	98	100	20
Range of passive movement (°):						
- abduction	160.0	160.0	10.0	64	88	60
- forward flexion	155.0	140.0	10.0	57	71	60
- extension	67.5	60.0	5.0	70	88	50
- external rotation	57.5	60.0	0.0	73	87	50
- internal rotation	110.0	100.0	10.0	99	100	20
Neck:						
- rotation (right)	70	70	0	86	98	25
- rotation (left)	70	70	-5	77	95	30
- flexion	60	40	5	60	91	30
- extension	50	50	0	67	98	40
- lateral flexion (r)	40	40	0	72	98	30
- lateral flexion (l)	40	30	0	65	88	40

The validity of the Southampton Examination Proforma (SEP) as compared with hospital clinic-based rheumatologists or hand surgeons was assessed in a total of 88 subjects. The clinic diagnosis or diagnoses were abstracted for the 43 subjects previously used in the reliability part of the study. The other 45 subjects were recruited from rheumatology and orthopaedic outpatient clinics at Southampton General Hospital and two other district general hospitals. This second group of patients was seen by a doctor in clinic, and if they had one or more upper limb complaints in the opinion of that doctor, they were invited to undergo the Southampton physical examination, performed by one of the two same observers. Table 16 shows the doctors' diagnoses for the 88 subjects, and these are compared with the diagnoses derived from the Southampton examination proforma in Table 17.

Shoulder diagnoses and carpal tunnel syndrome were the most common specific diagnoses given by the clinic doctors. A large proportion of patients was given a diagnosis of regional pain, and a number were given diagnoses not covered explicitly in the SEP.

Table 16: Clinic diagnoses for the study subjects

Diagnosis	**N**
Single diagnoses:	
Shoulder capsulitis	9
Bicipital tendinitis	1
Rotator cuff tendinitis	12
Lateral epicondylitis	8
Medial epicondylitis	0
Tenosynovitis	1
de Quervain's disease	5
Carpal tunnel syndrome	13
Cervical spondylosis	3
Fibromyalgia	0
Two diagnoses:	
Carpal tunnel syndrome & de Quervain's disease	1
Fibromyalgia & Cervical spondylosis	1
Shoulder capsulitis & cervical spondylosis	3
Shoulder capsulitis & other *(Both were OA)*	2
Lateral & medial epicondylitis	2
de Quervain's disease & other *(trigger finger)*	1
Carpal tunnel syndrome & pain *(neck pain)*	1
Three diagnoses:	
Shoulder capsulitis, lateral & medial epicondylitis	1
Other specific diagnosis *(OA, RA, trigger finger, swelling index MCP, demyelination, palindromic rheumatism, supraspinatus tear)*	8
Pain *(neck, neck & shoulder, arm, elbow, forearm, wrist)*	16

The SEP diagnoses were in total agreement with the clinic diagnoses in 24 (27%) subjects (the shaded cells in Table 17), and gave the clinic diagnosis plus extra diagnoses in a further 29 (33%) subjects. These extra diagnoses were predominantly second shoulder diagnoses alongside clinic diagnoses of shoulder capsulitis or rotator cuff tendonitis, or cervical spondylosis alongside a variety of clinic diagnoses (27 subjects). The other two subjects were given

i) olecranon bursitis by the SEP as well as cervical spondylosis by both the SEP and the clinic,

ii) rotator cuff tendonitis, cervical spondylosis and fibromyalgia by the SEP as well as shoulder capsulitis by both the SEP and the clinic.

Table 17: Extent of agreement between the SEP diagnosis and that of the hospital clinic

(Analysis based on 88 people)

Clinic Diagnosis	**SEP Diagnosis based on the nurse's examination**					
	Correct diagnosis	Correct diagnosis plus extra diagnosis/es	One diagnosis instead of two	Incorrect diagnoses only	Correct, incorrect and missed diagnoses	No diagnosis
Shoulder capsulitis (9)		9				
Bicipital tendinitis (1)		1				
Rotator cuff tendinitis (12)		12				
Lateral epicondylitis (8)	5			1		2
Medial epicondylitis (0)						
Tenosynovitis (1)						1
de Quervain's disease (5)	4					1
Carpal tunnel syndrome (13)	4	2		2		5
Cervical spondylosis (3)		1				2
Fibromyalgia (0)						
CTS & dQ (1)			1			
Cervical spond. & fibro. (1)			1			
SC & Cervical spond. (3)		2			1	
SC, LE & ME (1)					1	
SC & other (2)		2				
LE & ME (2)			2			
dQ & other *(trigger finger)* (1)						1
CTS & pain *(neck)* (1)						1
Other diagnosis (8)				4		4
Pain (16)				9		7

Proforma-derived diagnosis is that based on the first observer's examination of the patient. The cells are simple frequency counts. Blank cells have a count of zero

Nine subjects with only pain according to the clinic opinion had a specific diagnosis according to the SEP, seven of these being cervical spondylosis for subjects with neck pain. (All diagnoses were at least in the same region as the reported pain). One subject with 'other specific diagnosis' of RA/Chronic pain had a diagnosis of tenosynovitis from the SEP, one with palindromic rheumatism had the diagnoses of acromioclavicular joint dysfunction, shoulder capsulitis and rotator cuff tendonitis from the SEP, one with a supraspinatus tear had the diagnoses rotator cuff tendonitis and shoulder capsulitis from the SEP and another with neck osteoarthritis had a diagnosis of cervical spondylosis.

Four subjects had two clinic diagnoses, but only one of the diagnoses was picked up in the SEP. Of the remaining 18 subjects, 13 had no diagnosis from the SEP, but had clinic diagnoses that were designed to be discerned by the SEP (in particular carpal tunnel syndrome). Of the other five subjects, one had lateral epicondylitis according to the clinic, but cervical spondylosis according to the SEP. Two had carpal tunnel syndrome, but were given a diagnosis of cervical spondylosis plus shoulder capsulitis and cervical spondylosis plus lateral epicondylitis respectively by the SEP. Another subject had shoulder capsulitis and cervical spondylosis, but was given a diagnosis of shoulder capsulitis and rotator cuff tendonitis by the SEP. The final subject had shoulder capsulitis and lateral and medial epicondylitis according to the clinic, but shoulder capsulitis, bicipital tendonitis, rotator cuff tendonitis and lateral epicondylitis by the SEP.

Cervical spondylosis was commonly diagnosed by the SEP, but not by the clinic, as were multiple shoulder diagnoses in the presence of a single shoulder disorder according to the clinic. These two shortcomings formed the majority of discrepancies between the clinic diagnoses and those of the SEP. Cervical spondylosis was in practice based almost entirely on the presence of neck pain because the ranges of movement proposed to indicate restriction were met by the majority of subjects, regardless of whether or not they reported neck pain. Further investigation into the shoulder diagnoses (particularly to distinguish between shoulder capsulitis and rotator cuff tendonitis) from the Southampton physical examination was clearly indicated by the findings of this study. Carpal tunnel syndrome was missed by the SEP on nine out of the possible fifteen occasions, and medial epicondylitis was not diagnosed at all. The diagnosis of fibromyalgia, based solely on the presence of 11 out of 18 possible tender spots, was also disregarded. This was because agreed and validated criteria for fibromyalgia already exist [32], but cannot be used in the SEP because investigation into lower limb pain is not made. It was felt that the information concerning tender spots was not sufficient to make a diagnosis of fibromyalgia.

Thus further refinement of the diagnostic criteria was suggested by the findings of this study, especially in relation to shoulder capsulitis, rotator cuff tendonitis, cervical spondylosis and carpal tunnel syndrome.

1.2.3 Systematic data-driven classification methods

The HSE criteria set was based on expert experience gathered from a number of medical disciplines. As such, the criteria were based on *a priori* beliefs about how signs and symptoms coexist and that such constellations indicate particular conditions. An alternative approach to classification is one not initially based on medical expertise, but rather based purely on the way signs and symptoms are actually observed to cluster with each other in individuals. Such an approach has the potential to identify previously unidentified symptom-sign complexes, as well as to compare the resulting classification scheme with a medically driven one. Research might conclude, for example, that a data-driven classification scheme little resembles the classical symptom-sign profiles described in the medical literature: a finding reported by a group of investigators using cluster analysis to classify signs and symptoms of shoulder complaints in a general practice in the Netherlands [61].

Data-driven, but medically informed analysis may be another approach to classification or the refinement of diagnostic criteria which allows the researcher to investigate the validity of *a priori* assumptions in a more focused manner. Such analysis might be useful in finding cut points for 'normal' or 'abnormal' continuous measurements (such as ranges of movement) or in deciding whether to include a particular criterion in a case definition.

1.3 Study objectives

Neck and upper limb musculoskeletal disorders are common in the UK. They cause notable discomfort and are a major cause of time lost from work. They comprise a heterogeneous group of disorders that are largely poorly defined, and the confusion as to the use of diagnostic labels has impeded epidemiological research in this area. The main functions of diagnostic labels are to direct treatment and disease management and to facilitate disease prevention by building up a vocabulary that encapsulates the nature of different conditions, their natural history, and often, but not always, describes their underlying pathology. Reviews of classification criteria for neck and upper limb musculoskeletal disorders already available were published during the early 1990's, and these concluded that further work to resolve the confusion was urgently needed. In 1997 an HSE-convened workshop compiled a set of diagnostic criteria for eight of the most common conditions as a starting point for further research, from which the SEP was devised. It has been tested in a hospital setting and whilst it showed good repeatability, a few of the diagnostic criteria need some refinement. The SEP also requires evaluation in the community. Its validity should be tested against associated risk factors and prognosis, because 'gold standard' diagnoses do not exist or are not readily available for the majority of these disorders.

An alternative approach to classification is by data-driven methods (cluster analysis) which produce groupings of individuals based purely on observed data rather than on *a priori* considerations.

Thus a cluster analysis based on the widely varying physical profiles of signs and symptoms of upper limbs and necks from a large community population might bring an alternative understanding to the classification problem of musculoskeletal conditions. It would either confirm or contest the current thinking proposed in the medical literature on diagnostic classifications of musculoskeletal disorders on the upper limb and neck.

The SEP has been used in a community study in two areas of Southampton. It was anticipated that approximately 1400 examinations would be performed in total over the two-year study period amongst a representative sample of the Southampton working-age population. The focus of this thesis is to examine the symptom-sign neck and upper limb profiles of a large general population by cluster analysis of data collected in this community study.

The aims of this thesis are:

1) To classify and characterise each of the neck, shoulder, elbow and wrist/hand symptom-sign profiles amongst a working-age population from the UK using cluster analysis techniques.
2) To compare the resulting classifications with the HSE classifications, once these have been refined.
3) To validate both classification systems by associated disability, healthcare utilisation and risk factors.

CHAPTER 2: METHODS I

THE SOUTHAMPTON STUDY

2 Methods I: The Southampton Study

2.1 Population and subject selection

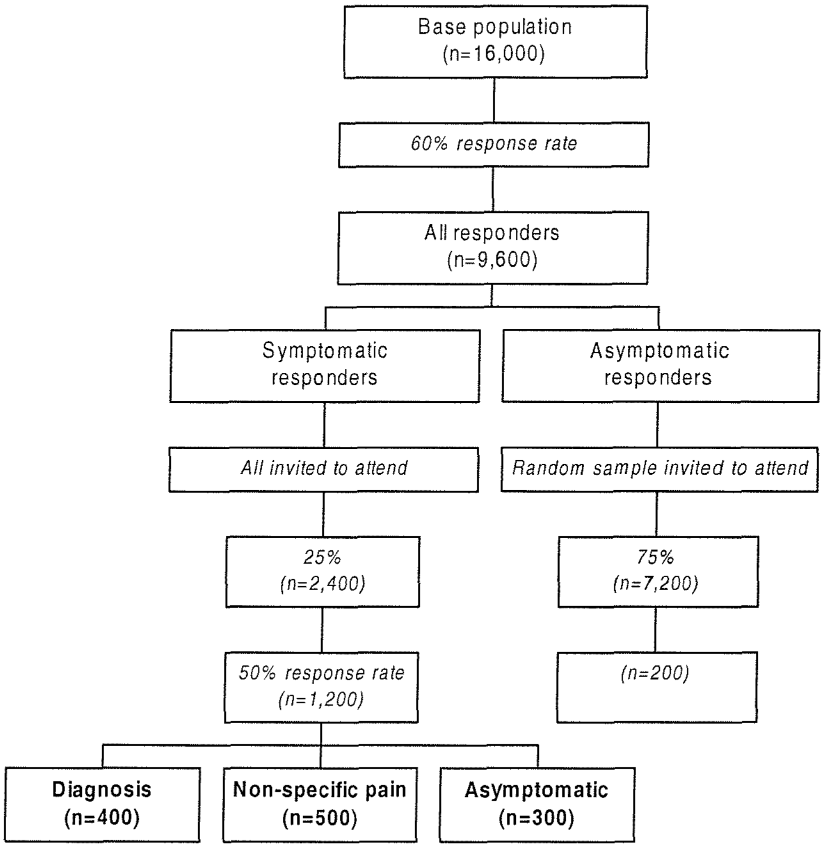
The study population comprised men and women of working age (25 – 64 years) registered at two general practices in Southampton. The first general practice register used (Hill Lane practice) contained 3620 subjects in the specified age bracket, of whom 43 were excluded by the practice. This was to avoid contacting subjects whom the practices felt should not be approached, for example due to recent bereavement or terminal illness. The second practice register (Bitterne practice) contained 6800 subjects aged 25 – 64 years, of whom 113 were excluded by the practice.

During May 1998 to January 1999 three batches of approximately 900 – 1500 survey questionnaires were sent to the 3577 subjects recruited from the Hill Lane practice register. The questionnaire covered demographic, occupational, lifestyle and health information, and explored upper limb or neck symptoms experienced in the last seven days. Reminder questionnaires were sent to non-responders after 4 – 6 weeks. Responders who had reported neck or arm pain or symptoms of numbness or tingling on the survey questionnaire were then sent a letter asking whether they would be willing to undergo a physical examination of their neck and arms, either at the general practice or in their own homes. If such a subject indicated that they would be willing to undergo the physical examination, a study research nurse contacted them to confirm a mutually convenient time for this to take place. A random sample of asymptomatic responders was also approached and invited to attend the physical examination. All physical examinations of subjects from the Hill Lane practice took place from May 1998 to June 1999.

Between June 1999 and August 2000 ten batches of approximately 300 – 1100 survey questionnaires were sent to the 6687 subjects recruited from the Bitterne practice register. Physical examinations for these subjects were completed between June 1999 and November 2000. Again, a random sample of asymptomatic responders was invited to attend the physical examination from this practice.

The original aim of the study was to examine 400 subjects who had at least one of the specific neck or upper limb disorders which the physical examination was designed to detect. The number of subjects required at each stage to fulfil this aim was estimated using response rates and prevalence of neck and arm symptoms (i.e. pain or numbness/tingling) in the community, based on previous experience (Figure 6).

Figure 6: Proposed study numbers



As the study progressed, it was clear that whilst the response rates were as expected, the prevalence of neck or arm symptoms in the community was much higher than had been anticipated. The percentage of those with symptoms from the questionnaire attending the physical examination also remained as predicted. Thus a much lower base population (less than 10,000 subjects) was actually needed to obtain the 2400 symptomatic responders, and hence the 400 responders with specific conditions attending the physical examination.

2.2 Data collected

2.2.1 Survey questionnaire

All questionnaires (Appendix III) were sent out with an identification number pre-printed on the form. This meant that the responder could be checked to be the expected responder (rather than another member of the household, for example), and that reminders could be sent to non-responders.

The first section of the questionnaire asked for general demographic information (age, sex, racial origin) on the subject. Height, weight, smoking habits (past and present), handedness and whether the subject had a paid job were also recorded in this section.

Section 2 explored subjects' current occupation. Those subjects in paid work were asked to report their job title and industry, and were questioned about mechanical tasks undertaken during an average working day. Psychosocial factors at work including work targets, bonus payments, deadlines, support from colleagues or immediate superior, and choice over ways of working were then explored. A global question on work satisfaction was also included. Finally, in the occupation section of the questionnaire, responders were asked whether they had ever changed from a job because of neck or upper limb problems, and if so, to detail the job and nature of the problems experienced. Responses regarding the physical demands of work were classified dichotomously on the questionnaire [73,74].

Section 3 comprised questions about participation in sports activities during the past year and about DIY or craft activities undertaken for more than 20 hours in the past year. The latter activities were only of interest if they involved the use of shoulder, arm or hand muscles.

Section 4 was concerned with certain areas of subjects' health, including diabetes, rheumatoid arthritis and previous broken bones in the hand, arm or collar bone. Women were asked about their use of the contraceptive pill and hormone replacement therapy. Questions about musculoskeletal problems in the neck, shoulders, elbows, wrists or hands followed. These included presence of pain lasting a day or longer in the last seven days, whether that pain impeded everyday activities and how long ago the pain first began. Subjects were then asked about symptoms of numbness or tingling in the fingers or thumbs, hands or arms lasting at least three minutes in the past week. Questions regarding musculoskeletal conditions were based on the standardised Nordic questionnaire [75], which has been found to have satisfactory accuracy for the purposes of epidemiologic screening and research [76-79]. The final nine questions asked about feelings of anxiety and vitality experienced in the past month, and constituted the anxiety and vitality sections of the mental health score from the SF-36 health questionnaire [80].

The questionnaire was completed by responders and returned in a pre-paid envelope.

Information obtained from the questionnaire was used to estimate the prevalence of arm and neck pain and numbness or tingling in the community, along with some assessment of associated disability and duration of pain. The remaining data were used to explore the relationship between occupational or leisure activities, psychosocial and psychological factors and reported pain at different sites.

2.2.2 Clinical interview

Subjects who agreed to participate in the second stage of the study, namely the physical examination, also underwent a clinical interview (Appendix II) which was administered by the research nurse immediately prior to the physical examination. Some attempt was made at this point by the research nurse to check individual survey questionnaires for completeness, clarity and logical answers. In particular, any reported cases of rheumatoid arthritis were verified to have been diagnosed by a doctor (by questioning the subject), and missing information was recovered where possible.

At the start of the interview, the nurse reminded the subject where they had reported pain or numbness/tingling in their survey questionnaire and asked them to show the nurse exactly where this pain had been. Nurses were given written instructions and a body diagram defining the neck, shoulder, elbow, forearm and wrist areas for the purposes of the study (Appendix I). They made a note of the correct anatomical locations of pain reported by subjects according to their guidelines if this differed from the subject's own report. Both the original report of pain and any changes made in location were noted and included on the final database.

For any area where a subject had previously reported pain or numbness or tingling, the research nurse now asked a series of questions. These included: the number of days in the past year for which the pain or numbness/tingling was felt; whether a doctor, physiotherapist, chiropractor or osteopath had been consulted in the last year about that pain or numbness/ tingling; and whether the subjects had received prescribed or non-prescribed medication, physiotherapy or manipulation, or an injection for that complaint. The subject was then asked whether they had taken time off work for that specific pain or numbness/ tingling in the past year, and if so, for how long. They were also asked if they had changed what they did at work because of this problem. Finally, for each location of pain or numbness/tingling, subjects were asked whether particular activities had been made difficult or impossible because of that complaint. Some core activities, such as sleeping and carrying bags were queried regardless of the area of pain; others were pertinent to particular locations (for example, difficulty getting things down from high shelves was only explored in conjunction with shoulder pain, whilst difficulty with writing and undoing jars was explored in association with wrist pain).

All subjects were asked whether they had had pain lasting a day or longer in the past seven days at the neck, right and left shoulder, right and left elbow, right and left wrist or hand areas. All subjects were asked whether they had experienced numbness /tingling lasting three minutes or longer in the past seven days in their fingers or thumbs, other parts of their hands or other parts of their arms. Any subjects who reported pain or numbness /tingling at a site that had not been previously reported in the survey questionnaire (i.e. a more recent complaint) were requested to give full information on the duration of the complaint, health care sought, time off work and associated disability as already described.

These questions were asked in order to obtain a profile of disability and the burden on health care associated with different areas of pain.

2.2.3 Physical examination

The physical examination (SEP) is presented in full in Appendix I and is discussed in Section 1.2.2. Physical signs (tenderness, swelling, ranges of movement, pain on resisted movement and the hand examination) were noted as they were observed by the research nurse at the examination. Physical symptoms (pain, numbness /tingling) were those reported by the subjects as occurring in the last seven days. Thus, the symptoms reported in the physical examination tallied with those reported at the clinical interview.

The neck examination was concerned with ranges of movement and was performed using a goniometer and Plurimeter-V [81], as were the ranges of movement at the shoulder. The wrist/hand examination included a Katz hand diagram which was shaded in by the research nurse as the subject described where their symptoms had occurred. The interpretation of the diagrams into categories of classical, probable, possible or unlikely carpal tunnel syndrome was left to the study rheumatologist. Electroneurometry (indicated on the final page of the examination) was not performed. No further equipment was required for the examination.

Any extra information that the research nurse was given (for example coexistent conditions, previous surgery or injury) was noted.

The physical examination data were used to diagnose musculoskeletal disorders according to the algorithm described in Section 1.2.2. These data were also used in this thesis in cluster analyses to investigate the different symptom and sign profiles observed.

2.2.4 Repeatability of the SEP in a community setting

The repeatability of the physical examination when performed in this community setting was investigated in 97 subjects, all of whom were responders in the first stage of the main study, and had agreed to attend the second stage. Subjects attending their examination on particular dates were asked whether they would be willing to undergo the examination twice with two different observers. None refused. The observers were one of two research nurses with a rheumatologist. Repeatability between the two research nurses was assessed in a further 18 subjects, recruited in a similar way to the above 97 subjects. In both investigations, the two physical examinations for each subject were spaced a few minutes apart. Only data from the physical examination performed by the first observer (allocated randomly) were included in the main study.

2.3 Data management

Prior to data entry onto computer, all data from subjects attending the clinical interview and physical examination were checked to be logical and complete by the study co-ordinator, research assistant or the study statistician. These checks were performed promptly following the physical examination so that any errors found might be resolved whilst the research nurse could still remember that particular subject. Some information was retrieved in this way. Any extra information added by the research nurse regarding other surgery, injury or coexistent conditions was added to a separate computer database at this stage.

Data were double entered onto computer Access files by data entry clerks, and the differences between the two entries were resolved by a third data entry clerk. Univariate data checks, such as missing values and range checks, were performed by computer staff and were resolved by research nurses or the research assistant. Occupation and social class codes were added using the Standardised Occupational Classification (SOC90)[82].

Study-specific checks were performed by the study statistician, and comprised logical checks both within and between the three parts of the study: survey questionnaire, clinical interview and physical examination. These included, for example, checking that female responders and not male responders had answered the oral contraception and hormone replacement therapy questions. Similarly, it was confirmed that if shoulder pain was reported at the survey questionnaire, then the questions regarding shoulder pain at the clinical interview were also answered (unless that shoulder pain turned out to be outside the shoulder region according to the study definition, in which case the research nurse noted this). Further range checks, such as those for shoulder or neck movement in the physical examination, were also performed. All study-specific checks were resolved by the research assistant or study statistician. Any resulting data changes were programmed by the statistician, but the original data were not irretrievably overwritten. The finalised database was locked by the statistician so that the data could not be inadvertently overwritten.

2.4 Computer programming

All study data were converted to SPSS [83] and STATA [84] data files following data entry. Data checking, initial data exploration, repeatability analysis and validity analysis were performed using SPSS and STATA statistical computer software.

Physical examination data were converted to ASCII files for the cluster analyses, which were performed using ClustanGraphics software [85].

2.5 Analysis plan

Data collected from the community repeatability study were analysed and presented similarly to the results gained in the hospital repeatability study (Section 1.2.2).

An initial exploration of the age and sex distribution of the responders in the main study as well as reported pain at the survey questionnaire and subsequently at the physical examination was made. Data from the subjects who completed the survey questionnaire, clinical interview and physical examination were used for the cluster analyses presented in this thesis.

Four separate cluster analyses were performed using the neck, shoulder, elbow and wrist/hand parts of the physical examination data. A preliminary cluster analysis was also performed on the Katz hand diagram data in order to investigate reported patterns of numbness and tingling. It was planned to incorporate the resulting clusters in the wrist/hand cluster analysis as a replacement for the Katz four-point classification, if it was felt that they provided a more informative grouping. Once a satisfactory cluster solution had been found in each of the four main analyses, the clusters were characterised and presented graphically. They were investigated to see how closely they resembled musculoskeletal conditions described in the medical literature, such as those summarised in Section 1.1.2 (using the HSE criteria).

For each of the four analyses, associated disability information gained from the clinical interview and risk factor information (occupational and psychosocial) gained from the survey questionnaires were used to validate the clusters.

CHAPTER 3: METHODS II

STATISTICAL CLASSIFICATION TECHNIQUES

3 Methods II: Statistical Classification Techniques

A variety of tools are available for detecting patterns of clustering in data statistically [86].

There are four major types of clustering method: hierarchical, partitioning, overlapping and ordination algorithms. Nevertheless, they all share a common framework, as discussed by Milligan and Cooper [87]. This involves

1) Selecting the entities to be clustered.
2) Selecting the variables to be used in the cluster analysis. These should contain sufficient information to permit the clustering of the objects.
3) Deciding whether to standardise the data, and if so, which procedure to use.
4) Selecting a similarity or dissimilarity measure. These measures reflect the degree of closeness or separation between objects.
5) Selecting a clustering method. This will largely depend on the type of cluster the researcher is looking for, (e.g. overlapping clusters, hyperspherical clusters or elongated clusters) since different methods have been designed to find different types of clusters.
6) Determining the number of clusters. Much has been written in the literature concerning this controversial problem.
7) Interpreting, testing and replicating the resultant cluster analysis. Interpretation within the context of the area of research; testing whether the clusters are significant, or just an arbitrary grouping of random noise data; replication in other data samples as a validation of the original results (or otherwise).

3.1 Selecting the entities to be clustered

The entities (in this case the necks/ shoulders/ elbows/ wrists on which the physical examination was performed) should be representative of the population of interest. Milligan and Cooper [87] suggested that the sample of entities should also be systematic rather than random, so that wide coverage of entity types or over-representation of less prevalent suspected groups can be engineered. This would ensure that such groups were detected in the cluster analysis, rather than being treated as outliers to another group. Thus, equal representation of different entity types is more important (particularly if the researcher wishes to characterise the whole spectrum of possible profiles) than retaining the prevalence distribution of different groups in the population. However, exploratory cluster analysis is aimed at discovering and characterising the groups observed in a population, and thus over-representation of certain profiles at the outset may not be possible.

For the classification of upper limb and neck symptoms, the limbs of all subjects attending the Southampton examination were considered, except for those with missing data in the physical examination. Thus, those subjects who had reported symptoms in the questionnaire were over-represented compared to those reporting no symptoms in the questionnaire, since only a small sample of subjects reporting no symptoms at baseline were invited for examination. This was desirable, to avoid the analysis being swamped by asymptomatic necks/limbs.

Incomplete examinations were rare (less than 3.3% of examinations had missing data at any of the neck/ shoulder/ elbow/ wrist locations), and a minority of these instances (less than 30%) occurred because the subject was unable to perform the required action due to pain or disability. It was reasoned that these subjects would be misrepresented in any cluster analysis that ignored their missing data, and might influence the resulting cluster analysis unduly. The remaining missing data were from examinations where only a small number of variables were incomplete, often alongside no report of pain, and no explanation for the missing values was given by the nurse. It was reasoned that these subjects represented randomly generated missing data, and that there was no gain in imputing values in order to include them in the analysis.

Different anatomical regions (neck, shoulder, elbow, wrist/hand) were analysed separately, using the examination data on both limbs of each subject.

3.2 Selecting the variables to be used in the cluster analysis

Selecting the variables to be used in a cluster analysis is a non-trivial task, and one which has been proved to have an important effect in cluster analysis. Milligan [87] states that 'the inclusion of even one irrelevant variable seriously reduced the extent of cluster recovery' (i.e. cluster identification), and strongly warns against including variables in an analysis indiscriminately. Fowlkes *et al.* [88] demonstrated this point in a simulated data set, and presented a method for choosing a subset of variables to be used in a cluster analysis from the original variables. Their procedure identifies those variables which provide the most evidence of clustering within the data, and is a computationally intensive stepwise procedure which runs in parallel with the actual clustering of the data.

Another way of reducing the original variable set is to employ dimension-reduction tools such as principal components analysis (PCA) or factor analysis, which create combinations of the original variables in order to retain information whilst reducing the number of variables.

From a mathematical viewpoint, using PCA or factor analysis is likely to influence the outcome of any cluster analysis performed on the transformed data. This is because reducing the number of variables representing different aspects of the data may diminish or accentuate particular features of the data, and thus highlight or obscure groupings within the data, as Cooper and Milligan describe [87]. This paper also suggests that the performance of the cluster analysis on such transformed variables offers no advantage over an analysis using standardised data (a far simpler procedure). Thus, dimension reduction via PCA or factor analysis may or may not be desirable prior to cluster analysis, but is more a matter for the researcher's judgement than of statistical practice. Cluster analyses have been performed on variables transformed by PCA or factor analysis, particularly in the field of psychology [89].

An indirect method of variable selection is that of using weighted variables. The weighting can take place at either the standardisation or the choice of distance measure phase in the cluster analysis process, and will be discussed under those headings.

The aim of cluster analysis in the current study was to explore the way in which signs and symptoms recorded at the SEP coexisted at each anatomical location. Thus, it was desirable to retain all of the original variables recorded on the SEP relating to each anatomical region in the cluster analysis for that region. General physical information such as height and weight recorded in the SEP was not used in any cluster analysis and nor was demographic information. Whilst these variables may give useful diagnostic information, we wished to assess the diagnostic scope of the physical examination data alone, and hence were able to use gender and age for external validation of the cluster analysis.

The only exception to this was in the cluster analysis for the neck, where information about pain experienced in the last seven days was recorded in the clinical interview rather than the physical examination.

3.3 Deciding whether to standardise the data, and if so, which procedure to use

Cluster analysis is founded on the similarity or dissimilarity of the entities to be clustered. The distance measure is therefore of utmost importance, and is a scalar unit reflecting all the variables in the multivariate dataset. However, variables used in the cluster analysis may vary widely in their range of values and the distance measure will be influenced by this. For example, in the SEP of the shoulder, signs and symptoms are all marked as absent or present except for the angles of movement, which range from 0° to 180° for abduction and forward flexion, 0° to 90° for extension and external rotation and 0° to 110° for internal rotation (possibly higher maxima in hypermobile subjects). If the angles of movement were each kept as untransformed values they would contribute to the overall distance measure based on differences unequally, since the difference between two subjects, one with maximum and the other with no shoulder movement, is twice as large for abduction (180°) as it is for external rotation (90°). Moreover, binary variables are coded as 0=absent and 1=present by convention, and any distance between subjects created by the binary observations clearly would be swamped by the distance due to different angles of movement. The individual researcher must decide whether to keep this inherent imbalance, or whether to make each variable equally represented (weighted) in the distance measure, by standardising the data.

One method for standardisation involves transforming the observed maximum and minimum values of every ordinal or numerical variable onto a linear scale from 0 to 1 by dividing by the range. Thus, as for binary variables, the maximum possible value and the maximum possible difference between subjects will always take a value of 1 for each variable. Similarly, the minimum possible value and minimum possible difference between subjects will always take a value of 0 for each variable. Such standardisation procedures have been shown to offer improved performance in terms of cluster recovery compared to untransformed data in a number of simulated datasets [90].

Another method for standardisation is to replace all variables by their z-scores (that is replacing a value x by $[x-\bar{x}/sd(x)]$ where $\bar{x}$ is the sample mean and $sd(x)$ is the sample standard deviation), which thus takes into account the distribution of the values rather than the range alone. This method will accentuate the range of values of a variable that does not vary widely compared to the other variables in the dataset and thus assign to that variable slightly greater importance. Fleiss and Zubin [91] argued that standardisation by dividing by total sample deviations can dilute the differences observed in discriminating variables, because the unobservable (to be discovered) group structure of the data has not been accounted for. Milligan and Cooper [90] conducted a large-scale simulation study to assess the performance of eight different standardisation procedures, and reached conclusions similar to those of Fleiss and Zubin. They also demonstrated that replacing original values by ranks performs poorly compared to standardisation by division by the range, z-scores and non-standardised values.

Other tailor-made standardisation procedures can be used as the researcher requires, but clearly require *a priori* rationale, and investigation into their mathematical properties.

Nominal variables do not need to be standardised in the same way as ordinal or numerical variables do, since calculation of the similarity or dissimilarity between two values of a nominal variable is performed differently. It may be necessary, however, to take the number of categories into account in any distance measure. This will be discussed in the next section.

In the analysis for this thesis, the intention was to make each variable equally represented in the distance measure (having no *a priori* reasons to do otherwise); thus it was necessary to scale down any non-binary observations in the SEP. Each non-binary observation was transformed by dividing the values by the range, as suggested by various authors already cited. This method also allows for intuitive interpretability of values and differences between subjects for each variable, since they all range from 0 to 1.

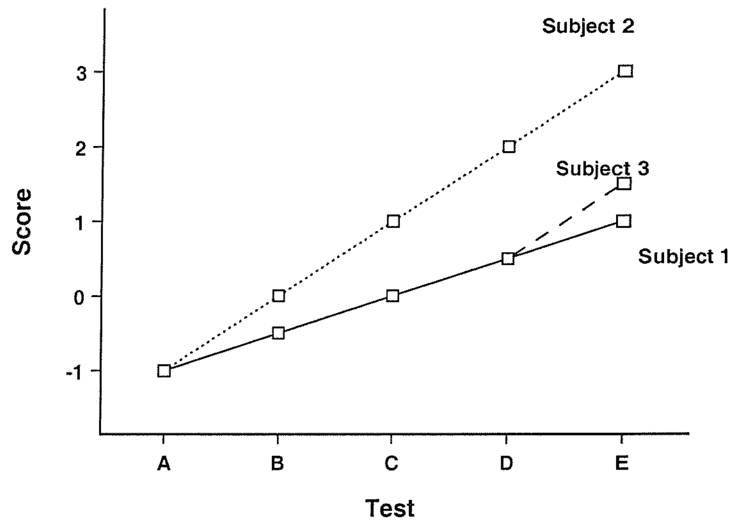
The ranges of movement for the neck and shoulder examinations were standardised using this method. No variables in the elbow or wrist/hand examinations required standardisation.

3.4 Selecting a similarity or dissimilarity measure

Cluster analysis transforms the data made up of n objects and p variables into an nxn triangular matrix representing the pairwise distances between every object and each of the other objects. Numerous similarity and dissimilarity measures have been proposed for use in cluster analysis. These fall into two categories: measures of distance (families of measures for binary, nominal, ordinal and numerical variables) and measures of correlation. The properties of correlation will be discussed first.

For each pair of objects with a vector of p variable values, the product-moment correlation coefficient is calculated. On close examination, however, this measure is difficult to interpret [91]. Figure 7, below shows three hypothetical profiles for tests "A" to "E".

Figure 7: Hypothetical profiles of five tests for three subjects



The vectors for the three profiles are Subject 1: (-1, -0.5, 0, 0.5, 1), Subject 2: (-1, 0, 1, 2, 3), Subject 3: (-1, -0.5, 0, 0.5, 1.5). Thus Subject 2's scores are twice those for Subject 1 plus one. The correlation between Profiles 1 and 2 is one (perfect concordance), even though they are not identical or even parallel. The correlation between Profiles 1 and 3 is 0.986, thus suggesting that these two profiles are less similar than Profiles 1 and 2. It is unlikely that this measure, used as a similarity index, would be desirable in any situation.

Measures of distance

These fall into two groups: measures denoting similarity and those denoting dissimilarity. Since both measures are bound by 0 and 1, and are considered to be 'opposite' measures, similarity, S, is defined as S=1-D (D=dissimilarity); one measure is easily calculated from the other. Distance measures all take the form of a sum of p components relating to the p variables.

Distance between subjects for binary variables

Table 18 shows the possible agreement for p binary variables between two subjects. The binary variables are assumed to represent a presence (+) or absence (-).

Table 18: Agreement for two binary variables

		Subject 1		
		+	-	
Subject 2	+	*a*	*b*	*a+b*
	-	*c*	*d*	*c+d*
		a+c	*b+d*	*a+b+c+d=p*

Subjects 1 and 2 both show positive traits in '*a*' variables, negative traits in '*d*' variables, and different traits in '*b+c*' variables. Distance measures for binary variables are made up of combinations of observed values *a, b, c* and *d* for each pair of subjects, and are usually expressed as similarity measures. The most common measures are presented in Table 19.

Table 19: Measures of similarity for binary variables

Name	Measure	Description
Simple matching coefficient	$\frac{a+d}{a+b+c+d}$	The proportion of total agreements from all possible agreements.
Russell and Rao	$\frac{a}{a+b+c+d}$	The proportion of total positive agreements from all possible agreements.
Jaccard's coefficient of community	$\frac{a}{a+b+c}$	Absence of traits (d) is not counted as either agreement or disagreement and is omitted from the numerator and denominator.
Kulczynski 1	$\frac{a}{b+c}$	The ratio of positive agreement to any disagreement.
Ochiai	$\frac{a}{\sqrt{(a+b)(a+c)}}$	The ratio of actual agreement observed versus expected agreement.

The most important variation in distance measures is due to the treatment of double absences, the value '*d*'. In many real situations double absences cannot be seen as a positive indication of similarity, but rather as a lack of information as to two subjects' agreement or otherwise. For example, the absence of a rare disease, or the absence of a species from two environments does not necessarily imply that objects are similar. Researchers must make their own judgement on this issue in the light of their data. More than one measure of similarity can be utilised in the overall distance measure so that the value *d* can be treated differently for different binary variables, as deemed appropriate.

Gower and Legendre [92] detailed and proved the mathematical properties of these measures, some of which may be required for certain clustering algorithms. These are noted in the next section when each algorithm is discussed.

<u>Distance between subjects for nominal categorical variables</u>
Agreement between nominal categorical variables needs to be treated differently from agreement between ordinal categorical or numerical variables. Hannappel and Piepho described a cluster analysis performed on environmental data consisting of six variables, four of which were nominal categorical, each having between 3 and 5 categories [93]. One of these variables described the types of sediment present at each of 100 environmental sites: fine grain in initial state, silt, mixed sediments and sand. The categories were given numerical codes 1, 2,3 and 4 respectively on the electronic dataset. Any distance measure that used these codes in their numerical sense would then imply that fine grain in initial state is more dissimilar from sand than silt is.

The authors suggested the use of the generalised M-coefficient, which assigns a value of one for any difference in sediment type, and zero otherwise. Over a whole range of nominal categorical variables, this coefficient is the sum of differences divided by the total number of variables and then expressed as a similarity rather than a dissimilarity measure (i.e. the percentage of concordance between two data vectors). A development of the generalised M-coefficient is the similarity coefficient of Hyvärinen, which additionally takes into account the number of categories all the variables have.

For example, if 6 categorical variables have *u, v, w, x, y, z* categories, the Hyvärinen similarity for the data vectors (4, 4, 3, **1**, 3, 2) and (4, 4, 3, **3**, 3, 2) is:

$$1-(\frac{(0*u)+(0*v)+(0*w)+(1*x)+(0*y)+(0*z)}{u+v+w+x+y+z})=1-\frac{x}{u+v+w+x+y+z}$$

because only the fourth variable differs between the two vectors, and it has x levels. Hence, agreement between variables with many categories is deemed more similar than agreement between variables with fewer categories, which is intuitively desirable. If all variables are binary, this similarity measure is equivalent to the simple matching coefficient.

Distance between subjects for ordinal and numerical variables

Ordinal and numerical variables are not distinguished from each other in terms of appropriate distance measures. Table 20 lists a few of the commonest distance measures, expressed in terms of dissimilarity.

Table 20: Measures of dissimilarity for ordinal and numerical variables

	Name	**Formula**
1	Squared Euclidean distance	$\frac{1}{p}\sum_{k=1}^{p}(x_{ik}-x_{jk})^2$
2	Weighted squared Euclidean distance	$\frac{1}{p}\sum_{k=1}^{p}(x_{ik}-x_{jk})^2/r_k^{\ 2}$
3	City block (weighted)	$\frac{1}{p}\sum_{k=1}^{p}\lvert x_{ik}-x_{jk}\rvert/r_k$
4	Minkowski distance (weighted)	$\frac{1}{p}\sum_{k=1}^{p}\left[(x_{ik}-x_{jk})^t/r_k^{\ t}\right]^{1/t}$
5	Canberra metric	$\frac{1}{p}\sum_{k=1}^{p}\frac{\lvert x_{ik}-x_{jk}\rvert}{\lvert x_{ik}+x_{jk}\rvert}$

Where p is the number of variables, x_{ik} and x_{jk} are the values for individuals i and j on the k^{th} variable, r_k is a weight assigned to the k^{th} variable, and t is the Minkowski parameter (assigned by the researcher).

The first four measures all belong to the Minkowski family of distance measures (the first three are special cases of the fourth) and are the most widely used measures of distance for cluster analysis. When weighted by the range of each variable, each of these distances is equivalent to one minus the simple matching coefficient, if all variables are binary.

The Canberra metric compares the difference between two subjects with their sum. A multitude of other, less common, distance measures are not discussed here.

The choice between different coefficients of dissimilarity is likely to depend on the requirements of the clustering algorithm.

Distance measures between subjects for mixed variable types

Gower proposed a general coefficient of similarity [94] in 1971, which was designed to be used on binary, categorical, ordinal and numerical data. It takes the form:

$$S_{ij} = \sum_{k=1}^{p} s_{ijk}\delta_{ijk} \Big/ \sum_{k=1}^{p} \delta_{ijk}$$

where S_{ij} is the similarity (or dissimilarity) between entities i and j, s_{ijk} is the distance between i and j for the k^{th} variable, and δ_{ijk} is an indicator variable denoting whether the distance between i and j is a valid comparison.

Table 21 presents the values of s_{ijk} and δ_{ijk} for binary variables. For binary variables alone, Gower's coefficient is equivalent to Jaccard's coefficient of community.

Table 21: Gower's coefficient for binary variables

	Value of K			
Individual i	+	+	-	-
j	+	-	+	-
s_{ijk}	1	0	0	0
δ_{ijk}	1	1	1	0

For categorical variables, s_{ijk} =1 if the two individuals i and j agree on the k^{th} variable, and s_{ijk} =0 otherwise. δ_{ijk} is always 1. This is equivalent to the generalised M-coefficient.

For continuous or ordinal variables, $s_{ijk} = 1-|x_{ik} - x_{jk}| / r_k$ where r_k is the range of the k^{th} variable. Again δ_{ijk} is always 1. This is equivalent to a weighted city block measure that has been expressed as a similarity.

This coefficient, as with the distance measures for single-variable type datasets, ranges between 0 and 1 with a value of 1 indicating complete agreement and 0 indicating maximal disagreement. It has the mathematical property of creating a similarity matrix which is 'positive semi-definite', a prerequisite for some clustering algorithms.

As previously noted, the family of Minkowski distance measures can be used for combined binary and numerical data.

Weighted distance measures

The nature of distance measures being in some form a sum of component parts lends itself well to the inclusion of weights. These may be constant weights applied to the distance components relating to certain variables, to allow the information on the variable to be accentuated or diminished. Such weights take the form

$$S_{ij} = \sum_{k=1}^{p} s_{ijk}\delta_{ijk}w_k \bigg/ \sum_{k=1}^{p} \delta_{ijk}w_k$$

where S_{ij} , s_{ijk} and δ_{ijk} are as for the Gower's general coefficient of similarity. w_k denotes the weight given to the k^{th} variable. The division by the sum of the weight/indicator product ensures that the overall similarity is still bounded by 0 and 1.

Weights can also be used on the comparison of values for a variable rather than a variable itself. This might be required to accentuate the agreement of the presence of a rare disease, whilst still retaining some level of agreement of its absence. Such a weighting takes the general form

$$S_{ij} = \sum_{k=1}^{p} s_{ijk}\delta_{ijk}w_k(x_{ik}, x_{jk}) \bigg/ \sum_{k=1}^{p} \delta_{ijk}w_k(x_{ik}, x_{jk})$$

where S_{ij} , s_{ijk} and δ_{ijk} are as for the Gower's general coefficient of similarity and $w_k(x_{ik}, x_{jk})$ is a function relating to the k^{th} variable for individuals i and j .

It can be seen that both forms of weighting entail explicit computer programming to specify the required weights, which (particularly in the latter case) may become prohibitively intensive with increasing numbers of subjects and variables.

A data-driven technique designed to amplify the grouped structure in the data via variable weighting (of the first kind) has been devised by De Soete [95] and has proved to be particularly effective in enhancing cluster recovery [96].

Missing values

The fact that distance measures are sums of component parts allows for data with missing values to be used in cluster analysis. Missing values for individuals i and j on any variable result in missing distance components, which can be simply dropped from both the numerator and denominator of the distance measure. This is equivalent to setting δ_{ijk} to zero in the previous equations. Drawbacks of using data containing missing values include a) loss of information, potentially leading to misclassification of subjects or even masking the underlying group structure, and b) loss of mathematical properties required for some clustering algorithms.

Although the distance measure should be chosen carefully and with the nature of the data in mind, Punj and Stewart [97] argued that errors made at this stage do not seem to be as serious as others in the clustering process.

For the neck and shoulder examinations, which included both binary and numerical variables, Euclidean distance was used. This is suitable for both binary and numerical variables and for a wide range of clustering algorithms, due to its favourable mathematical properties. The unweighted squared Euclidean distance (a widely used distance measure) was chosen, since the data had already been weighted by variable standardisation.

Two types of unweighted binary similarity measures were used in the elbow and Katz hand cluster analyses. These two measures were Jaccard's coefficient and the simple matching coefficient, being the commonest distance measures that handle double absences in data differently. Unweighted measures were chosen because there were no *a priori* reasons to treat any variables as more important than others.

Jaccard's coefficient of community and the simple matching coefficient were proposed as the distance measures for the wrist/hand examination data, if binary variables to denote the Katz hand diagram clusters were used alongside all the other binary data. If the ordinal Katz classification was used (the original classification), squared Euclidean distance was proposed as the distance measure.

Essentially the distance measures chosen for the cluster analyses were those reported to be most suitable for the types of variables in each dataset, and, given a choice, those most widely used and reported in the literature.

3.5 Selecting a clustering method

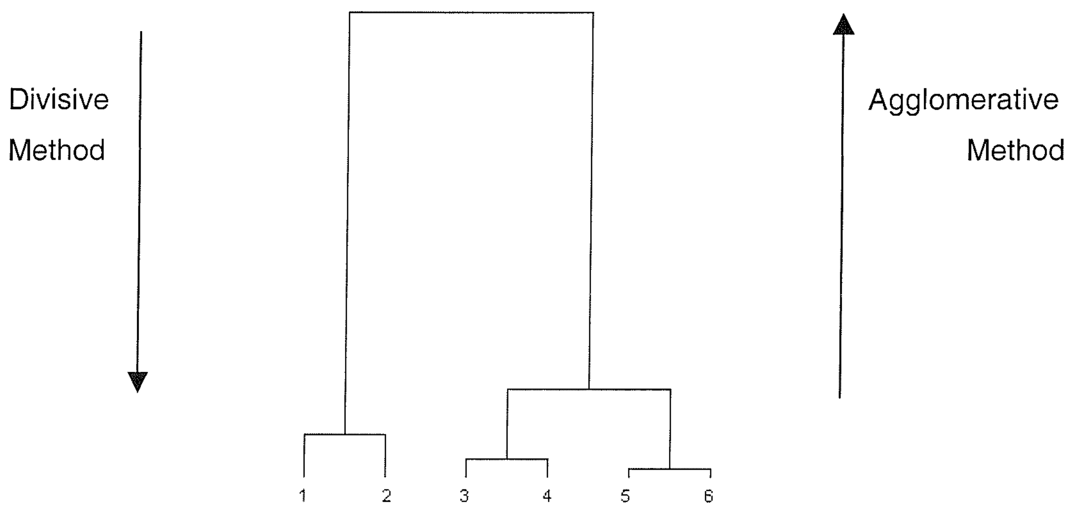
Four types of clustering methods exist: hierarchical, partitioning, overlapping and ordination.

Hierarchical methods

Hierarchical clustering is perhaps the most widely used clustering method [98], and is sequential. Agglomerative (gathering objects together) methods start with a number of single entities, *n*, each considered to be a separate cluster. At each stage two of the clusters are merged, and considered to become a single cluster. This process repeats until only one cluster containing all the entities exists. This routine generates a strictly nested hierarchy of *n* partitions and clusters which are non-overlapping. Once two elements have been joined together they cannot be subsequently separated. Divisive methods of hierarchical clustering reverse the agglomerative process, starting with one cluster containing all elements and ending with *n* clusters containing one entity each. Divisive methods require complex computation and as a result are used rarely [99].

Either hierarchical method can be represented graphically by a dendrogram (Figure 8). The dendrogram can be understood best as a picture of a hanging mobile [100] - the joins between objects indicate the clusters formed, and the height of the join from the base of the mobile indicates a measure of distance. The order from right to left of the objects when represented on a dendrogram is irrelevant.

Figure 8: Example of a dendrogram describing a hierarchical clustering of six objects



The hierarchical clustering method gives an entire clustering map from *n* objects to one cluster. The researcher then has to decide which level represents meaningful or homogeneous clusters. This will be discussed later.

Mathematically, the agglomerative process is as follows: the *nxn* triangular matrix of computed distances is examined, and the two most similar entities are joined together. Then the matrix must be recalculated for the *(n-1)*x*(n-1)* distance matrix, which in turn is examined for the two most similar entities (i.e. the minimum dissimilarity value in the matrix). A decision is then taken on how to calculate the distance between a single object and the two-object cluster just created (and for subsequent stages two clusters each of any size). A number of common methods exist for calculating the distance between two clusters (Table 22).

Table 22: Common methods for calculating distance between two clusters

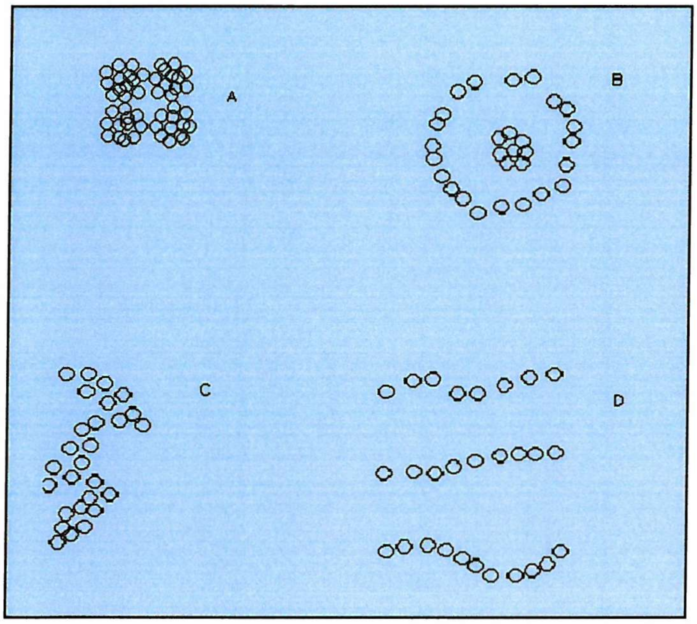
Single Linkage (Nearest neighbour)	$D_{RQ} = Min(d_{rq}, r \in R, q \in Q)$
Complete Linkage (Furthest neighbour)	$D_{RQ} = Max(d_{rq}, r \in R, q \in Q)$
Group Average (UPGMA)	$D_{RQ} = \frac{1}{(l*m)} \sum_{r=1}^{l} \sum_{q=1}^{m} d_{rq}$
Centroid method	$D_{RQ} = d(centroid_r, centroid_q)$
Median method	$D_{RQ} = \frac{1}{p} \sum_{k=1}^{p} d(median_{rk}, median_{qk})$
Ward's method (Increase in error sum of squares)	$D_{RQ} = \frac{1}{l+m} \sum_{e=1}^{l+m} d^2(e \in (R \cup Q), centroid_{RQ}) - [\frac{1}{l} \sum_{r=1}^{l} d^2(r, centroid_R) + \frac{1}{m} \sum_{q=1}^{m} d^2(q, centroid_Q)]$

Where D_{RQ} is the distance between cluster R, containing l elements r, and cluster Q containing m elements q, and d is the distance measure between elements, chosen by the researcher as described in Section 1.3.4.

Single and complete linkage use the minimum and maximum distances, respectively, between an existing entity in cluster R and one in cluster Q. They are straightforward to compute and have been widely used. Single linkage has a number of desirable properties [101] but has been shown to give the least successful cluster recovery when tested [101,102]. Single linkage tends to produce a small number of large heterogeneous clusters. This is because a single entity is more likely to be joined to an already existing large cluster than to another single entity or smaller cluster, simply because there are more entities (and thus chances) in a large cluster for the single object to be nearest to. This phenomenon is known as chaining. Clearly, if a dataset is believed to contain clusters with a chain-like nature, as in the examples in Figure 9, this will be the desirable cluster algorithm.

Complete linkage tends to produce hyperspherical clusters, but again this algorithm exhibits poor recovery of clusters when tested, (although not as poorly as the single linkage algorithm [102]). The group average, centroid and median algorithms all aim to recruit new members to a cluster that are close to the 'centre' of the cluster. The centroid and median algorithms are based on Euclidean geometry and therefore should be used with distance measures known to be Euclidean, which may restrict their use.

Figure 9: Types of data clusters



The cluster structures in B and D might be recovered well by a single linkage algorithm. The centroid algorithm might identify the clusters in A more accurately than single linkage, whilst Ward's method might recover the cluster structure of C most successfully. Since some members of A and C in different clusters can be seen to lie as close to each other as to members of the same cluster, single linkage would be unlikely to recover the correct cluster structure.

Ward's minimum variance [103] essentially finds groups which are as homogeneous as possible at each level and uses every element in each group rather than a summary statistic to define cluster distances. The algorithm searches for the next union of two clusters which results in the minimum increase in the sum of squared Euclidean distances between all cluster members and the cluster centroid, compared to the sum of distances seen in the two separate clusters. This algorithm is clearly based in Euclidean geometry, and should be used with a Euclidean distance. It has shown good recovery of clusters even in the presence of considerable background error data [102] and is therefore a popular choice.

Many other less common algorithms are available, but the literature and validation of the techniques is sparse.

Hierarchical methods have some logical and mathematical problems. Firstly, hierarchical clustering originated in the field of biology, where Linnean taxonomy into the graded sequence of class, order, genus, species and variety was appropriately described in a hierarchical model. The wider application of hierarchical techniques to other fields of study may not be suitable, and this should be a consideration in the choice of clustering method.

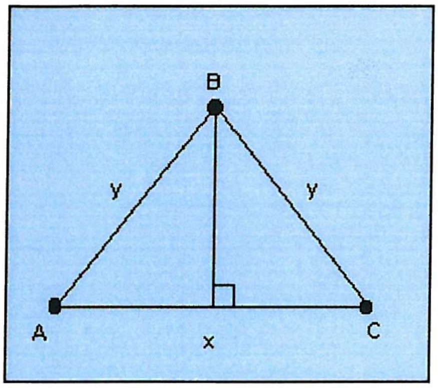
Secondly it is possible, using certain clustering algorithms, to produce sequential distance measures at each stage of the clustering process which are not monotonically increasing. Such procedures do not conform to the ultrametric inequality (so-named because it is a stronger condition for a metric distance than the usual triangle inequality)

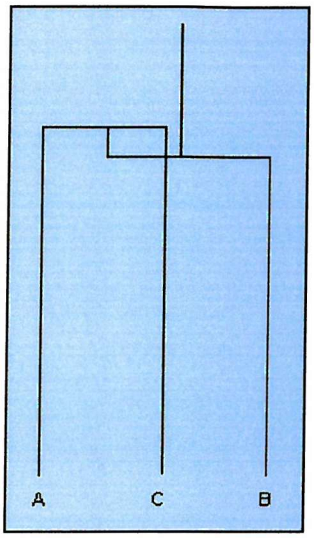
$$d_{ij} \leq Max(d_{ik}, d_{jk})$$

Where d is the distance at which the pairs of points i and j and k become part of the same cluster.

Monotonically increasing distance measures may be advantageous since, within such a framework, it is impossible for a cluster merged later in the hierarchy to have a lower distance value than that of a cluster merged earlier in the hierarchy. This supports the intuition that points which have the smallest distance apart should be merged first. Milligan [104] explains the concept of the ultrametric inequality and proves that single and complete linkage, group average and Ward's method all observe it. Some authors, however, have argued that there is no *a priori* reason to retain the ultrametric property. For example, the three points shown in Figure 10a form an (almost equilateral) isoceles triangle and ideally should be merged into one cluster at the same time. Point B is slightly further away from points A and C than they are from each other. Using single linkage, A would be joined to C at distance x, and B would then be joined to the cluster AC at a distance y. Using the centroid algorithm, on the other hand, A would be joined to C at distance x as before, but B would then be joined to cluster AC at a distance $\sqrt{y^2 - (1/2x)^2}$ (which will be less than x if x and y are similar). The dendrogram of this process would look like that shown in Figure 10b. Such a representation may be desirable to show that three points (or clusters) are almost equidistant from each other. The pairwise merging (or division) of points employed by hierarchical cluster analysis does not allow for larger groups (or more groups) to be formed in one stage.

Figures 10a and 10b: Clustering for points in an almost equilateral triangle, and the resulting dendrogram





The researcher must again decide whether retaining the ultrametric property by choosing an appropriate clustering algorithm is important for their analyses. It would seem, however, that if by using a non-ultrametric algorithm the clustering solution includes non-monotonically increasing distances, that hierarchical clustering may not be the most appropriate clustering method [105,106].

Partitioning methods

Partitioning methods (often referred to as k-means methods) produce non-overlapping clusters and only a single data partition (as opposed to hierarchical methods, which produce n partitions). Whilst some methods allow for a variable number of clusters to be formed, most require the researcher to specify this number, and herein lies the first drawback to most partitioning methods.

Having decided by some means on the number of clusters required (say *c*), partitioning methods require a starting partition to create the *c* clusters, or seed points which act as centroids of each cluster. Seed points may be random or user specified; partitions again may be random, or could use a hierarchical clustering solution as a starting point. Data elements are then assigned to clusters, whose centroids may or may not be immediately updated, depending on the complexity of the algorithm. Data elements may be assigned to clusters based simply on the nearest Euclidean distance between the point and a cluster centroid, or on more involved statistical criteria. This process of reallocating the data points is repeated until the solution converges and no more reallocation is required. (Note that convergence is not always guaranteed). Thus, points can be moved from one cluster to another throughout the process. Data outliers are usually forced to join one of the clusters.

The choice of seed points on a starting partition is important since it can affect the final solution. Random seeds or partitions may lead to a local optimum solution rather than the global optimum [107], with no way of knowing which of these two types of optima has been found. Milligan has shown [87] that using informative starting seeds leads to improved recovery of the clusters in simulated data of known structure, and that this includes the use of Ward's hierarchical algorithm as a starting partition. Specific k-means algorithms are given in Cormack's review of classification [108].

Partition methods involve two crucial decisions: determining the starting partition of the data, and determining the number of clusters. Answering these questions may become a little easier by considering a hierarchical method first.

Overlapping methods

Far fewer algorithms exist that allow for overlapping clusters than exist for either of the hierarchical or partition methods already discussed. Such algorithms often stem from graph theory, and none have been rigorously validated using simulated data with known cluster structure [87]. These methods will not be considered further.

Ordination methods

These methods aim to present a form of dimensional representation of the data, often based on fewer variables than in the original data. Methods such as factor analysis and multidimensional scaling fall into this category. A final cluster solution is not gained from these methods directly: rather, a subjective cluster solution is obtained based on the spatial representation of the data that these methods aim to provide. Such methods are seen by some authors to be dimension-finding tools rather than cluster-finding ones [87]. These methods will not be discussed further.

Since the cluster analyses of physical examination findings presented in this thesis were exploratory, it was decided to use hierarchical clustering initially and then refine the solutions in k-means analyses. This approach takes advantage of the complementary attributes of hierarchical and partitioning methods. Hierarchical methods are effective in preliminary analyses and for finding the likely number of clusters, but are inflexible in the re-assignment of individual entities to a different cluster. Partitioning methods are valuable to move entities between different clusters, but the number and original partition of the clusters needs to be pre-specified.

Within the broad group of hierarchical algorithms, Ward's method is seen as the most effective, being more robust to data error than other methods when tested and yielding homogeneous clusters, although it requires distance measures based in Euclidean geometry.

Ward's method was therefore used for the neck, shoulder and wrist/hand analyses, (where these examinations used squared Euclidean distance), and for the analyses of the elbow, Katz hand and wrist/hand examinations that had employed the simple matching coefficient (a Euclidean distance) as its similarity measure.

The group average algorithm was used in the elbow, Katz hand and wrist/hand analyses that had used a non-Euclidean distance measure (the Jaccard measure of community). The group average algorithm is not based in Euclidean geometry and displays better performance than other non-Euclidean algorithms.

The k-means (partitioning) procedure chosen compared the squared Euclidean distance between data points and the cluster centroids, and recomputed cluster centroids after each object was reallocated. Using this technique as a refining method allowed entities to be moved into different, more suitable clusters if necessary. However, because it used squared Euclidean distance, the results of the elbow, Katz hand and wrist/hand analyses by the group average algorithm may have been altered purely because of the changed distance measure, rather than to improve the hierarchical model. Thus any changes made in the k-means procedure for these analyses were identified and investigated.

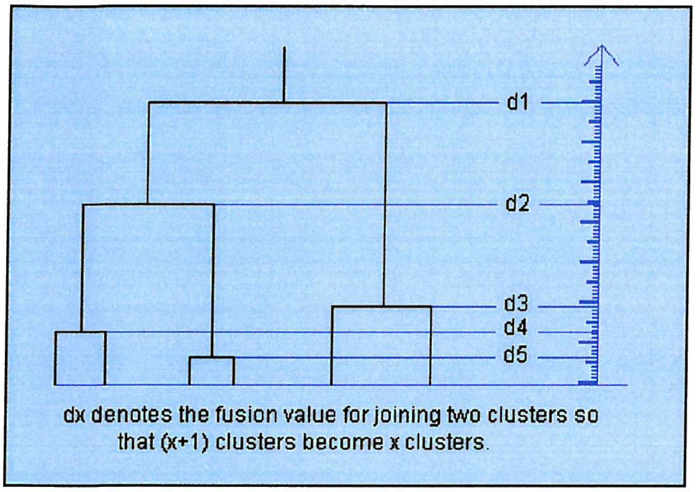
3.6 Determining the number of clusters

Until now, the number of clusters indicated by a clustering method has not been addressed, and clustering procedures themselves give virtually no information as to the number of clusters present in the data. Hierarchical methods simply give the complete range of clustering solutions, from n clusters to 1. Partition methods usually require specification of the number of clusters at the outset. This issue has been the focus of much criticism of clustering techniques in past years [87,109,110].

Numerous procedures for the determination of the number of clusters have been proposed, and when applied to hierarchical methods, are known as stopping rules (many are applicable to partition methods as well). Stopping rules that are independent of human judgement involve numerical calculations, the majority of which compare within, between or total cluster distances (or a combination) for successive numbers of clusters. The remaining rules consider the fusion values at successive levels in the hierarchy or use a variety of distributional assumptions, some of which have proved to be erroneous. (Note that fusion values are simply the distance between the pairs of clusters that are being joined together at a particular level of the hierarchy, and are calculated using the methods described in the previous section. They are visually represented by the vertical heights of clusters in a dendrogram.)

Milligan [109] tested the performance of 30 stopping rules on 432 artificial datasets with 2, 3, 4 or 5 distinct non-overlapping clusters in a simulation study. His findings were rather discouraging in that only the best two rules (both based on within and between cluster distances) could correctly identify the number of clusters in at least 90% of the datasets, and both of these are computationally intensive and unavailable on standard statistical software. The simplest stopping rule to compute, the stepsize criterion, examines the difference in fusion levels between successive hierarchy levels (Figure 11). It suggests that a large difference indicates that the data were overclustered before the last merger, and the optimal number of clusters is after that merger. This rule correctly identified the number of clusters in 63% of the datasets in Millgan's study, but tended to underestimate the number of clusters in the remaining datasets. Half of the stopping rules investigated identified the correct number of clusters in less than half of the datasets. It would therefore seem that independent stopping rules alone may not give a satisfactory answer to this problem.

Figure 11: Fusion levels presented on the dendrogram



It should be noted at this point that an important difference exists between investigating the correct number of clusters in a theoretical context and that in an exploratory applied context. In the former context, a correct number of clusters exists to be recovered by various cluster strategies since the clusters have been defined and created in a known way. In the latter context, however, no such underlying structure in known. In fact, whether clusters exist at all (in a mathematical sense) in the data becomes an irrelevant question since it is the nature of the clusters that is of interest, and it is the characterisation of the clusters that attributes them worth.

Thus the applied researcher may be more interested in finding the number of meaningful clusters, or clusters which show relevant 'important' differences between groups, rather than those that are mathematically important. Characterisation of the clusters, therefore, may be more helpful in deciding the number of clusters than the mathematical approach alone. Such an approach is clearly open to criticism, especially in the light of the fact that clustering methods will *always* produce clusters, whether or not they exist in a mathematically-defined way in the data, and 'human ingenuity is quite capable of providing a *post hoc* justification of dubious classifications' [100]. It would seem that for applied research, a combination of the mathematical and interpretative approach might be the safest and most informative tactic.

Consequently, it was decided to follow a pragmatic approach for the study analyses: a small number of clusters (up to ten), the further fusion of which occurred only with large differences in fusion values, would be initially investigated from the hierarchical cluster analysis. (In this situation 'large' is relative to the range of fusion values observed throughout the hierarchical process.) This process is essentially a modified stepsize criterion. This number of clusters was carried forward to the k-means analysis. If the characterisation of the resulting clusters indicated that the groups were either too heterogeneous to represent meaningful groups, or too homogeneous to represent important differences, the number of clusters was adjusted upwards or downwards respectively. Any changes in the number of clusters again took account of the fusion values at each level of the hierarchical clustering process. Thus, a combination of the mathematically simple stepsize criterion and a perhaps more informative approach for applied research was employed to determine the number of clusters for each of the cluster analyses.

3.7 Interpretation, testing and replicating

Once a cluster solution has been obtained, its use must be assessed. Interpretation of the clusters in the context of the field of research is the first stage of this process. Characterisation of the main features of each cluster is often presented effectively by graphical representation of key variables used in the cluster analysis for the different groups found [61,111,112].

Mathematical testing of the cluster solution (and of the methods used to obtain the solution) can be approached in a number of ways and should be the next stage in validating the cluster analysis. Those techniques that compare the cluster solution against the known structure of the data [113] cannot be used in exploratory cluster analysis since the data structure is unknown (and is to be investigated) at the outset.

Two further approaches to mathematical testing of the cluster solution involve:

a) replication of the cluster analysis in a second data sample coming from the same population as the original data sample, and
b) comparison of the cluster analysis in the same data using alternative methods, algorithms or distance measures.

Satisfactory results from such testing would suggest generalisability of the results, and stability of the cluster solution. It should be noted that these two considerations do not necessarily imply accuracy of the cluster solution in finding the true underlying group structure of the data. However, for exploratory analyses, and in conjunction with external validation of the results in context, these conditions will strongly indicate that a true representation of the group structure within the data has been found.

Replication of the cluster analysis in a second data sample is performed thus:

1) The cluster analysis in the original data sample A is performed, leading to the data partition A.
2) The cluster analysis is repeated in the second data sample B by the same method, leading to data partition B.
3) Data points in sample B are reclassified according to their resemblance to the clusters in partition A, leading to partition A*.
4) The partitions B and A* are compared.

Step 3 can be made in a number of ways [114]. Objects in sample B may be assigned to the nearest centroid of a cluster produced by partition A (nearest-centroid technique [115]), or to the cluster where their nearest data point is found (nearest neighbour technique). These two methods each have intuitively desirable properties: that each object will be assigned to the cluster which, as a whole, it is most similar to in the former case; and that objects which are most similar or even exactly the same will be assigned to the same cluster in the latter case. The latter method gave superior performance when tested in structured datasets.

A variety of indices for the comparison of two partitions of the same data (step 4) have been proposed. The Rand index, corrected Rand index, and Jaccard statistic have been used widely, and are measures of the agreement between two partitions [92,113]. The Rand index and Jaccard statistic can be calculated from the values given in Table 23. Frequency counts for each cell are computed from the $n(n-1)/2$ pairs of points in the dataset. Each pair of points can be assigned:

i) to the same cluster in both partition A* and partition B (value a in Table 23)

ii) to different clusters in both partition A* and partition B (value d in Table 23)

iii) to the same cluster in partition A* but to different clusters in partition B (value b in Table 23), or

iv) to different clusters in partition A* but to the same cluster in partition B (value c in Table 23)

The Rand index is computed as $(a+d)/[n(n-1)/2]$ (akin to the simple matching coefficient) and measures the proportion of agreement between the two partitions. The Jaccard statistic is computed as $a/(a+b+c)$, and was proposed because it was felt that d cannot be included as either agreement or disagreement of the partitions. The corrected Rand index takes account of the number of agreements that might be expected by chance and is described in detail by Hubert [92]. If the characteristics of the clusters created by Partitions B and A* are similar, a simple kappa statistic can be used to test the agreement between the two classifications rather than examining the concordance between pairs of observations.

Table 23: Classification of pairs of data points for two data partitions

		Partition B Pair in the same cluster	 Pair in different clusters	
Partition A*	Pair in the same cluster	*a*	*b*	*a+b*
	Pair in different clusters	*c*	*d*	*c+d*
		a+c	*b+d*	*a+b+c+d = n(n-1)/2*

Comparison of cluster analyses of the same data performed by different techniques can be also assessed by the Rand index or a similar statistic. Again, this may be helpful in assessing stability of the cluster solution, but can be also used to compare the performance of alternative clustering methods.

Sokal and Rohlf [98] describe a method for the comparison of dendrograms (i.e. the whole hierarchical cluster map) by transforming the information from the two comparison dendrograms into two data matrices, and performing the Pearson product-moment correlation coefficient between corresponding values (known as cophenetic correlations). The matrices are generated by replacing the distance between two data points *i* and *j* in a triangular distance matrix with the fusion value at which points *i* and *j* are placed in the same cluster in the hierarchical clustering process. A clear and unusual example of this technique is given by Lapointe and Legendre [116] in the classification of pure malt Scotch whiskies.

The final stage in assessing a cluster solution is made by returning to the data and examining the solution in the light of other relevant information not used in the cluster analysis. Analysis of variance (ANOVA) techniques, logistic regression or summary statistics can all be used to validate the cluster solution with other data available for the objects under study [111,112,117]. For example, a cluster solution identifying different diagnostic categories could be evaluated by comparing associated risk factors or prognosis between the clusters. It is this final stage which may tell the researcher the most about the applied use of a specific classification system.

The robustness of the clusters found in the analyses for this thesis was firstly tested by using different clustering techniques (hierarchical and k-means) and different similarity measures and algorithms. No formal comparisons were made between the different data partitions obtained, but any differences were reported descriptively.

Secondly, the analysis was replicated at each location by using the data from the Hill Lane practice examinations separately from the Bitterne practice examinations. The Hill Lane examinations were initially clustered, followed by the Bitterne examinations. The Bitterne examinations were then re-classified according their resemblance to the Hill Lane clusters, using the nearest-centroid technique (the available method in ClustanGraphics software). The different cluster solutions for the Bitterne examinations were compared using the Rand index, or Kappa statistic if appropriate. If these cluster solutions were in enough agreement to suggest that some robust clusters had been identified (even if not all of the clusters were identified in both cluster solutions), an overall cluster analysis was performed on all the available data.

Characteristics of the clusters formed from all the available data were presented graphically and described. These clusters were then compared with the medically driven diagnoses described in Section 1.2.2. Validation of both the clusters and the diagnoses was made by investigating associated self-reported disability, healthcare utilisation, psychological variables and putative risk factors, and will be described in detail in Chapter 7.

CHAPTER 4: RESULTS I

REPEATABILITY OF THE SEP IN THE COMMUNITY

4 Results I: Repeatability of the Southampton Examination Proforma in the Community

The repeatability of the Southampton examination proforma (SEP) was studied in the community [118]. Repeatability of the physical examination between the trained research nurses involved in the study and the study rheumatologist (who had trained the nurses) was investigated in 97 study participants, as described in Section 2.2.4. Repeatability between the two research nurses who carried out over 79% of the examinations during the study was also investigated. Each nurse examined eighteen study participants attending the physical examination.

Table 24 summarises the between-observer repeatability of physical signs included in the examination proforma on both left and right limbs between the research nurses and the rheumatologist. Cohen's kappa statistic (κ) indicates the measure of agreement observed above that expected by chance, and κ^a indicates a measure of agreement when both left and right limbs are considered together for each subject, as previously detailed (see Section 1.2.2).

Signs were observed with less agreement (lower kappa values) in the community compared to the observations on hospital clinic patients, and some individual signs were seen to be under-reported by the research nurses: a painful arc, Dupuytren's contracture and Heberden's nodes. Others were over-reported: shoulder tenderness, AC joint stress test, forearm, dorsal or palmar wrist tenderness, thumb base tenderness, swelling in the forearm/wrist, Phalen's test and pain on resisted thumb extension. As seen in the hospital repeatability study, shoulder and elbow signs were generally in better agreement between observers than the wrist/hand signs. The hand examination showed poor agreement for wrist tenderness, pain elicited on resisted wrist movement, and abnormal sensation in the index or little finger. As in the hospital study, κ^a was very similar to κ for all physical signs.

Table 24: The between-observer repeatability of physical signs according to the SEP in the community

Either nurse vs. rheumatologist

Signs	N	Nurse/Rheumatologist				$\kappa^{\dagger}$	κ^{a}
Shoulder		-/-	-/+	+/-	+/+		
Any shoulder tenderness	194	161	5	15	13	0.51	0.49
Shoulder pain on resisted elbow flexion	194	190	0	2	2	0.66	0.61
Shoulder pain on resisted forearm supination	194	189	3	1	1	0.32	0.44
Shoulder pain on resisted external rotation	194	184	4	3	3	0.44	0.46
Shoulder pain on resisted internal rotation	194	186	5	3	0	-0.02	-0.03
Shoulder pain on resisted abduction	194	170	9	10	5	0.29	0.33
Painful arc	194	179	7	3	5	0.47	0.48
Positive AC joint stress test	194	170	6	13	5	0.30	0.28
Elbow							
Lateral elbow tenderness	194	178	4	6	6	0.52	0.53
Medial elbow tenderness	194	176	4	5	9	0.64	0.59
Posterior elbow tenderness	194	190	0	3	1	0.40	0.39
Other elbow tenderness*	194	188	0	6	0	0	-0.02
Lateral elbow pain on resisted wrist extension	194	183	4	3	4	0.52	0.51
Medial elbow pain on resisted wrist flexion	194	184	2	4	4	0.56	0.59
Swelling posterior elbow*	194	193	0	1	0	0	0
Wrist and Hand							
Radial wrist tenderness	194	184	4	5	1	0.16	0.22
Medial wrist tenderness	194	190	2	2	0	-0.01	-0.01
Forearm, dorsal or palmar wrist tenderness	194	177	3	12	2	0.18	0.11
Thumb base tenderness	194	166	3	13	12	0.56	0.52
Any swelling in the forearm/wrist	194	177	2	9	6	0.50	0.53
Pain on resisted radial wrist movement	194	183	6	5	0	-0.03	-0.03
Pain on resisted medial wrist movement	194	189	3	2	0	-0.01	-0.01
Pain on resisted finger extension	194	188	3	1	2	0.49	0.44
Pain on resisted finger flexion	194	186	1	5	2	0.39	0.38
Muscle wasting (thenar) *	194	192	0	2	0	0	0
Muscle wasting (hypothenar)	194	194	0	0	0	1.00	1.00
Dupuytren's contracture	194	159	15	0	20	0.69	0.67
Heberden's nodes	194	89	24	14	67	0.60	0.54
Abnormal light touch:							
thumb	194	181	5	3	5	0.53	0.57
index finger	194	176	9	6	3	0.25	0.30
little finger	194	178	7	6	3	0.28	0.35
Positive Phalen's test	194	165	1	12	16	0.68	0.62
Positive Tinel's test	194	182	2	7	3	0.38	0.37
Weakness of thumb abduction	194	190	0	3	1	0.40	0.46
Weakness of thumb opposition*	194	191	0	3	0	0	-0.01
Pain on resisted thumb extension	194	181	1	9	3	0.36	0.26
Positive Finkelstein's test	194	177	7	6	4	0.35	0.37

* these items will have a kappa of zero by definition

† the standard error of κ ranged from 0.00 to 0.07 for all measurements, and from 0.00 to 0.10 for κ^{a}

Table 25 summarises the repeatability of the measurement of shoulder and neck movements in the community repeatability study. Findings were similar to those in the hospital study, with up to 23% of observations being at least 20° apart between the two observers in the shoulder movements, and up to 5% being at least 20° apart in the neck movements. There was a tendency for the rheumatologist to observe higher ranges of movement compared to the nurses, although this was only slight.

Table 25: Repeatability of measurements of movement in the SEP in the community
Based on 194 shoulders/97 necks

	Median (Nurse)	Median (Rheum.)	Median difference (Nurse - Rheum.)	% pairs within 10°	% pairs within 20°	Worst disagreement (°)
Shoulder:						
Range of active movement (°):						
- abduction	135	147	-10	51	77	50
- forward flexion	142.5	148	-5	57	82	50
- extension	50	56	-6	54	82	46
- external rotation	60	66	-1.5	55	88	43
- internal rotation	110	110	0	100	100	10
Range of passive movement (°):						
- abduction	140	148	-5	57	82	50
- forward flexion	145	148	0	54	85	50
- extension	52.5	56	-2	57	87	41
- external rotation	60	66	-2.5	56	87	43
- internal rotation	110	110	0	100	100	10
Neck:						
- rotation	65	65	0	80	95	30
- lateral flexion	40	40	3.5	73	97	26
- flexion	55	50	1	72	98	45
- extension	50	50	2	74	97	26

The impact of the low repeatability of some signs in the SEP was assessed by considering the diagnoses yielded from the SEP by the rheumatologist and by the nurses. Repeatability was good or excellent ($\kappa > 0.4$) for diagnoses of shoulder capsulitis, bicipital tendonitis, rotator cuff tendonitis, lateral epicondylitis, medial epicondylitis, de Quervain's tenosynovitis, carpal tunnel syndrome, fibromyalgia and cervical spondylosis. Repeatability was poor for tenosynovitis, as well as for acromioclavicular joint dysfunction, subacromial bursitis and olecranon bursitis. Of these, tenosynovitis caused more concern because of its higher prevalence of diagnosis (by either nurse or rheumatologist).

Table 26 summarises the repeatability of the SEP between the two research nurses. Only signs where a positive finding was reported by at least one of the nurses have been presented. In general agreement between the nurses was good or excellent, with four signs showing poor agreement: medial elbow pain on resisted wrist flexion, thumb base tenderness, positive Tinel's test and positive Finkelstein's test. The repeatability of the measurements of shoulder and neck movements between the two nurses was similar to that presented in Table 25: one nurse showed a slight tendency for higher ranges of movement compared to the other, and 69% of shoulder movements were measured to within 20° of each other. Neck movements were all measured to within 25° of each other.

Table 26: The between-nurse repeatability of physical signs according to the SEP in the community

Signs	N	Nurse 1 / Nurse 2				κ†	κ[a]
Shoulder		- / -	- / +	+ / -	+ / +		
Any shoulder tenderness	36	23	4	3	6	0.50	0.38
Shoulder pain on resisted elbow flexion	36	34	0	0	2	1.00	1.00
Shoulder pain on resisted forearm supination*	36	34	0	2	0	0	0
Positive AC joint stress test	36	31	0	3	2	0.53	0.52
Elbow							
Lateral elbow tenderness*	36	34	0	2	0	0	0
Medial elbow tenderness	36	30	2	2	2	0.44	0.52
Other elbow tenderness	36	34	0	1	1	0.65	0.64
Lateral elbow pain on resisted wrist extension*	36	34	0	2	0	0	0
Medial elbow pain on resisted wrist flexion	36	33	2	1	0	-0.04	-0.03
Swelling posterior elbow	36	35	0	0	1	1.00	1.00
Wrist and Hand							
Radial wrist tenderness	36	32	1	1	2	0.64	0.51
Forearm, dorsal or palmar wrist tenderness	36	33	1	1	1	0.47	0.46
Thumb base tenderness	36	28	5	1	2	0.32	0.28
Any swelling in the forearm/wrist	36	30	3	0	3	0.63	0.67
Pain on resisted finger extension*	36	34	2	0	0	0	0
Muscle wasting (thenar)*	36	35	1	0	0	0	0
Heberden's nodes	36	29	0	1	6	0.91	0.88
Abnormal light touch – little finger	36	33	1	0	2	0.79	0.78
Positive Phalen's test	36	27	1	5	3	0.41	0.37
Positive Tinel's test	36	31	2	2	1	0.27	0.21
Weakness of thumb abduction*	36	34	2	0	0	0	0
Weakness of thumb opposition	36	34	0	1	1	0.65	0.54
Pain on resisted thumb extension	36	32	1	1	2	0.64	0.51
Positive Finkelstein's test	36	32	0	3	1	0.37	0.33

* these items will have a kappa of zero by definition

† the standard error of κ ranged from 0.00 to 0.17 for all measurements, and from 0.00 to 0.24 for κ[a]

4.1 Conclusions

The repeatability of the SEP, when tested in a community setting produced mixed findings. A lower level of repeatability compared to that seen in the hospital setting was evident, and this may reflect a higher prevalence of indistinct conditions and ambiguous signs in the community compared to those referred to the hospital by their general practitioner. Repeatability of shoulder and elbow signs was generally better than that for wrist/hand signs, as was previously observed in the hospital repeatability study. The diagnosis of tenosynovitis gave the most cause for concern, having poor repeatability ($\kappa \leq 0.4$) between the nurses and rheumatologist.

The repeatability between the nurses and rheumatologist was assessed in subjects recruited from the community survey attending the second stage of the study. Appointments were made at the subjects' convenience, and the days on which repeatability was assessed were determined purely by the rheumatologist's schedule. Both nurse and rheumatologist had the original survey questionnaire and the nurse interview (conducted by one observer only) with them when they conducted the physical examination, and both therefore knew where pain had been reported by the subject. Thus, this study accurately reflected the procedure used in the second stage of the main study, and the subjects attending it. The order in which the nurse or rheumatologist examined subjects was determined randomly so that systematic bias due to order effects was minimised.

The between-nurse repeatability assessment was conducted in a similar fashion to the nurse/rheumatologist one. However, the study sample was small, and since many signs are rare, it was inevitable that some of them would not be present in any subjects in this study. In total 13 signs could not be accurately assessed (other than to note that both nurses agreed on their absence) because they were reported by neither nurse.

Little has been published on the repeatability of individual physical signs such as those used in the SEP. A small study of twelve subjects reported a similar level of repeatability for Phalen's test (κ=0.65) and a higher level of repeatability for Tinel's test (κ=0.78) [119]. The study was conducted in subjects with suspected carpal tunnel syndrome, and a higher reported level of repeatability might therefore be expected.

Interobserver differences in ranges of movement at the shoulder were similar to those reported by Croft *et al.* [72], and the repeatability of shoulder diagnoses has been reported at low levels (κ=-0.03 – 0.48 [120]) and at very high levels (κ=0.875)[121] in subjects complaining of shoulder pain. The latter study, however, comprised only 21 subjects.

CHAPTER 5: RESULTS II

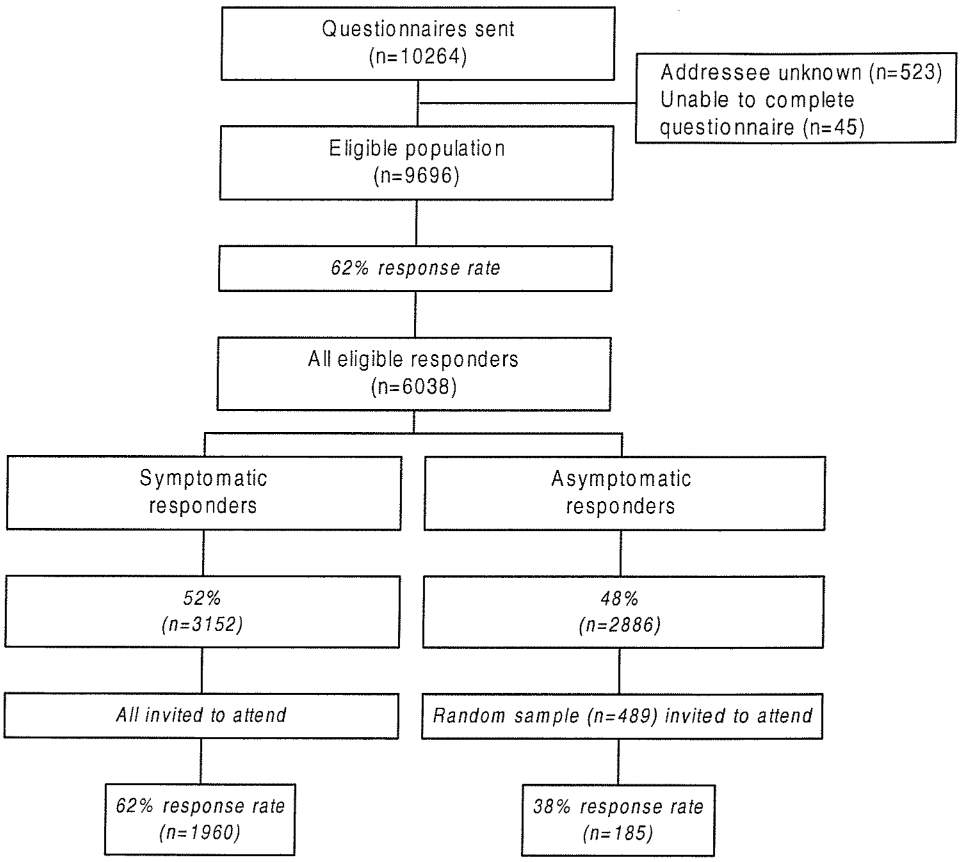
SUBJECTS IN THE STUDY AND REPORTED SYMPTOMS

5 Results II: Subjects in the Study and Reported Symptoms

5.1 Subjects in the main study

The initial response rate to the survey questionnaire was as expected (62%, compared to the expected 60% in Section 2.1, Figure 12). Just over half of the responders (52%) reported pain, numbness or tingling occurring in the previous seven days. This prevalence was the same to within 2% in the first time responders (75% of all responders) compared to those who only returned their questionnaire after a reminder had been sent to them. This prevalence was clearly far higher than was originally anticipated.

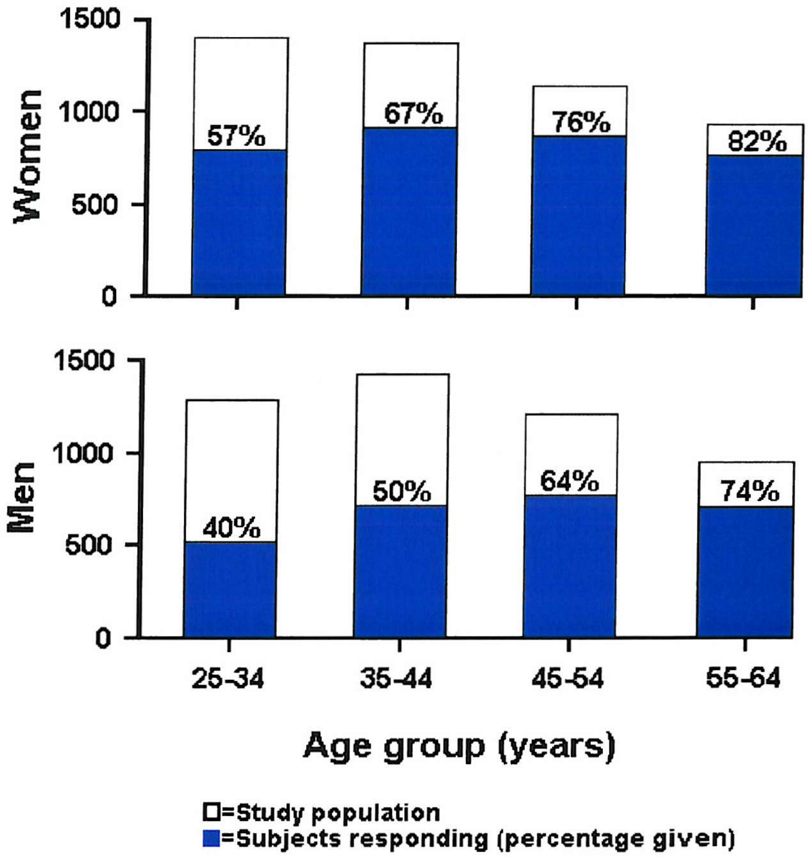
Figure 12: Subjects contacted in the Study



A higher proportion of the symptom positive responders was prepared to attend the physical examination and clinical interview than had been expected (62% compared to the projected 50%), but a much smaller proportion of the asymptomatic responders was prepared to participate in this part of the study (185 subjects of 489 approached, 38%).

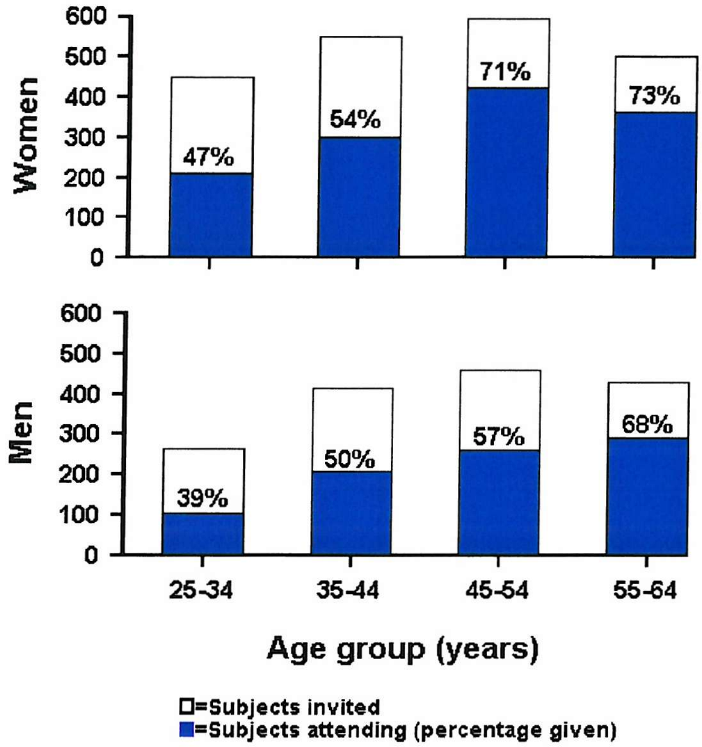
Figure 13 summarises the age and sex distribution of the study population and the questionnaire responders. Older subjects were much more likely to respond to the questionnaire than younger ones, as were women. Most responders (95%) were European, reflecting the catchment area of the general practices. The remaining 5% were predominantly Indian.

Figure 13: Age distribution of study population and responders, by sex



Similarly, men in the youngest age group were under-represented in those subjects attending the second stage of the study (the clinical interview and physical examination), whilst the older two age bands were over-represented in both the sexes compared to the demography of the second stage invitees (Figure 14).

Figure 14: Age distribution of subjects invited and attending the clinical interview and physical examination, by sex



Symptomatic subjects attended the second stage of the study between 1 and 384 days after the receipt of their questionnaire. Of these subjects, 81% were seen within two months of receiving their questionnaire, and 97% were seen within four months. There was a longer delay in seeing asymptomatic subjects after receiving their survey questionnaire because in the Hill Lane practice these subjects were seen after those reporting symptoms. Half of these subjects were seen between 4 months and one year after returning the survey questionnaire.

5.2 <u>Symptoms reported at the survey questionnaire phase</u>

Symptoms were reported more often by women than men, and by older subjects (Table 27). The prevalences of elbow pain and numbness and/or tingling were lower at younger ages in both men and women; all symptoms showed some increase in prevalence with age. Women had a higher prevalence of each symptom, with the exception of elbow pain, which was reported equally by men and women. Numbness and/or tingling were the most commonly experienced symptoms amongst responders (27%), followed by neck, shoulder and wrist/hand pain. The pain was bilateral in 35%

of those with shoulder pain, 28% of those with elbow pain, and 42% of those with wrist/hand pain.

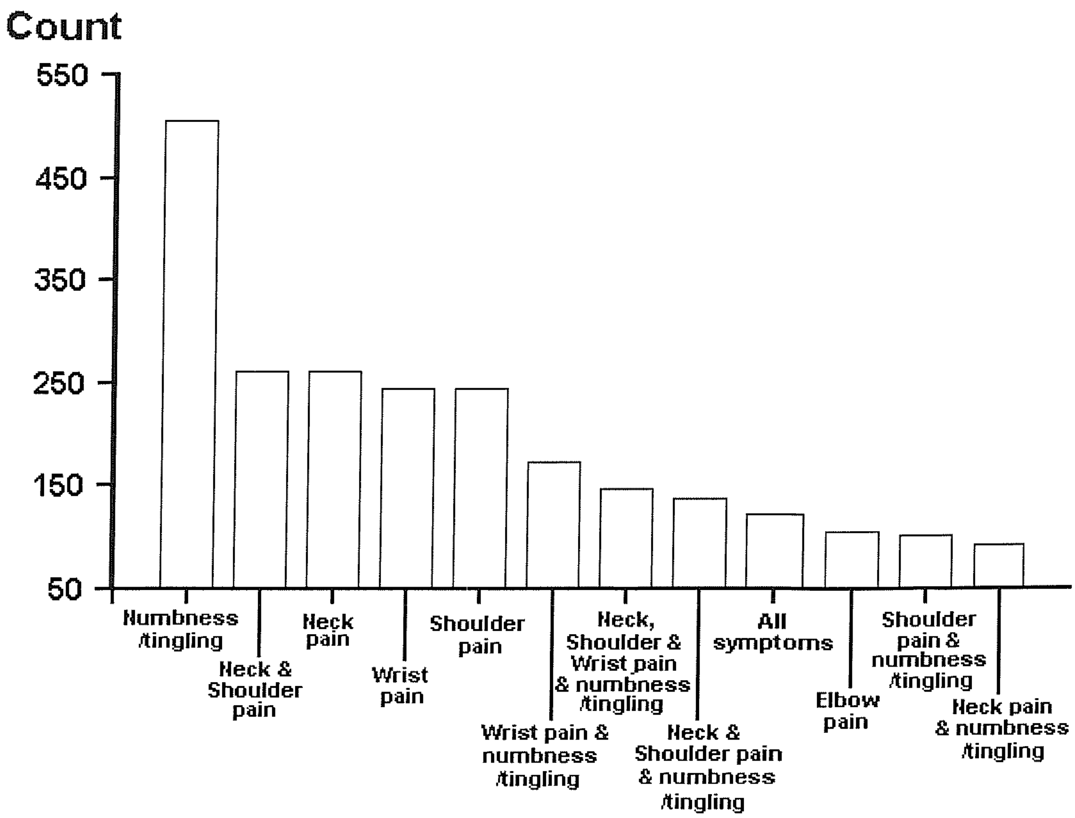
Table 27: Prevalence of symptoms lasting a day or longer in the past week by age and sex

	N	Pain N (%) Neck	Shoulder(s)	Elbow(s)	Wrist(s)/ Hand(s)	Numbness/ Tingling N (%) Arms/ Hands/ Fingers
MEN						
25 – 34	514	84 (16)	76 (15)	21 (4)	83 (16)	105 (20)
35 – 44	716	152 (21)	143 (20)	88 (12)	124 (17)	170 (24)
45 – 54	766	169 (22)	163 (21)	111 (14)	121 (16)	199 (24)
55 – 64	700	156 (22)	187 (27)	90 (13)	172 (25)	202 (29)
WOMEN						
25 – 34	796	165 (21)	176 (22)	28 (4)	121 (15)	165 (21)
35 – 44	910	220 (24)	215 (24)	73 (8)	155 (17)	254 (28)
45 – 54	870	270 (31)	252 (29)	130 (15)	242 (28)	287 (33)
55 – 64	766	219 (29)	235 (31)	95 (12)	250 (33)	242 (32)
Total	6038	1435 (24)	1447 (24)	636 (11)	1268 (21)	1624 (27)

Nurse checks for the anatomical location of reported symptoms in the questionnaire suggested that there was some discrepancy between responders' definitions of the shoulder region, and that used in this study. Shoulder pain reported at the questionnaire was reclassified as no pain in 12% of cases because the pain reported was outside the neck or upper limb area, and a further 2% of shoulder pain was reclassified as neck pain. Elbow and wrist/hand pain and numbness/tingling were reported more accurately at the questionnaire, with less than 2% of cases being reclassified for each of these symptoms following the nurse interview.

Numbness/tingling alone was the most common pattern, reported by 16% of symptom-positive responders, whilst pain at a single site (neck, shoulder, elbow or wrist/hand) accounted for a further 27% of symptom-positive subjects (Figure 15). The most common co-existing symptoms were neck and shoulder pain (260 subjects, 8%), wrist/hand pain and numbness/tingling (172 subjects, 5%) and neck, shoulder and wrist pain with numbness/tingling (146 subjects, 5%). Three or more of neck, shoulder, elbow, wrist/hand pain and/or numbness/tingling were reported by 917 subjects (29%) and 121 subjects reported pain at all four sites as well as numbness/tingling.

Figure 15: Patterns of symptoms reported in the questionnaire



Difficulty with everyday activities caused by pain was experienced at each of the four sites in similar proportions. Difficulties were reported more frequently by women and in older age groups (Table 28). No questions regarding disability associated with numbness or tingling were asked at the survey questionnaire.

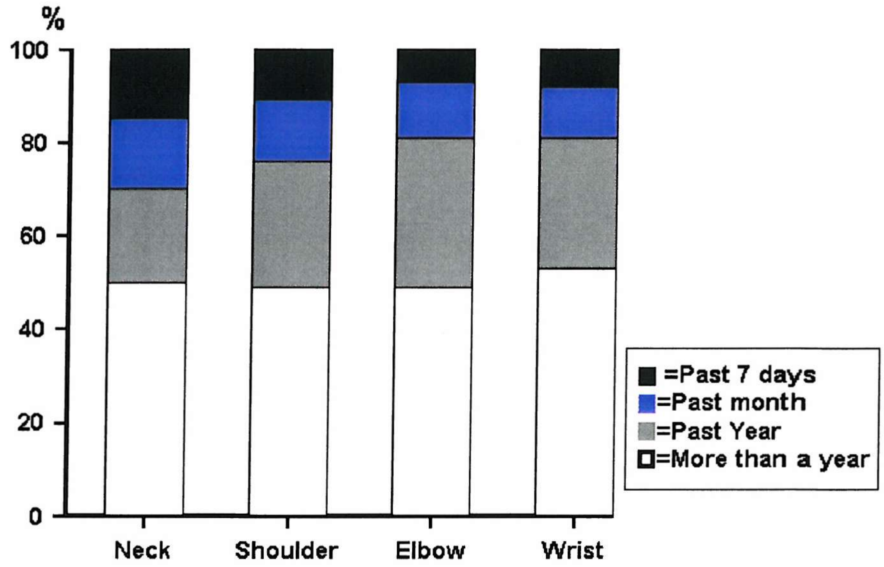
Table 28: Proportion of pain causing difficulty with everyday activities

		Pain causing difficulty with everyday activities N (%*)			
	N	Neck	Shoulder(s)	Elbow(s)	Wrist(s)/ Hand(s)
MEN					
25 – 34	514	49 (58)	34 (45)	10 (48)	43 (52)
35 – 44	716	90 (59)	62 (43)	53 (60)	76 (61)
45 – 54	766	120 (71)	90 (55)	75 (68)	79 (65)
55 – 64	700	111 (71)	131 (70)	66 (73)	127 (74)
WOMEN					
25 – 34	796	117 (71)	90 (51)	16 (57)	85 (70)
35 – 44	910	159 (72)	125 (58)	42 (58)	106 (68)
45 – 54	870	199 (74)	138 (55)	96 (74)	173 (71)
55 – 64	766	160 (73)	152 (65)	69 (73)	192 (77)
Total	6038	1005 (70)	822 (57)	427 (67)	881 (69)

*Percentages are of all pain, whether or not it caused difficulties (refer to Table 27)

For each location, between 49% and 53% of subjects reported their pain as starting over a year previously, whilst only 7% – 15% reported that the pain had started in the past week (Figure 16). There were no differences in the duration of reported pain between men and women, and a slight tendency for the longer duration of pain to be reported in the older age groups.

Figure 16: Duration of pain reported at the survey questionnaire



Symptoms reported at the clinical interview and physical examination

Only 526 (25%) of the 2145 subjects who attended the clinical interview and physical examination had symptoms in precisely the same place (and side) as previously reported at the survey questionnaire. Just under a third of these (30%) were subjects who had reported no symptoms at the questionnaire (and remained asymptomatic), whilst 46% were subjects with a single site of pain or numbness/tingling alone. The remaining 22% had multiple sites of symptoms, although no particular pattern of symptoms was predominant.

There was a slight tendency for subjects reporting pain of longer duration at the questionnaire to continue reporting that pain at their clinical interview (Table 29). However, the time lag between filling in the survey questionnaire and attending the nurse appointment did not appear to affect whether subjects reported the same symptoms at the clinical interview as at their questionnaire. Some subjects reported the same positive symptoms after a time lag of over 6 months, whilst others did not report the same symptoms even two weeks later. The median time lag was higher in those who reported the same symptoms than in those who did not (44 days versus

35 days). The tendency to report the same symptoms at the questionnaire as at the nurse appointment was also unaffected by age or sex.

Table 29: Duration of reported pain at the survey questionnaire compared with pain at the clinical interview

Pain	Pain duration (reported at questionnaire)	Pain at questionnaire and clinical interview (%)	Pain at questionnaire only (%)
Neck	In the past week	9	18
	Longer than a week	8	17
	Longer than a month	21	19
	Longer than a year	62	46
Shoulder	In the past week	7	11
	Longer than a week	9	14
	Longer than a month	25	28
	Longer than a year	60	47
Elbow	In the past week	6	6
	Longer than a week	9	15
	Longer than a month	33	34
	Longer than a year	52	45
Wrist	In the past week	6	8
	Longer than a week	8	12
	Longer than a month	25	29
	Longer than a year	61	51

Considering individual sites of symptoms, numbness/tingling was the most transient with 30% of all subjects (N=2145) reporting resolved numbness/tingling by the second stage of the study and 2% of subjects presenting with new numbness/tingling. Thus 68% of subjects reported numbness/tingling in the same way at both the survey questionnaire and at the clinical interview. Neck, shoulder and wrist symptoms were each reported unchanged at the clinical interview from the questionnaire in 73%, 77% and 71% of subjects respectively, and elbow pain was reported identically at the two stages of the study in 86% of subjects. The majority of changes in the reporting of symptoms were due to resolution of symptoms, rather than because new symptoms had occurred. Overall, 31% of the previously symptom-positive subjects reported no symptoms at the nurse interview.

5.3 Summary

In total 9696 men and women of working age were contacted with the survey questionnaire, of whom 62% responded. Young men were found to be the least likely group to respond to the initial questionnaire, a finding similar to that of other authors carrying out musculoskeletal pain surveys using general practice registers as their study population [122]. A total of 2145 physical examinations and nurse interviews were performed, 1960 amongst symptom-positive responders at the questionnaire, and 185 amongst asymptomatic responders. Again, young men were the least likely to participate in this stage of the study.

The prevalence of reported pain, numbness or tingling was 52%, and this was the same in those subjects who responded immediately and in those who responded following a reminder letter.

Elbow pain was reported half as often as all other symptoms (neck, shoulder and wrist pain, numbness/tingling) which were all reported by around 24% of responders at baseline. Over a quarter of responders reported at least three of the five possible symptoms. Prevalences of reported pain amongst responders were similar to those reported in other studies [7-9,21] although the prevalence of neck pain was at the higher end of reported prevalences [6]. Prevalences of all symptoms were higher in women (except for elbow pain, which was reported equally by men and women) and with older age, as has been previously reported [1]. Reported pain was of over a year's duration in a significant proportion of those reporting pain (upwards of 49%), whilst pain which had started in the previous week accounted for less than 11% of all pain reported. However, a comparison of baseline reported symptoms with symptoms reported at the second stage of the study demonstrated the transience of these conditions: 32% of subjects reported numbness or tingling differently at the time of the nurse interview compared to baseline, as did 14% – 29% of subjects regarding each of their neck, shoulder, elbow and wrist/hand pain. The majority of these differences were due to resolution of symptoms, and just under a third of the baseline symptom-positive subjects were completely asymptomatic at the nurse interview.

CHAPTER 6: RESULTS III

CLUSTER ANALYSIS OF SIGNS AND SYMPTOMS RECORDED AT THE SEP

6 Results III: Cluster analysis of Signs and Symptoms Recorded at the SEP

2145 subjects underwent the nurse interview and physical examination, and data from the latter were analysed using cluster analysis techniques. Each location on the SEP was examined in turn. Cluster analyses were initially performed on data from the Hill Lane practice only, and replication of the clusters was tested in data generated from the Bitterne general practice. A final cluster solution based on all of the available data was investigated if at least some of the clusters found in the Hill Lane and Bitterne analyses appeared to be robust.

6.1 The neck

Physical examination at the neck consisted of a question about pain lasting a day or longer in the past seven days, and ranges of movement in six directions (Appendix I). Of the 2145 subjects examined, four had missing neck examination data and were excluded from the analysis. Two of these had neck extension range of movement missing, one was wheelchair-bound and had no neck movements measured, and one had three missing neck movements because they were too painful to perform.

700 subjects with complete neck examination data were recruited from the Hill Lane practice, and data from these subjects were analysed first.

The seven variables from the neck examination were standardised and the squared Euclidean distance was calculated for every pair of subjects, as proposed in Sections 3.3 and 3.4. Ward's hierarchical method of cluster analysis was performed (Section 3.5), yielding the dendrogram presented in Figure 17a.

Two main clusters were identified (represented by the two long vertical lines in Figure 17a) and each of these was split into two clusters (the four short vertical lines). Note that the order from right to left of the clusters is arbitrary. No discernible sub-clusters were evident beyond these four. The comparatively large step up in fusion value from four clusters to three seen in Figure 17b confirmed that the corresponding fusion (the green and blue clusters in Figure 17a) should not be made, whilst the comparatively flat gradient from five clusters to four (corresponding in this case to the splitting into two of the yellow cluster in Figure 17a) suggested that no subdivision of the four clusters was necessary (as described in Section 3.6).

Figure 17a: Dendrogram of Ward's hierarchical clustering on Hill Lane data neck examination observations

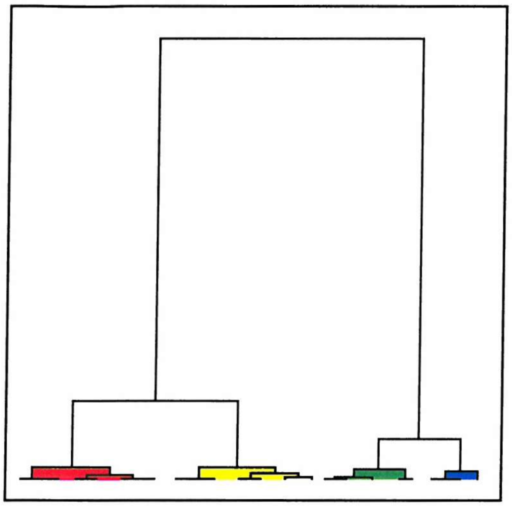
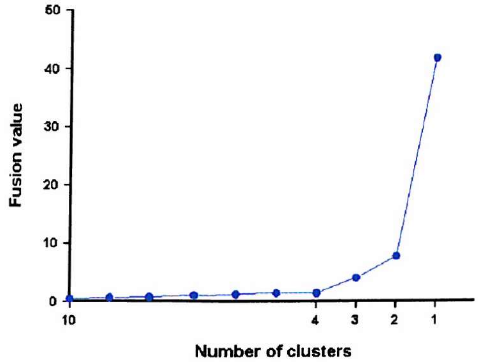


Figure 17b: Fusion values for successive neck cluster joins



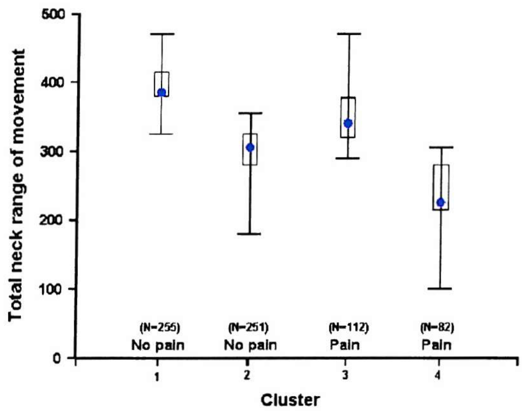
The k-means procedure was employed to refine these four clusters, and 59 (8%) subjects moved into a different cluster. These moves were all between Clusters 1 and 2, coloured red and yellow respectively (20 moves), or between Clusters 3 and 4, coloured green and blue respectively (39 moves).

6.1.1 Characterisation of the neck clusters (Hill Lane data only)

Clusters 1 and 2 contained all and only subjects with no neck pain at the examination. Clusters 3 and 4 contained the subjects reporting neck pain. As the dendrogram in Figure 17a indicates, neck pain was the dominant distinguishing factor between the two overriding clusters.

Figure 18 shows a box-and-whisker plot of the total sum of all six ranges of neck movement for the four clusters. Median values are represented by the blue dot, interquartile ranges by the boxes, and maximum and minimum values by the outer horizontal lines. Clusters 2 and 4 had lower overall movement compared to Clusters 1 and 3, although there was some overlap: 117 subjects in Clusters 1 and 2 had overall neck movement in the range 325° to 355°, and 22 subjects in Clusters 3 and 4 had overall neck movement in the range 290° to 305°. Exploration of the individual neck movements showed that they were highly correlated, with pairwise correlations of 0.367 to 0.782. There was a lot more overlap of ranges of individual neck movements between the clusters, suggesting that it was the overall restriction in neck movement that distinguished subjects in Cluster 1 from those in Cluster 2, and those in Cluster 3 from those in Cluster 4. The subjects who were moved between clusters by the k-means procedure tended to be those with a range of neck movement in the overlapping regions between Clusters 1 and 2, and Clusters 3 and 4.

Figure 18: Box-and-whisker plot showing the sum of neck ranges of movement for the four clusters



Cluster 1 (N=255) thus represented subjects with neither symptoms nor signs at the neck;
Cluster 2 (N=251) identified subjects with no pain, but some restriction in movement;
Cluster 3 (N=112) comprised subjects with pain but little or no restriction; and
Cluster 4, the smallest group (N=82), consisted of subjects with the most severe neck conditions of both pain and restricted movement.

6.1.2 <u>Replication of the neck clusters (Bitterne data only)</u>

Neck examination data from the Bitterne practice (1441 subjects with non-missing data) were clustered using the same methods described for the Hill Lane data. The resultant dendrogram (Figures 19a and 19b, note the break in the y-axis) clearly indicated two main clusters, each divided into two smaller clusters, as seen in the Hill Lane analysis.

The k-means procedure moved 249 (17%) subjects into another cluster: 199 from Cluster 2 (yellow) to Cluster 1 (red) and 50 from Cluster 3 (green) to Cluster 4 (blue).

Clusters 1 and 2 contained subjects with no neck pain whilst Clusters 3 and 4 contained those reporting neck pain. Subjects in Clusters 1 and 3 generally had a higher range of neck movement than those in Clusters 2 and 4. The high percentage of subjects who were moved into a different cluster by the k-means procedure suggested that whilst the two overriding clusters (distinguished by the presence or absence of neck pain) were mathematically robust, the subdivision of those clusters (according to the range of neck movement) was less so.

Figure 19a: Dendrogram of Ward's hierarchical clustering on Bitterne data neck examination observations

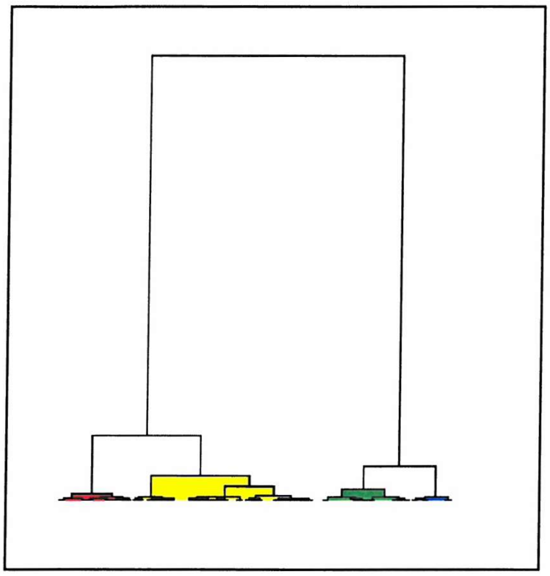
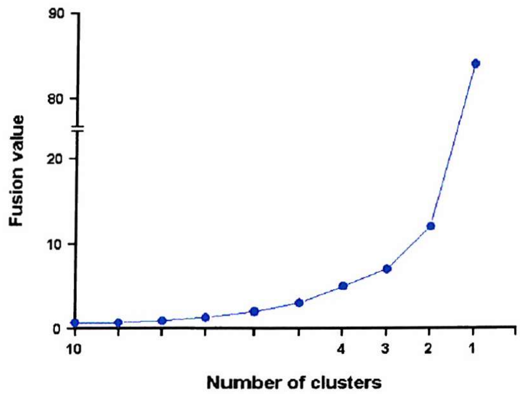


Figure 19b: Fusion values for successive neck cluster joins



The same examination data were then clustered by the k-means procedure, using the cluster centroids from the Hill Lane clusters as the seed points, as proposed in Section 3.7. (This procedure is a standard function in ClustanGraphics software.) The two cluster solutions for the Bitterne data were contrasted (Table 30) and the kappa statistic was used to formally compare their agreement, because the clusters yielded in the two solutions were directly analogous.

The observed agreement was 89% whilst the expected agreement was 31%, yielding a κ value of 0.84. This value indicates a high level of agreement, although it is clear from both general practice populations that the data suggest a continuum across the spectrum of range of neck movement rather than distinct clusters, both in those subjects with, and without neck pain.

Table 30: Comparison of the Bitterne neck clusters obtained by two different methods

	Hierarchical (Ward's method) clustering plus k-means:				
	1: Asympt-omatic	2: Restricted ROM* only	3: Pain only	4: Pain & restricted ROM*	N
K-means clustering using Hill Lane seed points:					
1: Asymptomatic	396	1	0	0	397
2: Restricted ROM* only	108	551	0	0	659
3: Pain only	0	0	175	0	175
4: Pain & restricted ROM*	0	0	53	157	210
N	504	552	228	157	1441
		κ=0.84			

*ROM = range of movement

6.1.3 Cluster analysis of the neck examination data from the whole population

The whole data were lastly clustered as one population (2141 subjects), using the methods previously employed, and the final four-cluster solution was obtained (Figures 20a and 20b, note the break in the y-axis).

Figure 20a: Dendrogram of Ward's hierarchical clustering on both Hill Lane and Bitterne data neck examination observations

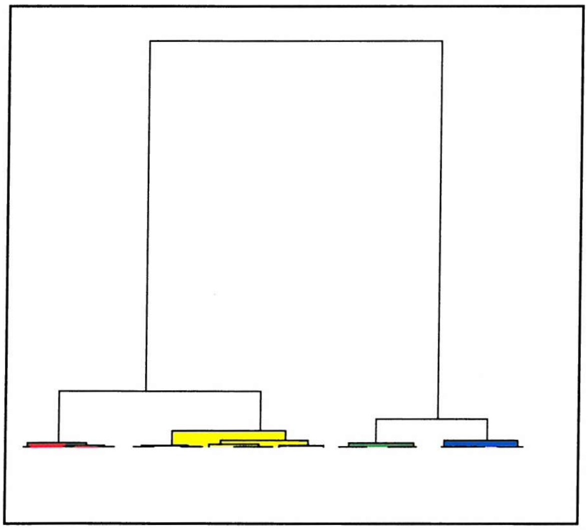
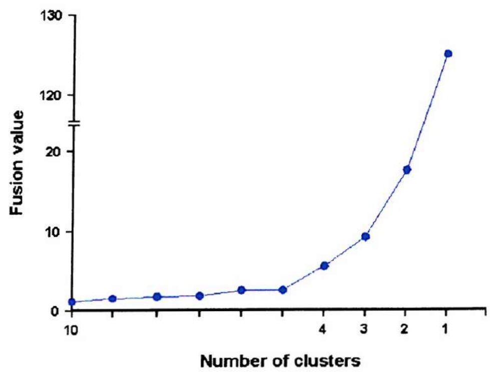


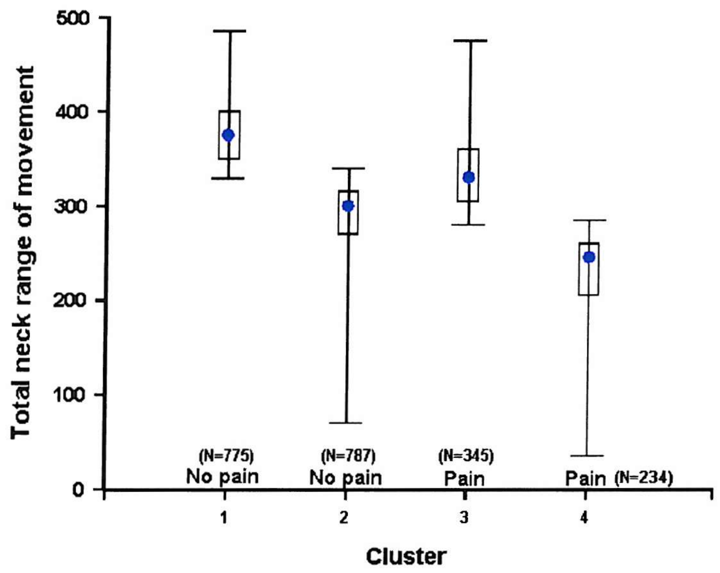
Figure 20b: Fusion values for successive neck cluster joins



The k-means procedure moved 252 in Clusters 1 and 2, and 66 subjects in Clusters 3 and 4. Again, Clusters 1 and 2 contained subjects with no neck pain, whilst Clusters 3 and 4 contained subjects reporting neck pain. Clusters 1 and 3 had higher ranges of neck movement compared to Clusters 2 and 4 (Figure 21, showing medians, interquartile ranges, maxima and minima).

The final four-cluster solution for the neck examination data was therefore:
Cluster 1 (N=775) - subjects with neither symptoms nor signs at the neck;
Cluster 2 (N=787) - subjects with no pain, but some restriction in movement;
Cluster 3 (N=345) - subjects with pain but little or no restriction;
Cluster 4 (N=234) - subjects with the most severe neck conditions of both pain and restricted movement.

Figure 21: Box-and-whisker plot showing the sum of neck ranges of movement for the finalised neck clusters



The demographic characteristics of the subjects were compared between the four neck clusters (Table 31). These variables were not included in the cluster analysis. Higher proportions of women were seen in the two pain-positive clusters, whilst those subjects in clusters characterised by restricted range of movement were predominantly aged 45 – 64 years. Clusters 3 and 4, which involved subjects with neck pain at the examination, contained the highest proportion of subjects reporting neck pain at baseline, as might be expected. High proportions of these subjects also reported difficulty carrying out daily activities, and pain lasting over a year previously. None of these three baseline characteristics clearly distinguished between those with and those without restricted neck movement.

Table 31: Demographic and baseline characteristics in the four neck clusters

Characteristics	**Cluster**			
	1: Asympto -matic	**2: Restricted ROM* only**	**3: Pain only**	**4: Pain & restricted ROM***
N	775	787	345	234
Colour in dendrogram	Red	Yellow	Green	Blue
Demographic characteristics:				
% female	59	56	70	64
% 25 – 34 yrs	23	6	20	3
% 35 – 44 yrs	31	18	29	11
% 45 – 54 yrs	27	33	31	40
% 55 – 64 yrs	19	43	20	46
Report of neck pain at baseline:				
% neck pain present	23	39	79	88
% activities difficult/impossible	13	25	54	76
% neck pain > 1 yr	7	20	44	62

*ROM = range of movement

6.1.4 Comparison of the neck clusters with the HSE neck diagnoses

The SEP was designed to diagnose cervical spondylosis based on the presence of neck pain and restricted neck movement. However, although cut points were suggested for thresholds of restricted neck movement, these were not based on a general UK population sample. The neck examination data from this study were therefore used to investigate ranges of neck movement in association with age, sex and neck pain [123]. Findings suggested that a) neck movement in all directions was inversely correlated with age; b) there was little difference between the sexes in neck movement; and c) there was a statistically significant drop in neck ranges of movement in association with neck pain, even when any age or sex effect had been accounted for, although this reduction in movement was small (approximately 7° in each direction). It was concluded from this work that any definition for restricted range of movement needed to be sufficiently strict (and thus specific, rather than sensitive) in order to maximise the chance of identifying true underlying pathology. Therefore, a range of neck movement of less than two standard deviations below the mean in any direction, within each 10-year age band and sex strata was considered to be restricted neck movement. Using this definition, cervical spondylosis was diagnosed, and compared with the neck examination clusters (Table 32).

Table 32: Prevalence of cervical spondylosis amongst the four neck clusters

Diagnosis (%)	**Cluster**			
	1: Asympto -matic	**2: Restricted ROM* only**	**3: Pain only**	**4: Pain & restricted ROM***
N	775	787	345	234
Colour in dendrogram	Red	Yellow	Green	Blue
No diagnosis	100	100	97	58
Cervical Spondylosis	0	0	3	42

*ROM = range of movement

All diagnoses of cervical spondylosis were made in necks from Clusters 3 and 4, i.e. those clusters characterised by neck pain. Nine subjects (3%) from Cluster 3 had a diagnosis of cervical spondylosis, even though this cluster was generally characterised by a high range of movement. Six of these subjects were in the youngest age group, with a total range of neck movement towards the lower end of the cluster's distribution (Figure 21). Their ranges of movement may have seemed high, but when compared to others of their age and sex, they were considered restricted. Subjects from Cluster 4 with a diagnosis of cervical spondylosis (42%) spanned the different age and sex strata, although there were more women, and more older subjects with the diagnosis.

6.2 The shoulder

Physical examination at the shoulder covered the location of reported pain, location of any tenderness, pain on resisted movements, specific diagnostic tests (painful arc) and active and passive ranges of movement in five directions (Appendix I). Of the 4290 shoulders examined (right and left on each subject), 100 had missing data and were excluded from the analysis. Of the missing examinations 28% occurred because a movement or clinical test had not been performed due to pain. The remaining 72% were due to one or two observations having been missed out by mistake.

700 subjects with complete examination data (1397 shoulders) were recruited from the Hill Lane practice and were analysed first.

The 33 variables were standardised, and squared Euclidean distances were computed. Ward's hierarchical method was performed, yielding the dendrogram presented in Figure 22a. Five main clusters were identified, one of which was clearly separate from the other four (coloured in red). As with the cluster analysis of the neck examination data, the step up in fusion values from five clusters to four (Figure 22b, note the break in the y-axis) confirmed that the corresponding fusion (the cyan and blue clusters in Figure 22a) should not be made, whilst the flat gradient from six clusters to five suggested that no further subdivision (in this case, of the orange cluster in Figure 22a) was necessary.

Figure 22a: Dendrogram of Ward's hierarchical clustering on Hill Lane data shoulder examination observations

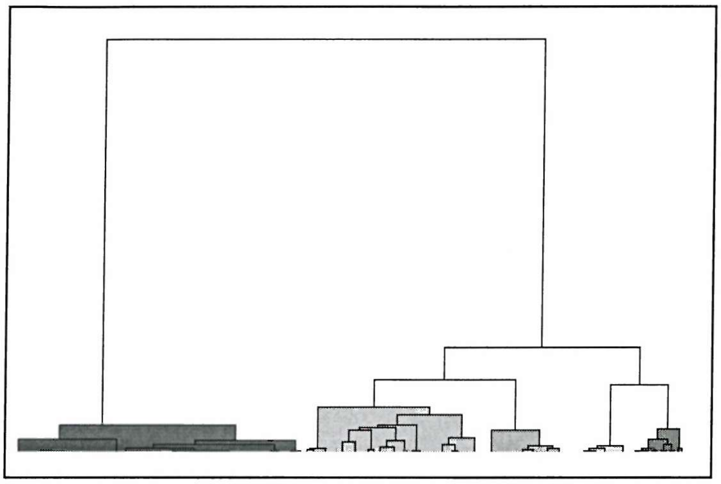
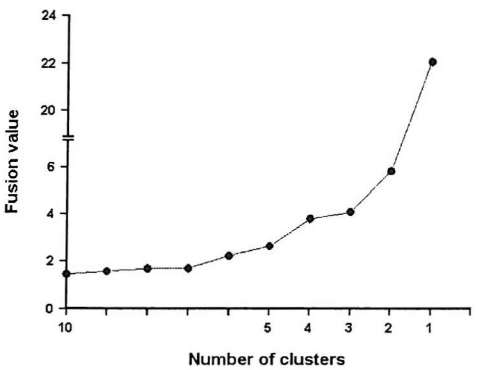


Figure 22b: Fusion values for successive shoulder cluster joins



The k-means procedure refined this five-cluster solution by moving 54 (3.8%) shoulders into other clusters; 44 of these were from Cluster 2 (the orange cluster in Figure 22a).

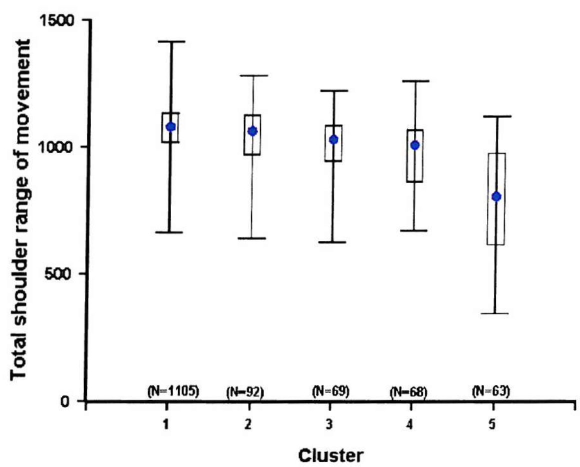
6.2.1 <u>Characterisation of the shoulder clusters (Hill Lane data only)</u>

The prevalence of symptoms and signs in each shoulder cluster was explored (Table 33), and Figure 23 shows a box-and-whisker plot of the total sum of all ten ranges of movement (active and passive movement in five directions) for the five clusters. Ranges of individual movements gave similar graphs to this one (medians, interquartile ranges and maxima and minima are plotted).

Table 33: Prevalence of symptoms and signs in the Hill Lane shoulder clusters

Characteristics	**Cluster**				
	1: 'Normal'	**2: Mixed signs**	**3: Pain only**	**4: AC signs**	**5: Severest involvement**
N	**1105**	**92**	**69**	**68**	**63**
Colour in dendrogram	Red	Orange	Green	Cyan	Blue
Any pain:	*0*	*44*	*100*	*15*	*100*
Deltoid		7	42	1	62
Anterior		2	13	0	51
AC joint		0	6	10	51
Subacromial bursa		7	6	0	22
Diffuse		10	6	0	5
Posterior		21	45	3	38
Other		0	0	0	0
Any tenderness:	*0*	*100*	*3*	*100*	*100*
Deltoid		11	3	6	22
Anterior		28	0	12	27
AC joint		0	0	87	40
Subacromial bursa		23	0	9	21
Diffuse		10	0	3	0
Posterior		36	0	1	13
Other		0	0	0	0
Other signs:	*3*	*29*	*26*	*78*	*90*
Pain on resisted:					
- Elbow flexion	0.6	8	3	10	17
- Forearm supination	0.2	4	1	1	6
- External rotation	0.7	13	12	4	19
- Internal rotation	0	8	9	0	21
- Abduction	1	12	19	16	52
AC joint stress test	0.8	5	10	75	68
Painful arc	0	0	1	1	10

Figure 23: Box-and-whisker plot showing the sum of shoulder ranges of movement for the five clusters



The first cluster, clearly separate from the others in Figure 22a, was the largest (N=1105) and contained shoulders with no pain, no tenderness, few other signs and a high range of movement.

Cluster 2 (N=92) identified a group with mixed characteristics: 44% had pain at a number of different sites, particularly at the posterior region of the shoulder, all had tenderness (predominantly at one site only) most commonly at the subacromial bursa, anterior or posterior shoulder. Just under a third of shoulders also displayed other positive signs. These shoulders displayed a lowered range of movement overall from the shoulders in Cluster 1, but this was only slight.

Cluster 3 (N=69) contained shoulders with pain (predominantly at one site) mainly in the deltoid or posterior areas. Deltoid tenderness was present in only 3% of these shoulders. A quarter of these shoulders had other positive signs, mostly pain on resisted abduction. The ranges of shoulder movement were slightly lower than those seen in either of Clusters 1 or 2.

Cluster 4 (N=68) identified shoulders mostly without pain (10% had AC pain) and with AC tenderness (87%). 75% also had a positive AC joint stress test, and 16% showed pain on resisted abduction. The ranges of movement were similar to those in Cluster 3.

Cluster 5 (N=63) comprised shoulders with pain (at multiple sites) and tenderness (predominantly at one site). Other positive signs were reported in 90% of these shoulders. The lowest ranges of shoulder movement were seen in this cluster.

6.2.2 Replication of the shoulder clusters (Bitterne data only)

Shoulder examination data from the Bitterne practice (2793 shoulders with non-missing examinations) were clustered using Ward's hierarchical cluster algorithm on squared Euclidean distances, as for the Hill Lane data. The dendrogram yielded by this analysis (Figures 24a and 24b) suggested that four clusters should be considered, rather than the five observed in the Hill Lane data.

The k-means procedure refined the cluster analysis by making 97 (3.5%) moves. Over 70% of these moved shoulders into and out of Cluster 2 (coloured yellow in Figure 24a).

Figure 24a: Dendrogram of Ward's hierarchical clustering on Bitterne data shoulder examination observations

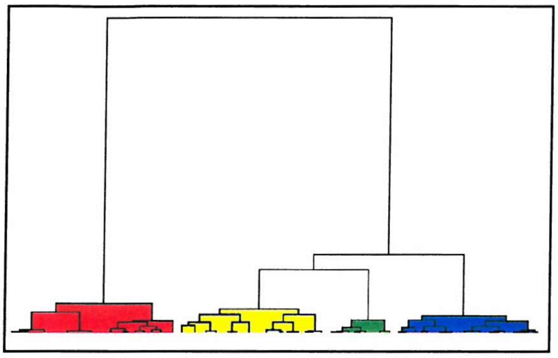
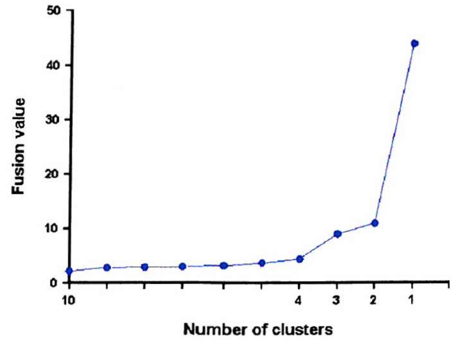


Figure 24b: Fusion values for successive shoulder cluster joins



The same examination data were then clustered by the k-means procedure, using the cluster centroids from the five Hill Lane clusters as the seed points. The two cluster solutions for the Bitterne data were compared using the kappa statistic and Rand Index, as described in Section 3.7 (Table 34).

The observed agreement between the clusters (the numbers in bold) was 94%, whilst the expected was 68%. Thus κ was 0.81, indicating excellent agreement, and replication of the cluster solution. The agreement between each pair of shoulders as to whether they were placed in the same cluster, or in different clusters was 99.0% according to the Rand index and 98.6% according to the Jaccard statistic. The Rand index measures the percentage of times that the two cluster solutions agree in their placement of pairs of shoulders, either to the same cluster, or to different clusters. The Jaccard statistic discounts the pairs of shoulders placed in different clusters by both of the cluster solutions, considering them to be evidence of neither agreement nor disagreement. The Rand index / Jaccard statistic approach to quantifying agreement (and thus replication) between cluster solutions is less stringent than the kappa statistic in that it does not require explicit one-to-one matching of the clusters in the two solutions. However, it may give a more accurate estimate of replicability when the two cluster solutions show very little resemblance, and one-to-one matching of the clusters in inappropriate. It should be noted that the Rand index and

Jaccard statistic give a percentage estimate of agreement, unlike the kappa statistic, which measures the agreement over that expected by chance.

Table 34: Comparison of the Bitterne shoulder clusters obtained by two different methods

K-means clustering using Hill Lane seed points:	**Hierarchical (Ward's method) clustering plus k-means:**

Kappa statistic

Clusters:	1	2	3	4	N
Colour in dendrogram	Red	Yellow	Green	Blue	
1: 'Normal'	**2290**	0	0	4	2294
2: Mixed signs	0	**108**	39	2	149
3: Pain only	0	0	**117**	3	120
4: AC signs	0	79	0	23	102
5: Severest involvement	0	2	14	**112**	128
N	2290	189	170	144	2793
		κ=0.81			

Rand index/Jaccard statistic

Pairs of shoulders:	In the same cluster	In different clusters	N
In the same cluster	2,643,862	17,654	2,661,516
In different clusters	19,470	1,218,042	1,237,512
N	2,663,332	1,235,696	3,899,028
	Rand index=99.0% , Jaccard statistic=98.6%		

Rand index/Jaccard statistic*

Pairs of shoulders:	In the same cluster	In different clusters	N
In the same cluster	22,957	8,494	31,451
In different clusters	19,470	75,332	94,802
N	42,427	83,826	126,253
	Rand index=77.9% , Jaccard statistic=45.1%		

*Removing the 2290 shoulders in the 'Normal' clusters according to both clustering algorithms

The Rand index and Jaccard statistic were heavily influenced by the large 'Normal' cluster, which contributed 2,620,905 pairs of shoulders being in the same cluster according to both cluster solutions. Removing this group gave a more accurate indication of the replicability of the other (arguably more interesting) clusters identified. The Rand index for this analysis (bottom of Table 34) was 77.9% and the Jaccard statistic was reduced to 45.1%.

From both analytical approaches, it was clear that whilst some of the clusters were robust, others were less so, with the 'Mixed signs' and 'AC signs' clusters identified in the Hill Lane data being identified as one, more heterogeneous, cluster (coloured yellow in Figure 24a) in the Bitterne data.

6.2.3 Cluster analysis of the shoulder examination data from the whole population

The shoulder examination data generated from both general practices (N=4190) were clustered as before. The four-cluster hierarchical solution yielded (Figures 25a and 25b, note the break in the y-axis) was more similar to that produced by the Bitterne data alone than the one produced by the Hill Lane data alone, probably because there were twice as many data points from the Bitterne practice as from the Hill Lane practice. The k-means procedure refined the cluster solution by moving 151 (3.6%) shoulders into another cluster, 124 of which were moves between the green and blue clusters in Figure 25a.

Figure 25a: Dendrogram of Ward's hierarchical clustering on both Hill Lane and Bitterne data shoulder examination observations

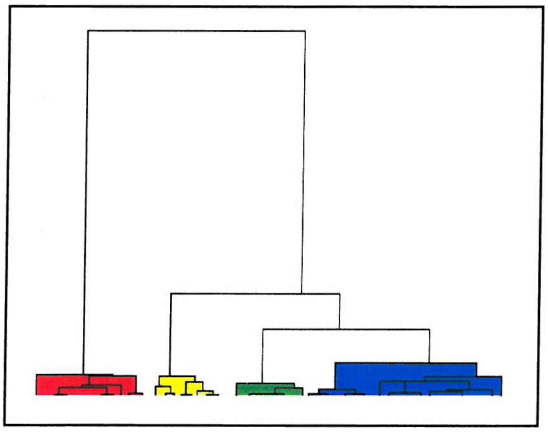
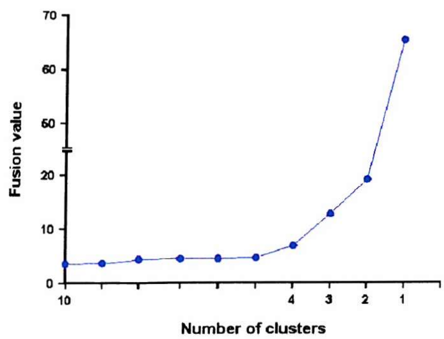


Figure 25b: Fusion values for successive shoulder cluster joins



The shoulder clusters obtained from cluster analysis using all the available data were characterised more distinctly by the severity of the shoulder complaint, and the presence of pain than by the anatomical location of any symptoms or signs (Table 35 and Figure 26). Cluster 1 (N=3396) comprised shoulders with no pain, no tenderness and few other positive signs. These shoulders also displayed the greatest range of shoulder movement, although the difference compared to Clusters 2 and 3 was only slight. Cluster 2 (N=287) identified shoulders without pain, (except for 2% of them with AC joint or subacromial bursa pain) with tenderness at one or two sites. Half of them had other positive signs, mainly pain on resisted abduction or a positive AC joint stress test.

Figure 26: Box-and-whisker plot showing the sum of shoulder ranges of movement for the four clusters

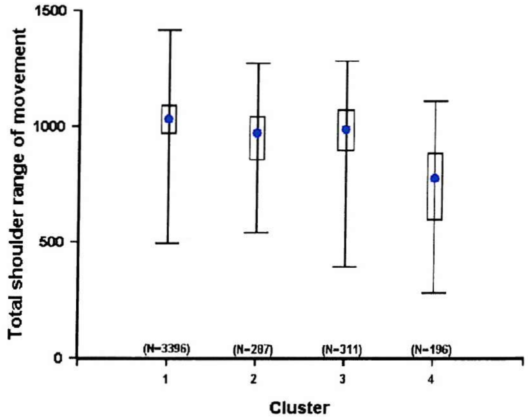


Table 35: Prevalence of symptoms and signs in the final shoulder clusters

Characteristics	Cluster			
	1: 'Normal'	2: Signs	3: Pain	4: Severest involvement
N	**3396**	**287**	**311**	**196**
Colour in dendrogram	Red	Yellow	Green	Blue
Any pain:	*0*	*2*	*100*	*90*
Deltoid		0	41	53
Anterior		0	18	24
AC joint		2	8	32
Subacromial bursa		0.4	9	15
Diffuse		0	8	2
Posterior		0	50	36
Other		0	0	0
Any tenderness:	*0*	*100*	*41*	*95*
Deltoid		11	4	30
Anterior		22	11	20
AC joint		42	2	57
Subacromial bursa		24	11	34
Diffuse		2	3	1
Posterior		20	15	20
Other		0.7	0.3	0
Other signs:	*5*	*46*	*37*	*98*
Pain on resisted:				
- Elbow flexion	0.5	5	4	24
- Forearm supination	0.2	2	0.6	7
- External rotation	0.5	5	10	33
- Internal rotation	0.3	2	5	33
- Abduction	2	15	16	69
AC joint stress test	2	34	9	86
Painful arc	0.5	3	5	8

Cluster 3 (N=311) identified shoulders with pain predominantly at the deltoid or posterior regions of the shoulder; 41% also had shoulder tenderness at a variety of different anatomical sites. Just over a third of these shoulders had other positive signs, predominantly pain on resisted abduction or external rotation. Cluster 4, the smallest group identified (N=196) was characterised by multiple sites of pain, tenderness and/or other positive signs. All shoulders in this group had at least three positive findings (signs or symptoms), and over half of them had six or more. The most common positive findings in this cluster were the AC joint stress test (86%), pain on resisted abduction (69%), AC joint tenderness (57%) and deltoid pain (53%). These shoulders had the lowest ranges of movement, and had a median movement that was 255° lower than that in Cluster 1 shoulders.

Table 36 presents the demographic characteristics of subjects in the four shoulder clusters, history of shoulder pain from the baseline questionnaire and the proportion of pairs of shoulders within each individual falling into the same, or another, cluster. A high proportion of subjects in Cluster 2 ('Signs') were women, and older subjects' shoulders were seen proportionately most frequently in Cluster 4.

Table 36: Demographic and baseline characteristics in the four shoulder clusters

Characteristics	**Cluster**			
	1: 'Normal'	**2: Signs**	**3: Pain**	**4: Severest involvement**
N	**3396**	**287**	**311**	**196**
Colour in dendrogram	Red	Yellow	Green	Blue
Demographic characteristics:				
% female	59	75	56	62
% 25 – 34 yrs	16	10	15	5
% 35 – 44 yrs	25	21	17	12
% 45 – 54 yrs	32	33	29	36
% 55 – 64 yrs	28	36	39	46
Report of shoulder pain at baseline:				
% Shoulder pain present	11	39	82	83
% activities difficult/impossible	7	28	54	70
% shoulder pain > 1 yr	5	22	46	54
Laterality - Number (%) of shoulders whose pair:				
Was in the same cluster	2956 (87%)	106 (37%)	114 (37%)	56 (29%)
Was in the 'Normal' cluster	-	141 (49%)	175 (56%)	95 (48%)
Was in a different cluster (not 'Normal')	411 (12%)	38 (13%)	21 (7%)	39 (20%)
Was not in the analysis	29 (1%)	2 (1%)	1 (0.3%)	6 (4%)

Clusters 3 and 4, which involved shoulder pain at the examination, contained the highest proportions of shoulders with pain reported at the questionnaire, pain reported at the questionnaire making activities difficult or impossible and pain of duration longer than a year previous to the questionnaire. However, a substantial proportion of shoulders showed these characteristics in Cluster 2, the 'Signs' group. This may suggest a group of shoulders with resolving disorders, late-stage shoulder capsulitis or possibly the early stages of shoulder disorders.

Symptom-sign profiles were most frequently unilateral (411 subjects) or similar enough within subject to be in the same cluster (1616 subjects, of whom 1478 had two 'normal' shoulders in Cluster 1). Only 49 subjects had shoulders that were in Clusters 2, 3 or 4 and were in different clusters to each other. 38 subjects had only one shoulder with complete examination data.

6.2.4 Comparison of the shoulder clusters with the HSE shoulder diagnoses

All clinically-driven diagnoses were seen in Clusters 2, 3 and 4 (Table 37).

Table 37: Prevalence of diagnoses amongst the four shoulder clusters

Diagnoses (%)	**Cluster**			
	1: 'Normal'	**2: Signs**	**3: Pain**	**4: Severest involvement**
N	**3396**	**287**	**311**	**196**
Colour in dendrogram	Red	Yellow	Green	Blue
None	100	98	21	10
Rotator cuff tendonitis (RT)	0	0	4	1
Bicipital tendonitis (BT)	0	0	1	0
Shoulder capsulitis (SC)	0	1	45	13
Acromioclavicular joint disorder (AC)	0	0.3	0	0.5
Subacromial bursitis (SAB)	0	0.3	0.3	0
RT & SC	0	0	22	38
RT & SAB	0	0	0.3	0
BT & SC	0	0	1	1
SC & AC	0	0	0	6
SC & SAB	0	0	3	1
RT & SC & BT	0	0	0.3	6
RT & SC & AC	0	0	0	13
RT & SC & SAB	0	0	2	7
SC & AC & SAB	0	0	0	0.5
SC & BT & SAB	0	0	0	1
RT & SC & BT & AC	0	0	0	0.5

Cluster 2 ('Signs') contained very few shoulders with a clinical diagnosis, because so few of them had pain at any site. No acromioclavicular joint disorder and hardly any subacromial bursitis diagnoses were seen in Cluster 3 ('Pain'), because these diagnoses required tenderness to be present. Cluster 4 contained shoulders with single diagnoses (14.5%), multiple diagnoses (74%), and no diagnosis (10%).

As suggested by both the cluster analysis findings and the pattern of clinical diagnoses, there was little distinction between different shoulder disorders following non-invasive examination. In particular, rotator cuff tendonitis was rarely diagnosed without an accompanying diagnosis of shoulder capsulitis. The clusters formed indicated that a grading of severity of shoulder involvement, from no involvement (neither symptoms nor signs), through no symptoms only signs, symptoms and some signs, to multiple symptoms and signs, was the key distinguishing feature of these shoulders.

6.3 The elbow

Physical examination of the elbow consisted of the location of reported pain, location of any tenderness, pain on resisted movements, crepitus and swelling (Appendix I). Of the 4290 elbows examined, 27 had missing data and were excluded from the analysis (7 of these occurred because a movement had not been performed due to pain; the other 20 were due to an observation being missed out by mistake).

703 patients were recruited, giving a total of 1404 elbows with complete information from the Hill Lane practice. These data were analysed first.

The 17 variables were analysed in two ways, as proposed in Sections 3.3 and 3.4:
1) Using squared Euclidean distance and Ward's method, and
2) Using the Jaccard measure of community and group average linkage.

Ward's method on squared Euclidean distance identified six main clusters (Figure 27a), with one (coloured red) being clearly separate from the other five. The fusion values for this cluster solution (Figure 28) showed a clear step-up from six clusters to five, but not from seven clusters to six. The group average method on the Jaccard measure of community identified one main cluster and a further four minor groups (with 12, 22, 31 and 5 members in the blue, cyan, pink and orange clusters in Figure 27b respectively). Four of the five groups corresponded closely in character to four of those identified by Ward's method (denoted by corresponding colours in Figures 27a and 27b). It should be noted that the fusion distances in these two dendrograms are not comparable. It was decided therefore to continue analysis with the Ward's method solution, which gave a clearer group structure. The failure of the cluster analysis using the Jaccard measure of community and group average linkage to produce meaningful clusters may have been due to a combination of factors: the distance measure used would have inflated the dissimilarities between the elbow profiles compared to squared Euclidean distance (by ignoring the highly prevalent agreements in absences of traits), and the group average linkage method is less stringent than Ward's method is in creating tightly homogeneous clusters. The k-means procedure refined the six-cluster solution by making 5 (0.4%) inter-cluster moves.

Figures 27a and 27b: Dendrograms of a) Ward's and b) group average linkage hierarchical clustering on Hill Lane data elbow examination observations

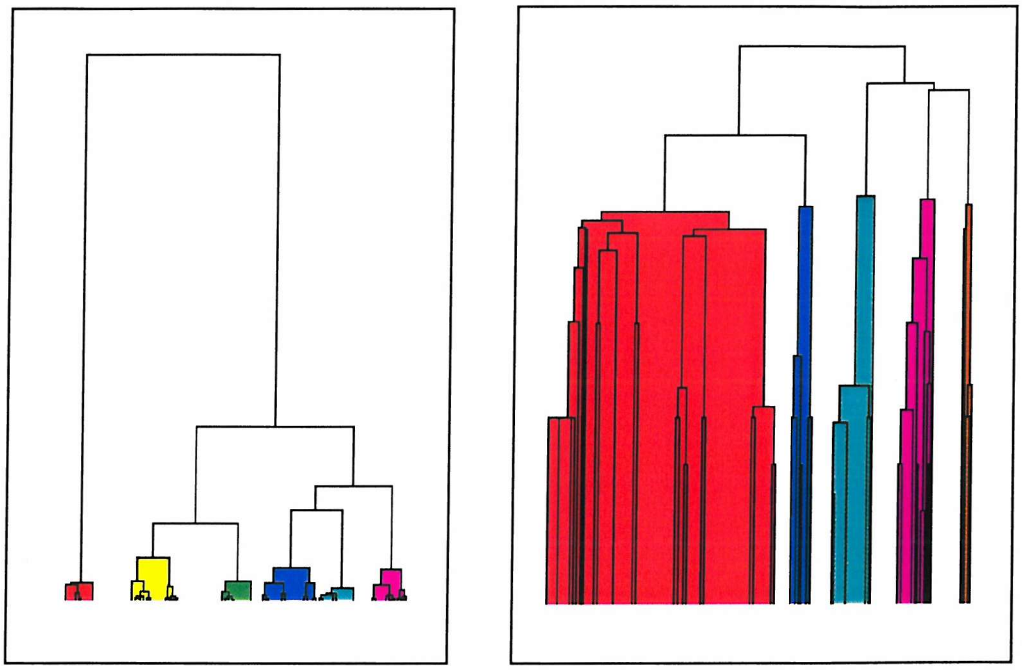
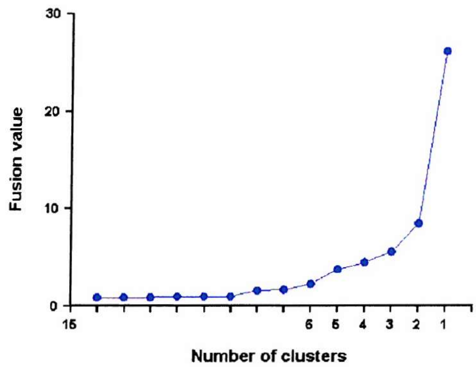


Figure 28: Fusion values for successive elbow cluster joins



6.3.1 *Characterisation of the elbow clusters (Hill Lane data only)*

Cluster 1 (N=1221) identified elbows with minimal symptoms and signs (Table 38). Cluster 2 (N=47) comprised elbows with no pain (except at the antecubital fossa in 6% of them), tenderness (predominantly at one site each) at sites except the lateral elbow, minimal pain on resisted movement, and some swelling (11%). This cluster was the most heterogeneous (the yellow cluster in Figure 27a). Cluster 3 (N=47) was characterised by elbows with lateral tenderness and some lateral pain on resisted movement. Lateral pain was not present. Cluster 4 (N=37) comprised elbows with various sites of pain and some posterior tenderness. All elbows with posterior tenderness also had posterior pain and included one elbow with swelling over the posterior joint as well. This group of elbows that mostly showed symptoms but not signs was heterogeneous with respect to the site of pain and presence of posterior elbow signs (the blue cluster in Figure 27a). Cluster 5 (N=23) was the smallest group and identified elbows with medial (alongside some posterior) signs and symptoms. The final cluster, Cluster 6 (N=29) was characterised by elbows with lateral alongside some medial and posterior signs and symptoms.

Table 38: Prevalence of symptoms and signs in the Hill Lane elbow clusters

Characteristics	**Cluster**					
	1: 'Normal'	**2: Signs, not lateral**	**3: Lateral signs**	**4: Pain/ posterior signs**	**5: Medial symptoms and signs**	**6: Lateral symptoms and signs**
N	**1221**	**47**	**47**	**37**	**23**	**29**
Colour in dendrogram	Red	Yellow	Green	Blue	Cyan	Pink
Any pain:	*0*	*6*	*6*	*100*	*100*	*100*
Lateral		0	0	30	4	100
Medial		0	2	24	96	21
Posterior		0	2	51	17	10
Antecubital fossa		6	2	3	0	0
Other		0	0	3	4	0
Any tenderness:	*0*	*100*	*100*	*14*	*100*	*100*
Lateral		0	100	0	4	97
Medial		57	11	0	100	17
Posterior		15	4	14	9	3
Antecubital fossa		28	2	0	0	3
Other		6	0	0	4	0
Pain on resisted movement:*	*0.3*	*2*	*26*	*0*	*48*	*59*
Lateral elbow on wrist extension	0.3	0	26	0	0	59
Medial elbow on wrist flexion	0.1	2	0	0	48	14
Swelling posterior joint	1	11	0	3	0	14

*crepitus variables are not presented because there were so few positive responses

6.3.2 Replication of the elbow clusters (Bitterne data only)

Elbow examination data from the Bitterne practice (2859 elbows with non-missing examinations) were clustered using the two methods already employed on the Hill Lane data. As before, Ward's hierarchical method on squared Euclidean distance gave a more clearly structured solution of seven clusters (Figures 29a and 30, note the break in the y-axis).

Figures 29a and 29b: Dendrograms of a) Ward's and b) group average linkage hierarchical clustering on Bitterne data elbow examination observations

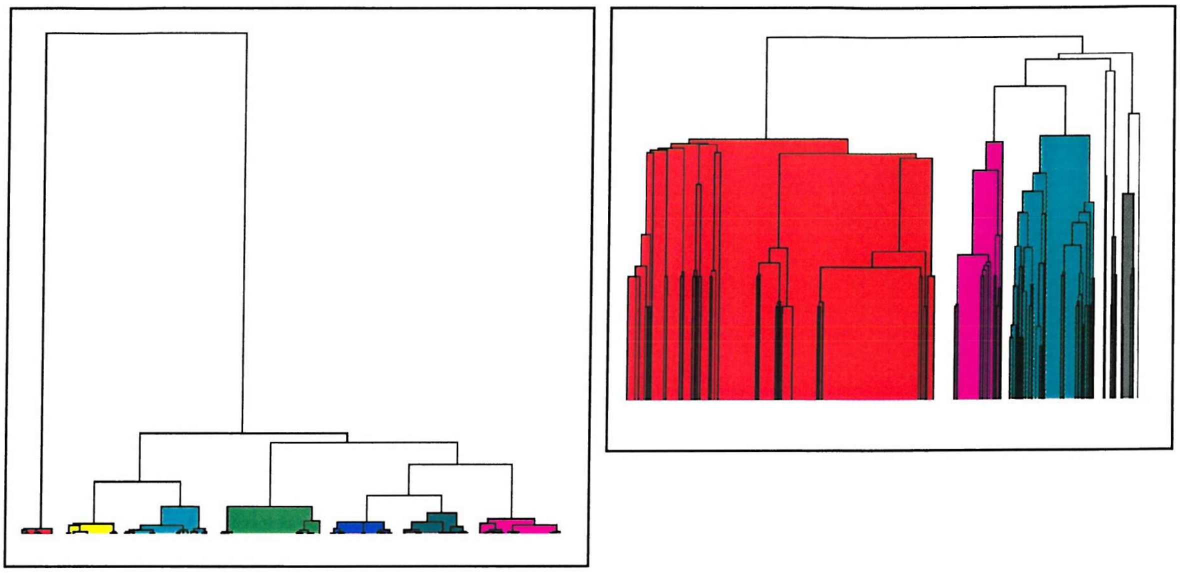
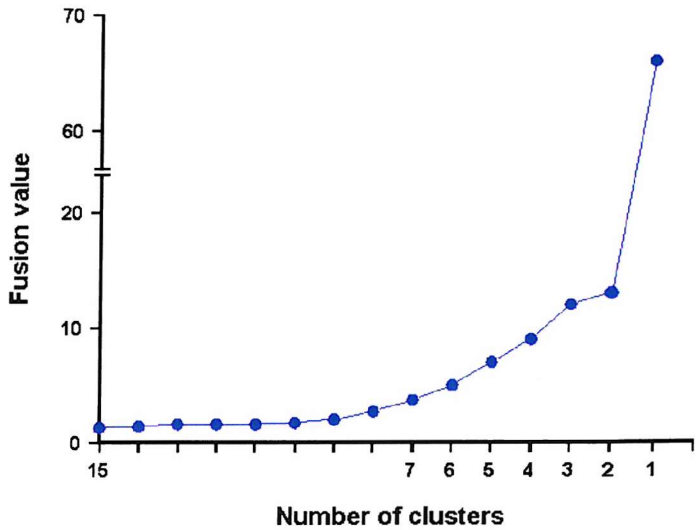


Figure 30: Fusion values for successive elbow cluster joins



The group average linkage algorithm on Jaccard's measure of community yielded three main clusters corresponding to three observed by Ward's method (the red, pink and cyan clusters in Figure 29b) and a further four clusters containing 2, 6, 12 and 1 members (the grey clusters). The k-means procedure refined the Ward's algorithm seven-cluster solution by moving 29 (1%) elbows into another cluster.

Table 39: Comparison of the Bitterne elbow clusters obtained by two different methods

K-means clustering using Hill Lane seed points: / **Hierarchical (Ward's method) clustering plus k-means:**

Kappa statistic

Clusters:	1	2	4	5	3	7	6	N
Colour in dendrogram	Red	Yellow	Green	Blue	Cyan	Pink	Dark Green	
1: 'Normal'	**2460**	0	0	0	0	0	0	2460
2: Signs, not lateral	0	**54**	0	0	0	0	31	85
3: Lateral signs	0	21	**112**	0	0	5	1	139
4: Pain/posterior signs	0	0	0	**45**	0	0	9	54
5: Medial symptoms and signs	0	0	0	0	**46**	0	0	46
6: Lateral symptoms and signs	0	0	0	0	14	**60**	1	75
N	2460	75	112	45	60	65	42	2859

κ=0.89

Rand index/Jaccard statistic

Pairs of elbows:	In the same cluster	In different clusters	N
In the same cluster	3,036,824	6,148	3,042,972
In different clusters	2,438	1,040,101	1,042,539
N	3,039,262	1,046,249	4,085,511

Rand index=99.8% , Jaccard statistic=99.7%

Rand index/Jaccard statistic*

Pairs of elbows:	In the same cluster	In different clusters	N
In the same cluster	12,254	6,148	18,402
In different clusters	2,438	58,561	60,999
N	14,692	64,709	79,401

Rand index=89.2% , Jaccard statistic=58.8%

*Removing the 2460 elbows in the 'Normal' clusters according to both clustering algorithms

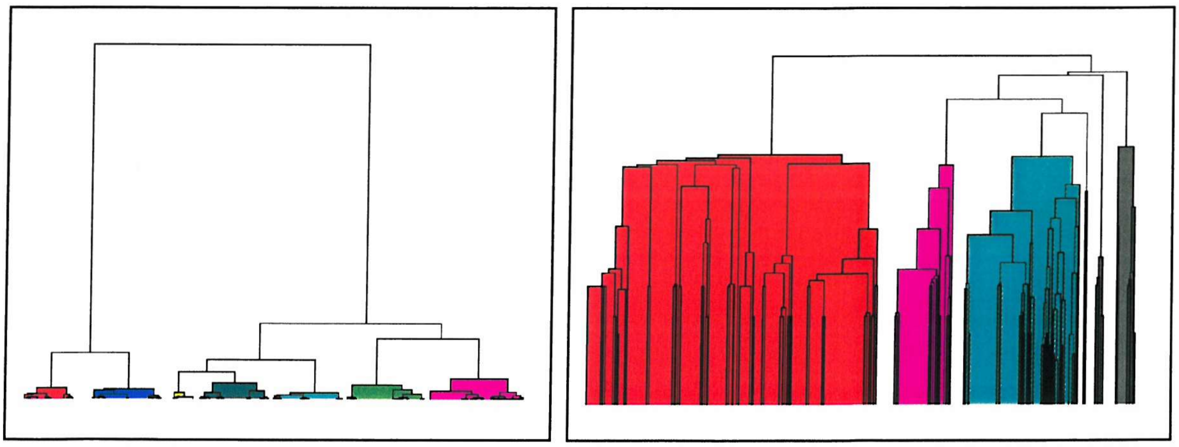
The same elbow examination data were analysed using the k-means procedure, with the six Hill Lane cluster centroids as the seed points. The two cluster solutions for the Bitterne data were compared using the kappa statistic and Rand index (Table 39).

The observed agreement between the clusters (numbers in bold) was 97%, whilst the expected was 74%, yielding a κ of 0.89. The agreement between each pair of elbows as to whether they were placed in the same cluster, or in different clusters was 99.8% according to the Rand index, and 99.7% according to the Jaccard statistic. These values were adjusted to 89.2% and 58.8% when the 2460 elbows in the 'Normal' cluster had been removed. All of these indices suggested a high level of agreement, and thus replicability of the clusters. In particular, the 'Normal', 'Medial symptoms and signs' and 'Lateral symptoms and signs' were the only clusters to be consistently identified by both Ward's and the group average linkage methods in both the Hill Lane and the Bitterne datasets (the red, cyan and pink clusters in Figures 27a, 27b, 29a and 29b). The least robust cluster from the Hill Lane analysis was Cluster 2 ('Signs, not lateral'), the characteristics of which were split into two groups in the Bitterne analysis: one distinguished by medial elbow tenderness and the other by posterior or antecubital fossa signs and symptoms (the yellow and the dark green clusters in Figure 29a).

6.3.3 <u>Cluster analysis of the elbow examination data from the whole population</u>

The elbow examination data from the whole population were clustered using the same methods previously employed, yielding the dendrograms presented in Figures 31a and 31b.

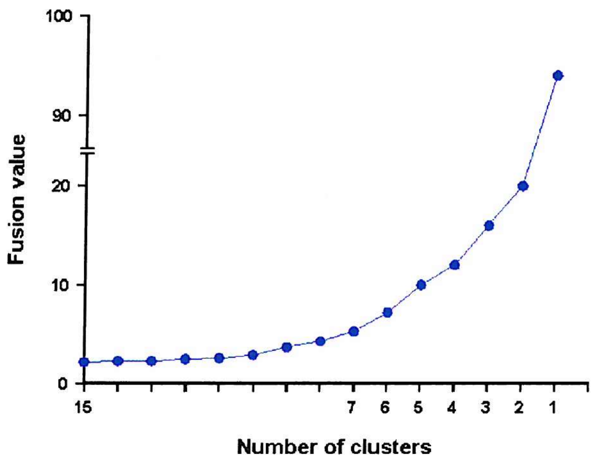
Figures 31a and 31b: Dendrograms of a) Ward's and b) group average linkage hierarchical clustering on both Hill Lane and Bitterne data elbow examination observations



As seen in the previous elbow cluster analyses, Ward's method on squared Euclidean distance produced a clearer hierarchical structure than that observed by group average linkage, although the three most stable clusters were replicated (the red, pink and cyan clusters). As the fusion values suggested (Figure 32, note the break in the y-axis), a seven-cluster solution was evident, reflecting the Bitterne data cluster solution. In this analysis, however, the 'Normal' elbows (the red cluster) were not joined to all of the other clusters at the last hierarchical stage, but joined with the blue cluster of elbows first.

The k-means procedure on the Ward's method cluster analysis moved 27 elbows into other clusters, with at least 4 elbows being moved into or out of any one cluster.

Figure 32: Fusion values for successive elbow cluster joins



The characteristics of these clusters are presented in Table 40. Cluster 1 (N=3681) was the largest cluster, and contained elbows with no pain, no tenderness and few other positive signs. Cluster 2 (N=77) was more closely associated with Cluster 1 than the other clusters according to the dendrogram in Figure 31a, mainly because of the lack of tenderness seen in these elbows. It comprised elbows with a variety of sites of pain (predominantly at one site only) and minimal other positive signs. Clusters 3, 4, 5, 6 and 7 all had tenderness, either at the lateral elbow (Clusters 6 and 7) or elsewhere (Clusters 3, 4 and 5). Cluster 3 (N=84) was characterised by medial elbow tenderness and some pain over the medial elbow on resisted wrist flexion. Cluster 4 (N=69) identified elbows with posterior elbow or antecubital fossa signs and symptoms, and Cluster 5 (N=66) included elbows with medial elbow symptoms and signs. Cluster 6 (N=174) was the second largest group and comprised elbows with lateral elbow signs, whilst elbows in Cluster 7 (N=112) additionally had lateral elbow pain.

Table 40: Prevalence of symptoms and signs in the final elbow clusters

Characteristics	**Cluster**						
	1: 'Normal'	**2: Pain**	**3: Medial signs**	**4: Posterior /AF**	**5: Medial symptoms and signs**	**6: Lateral signs**	**7: Lateral symptoms and signs**
N	**3681**	**77**	**84**	**69**	**66**	**174**	**112**
Colour in dendrogram	Red	Blue	Yellow	Dark Green	Cyan	Green	Pink
Any pain:	*0*	*100*	*0*	*42*	*100*	*0*	*100*
Lateral		34		4	5		88
Medial		29		1	97		21
Posterior		35		20	9		25
Antecubital fossa		6		22	2		4
Other		5		1	6		3
Any tenderness:	*0*	*0*	*100*	*100*	*100*	*100*	*100*
Lateral			4	3	11	100	99
Medial			100	0	100	13	23
Posterior			2	36	5	4	13
Antecubital fossa			4	58	2	1	3
Other			0	3	3	2	2
Pain on resisted movement:*	*0.3*	*6*	*26*	*7*	*50*	*26*	*54*
Lateral elbow on wrist extension	0.2	5	2	4	5	26	50
Medial elbow on wrist flexion	0.1	1	26	4	47	5	20
Swelling posterior joint	0.4	3	6	12	2	2	5

*crepitus variables are not presented because there were so few positive responses.

The demographic characteristics of subjects in the seven elbow clusters, history of elbow pain from the baseline questionnaire and the proportion of pairs of elbows within each individual falling into the same, or another, cluster were investigated (Table 41). Similar proportions of men and women were seen in each elbow cluster, although women were slightly more likely to have elbows with medial symptoms or signs (Clusters 3 and 5). The age distribution did not vary widely between the clusters, except that virtually no subjects of age 25 – 34 had elbows with medial or lateral signs or symptoms (Clusters 3, 5, 6 and 7). Subjects with these signs or symptoms were more likely to be age 45 – 64, compared to those subjects whose elbows had minimal abnormality.

Clusters 2, 4, 5 and 7, those identifying elbows with pain at the examination had the highest proportions of subjects with a history of elbow pain at baseline. Clusters 3 and 6 may represent elbows with resolving conditions, especially in those with long-lasting previous pain reported at the baseline questionnaire.

Symptom-sign profiles were most frequently unilateral (329 subjects) or similar enough within subject to be in the same cluster (1761 subjects, of whom 1671 had two 'normal' elbows). Only 33 subjects had two non-normal elbows in different clusters from each other, and there was no discernible pattern to these. The remaining 17 subjects included in the analysis contributed only one elbow with complete examination data.

Table 41: Demographic and baseline characteristics in the seven elbow clusters

Characteristics	Cluster						
	1: 'Normal'	2: Pain	3: Medial signs	4: Posterior /AF	5: Medial symptoms and signs	6: Lateral signs	7: Lateral symptoms and signs
N	**3681**	**77**	**84**	**69**	**66**	**174**	**112**
Colour in dendrogram	Red	Blue	Yellow	Dark Green	Cyan	Green	Pink
Demographic characteristics:							
%female	61	52	65	54	64	60	52
% 25 – 34 yrs	16	12	8	10	2	8	3
% 35 – 44 yrs	24	30	12	25	24	22	20
% 45 – 54 yrs	31	34	31	28	45	38	49
% 55 – 64 yrs	30	25	49	38	29	32	29
Report of elbow pain at baseline:							
% pain present	5	69	37	58	83	30	79
% activities difficult /impossible	3	53	27	39	67	19	69
% pain > 1 yr	2	35	21	23	56	15	40
Laterality – Number (%) of elbows whose pair was:							
In the same cluster	3342 (91)	32 (42)	18 (21)	16 (23)	22 (33)	58 (33)	34 (30)
In the 'Normal' cluster	-	39 (51)	54 (64)	44 (64)	29 (44)	104 (60)	59 (53)
In a different cluster (not 'Normal')	329 (9)	6 (8)	11 (13)	8 (12)	13 (20)	11 (6)	17 (15)
Not in the analysis	10 (0.3)	0	1 (1)	1 (1)	2 (3)	1 (0.6)	2 (2)

6.3.4 Comparison of the elbow clusters with the HSE elbow diagnoses

All clinically-driven diagnoses were seen in Clusters 4, 5 and 7, i.e. those elbows with pain and tenderness at the examination (Table 42). Cluster 4 contained only three elbows (6%) with a clinical diagnosis, olecranon bursitis, although a further seven had signs and symptoms at the posterior elbow. Cluster 5 contained 29 elbows

(44%) with a clinical diagnosis of medial epicondylitis, and 35 (53%) elbows with medial signs and symptoms but no diagnosis. The remaining two elbows had multiple medial signs but no medial pain and therefore were not diagnosed with medial epicondylitis. Cluster 7 contained 40 elbows (36%) with a clinical diagnosis of lateral epicondylitis, a further ten (9%) with diagnoses of lateral and medial epicondylitis, two (2%) with a diagnosis of medial epicondylitis and three (3%) with a diagnosis of olecranon bursitis. 104 (93%) elbows in this cluster had multiple symptoms and/or signs at the lateral elbow, and the other eight elbows had lateral elbow tenderness as well as pain at other sites. As suggested by both the cluster analysis findings and the pattern of clinical diagnoses, signs and symptoms at the elbow were grouped primarily by site, and thus the two classification schemes were in strong agreement.

Table 42: Prevalence of diagnoses amongst the seven elbow clusters

Diagnoses (%)	**Cluster**						
	1: 'Normal'	**2: Pain**	**3: Medial signs**	**4: Posterior /AF**	**5: Medial symptoms and signs**	**6: Lateral signs**	**7: Lateral symptoms and signs**
N	**3681**	**77**	**84**	**69**	**66**	**174**	**112**
Colour in dendrogram	Red	Blue	Yellow	Dark Green	Cyan	Green	Pink
None	100	100	100	94	56	100	51
Lateral epicondylitis	0	0	0	0	0	0	36
Medial epicondylitis	0	0	0	0	44	0	2
Olecranon bursitis	0	0	0	6	0	0	3
Lateral and medial epicondylitis	0	0	0	0	0	0	9

6.4 The hand and wrist

Physical examination of the hand and wrist covered the location of reported pain, tenderness and swelling, pain in the wrist or thumb on resisted movements, presence of Heberden's nodes, Dupuytren's contracture, muscle wasting, thumb weakness or abnormal sensation, and specific tests (Phalen's, Tinel's and Finkelstein's). The Katz diagram classification and any report of sleep disturbance due to numbness or tingling (from the nurse's interview) were also included.

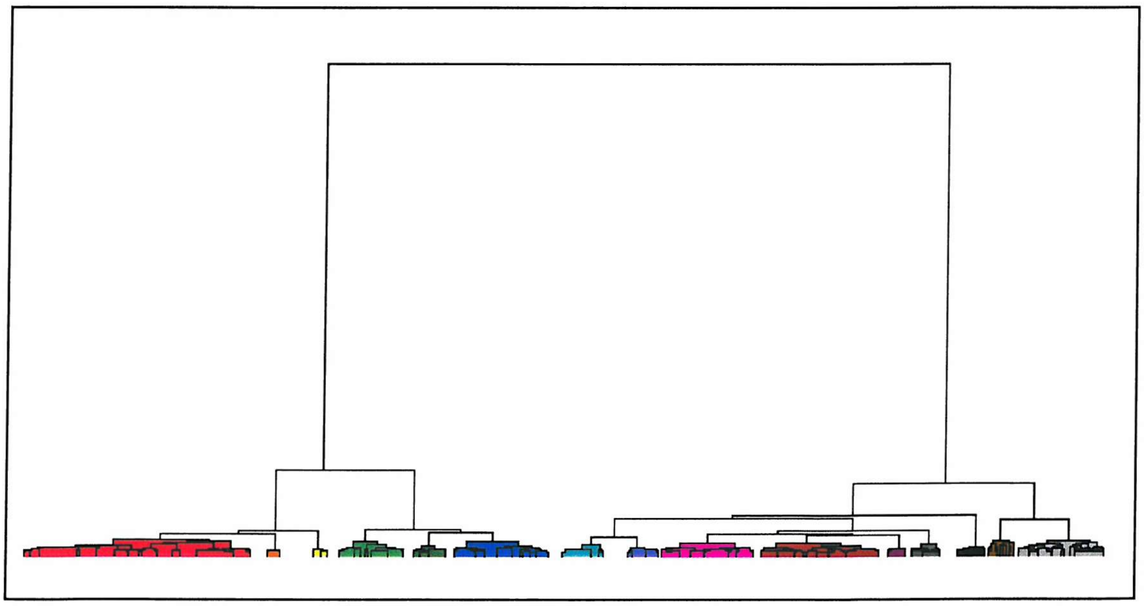
A preliminary cluster analysis was performed on the Katz hand diagram data alone (Appendix IV). The results of this analysis were incorporated into the cluster analysis of the hand and wrist examination data, in the form of seven binary variables denoting the presence or absence of numbness or tingling at the following sites: 1) little or ring fingers, 2) middle or index fingers, 3) thumb, 4) palm or dorsum, 5) palmar side, 6) dorsal side, 7) middle or proximal phalanges.

Of the 4290 hands examined, 141 had missing data and were excluded from the analysis (10% of these were because a movement had not been performed due to pain; the rest were due to one or two observations having been omitted by mistake).

699 subjects with complete examination data (1392 wrists/hands) were recruited from the Hill Lane practice and were analysed first.

The 57 variables were analysed using the two methods previously employed in the elbow cluster analyses (squared Euclidean distance with Ward's method, and Jaccard's coefficient of community with group average linkage). The dendrograms yielded by these methods (Figures 33a and 33b) suggested once again that the first method was more effective at recovering a distinct group structure than the second. It was also more mathematically robust (the k-means procedure made twice as many moves in refining the second method's 10-cluster solution than in refining the first's 15-cluster solution). The group average 10-cluster solution consisted of five large clusters of 44, 37, 29, 899 and 350 hands, and five minor clusters containing 9, 7, 11, 3 and 3 hands each.

Figure 33a: Dendrogram of Ward's hierarchical clustering on Hill Lane data wrist/hand examination observations



The dendrogram for the 15-cluster solution obtained from Ward's method indicated that there were two main clusters, each divided into two further clusters. Beyond these four sub-clusters a 7-cluster or 15-cluster solution was suggested (Figure 34, note the axis break). The 7-cluster solution was investigated initially, but four of the clusters were too heterogeneous to give clear discrimination and characterisation. The 15-cluster solution was therefore explored and characterised. The k-means procedure refined this solution by making 114 (8.2%) inter-cluster moves, two-thirds of which were made between clusters that would have been amalgamated in the 7-cluster solution (i.e. between Clusters 1, 2 and 3; 4, 5 and 6; 7 and 8; or 9, 10, 11 and 12).

Figure 33b: Dendrogram of group average linkage hierarchical clustering on Hill Lane data wrist/hand examination observations

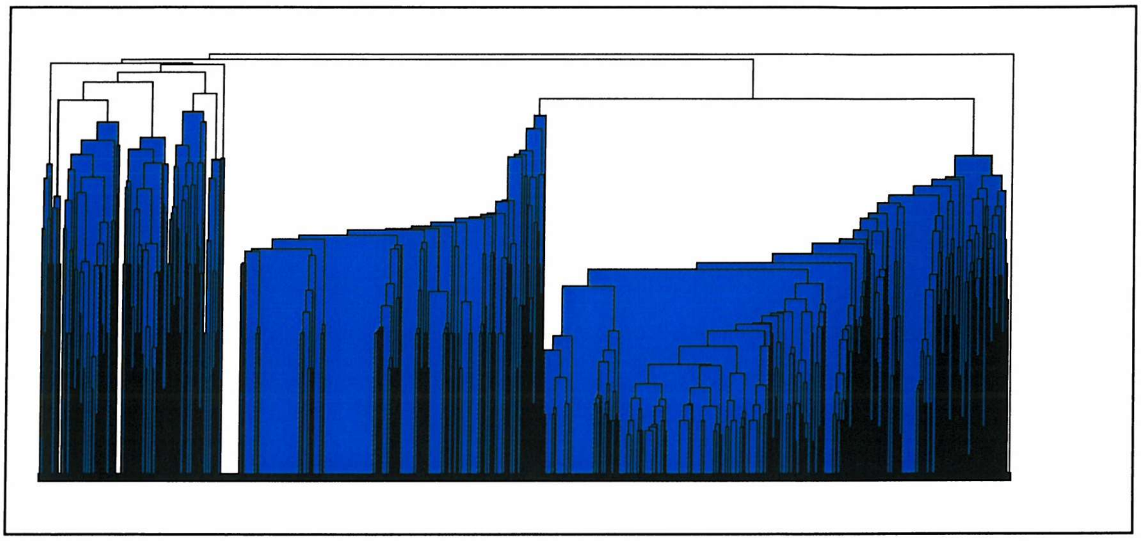
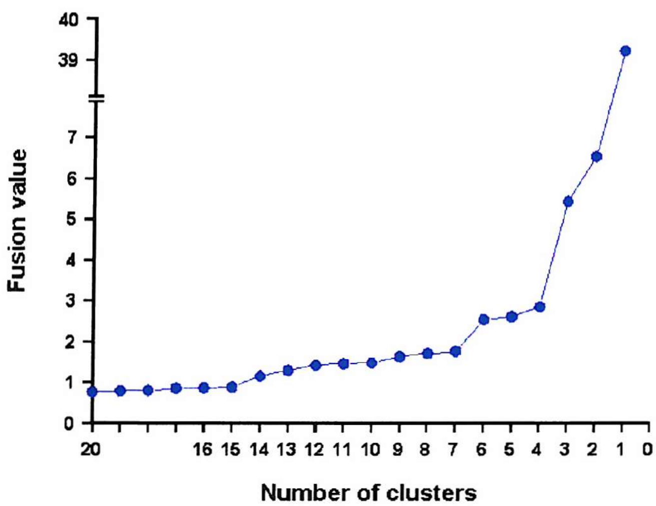


Figure 34: Fusion values for successive wrist/hand cluster joins



6.4.1 Characterisation of the wrist/hand clusters (Hill Lane data only)

Cluster 1 (N=761) was the largest and identified wrists/hands with minimal abnormalities (Tables 43a and 44a). Cluster 2 (N=59) comprised wrists/hands with a positive Phalen's test, but no accompanying symptoms and few other signs. Cluster 3 (N=81) was characterised by the presence of Heberden's nodes, but no symptoms and minimal other signs.

Cluster 4 (N=47) was characterised by pain and signs at the radial aspect of the wrist and thumb (emboldened in Table 43a) and 21% of these hands had Heberden's nodes. There were few other signs or symptoms and no numbness or tingling. Cluster 5 (N=35) identified hands with finger joint pain and accompanying signs (emboldened in Table 43a) with few other signs or symptoms and no numbness or tingling. Cluster 6 (N=60) identified wrists/hands mostly with one site of wrist pain and some accompanying tenderness, swelling pain on resisted movement or other musculoskeletal sign (Tables 43b and 44b).

These first six clusters were all characterised by an absence of numbness or tingling.

Cluster 7 (N=31) comprised wrists/hands with few musculoskeletal signs or symptoms (Table 43b), but numbness or tingling in the palmar aspect of the fingers (Table 44b). Note that the areas of numbness and tingling were defined according to the Katz hand clusters, not by individual wrists/hands. Just under half of these hands had a positive Phalen's test, and 29% of these subjects reported disturbed sleep due to numbness or tingling. Cluster 8 (N=44) contained wrists/hands with few musculoskeletal signs or symptoms, but with numbness or tingling in the distal phalanges only. Just over a third of these wrists/hands also had a positive Phalen's test, and 18% of these subjects reported disturbed sleep due to numbness or tingling.

Table 43a: Prevalence of musculoskeletal symptoms and signs in the first five Hill Lane wrist/hand clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Phalen**	**3: Heberden**	**4: Radial wrist & thumb**	**5: Joint**
N	**761**	**59**	**81**	**47**	**35**
Colour in dendrogram	Red	Orange	Yellow	Light green	Dark green
Any pain:	*0*	*0*	*0*	*100*	*100*
Dorsal forearm				6	0
Palmar forearm				2	0
Dorsal wrist				6	0
Palmar wrist				4	3
Radial wrist				**36**	3
Medial wrist				2	6
Thumb base				**96**	29
Finger joints				4	**100**
Other (mostly hand)				4	6
Any tenderness:	*4*	*14*	*11*	*77*	*63*
Dorsal forearm	0.1	2	1	4	0
Palmar forearm	0.1	2	0	2	0
Dorsal wrist	0.1	2	5	4	0
Palmar wrist	0	2	0	0	0
Radial wrist	0.7	3	1	**26**	3
Medial wrist	0.1	0	0	2	0
Thumb base	3	5	5	**72**	11
Finger joints	0.3	2	1	6	**51**
Other (mainly hand)	0.3	0	0	4	9
Any swelling:	*2*	*3*	*6*	*36*	*40*
Dorsal forearm	0.2	0	1	4	*0*
Palmar forearm	0	0	0	0	0
Dorsal wrist	0.1	2	1	4	0
Palmar wrist	0	0	0	0	0
Radial wrist	0.2	0	0	**15**	0
Medial wrist	0	0	1	2	0
Thumb base	0.4	0	0	**13**	6
Finger joints	2	2	4	2	**37**
Other (mainly hand)	0	2	1	0	0
Pain on resisted movement:	*3*	*7*	*10*	*51*	*29*
Radial wrist	0.8	2	2	**15**	6
Medial wrist	0.1	0	0	2	3
Finger extension	0.5	0	1	**9**	9
Finger flexion	0.3	4	1	4	14
Thumb extension	2	2	9	**36**	6
Other hand signs:					
Dupuytren's contracture	0.7	0	6	2	6
Heberden's nodes	0	0	**100**	21	**51**
Finkelstein's test positive	1	2	4	**19**	0
Thenar muscle wasting	0.1	2	1	11	0
Hypothenar muscle wasting	0.5	0	1	4	0
Weakness of thumb abduction	1	0	6	**21**	0
Weakness of thumb opposition	2	3	4	**11**	20

Table 44a: Prevalence of sensorineural symptoms and signs in the first five Hill Lane wrist/hand clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Phalen**	**3: Heberden**	**4: Radial wrist & thumb**	**5: Joint**
N	**761**	**59**	**81**	**47**	**35**
Colour in dendrogram	Red	Orange	Yellow	Light green	Dark green
Any numbness/ tingling (according to Katz hand cluster definition):					
Little/ ring finger	0	0	0	0	0
Middle/ index finger	0	0	0	0	0
Thumb	0	0	0	0	0
Palm/ Dorsum	0	0	0	0	0
Palmar Aspect	0	0	0	0	0
Dorsal Aspect	0	0	0	0	0
Proximal/ Middle phalanges	0	0	0	0	0
Sensorineural signs:					
Abnormal sensation:					
thumb	0.5	0	2	0	0
Index finger	0.8	2	7	0	0
little finger	0.7	2	4	0	0
Phalen's test positive	0	**100**	11	11	9
Tinel's test positive	0.9	5	2	4	6
Sleep disturbed because of numbness/tingling	0	2	0	0	0

Cluster 9 (N=74) identified hands with numbness or tingling in most of the hand except the thumb, 45% had a positive Phalen's test and 28% of these subjects reported disturbed sleep due to numbness or tingling. These wrists/hands had few musculoskeletal symptoms or signs. Cluster 10 (N=83) comprised hands with numbness or tingling in all regions of the hand. Nearly half of these had a positive Phalen's test and 43% of these subjects reported disturbed sleep due to numbness or tingling. As for Clusters 7, 8 and 9, Cluster 10 had few musculoskeletal symptoms or signs.

Cluster 11 (N=19) was the second smallest cluster identified, and contained hands with numbness or tingling in the middle and index fingers and thumb only (Table 44c). Over a third of these hands had a positive Phalen's test, 26% had Heberden's nodes and 17% had abnormal sensation in the index finger. Again, this group of wrists/hands had minimal musculoskeletal symptoms or signs (Table 43c). Cluster 12 (N=23) exhibited the highest prevalence of sensorineural signs accompanying numbness and tingling in most regions of the hand. Abnormal sensation in the thumb, index and little fingers was seen in over three-quarters of these hands, 78% had a positive Phalen's test and 52% had a positive Tinel's test. Over a third of these subjects had disturbed sleep due to numbness or tingling and 26% of these hands had Heberden's nodes.

Table 43b: Prevalence of musculoskeletal symptoms and signs in the second five Hill Lane wrist/hand clusters

Characteristics	**Cluster**				
	6: Wrist pain	**7: NT – palmar fingers**	**8: NT – distal**	**9: NT – all not thumb**	**10: NT – all**
N	**60**	**31**	**44**	**74**	**83**
Colour in dendrogram	Blue	Cyan	Purple	Pink	Dark red
Any pain:	*100*	*13*	*11*	*5*	*18*
Dorsal forearm	5	3	2	0	1
Palmar forearm	3	0	2	0	0
Dorsal wrist	**40**	10	5	1	6
Palmar wrist	**40**	3	0	3	2
Radial wrist	**32**	0	0	1	5
Medial wrist	**23**	0	0	0	0
Thumb base	3	0	5	0	4
Finger joints	0	0	2	0	2
Other (mostly hand)	10	0	5	1	1
Any tenderness:	*52*	*23*	*16*	*12*	*16*
Dorsal forearm	2	10	5	0	5
Palmar forearm	3	6	5	1	0
Dorsal wrist	**18**	19	2	4	2
Palmar wrist	**10**	0	0	1	2
Radial wrist	**8**	0	0	1	4
Medial wrist	**8**	0	0	1	0
Thumb base	3	0	9	3	4
Finger joints	0	0	0	1	1
Other (mainly hand)	2	0	0	0	0
Any swelling:	*23*	*3*	*5*	*3*	*6*
Dorsal forearm	2	0	0	0	*2*
Palmar forearm	0	0	0	0	0
Dorsal wrist	8	0	0	1	0
Palmar wrist	5	3	0	0	0
Radial wrist	8	0	0	0	1
Medial wrist	7	3	0	0	1
Thumb base	0	0	5	0	1
Finger joints	0	0	0	1	1
Other (mainly hand)	2	0	0	0	0
Pain on resisted movement:	*32*	*10*	*2*	*7*	*7*
Radial wrist	**15**	3	0	3	2
Medial wrist	**12**	6	0	1	2
Finger extension	**5**	0	0	0	2
Finger flexion	**8**	0	2	0	2
Thumb extension	10	6	0	4	2
Other hand signs:					
Dupuytren's contracture	0	0	5	0	0
Heberden's nodes	13	6	7	12	16
Finkelstein's test positive	5	3	0	5	0
Thenar muscle wasting	2	0	0	0	2
Hypothenar muscle wasting	2	0	5	0	1
Weakness of thumb abduction	3	3	0	1	5
Weakness of thumb opposition	8	3	0	4	7

Table 44b: Prevalence of sensorineural symptoms and signs in the second five Hill Lane wrist/hand clusters

Characteristics	**Cluster**				
	6: Wrist pain	**7: NT – palmar fingers**	**8: NT – distal**	**9: NT – all not thumb**	**10: NT – all**
N	**60**	**31**	**44**	**74**	**83**
Colour in dendrogram	Blue	Cyan	Purple	Pink	Dark red
Any numbness/ tingling (according to Katz hand cluster definition):					
Little/ ring finger	0	**100**	**100**	**100**	**100**
Middle/ index finger	0	**100**	**100**	**84**	**100**
Thumb	0	**0**	**0**	**0**	**100**
Palm/ Dorsum	0	**0**	**0**	**100**	**100**
Palmar Aspect	0	**100**	**100**	**69**	**100**
Dorsal Aspect	0	**0**	**100**	**100**	**65**
Proximal/ Middle phalanges	0	**100**	**0**	**100**	**100**
Sensorineural signs:					
Abnormal sensation:					
thumb	0	3	0	0	1
Index finger	2	6	2	1	1
little finger	0	6	0	5	1
Phalen's test positive	8	**45**	**36**	**45**	**47**
Tinel's test positive	2	16	9	4	8
Sleep disturbed because of numbness/tingling	0	**29**	**18**	**28**	**43**

Cluster 13 (N=30) was the last cluster identified with symptoms of numbness or tingling only: in this group they were in the palm or dorsum only and were accompanied by few other signs or symptoms.

The final two clusters were characterised by both pain and numbness or tingling, along with musculoskeletal and sensorineural signs. Cluster 14 (N=15) was the smallest cluster identified and contained wrists/hands with multiple sites of wrist pain as well as some thumb, forearm and joint pain. All wrists/hands had sites of tenderness and most had swelling and multiple sites of pain on resisted movement. Over half of these wrists/hands had Heberden's nodes or a positive Finkelstein's test. Muscle wasting and weakness of thumb movement were also seen most commonly in this cluster. Sensorineural signs were common in this group, with sleep disturbance being the most prevalent (60%) followed by a positive Tinel's test (40%) and a positive Phalen's test (33%). Symptoms of numbness or tingling were common, although there was a less clear pattern of these symptoms than in other clusters. Cluster 15 (N=30) was characterised by musculoskeletal signs and symptoms mainly associated with the radial aspect of the wrist and thumb (emboldened in Table 43c) alongside a mixed profile of numbness and tingling. The most common sensorineural signs of a positive Phalen's test and disturbed sleep were again evident in this group of hands with prevalences of 43% and 33% respectively.

Table 43c: Prevalence of musculoskeletal symptoms and signs in the third five Hill Lane wrist/hand clusters

Characteristics	**Cluster**				
	11: NT – thumb, *index*, middle	**12: NT – all plus signs**	**13: NT – hand only**	**14: All**	**15: Radial wrist, thumb and NT**
N	**19**	**23**	**30**	**15**	**30**
Colour in dendrogram	Dark purple	Dark Grey	Black	Brown	Light Grey
Any pain:	*5*	*13*	*17*	*100*	*100*
Dorsal forearm	0	0	0	20	7
Palmar forearm	0	0	3	**40**	0
Dorsal wrist	0	9	0	**100**	17
Palmar wrist	0	13	3	**93**	10
Radial wrist	0	0	7	**93**	**43**
Medial wrist	5	0	0	**100**	20
Thumb base	0	0	0	**47**	**93**
Finger joints	0	0	7	20	17
Other (mostly hand)	0	0	0	0	7
Any tenderness:	*0*	*17*	*3*	*100*	*83*
Dorsal forearm		0	0	0	0
Palmar forearm		4	0	7	7
Dorsal wrist		4	0	13	3
Palmar wrist		4	0	20	3
Radial wrist		0	0	**33**	**27**
Medial wrist		0	0	**67**	3
Thumb base		13	0	**53**	**63**
Finger joints		0	0	13	10
Other (mainly hand)		4	3	0	3
Any swelling:	*0*	*0*	*7*	*87*	*23*
Dorsal forearm			3	13	0
Palmar forearm			0	0	3
Dorsal wrist			0	27	7
Palmar wrist			0	27	0
Radial wrist			0	**33**	10
Medial wrist			0	**47**	7
Thumb base			0	0	7
Finger joints			7	20	10
Other (mainly hand)			0	0	0
Pain on resisted movement:	*0*	*4*	*3*	*87*	*57*
Radial wrist		0	0	**60**	**23**
Medial wrist		0	0	**67**	10
Finger extension		0	0	**53**	0
Finger flexion		0	0	**40**	3
Thumb extension		4	3	**60**	**47**
Other hand signs:					
Dupuytren's contracture	0	4	0	7	0
Heberden's nodes	**26**	**26**	13	**60**	17
Finkelstein's test positive	0	0	3	**60**	17
Thenar muscle wasting	5	4	0	**27**	10
Hypothenar muscle wasting	0	13	10	**33**	3
Weakness of thumb abduction	5	4	0	**47**	**27**
Weakness of thumb opposition	0	9	3	**33**	**27**

Table 44c: Prevalence of sensorineural symptoms and signs in the third five Hill Lane wrist/hand clusters

Characteristics	Cluster				
	11: NT – thumb, index, middle	12: NT – all plus signs	13: NT – hand only	14: All	15: Radial wrist, thumb and NT
N	19	23	30	15	30
Colour in dendrogram	Dark purple	Dark Grey	Black	Brown	Light Grey
Any numbness/ tingling (according to Katz hand cluster definition):					
Little/ ring finger	0	96	0	80	87
Middle/ index finger	100	100	0	73	83
Thumb	100	70	0	33	17
Palm/ Dorsum	0	70	100	93	53
Palmar Aspect	100	100	100	73	90
Dorsal Aspect	100	83	100	100	80
Proximal/ Middle phalanges	100	78	0	87	67
Sensorineural signs:					
Abnormal sensation:					
thumb	5	78	10	7	3
Index finger	17	91	7	13	3
little finger	0	87	20	7	3
Phalen's test positive	37	78	10	33	43
Tinel's test positive	11	52	0	40	10
Sleep disturbed because of numbness/tingling	11	35	17	60	33

6.4.2 Replication of the wrist/hand clusters (Bitterne data only)

The wrist/hand examination data from the Bitterne practice (2757 wrists/hands with complete information) were clustered using the two methods previously employed on the Hill Lane analysis. Again, the group average 10-cluster solution (Figure 35b) was rejected in favour of the Ward's method 13-cluster solution (Figures 35a and 36, note the break in the y-axis), being less mathematically robust and needing 60% more moves in the k-means refinement procedure than the Ward's method solution. The group average 10-cluster solution consisted of five main clusters containing 1951, 563, 62, 94 and 35 wrists/hands, and five smaller clusters containing 11, 7, 2, 15 and 17 wrists/hands. The dendrogram yielded by Ward's method was similar to that seen for the Hill Lane data (Figure 33a): there were two main overriding clusters which were subdivided into either a 6-cluster solution or a 13-cluster solution. The 13-cluster solution was chosen for further exploration because the 6-cluster solution did not produce clusters with distinct characteristics. For the 13-cluster solution, 392 (14.2%) inter-cluster moves were made by the k-means procedure, 120 of which were from Cluster 5 (coloured light blue in Figure 42a) into Cluster 1 (red). Another 70 inter-cluster moves were from Cluster 5 into Cluster 3 (coloured yellow).

Figure 35a: Dendrogram of Ward's hierarchical clustering on Bitterne data wrist/hand examination observations

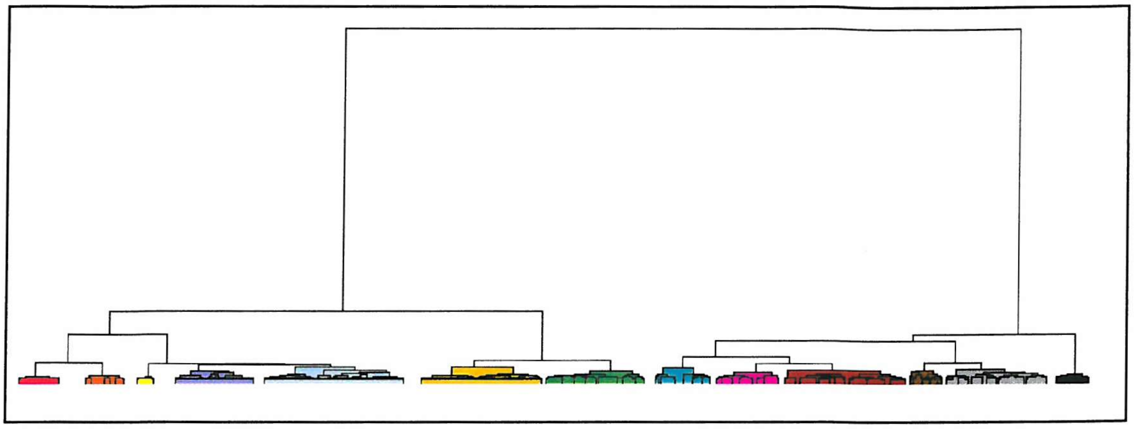


Figure 35b: Dendrogram of group average linkage hierarchical clustering on Bitterne data wrist/hand examination observations

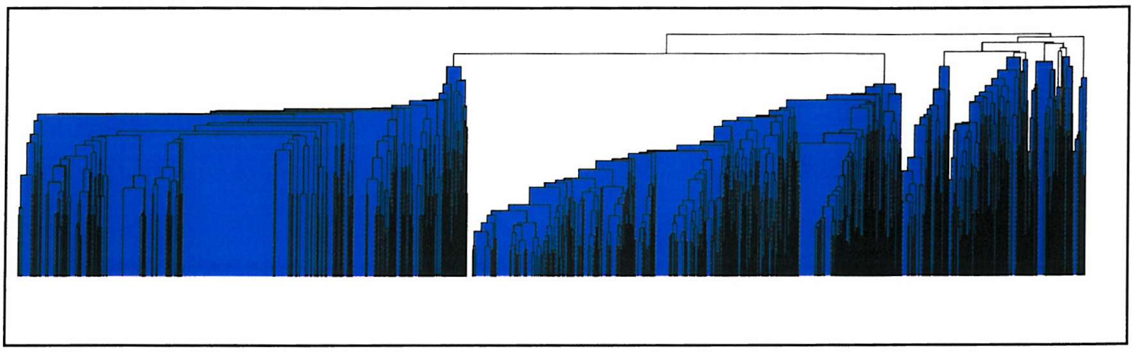
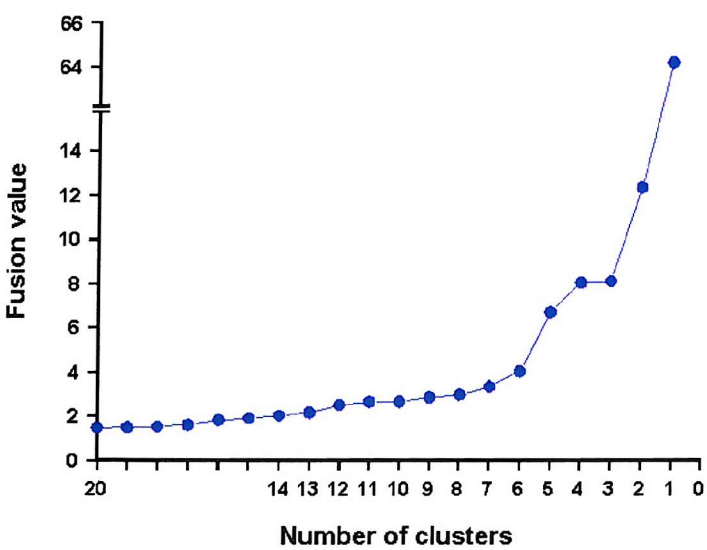


Figure 36: Fusion values for successive wrist/hand cluster joins



The same wrist/hand data were then analysed using the k-means procedure with the 15 Hill Lane cluster centroids as the seed points. The two cluster solutions for the Bitterne data were compared using the kappa statistic and Rand index (Table 45).

Table 45: Comparison of the Bitterne wrist/hand clusters obtained by two different methods

K-means clustering using Hill Lane seed points: **Hierarchical (Ward's method) clustering plus k-means:**

Kappa statistic

Clusters:	1	2	3	7	4	6	8	N
Colour in dendrogram	Red	Oran-ge	Yellow	Light Green	Lilac	Dark Yell-ow	Cyan	
1: "Normal"	**1352**	0	0	0	26	0	0	1418
2: Phalen	0	**104**	0	0	0	0	0	105
3: Heberden	0	21	**294**	0	70	0	0	429
4: Radial wrist & thumb	0	0	0	**67**	0	11	0	80
5: Joint	0	0	0	9	**1**	62	0	75
6: Wrist pain	0	2	0	0	0	**76**	0	78
7: NT – palmar fingers	0	0	0	0	0	0	**82**	85
8: NT – distal	0	0	0	0	0	0	14	15
9: NT – all not thumb	0	0	0	0	0	0	0	96
10: NT – all	0	0	0	0	0	0	0	153
11: NT – thumb, index, middle	0	0	0	0	0	0	0	24
12: NT – all plus signs	0	0	0	0	0	0	6	21
13: NT – hand only	0	0	0	0	0	0	0	85
14: All	0	0	0	0	0	0	0	17
15: Radial wrist, thumb and NT	0	0	0	0	0	0	7	76
N	1352	127	294	76	97	149	109	2757

Clusters (cont.):	5	9	10	13	11	12	N
Colour in dendrogram	Light Blue	Pink	Dark Red	Black	Brown	Light Grey	
1: "Normal"	40	0	0	0	0	0	1418
2: Phalen	0	0	0	0	1	0	105
3: Heberden	44	0	0	0	0	0	429
4: Radial wrist & thumb	2	0	0	0	0	0	80
5: Joint	1	0	0	2	0	0	75
6: Wrist pain	0	0	0	0	0	0	78
7: NT – palmar fingers	0	0	0	0	1	2	85
8: NT – distal	**0**	0	0	0	1	0	15
9: NT – all not thumb	0	**92**	0	0	2	2	96
10: NT – all	0	1	**144**	0	1	7	153
11: NT – thumb, index, middle	0	0	23	0	0	1	24
12: NT – all plus signs	0	2	9	0	1	3	21
13: NT – hand only	0	0	0	**85**	0	0	85
14: All	0	0	0	3	**14**	0	17
15: Radial wrist, thumb and NT	0	8	3	2	7	**49**	76
N	87	103	179	92	28	64	2757

κ=0.80

Table 45 (continued): Comparison of the Bitterne wrist/hand clusters obtained by two different methods

K-means clustering using Hill Lane seed points:	**Hierarchical (Ward's method) clustering plus k-means:**

Rand index/Jaccard statistic

Pairs of hands:	In the same cluster	In different clusters	N
In the same cluster	996,565	141,197	1,137,762
In different clusters	23,706	2,637,678	2,661,384
N	1,020,271	2,778,875	3,799,146
Rand index=95.7% , Jaccard statistic=85.8%			

Rand index/Jaccard statistic*

Pairs of hands:	*In the same cluster*	*In different clusters*	*N*
In the same cluster	83,289	51,965	135,254
In different clusters	23,706	827,350	851,056
N	106,995	879,315	986,310
Rand index=92.3% , Jaccard statistic=52.4%			

*Removing the 1352 wrists/hands in the 'Normal' clusters according to both clustering algorithms

The observed agreement between the clusters (numbers in bold) was 86%, whilst the expected was 28%, yielding a κ of 0.80. The agreement between every pair of hands as to whether they were placed in the same cluster or in different clusters from each other was 95.7% according to the Rand index, and 85.8% according to the Jaccard statistic. These values were adjusted to 92.3% and 52.4% respectively when the 1352 hands in the 'Normal' cluster according to both clustering algorithms had been removed.

These indices suggested a reasonable level of agreement, and thus replicability of the clusters. A total of ten clusters were identified in both the clustering solutions, and another two identified in the Hill Lane dataset, 'Joint' and 'Wrist pain' were amalgamated into one cluster in the direct clustering of the Bitterne dataset (coloured dark yellow in Figure 35a). Three clusters identified in Hill Lane 'Numbness or tingling in the distal fingers only', 'Numbness or tingling in the thumb, index and middle fingers' and 'Numbness or tingling throughout the hand plus sensorineural signs' were not identified in the direct clustering of the Bitterne dataset. Wrists/hands in those clusters according to the Hill Lane cluster solution were assigned to other clusters characterised by numbness or tingling. Two clusters were identified by direct clustering of the Bitterne dataset but not in the Hill Lane cluster solution: one characterised by finger joint tenderness and swelling and Heberden's nodes (the lilac coloured cluster in Figure 35a), the other by thumb base tenderness, Heberden's nodes, some thumb base swelling and pain on thumb extension (the light blue cluster in Figure 35a).

6.4.3 Cluster analysis of the wrist/hand examination data from the whole population

The wrist/hand data from the whole population (N=4149) were clustered using Ward's method on squared Euclidean distance and group average linkage on Jaccard's coefficient of community. Ward's method yielded 14 clusters (Figures 37a and 38, notice the breaks in the y-axis). Both the Hill Lane and Bitterne individual analyses had identified ten of these clusters, a further three were seen in the Hill Lane analysis only and one was observed in the Bitterne analysis only (note the corresponding colours in Figures 33a, 35a and 37a).

Group average linkage produced a 13-cluster solution (Figure 37b) consisting of five large clusters (of sizes 2842, 919,124, 101 and 80 wrists/hands) and eight smaller clusters (of sizes 21, 18, 14, 10, 8, 6, 3 and 3 wrists/hands). This solution was mathematically unstable, requiring 1062 (25.6%) inter-cluster moves to refine it, and retaining the characteristics of only seven of the 13 clusters originally identified. The 14-cluster solution yielded by Ward's method was refined by the k-means procedure, which made 610 (14.7%) inter-cluster moves, 422 of which were from Cluster 3 to Clusters 1 or 4 (coloured light blue, red and yellow in Figure 37a).

Figure 37a: Dendrogram of Ward's hierarchical clustering on both Hill Lane and Bitterne data wrist/hand examination observations

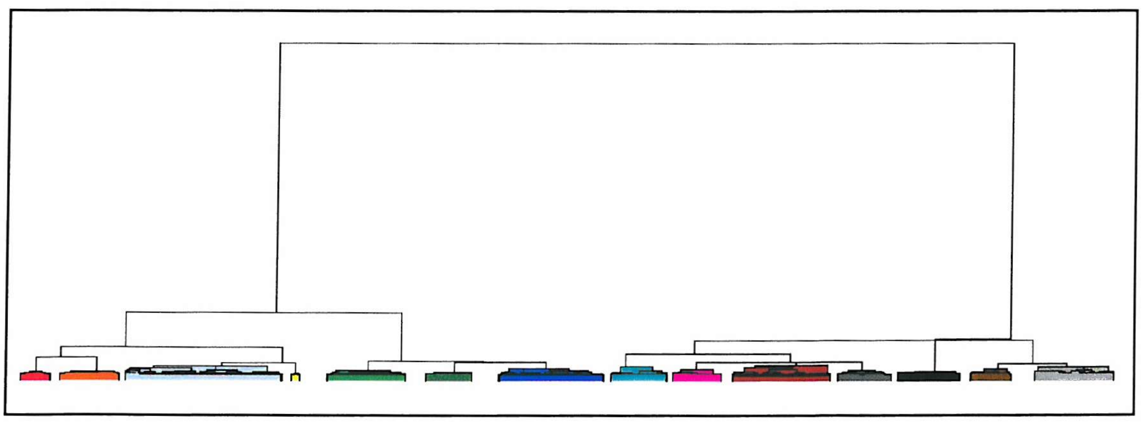


Figure 37b: Dendrogram of group average linkage hierarchical clustering on both Hill Lane and Bitterne data wrist/hand examination observations

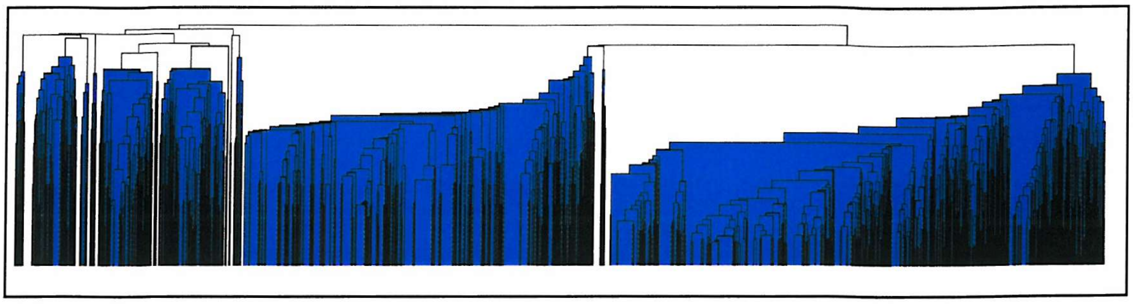
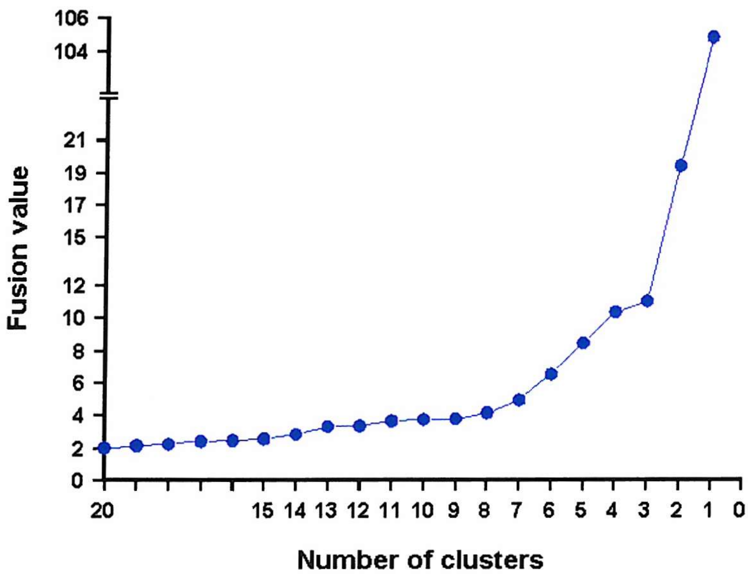


Figure 38: Successive fusion values for the wrist/hand clusters



The characteristics of the finalised wrist/hand clusters are presented in Appendix V. Clusters 1 and 2 (N=2115 and N=194 respectively) were characterised, as before, by minimal abnormality and by a positive Phalen's test only. Cluster 3 (N=123) comprised wrists/hands with thumb tenderness, 47% had Heberden's nodes and 24% had pain on thumb extension. Cluster 4 (N=428) identified wrists/hands with Heberden's nodes and few other signs or symptoms.

Clusters 5, 6 and 7 contained wrists/hands with pain and musculoskeletal signs but minimal sensorineural signs or symptoms. Cluster 5 (N=140) was characterised by radial wrist and thumb involvement and 35% had Heberden's nodes. Cluster 6 (N=102) identified hands with finger joint involvement, whilst Cluster 7 (N=126) identified wrists/hands with predominantly wrist pain and some accompanying tenderness, swelling or pain on resisted movement.

The remaining clusters all had symptoms of numbness or tingling. Cluster 8 (N=173) comprised wrists/hands with few musculoskeletal signs or symptoms (other than Heberden's nodes), but numbness or tingling in the fingers, particularly on the palmar aspect. Of these, 32% also had a positive Phalen's test. Cluster 9 (N=179) identified wrists/hands with numbness or tingling at most sites other than the thumb, and 38% had a positive Phalen's test. Cluster 10 (N=268) was characterised by numbness or tingling at all sites, 41% had a positive Phalen's test and 39% of these subjects reported disturbed sleep due to numbness or tingling. Both Clusters 9 and 10 had few musculoskeletal signs or symptoms other than Heberden's nodes.

Cluster 11 (N=48) was the second smallest cluster, and identified wrists/hands with numbness or tingling accompanied by abnormal sensation in the thumb, index or little fingers. Signs of Phalen's test, Tinel's test and disturbed sleep were also in evidence (56%, 29% and 38% respectively). Cluster 12 (N=118) contained wrists/hands with numbness or tingling in the palm and dorsum only. Just over a fifth of the subjects in this group reported disturbed sleep due to numbness or tingling, and only 8% of these wrists/hands had a positive Phalen's test.

Cluster 13 (N=39) was the smallest cluster identified, and comprised wrists/hands with multiple musculoskeletal and sensorineural symptoms and signs. Cluster 14 (N=96) was characterised by musculoskeletal signs and symptoms, particularly at the thumb and numbness or tingling mostly in the fingers. Half of these wrists/hands had a positive Phalen's test, or else the subjects reported disturbed sleep due to numbness or tingling.

The demographic characteristics of subjects in these wrist/hand clusters, history of numbness, tingling and wrist pain at baseline and laterality were explored (Tables 46a, 46b, 46c). A higher proportion of women was seen in nearly all of the non-'Normal' clusters compared to the 'Normal' one, in particular Clusters 6 ('Joint') and 13 ('All'), with 73% and 87% respectively. Nine of the non-'Normal' clusters had fewer than 10% of subjects in the lowest age band, with Clusters 3 and 4 having only 1% and Cluster 13 having 0%. These clusters had an excess of subjects in the highest age band compared to the 'Normal' cluster.

Table 46a: Demographic and baseline characteristics in the fourteen wrist/hand clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Phalen**	**3: Thumb signs**	**4: Heberden**	**5: Radial wrist & thumb**
N	**2115**	**194**	**123**	**428**	**140**
Colour in dendrogram	Red	Orange	Light Blue	Yellow	Light Green
Demographic characteristics:					
% female	56	68	71	68	71
% 25 – 34 yrs	21	17	1	1	6
% 35 – 44 yrs	29	32	18	6	9
% 45 – 54 yrs	29	34	29	35	31
% 55 – 64 yrs	21	16	52	58	54
Report of numbness or tingling in the fingers or hand or wrist pain at baseline*:					
% numbness/tingling present	28	51	32	32	30
% wrist pain present	14	25	48	21	76
% activities difficult /impossible	8	15	33	15	61
% pain > 1 yr	7	12	33	10	51
Laterality – Number (%) of hands whose pair was:					
In the same cluster	1640 (78)	84 (43)	46 (37)	234 (55)	66 (47)
In the 'Normal' cluster	-	54 (28)	29 (24)	69 (16)	34 (24)
In a different cluster (not 'Normal')	449 (21)	55 (28)	45 (37)	112 (26)	39 (28)
Not in the analysis	26 (1)	1 (1)	3 (2)	13 (3)	1 (1)

* Numbness or tingling at baseline was not reported separately for the right and left side

Whilst symptoms of numbness or tingling reported at baseline were common (a prevalence of at least 28% in each cluster), those clusters characterised by patterns of numbness or tingling at the physical examination had far higher baseline prevalences (80% and over). Similarly, those clusters characterised by wrist/hand pain (Clusters 5, 6, 7, 13 and 14) had higher proportions of subjects with a history of wrist/hand pain at baseline.

Table 46b: Demographic and baseline characteristics in the fourteen wrist/hand clusters

Characteristics	**Cluster**				
	6: Joint	**7: Wrist**	**8: NT – palmar fingers**	**9: NT – all not thumb**	**10: NT – all**
N	**102**	**126**	**173**	**179**	**268**
Colour in dendrogram	Dark green	Blue	Cyan	Pink	Dark red
Demographic characteristics:					
% female	73	62	54	60	68
% 25 – 34 yrs	8	25	5	13	9
% 35 – 44 yrs	16	29	23	28	21
% 45 – 54 yrs	29	25	45	31	39
% 55 – 64 yrs	47	21	28	27	31
Report of numbness or tingling in the fingers or hand or wrist pain at baseline*:					
% numbness/tingling present	33	33	88	87	91
% wrist pain present	75	81	30	25	24
% activities difficult /impossible	53	58	26	20	19
% pain > 1 yr	45	42	18	14	15
Laterality – Number (%) of hands whose pair was:					
In the same cluster	34 (33)	40 (32)	110 (64)	96 (54)	166 (62)
In the 'Normal' cluster	24 (24)	50 (40)	24 (14)	47 (26)	46 (17)
In a different cluster (not 'Normal')	42 (41)	36 (29)	37 (21)	34 (19)	52 (19)
Not in the analysis	2 (2)	0	2 (1)	2 (1)	4 (1)

* Numbness or tingling at baseline was not reported separately for the right and left side

Symptom-sign profiles were most frequently similar enough within subject to be in the same cluster (1323 subjects, of whom 820 had two 'normal' wrists/hands) or unilateral (449 subjects). A further 273 subjects had two non-normal wrists/hands in different clusters from each other, although there was no discernible pattern to these. The remaining 59 subjects included in the analysis contributed only one wrist/hand with complete examination data.

Table 46c: Demographic and baseline characteristics in the fourteen wrist/hand clusters

Characteristics	**Cluster**			
	11: NT – all plus signs	**12: NT – hand only**	**13: All**	**14: Thumb and NT**
N	**48**	**118**	**39**	**96**
Colour in dendrogram	Dark Grey	Black	Brown	Light Grey
Demographic characteristics:				
% female	52	46	87	68
% 25 – 34 yrs	8	20	0	5
% 35 – 44 yrs	17	19	18	17
% 45 – 54 yrs	46	27	46	29
% 55 – 64 yrs	29	34	36	49
Report of numbness or tingling in the fingers or hand or wrist pain at baseline*:				
% numbness/tingling present	96	84	90	84
% wrist pain present	44	36	79	80
% activities difficult /impossible	40	20	69	68
% pain > 1 yr	25	22	64	56
Laterality – Number (%) of hands whose pair was:				
In the same cluster	22 (46)	48 (41)	22 (56)	38 (40)
In the 'Normal' cluster	9 (19)	46 (39)	3 (8)	14 (15)
In a different cluster (not 'Normal')	17 (35)	23 (19)	12 (31)	42 (44)
Not in the analysis	0	1 (1)	2 (5)	2 (2)

* Numbness or tingling at baseline was not reported separately for the right and left side

6.4.4 Comparison of the wrist/hand clusters with the HSE wrist/hand diagnoses

The SEP was designed to diagnose carpal tunnel syndrome on the basis of pain, paraesthesia or sensory loss in a median nerve distribution with an accompanying sign: Tinel's test positive, Phalen's test positive, disturbed sleep in the last 7 days due to numbness or tingling, wasted thenar eminence, weakness of thumb abduction or weakness of thumb opposition. Paraesthesia in a median nerve distribution was originally defined as self-reported numbness or tingling in two of the thumb, index and middle fingers but not the palm or dorsum. No explicit mention was made of the ring or little fingers. Sensory loss in a median nerve distribution was defined as abnormal sensation (diminished or absent) in the thumb and index finger, but not in the little finger. Pain in a median nerve distribution was defined as reported pain on the palmar side of the hand only and not including that located at finger joints. This rather weak definition was used because more detailed information about the location of hand pain was not collected in this study.

In a hospital setting, the majority of subjects' diagnoses with carpal tunnel syndrome based on these criteria arose from symptoms of numbness or tingling rather than from sensory loss or pain. However, the validity of these criteria was poor, and further investigation was indicated (Section 1.2.2.). Data on patterns of numbness or tingling from this study were therefore explored in relation to other clinical findings suggestive of carpal tunnel syndrome and known risk factors [124].

The findings of this work indicated that there was a small group of subjects with one or both hands experiencing numbness or tingling in the majority or all of the palmar aspect of the thumb, index and middle fingers, with no little finger, dorsal aspect of the hand or palm involvement. This group of subjects may have had numbness or tingling in the palmar aspect of their ring finger(s). Such hands had a higher prevalence of positive Phalen's and Tinel's tests than other patterns of numbness or tingling, and were the only ones to be statistically significantly associated with occupational repetitive wrist and finger movements compared to hands with no numbness or tingling. These hands with extensive numbness and tingling in the median nerve distribution were associated with neither a higher prevalence of neck pain, nor with lower mental health or vitality scores according to the SF-36 questionnaire compared to hands with no numbness or tingling.

It was concluded from this work that the more stringent criterion of numbness or tingling in the majority or all of the palmar aspect of the thumb, index and middle fingers, with no little finger, dorsal aspect of the hand or palm involvement contributed to an empirically based case definition which was more discriminatory in a community setting than the original criterion (keeping the accompanying criteria for clinical signs the same as before) and that this definition should be adopted in this study. Symptoms of sensory loss and pain in a median nerve distribution were defined as they had been originally, partly because they led to few additional cases of carpal tunnel syndrome, and because more accurate data on the distribution of these symptoms were not available.

Diagnoses of osteoarthritis (OA) of the thumb base and of the finger joints (distal interphalangeal joints, DIP) were not part of the HSE criteria set, but the hand examination allowed for these diagnoses to be made. Symptomatic DIP OA was defined as Heberden's nodes present alongside pain in the DIPs, and thumb base OA was defined as pain and tenderness located at the thumb base.

These three diagnoses, along with tenosynovitis and de Quervain's disease were compared with the wrist/hand examination clusters (Tables 47a, 47b and 47c).

Table 47a: Prevalence of diagnoses amongst the first five wrist/hand clusters

Diagnoses (%)	**Cluster**				
	1: "Normal "	**2: Phalen**	**3: Thumb signs**	**4: Heberden**	**5: Radial wrist & thumb**
N	**2115**	**194**	**123**	**428**	**140**
Colour in dendrogram	Red	Orange	Light Blue	Yellow	Light Green
None	100	100	100	99.8	23.6
de Quervain's disease					3.6
Tenosynovitis					0.7
CTS*				0.2	
DIP OA					0.7
Thumb base OA					52.9
DIP + thumb base OA					6.4
dQ* + thumb base OA					4.3
Tenosynovitis + thumb base OA					7.1
dQ* +DIP + thumb base OA					0.7

*CTS=carpal tunnel syndrome, dQ=de Quervain's disease of the wrist

Table 47b: Prevalence of diagnoses amongst the second five wrist/hand clusters

Diagnoses (%)	**Cluster**				
	6: Joint	**7: Wrist**	**8: NT – palmar fingers**	**9: NT – all not thumb**	**10: NT – all**
N	**102**	**126**	**173**	**179**	**268**
Colour in dendrogram	Dark green	Blue	Cyan	Pink	Dark red
None	39.2	82.5	92.5	99.4	93.7
de Quervain's disease	1.0	1.6			
Tenosynovitis		15.1	1.2		0.4
CTS*	1.0		4.6	0.6	6.0
DIP OA	52.9		1.7		
Thumb base OA	1.0				
dQ* + tenosynovitis		0.8			
DIP + thumb base OA	3.9				
Tenosynovitis + DIP OA	1.0				

*CTS=carpal tunnel syndrome, dQ=de Quervain's disease of the wrist

All clinical diagnoses of de Quervain's disease were made in Clusters 13 (16 cases), 5 (12 cases), 14 (5 cases), 7 (3 cases), 6 and 11 (1 case each). Diagnoses of tenosynovitis were predominantly made in Clusters 13 (28 cases), 7 (20 cases) and 5 (11 cases) with Clusters 14, 12, 8, 6, 10 and 11 having a further 5, 3, 2, 1, 1 and 1 cases respectively. Thus musculoskeletal disorders of the wrist were concentrated in

Cluster 13 'All', Cluster 5 'Radial wrist and thumb', Cluster 7 'Wrist' and Cluster 14 'Thumb and numbness or tingling'.

Clinical diagnoses of painful DIP osteoarthritis (OA) were predominantly made in Cluster 6 'Joint' (59 cases), with a further 18 cases in Cluster 14, 11 cases in Cluster 5, five cases in each of Clusters 11 and 13, three in Cluster 8 and one case in Cluster 12. Diagnoses of thumb base OA were made in Cluster 5 'Radial wrist and thumb' (100 cases), Cluster 14 (49 cases), Cluster 13 (24 cases), with Clusters 6, 12 and 11 having a further 5, 4 and 2 cases respectively. DIP and thumb base OA were not generally diagnosed together: 22 hands were diagnosed with both, whilst 242 hands were given one or the other diagnosis only. Overall, DIP OA was concentrated in Cluster 6 'Joint' and Cluster 14 'Thumb and numbness or tingling' whilst thumb base OA was concentrated in Cluster 5 'Radial wrist and thumb', Cluster 13 'All' and Cluster 14.

Table 47c: Prevalence of diagnoses amongst the last four wrist/hand clusters

Diagnoses (%)	**Cluster**			
	11: NT – all plus signs	**12: NT – hand only**	**13: All**	**14: Thumb and NT**
N	**48**	**118**	**39**	**96**
Colour in dendrogram	Dark Grey	Black	Brown	Light Grey
None	70.8	92.4	7.7	31.3
de Quervain's disease	2.1			1.0
Tenosynovitis		2.5	17.9	1.0
CTS*	14.6	0.9		1.0
DIP OA	8.3	0.9		14.6
Thumb base OA	2.1	3.4	7.7	39.6
dQ* + tenosynovitis			7.7	
DIP + thumb base OA				1.0
dQ* + DIP OA			5.1	
dQ* + thumb base OA			5.1	2.1
CTS* + thumb base OA				2.1
Tenosynovitis + thumb base OA			23.1	2.1
dQ* + DIP + thumb base OA			2.6	1.0
dQ* + CTS + thumb base OA				1.0
Tenosynovitis + dQ* + thumb base OA			17.9	
Tenosynovitis + DIP + thumb base OA	2.1		2.6	2.1
Tenosynovitis + dQ* + DIP + thumb base OA			2.6	

*CTS=carpal tunnel syndrome, dQ=de Quervain's disease of the wrist

Carpal tunnel syndrome was diagnosed in 8 of the 14 clusters: predominantly in Cluster 10 (16 cases), Cluster 8 (8 cases) and Cluster 11 (7 cases), with Cluster 14 containing four more cases and Clusters 4, 6, 9 and 12 having one case each. Whilst the diagnoses of CTS in Clusters 8, 10 and 14 were based on symptoms of numbness or tingling in the majority of the median nerve distribution according to the revised diagnostic criteria, all diagnoses made in Cluster 11 were based on symptoms of sensory loss in the median nerve distribution. The single diagnoses of carpal tunnel syndrome made in Clusters 4 and 12 were also based on symptoms of sensory loss accompanied by either weakness of thumb opposition or disturbed sleep due to hand symptoms. The diagnoses made in Clusters 6 and 9 were based on palmar hand pain alongside either weakness of thumb opposition or positive Phalen's and Tinel's tests. Thus carpal tunnel syndrome was concentrated in Cluster 10 'Numbness/tingling – all', Cluster 8 'Numbness/tingling – palmar fingers' and Cluster 11 'Numbness/tingling – all plus signs'. Carpal tunnel syndrome was not often diagnosed in the presence of any of the other four wrist/hand diagnoses: it was diagnosed once alongside de Quervain's disease and thumb base OA, once alongside thumb base OA and 37 times alone.

Cluster 13 'All' identified wrists/hands of whom the majority had a clinical diagnosis (7.7% had no diagnosis), and two- thirds (66.7%) had two or more wrist/hand diagnoses. There were no diagnoses of carpal tunnel syndrome in this cluster. Wrists/hands with no diagnosis were characterised by multiple sites of tenderness, pain on resisted movements and other hand signs, but either did not report pain, or had pain that did not lead to a clinical diagnosis (Appendix V, Tables A13 and A14).

Around 70-80% of Clusters 5 'Radial wrist and thumb' and 14 'Thumb and numbness/ tingling' had clinical diagnoses, and in each cluster these were a mixture of the different musculoskeletal wrist/hand diagnoses, with few cases of CTS. Both Clusters had more thumb base OA with some DIP OA, de Quervain's and tenosynovitis diagnoses.

Over half of Cluster 6 'Joint' had a diagnosis of DIP OA, whilst 40% of them had no diagnosis. These subjects without a diagnosis had pain and tenderness in their finger joints, but no Heberden's nodes.

Each of the five diagnoses were present amongst the wrists/hands in Cluster 11 'Numbness/ tingling – all plus signs', although cases of carpal tunnel syndrome were the most frequent. Just over 70% of the wrists/hands in this cluster had no diagnosis, however.

Only 17.5% of the wrists/hands in Cluster 7 'Wrist' had a clinical diagnosis, and these were nearly all cases of tenosynovitis. The rest of the cluster was characterised by wrist pain, but with a variety of accompanying signs that did not indicate any clinical diagnosis.

The remaining clusters all contained under 10% of wrists/hands with a clinical diagnosis. The diagnoses made in these clusters were either carpal tunnel syndrome (Clusters 10 and 8) or were small numbers of diagnoses that were not necessarily reflective of the main characteristics of the cluster (Clusters 4, 9 and 12).

The cluster analysis solution was driven by both the location and nature of symptoms and signs, as were the clinical diagnoses, and thus they showed some agreement: both classification schemes generally differentiated musculoskeletal from sensorineural patterns of disease; both classified symptoms of DIP OA separately from thumb base OA; and both discriminated between OA and tenosynovitis (the exception being Clusters 5 and 14). Neither schemes separated de Quervain's disease of the wrist from tenosynovitis very well, with 32% of de Quervain's diagnoses being made alongside a diagnosis of tenosynovitis, and both diagnoses occurring separately and together in Clusters 5, 13 and 14. Thus tenosynovitis at the radial aspect of the wrist and that located at other regions of the wrist either occurred concurrently, or could not be differentiated by either of these two classification schemes.

Clearly, cluster analysis at the wrist/hand identified more symptom-sign profiles than there were clinical diagnoses, and this was partly a reflection of the complexity of the analysis (using 57 wrist/hand examination variables) and partly a reflection that the wrist/hand clinically-driven diagnoses did not comprise an exhaustive set of possible conditions. It is also of note that although some of the clusters identified tallied well with the clinically-driven diagnoses explored, large proportions of those clusters still did not get a diagnosis according to the criteria (such as Clusters 6 and 7).

6.5 Summary

Cluster analysis of the neck signs (ranges of movement) and symptoms (pain) identified four main groups characterised by no pain and a normal range of movement; pain and a normal range of movement; no pain and a lower range of movement; pain and a lower range of movement. The full range of high through to low ranges of movement was seen in both the pain-positive and the pain-negative subjects. Diagnoses of cervical spondylosis were mainly seen in the 'pain and a low range of movement' cluster.

Four main shoulder profiles were identified by the cluster analysis, based primarily on the severity of the shoulder disorder rather than on clinically recognised underlying pathology. Clinically-driven diagnoses did not make this latter distinction either, suggesting that the examination proforma was not specific enough to distinguish different pathologies, or that such discrimination could not be made in this setting. There was little discrimination between the ranges of shoulder movement seen in the pain-positive and pain-negative shoulder clusters.

Cluster analysis performed on signs and symptoms recorded at the elbow identified groups of disorders that corresponded well with the clinical entities of lateral and medial epicondylitis, and were mathematically robust when alternative cluster analysis techniques were employed. A pattern of signs and symptoms suggestive of olecranon bursitis was also seen, although this was less clear. Four other clusters were identified which had no associated clinical diagnoses: one was characterised by pain only, another by signs located at the medial epicondyle, another by signs at the lateral epicondyle, and the other by an absence of symptoms or signs.

The wrist/hand cluster analysis identified fourteen groups based on location and nature of symptoms and signs. Wrist/hands with sensorineural symptoms and signs were largely distinct from those with musculoskeletal disorders. The five main groups identifying musculoskeletal symptoms and signs only were predominantly based on constellations of symptoms and signs at particular sites (thumb, joints, wrist, radial wrist and thumb), with only one cluster based on a single sign (Heberden's nodes). One of the six clusters characterised by sensorineural disturbance only was based on a single sign (Phalen's test), whilst the others were characterised by a variety of patterns of numbness or tingling accompanied by some sensorineural signs. Two clusters were characterised by both musculoskeletal and sensorineural symptoms and signs. Diagnoses of tenosynovitis and de Quervain's disease of the wrist were not separated by either the cluster analysis or the clinically-driven diagnoses. Carpal tunnel syndrome was diagnosed amongst a variety of patterns of sensory disturbance, and was not frequently diagnosed alongside other disorders of the wrist/hand. DIP and thumb base OA were distinguished from each other by both classification systems.

Replication of the clusters was tested in a second population and the majority of clusters was found to be mathematically robust (the shoulder and wrist/hand analyses being the least robust). Final cluster analyses were performed on upwards of 4149 limbs (2141 necks). The majority of the clusters identified made clinical sense and represented a wide spectrum of disease. In particular, groups of 'mild' disease involvement were identified at each anatomical location, as well as a group of 'normal' limbs/necks at each site. More severe disease involvement was also represented, and the cluster analyses were sensitive to these cases. Where possible, analyses took into account the greater influence of some variables over others, and employed a refining method to the original hierarchical analysis. Alternative methods of analysis were compared.

CHAPTER 7: RESULTS IV

VALIDITY OF THE CLUSTERS AND MEDICAL DIAGNOSES

7 Results IV: Validity of the Clusters and Medical Diagnoses

Having characterised and established the repeatability of the clusters, their validity alongside that of the medical diagnoses was investigated. In the absence of a gold standard classification system and diagnostic criteria, the construct validity of these disease categories was assessed. It was hypothesised that different valid categories of musculoskeletal disease have

i) varying levels of associated morbidity (measured in this study in terms of disability and healthcare utilisation)

ii) different causal pathways. A number of physical occupational and psychosocial factors have been indicated as putative risk factors for musculoskeletal disease, as well as individual psychological factors. In this study occupational physical activities, work support, demand and control and SF-36 measures of vitality and mental health were collected at the baseline questionnaire, and were explored.

7.1 Form of model of analysis

Analyses were performed separately for the neck, shoulder, elbow and wrist/hand. The external measures of disability, healthcare utilisation, putative risk factors and individual psychological factors were treated as dependent binary 'outcome' measures in logistic regression models, with the medical diagnoses and clusters as independent 'predictors' in separate analyses. In the cases of disability and healthcare utilisation this was intuitively the correct form of relationship to model. For putative risk factors, however, it would be usual to think of them as possible predictors of musculoskeletal disorders, not the other way around. In the case of psychological profile, the association could be argued to work in either direction. However, analyses in a cross-sectional study setting investigate association rather than inferring causation, and thus it was acceptable to consider the relationship between different classifications of musculoskeletal disorders and putative risk factors using this model. In addition, odds ratios have the mathematical property of being symmetrical, so we would expect the odds ratio of being exposed to a certain risk factor given a particular class of musculoskeletal disorder to be similar (but not identical if adjustment factors were included in the model) to the odds ratio of having a particular class of disorder given exposure to the risk factor.

7.2 Outcome Measures

Self-reported disability in the previous 12 months due to pain or numbness/ tingling was recorded in the clinical interview and was site-, but not side- specific (Appendix II). Disability was defined as a report that any of a specific list of activities (different for each site) was impossible. It is of note that subjects only answered these questions if they also reported pain or numbness/ tingling at the postal questionnaire or clinical interview at the relevant site. Thus subjects who did not report symptoms on either occasion could not report symptom-related disability and were assumed to have no disability at that site, whereas subjects who did not have symptoms at the time of the examination could still report disability due to recently resolved pain or numbness/ tingling.

Similarly, subjects reported their presentation to a doctor and any prescribed treatment they received (medication, injection or operation) in the last 12 months due to symptoms at each site but not according to side. Other aspects of healthcare utilisation were recorded at the clinical interview, but have not been investigated in this thesis.

These three morbidity measures could not be assessed among those subjects who responded to the postal questionnaire only, and such subjects were not included in the analyses.

Physical activities at work were recorded in the postal questionnaire (Appendix III). These were investigated in relation to the site most likely to be relevant. Thus, having the neck bent forward for longer than two hours or twisted for longer than half an hour per day was investigated in relation to the clusters and medical diagnoses of the neck; holding hands above shoulder height for longer than one hour per day, and carrying weights on one shoulder or carrying weights of 5kg or more in one hand were investigated in relation to the shoulder clusters and medical diagnoses; repeatedly bending and straightening the elbow for longer than one hour per day was investigated in relation to the elbow clusters and medical diagnoses; and using a keyboard for more than four hours each day, and performing tasks involving repetitive movements of the wrists or fingers for longer than four hours in total per day were investigated in relation to the wrist/hand clusters and medical diagnoses.

Whether the right, left or both sides were exposed to these mechanical activities was not recorded.

Work demand, support and control were also investigated in the postal questionnaire. High work demand was defined as having a target number of articles or tasks to make or finish in the day, or receiving bonus payments if more than an agreed number of items were produced in the day or receiving payment according to the number of items produced in the day. Poor work support was defined as support from colleagues and immediate supervisor seldom or never being offered when there were difficulties at work. Poor work control was defined as being in the bottom quarter of the distribution of an overall control score which was made up of the summation of the following individual scores (possible responses were 'often', 'sometimes', 'seldom', 'never/almost never'): how often there was a choice in being able to decide how to do the work, how often there was a choice in being able to decide what to do at work, and how often there was a choice in being able to decide work timetables and breaks. These three aspects of psychosocial work environment were investigated in relation to the clusters and medical diagnoses of all four anatomical sites.

The putative mechanical and psychosocial occupational risk factors were assessed in all the postal questionnaire responders in addition to those who underwent the physical examination, and those subjects who did not report symptoms at the postal questionnaire and were not invited to attend the physical examination were included in these analyses as a distinct control group.

Poor mental health and vitality were investigated using questions from the SF-36 [80] (Appendix III) and were defined as the lowest third of the scores in each distribution. As for the occupational data, these measures were assessed in the postal questionnaire and thus a control group of asymptomatic postal questionnaire responders was included in the analyses.

7.3 Categorisation of the clusters and medical diagnoses within subject

The cluster analyses and medical diagnoses assessed each shoulder, elbow and wrist/hand separately, and produced limb-based rather than person-based classifications. However, the data used to validate the clusters were person specific, and thus, the limb-based classifications needed to be transformed into person-specific categories. These will be described at each location in turn.

7.3.1 The neck

The already person-specific neck classifications were refined in order to reflect the interrelation between the neck, shoulder, elbow and wrist/hand. The baseline 'Normal' neck cluster was divided into two mutually exclusive groups: those subjects with no neck abnormalities alongside no abnormalities at any of the other three sites (two sides at each site), and those subjects with no neck abnormalities but at least one other site being in a non-normal cluster. Thus five exclusive and exhaustive person-based neck cluster classifications were investigated for their validity:

1 - Normal at all sites (N=291)
2 - Normal at the neck, but abnormal at another site (N=484)
3 - Restricted range of neck movement (regardless of the status of other sites) (N=787)
4 - Neck pain (regardless of the status of other sites) (N=345)
5 - Neck pain and restricted range of neck movement (regardless of the status of other sites) (N=234).

The medical diagnosis at the neck was also refined, with those subjects not having a diagnosis of cervical spondylosis being subdivided into three mutually exclusive groups:

1 - a group with no diagnoses and no symptoms or clinically significant signs at any of the sites (N=443);
2 - a group with no cervical spondylosis and no pain at the neck, but with some symptoms, clinically significant signs or a specific diagnosis at another site (N=1119);
3 - a group with neck pain but without the diagnosis of cervical spondylosis (i.e. non-specific neck pain), regardless of symptoms, signs or diagnoses at other sites (N=472).

This categorisation was deemed to be in line with a clinically-based assessment of common musculoskeletal neck disorders.

For those analyses that included the postal questionnaire asymptomatic subjects a baseline 'control' category was used in both the cluster classification and the medically based categorisation.

7.3.2 The shoulder

The limb-based shoulder cluster analysis identified four groups: 'Normal', 'pain', 'signs' and 'pain with signs'. Of the 2114 subjects who were included in this analysis, 1616 (76%) had symmetrical profiles, 411 (19%) had unilateral shoulder disorders (one shoulder was in the 'Normal' cluster), 38 (2%) had only one shoulder included in the analysis, and 49 (2%) had shoulders in different non-normal clusters (Table 36). Since the clusters were differentiated by severity, it seemed reasonable to use the more severe shoulder for the overall subject classification. Thus the 411 subjects with unilateral shoulder disorders took on their non-normal shoulder's cluster. The 38 with only one shoulder in the analysis took on that shoulder's classification. Of the 49 subjects with each shoulder in different non-'normal' clusters, 39 of them had one shoulder in the 'pain and signs' cluster, and this worse classification (which still reflected the characteristics of the less severe shoulder, but added to them the characteristics of the more severe shoulder) was used. The remaining 10 subjects all had one shoulder in the 'pain' cluster and the other shoulder in the 'signs' cluster, and as a whole person, these subjects were classified as 'pain and signs', reflecting the profiles of both shoulders.

Those subjects in the 'normal' group were divided into those who had normal profiles at every site (the identical group of subjects to those in the baseline neck cluster), and those who had normal profiles at each shoulder, but a non-normal profile at other sites. Thus five exclusive and exhaustive person-based shoulder cluster classifications were investigated for their validity:

1 - Normal at all sites (N=291)
2 - Normal at both shoulders, but abnormal at another site (N=1216)
3 – Shoulder signs (regardless of the status of other sites) (N=196)
4 - Shoulder pain (regardless of the status of other sites) (N=233)
5 – Shoulder pain and signs (regardless of the status of other sites) (N=178).

Five medically derived shoulder diagnoses were used in the study. There was a lot of overlap in the diagnoses (Table 37), especially between rotator cuff tendonitis and shoulder capsulitis, and there were only a few shoulders with bicipital tendonitis (N=25, of which 3 were diagnosed without any other shoulder diagnosis), acromioclavicular joint disorder (N=38, of which 2 were diagnosed alone) and subacromial bursitis (N=37, of which 2 were diagnosed alone). Due to the overlap, an overall classification of 'any shoulder diagnosis' was investigated rather than the individual diagnoses. The remaining subjects were classified according to whether they had 'non-specific' shoulder pain (with or without shoulder signs), 'non-specific' clinically significant signs (those signs which contributed to diagnostic criteria) without shoulder pain, or no clinically significant shoulder signs or symptoms. Notice that a hierarchy of pain being more severe, or more 'important' than signs was therefore assumed.

Those subjects with no shoulder signs or symptoms were further divided into a group with no diagnoses and no symptoms or clinically significant signs at any of the sites (the identical group of subjects to those in the baseline neck medical diagnosis group) and a group without a shoulder diagnosis and no pain or clinically significant signs at either shoulder, but some symptoms, clinically significant signs or specific diagnosis at another site. As in the shoulder cluster categorisation, the worst shoulder was used for each subject. The numbers of subjects in each category were as follows:

Category 1 - no diagnoses and no symptoms or clinically significant signs at any of the sites - N=443;
Category 2 - no shoulder diagnosis and no pain or clinically significant signs at either shoulder, but some symptoms, clinically significant signs or specific diagnosis at another site - N=1000;
Category 3 - non-specific shoulder signs only, regardless of symptoms, signs or diagnoses at other sites - N=251;
Category 4 – non-specific shoulder pain, regardless of symptoms, signs or diagnoses at other sites - N=67;
Category 5 – specific shoulder diagnosis – N=353.

For those analyses that included the postal questionnaire asymptomatic subjects the baseline 'control' category was used in both the cluster classification and the medically based categorisation.

7.3.3 The elbow

Seven limb-based elbow clusters were identified: 'Normal', 'Pain only' (lateral, medial or posterior), 'Medial signs', 'Posterior /antecubital fossa', 'Medial symptoms and signs', 'Lateral signs', and 'Lateral symptoms and signs'. Of the 2140 subjects included in the cluster analysis, 1761 (82%) had symmetrical profiles, 329 (15%) had unilateral elbow disorders (one elbow was in the 'normal' cluster), 17 (0.8%) had only one elbow included in the analysis and 33 (2%) had elbows in different non-'normal' clusters (Table 41). As for the shoulder classifications, the worse elbow cluster was used to classify those subjects with unilateral elbow disease involvement, or missing data. Of the 33 subjects with two different non-'normal' elbow profiles, three had medial signs in one elbow alongside medial signs and symptoms in the other, and four had lateral signs in one elbow alongside lateral signs and symptoms in the other. The more severe elbow classifications were used for these seven subjects. The remaining 26 subjects were given a new classification denoting their mixed bilateral profile of elbow disease involvement.

The subjects in the 'normal' group were divided into those who had normal profiles at every site (as before in the neck and shoulder data-driven subject classifications), and those who had normal profiles at both elbows, but a non-normal profile at other sites. Thus nine exclusive and exhaustive person-based elbow cluster classifications were investigated for their validity:

1 - Normal at all sites (N=291)
2 – Normal at both elbows, but abnormal at another site (N=1390)
3 – Elbow pain only (N=55)
4 – Medial elbow signs (N=64)
5 – Posterior/ Antecubital fossa (N=53)
6 - Medial symptoms and signs (N=45)
7 - Lateral signs (N=134)
8 - Lateral symptoms and signs (N=82)
9 – Mixed bilateral profile (N=26).

The medical diagnoses were medial epicondylitis (N=41 elbows), lateral epicondylitis (N=50) and olecranon bursitis (N=7) (Table 42). Due to the small number of subjects who fulfilled criteria for any of these, an overall classification of 'any elbow diagnosis' was used. The remaining subjects were classified according to whether they had non-specific elbow pain (with or without elbow signs), non-specific clinically significant signs (those signs which contributed to diagnostic criteria) without elbow pain, or no clinically significant elbow signs or symptoms.

Those subjects with no clinically significant elbow signs or symptoms were divided into two groups: one with no diagnoses and no symptoms or clinically significant signs at any of the sites (as in the neck and shoulder medical diagnosis categorisations) and a group without an elbow diagnosis and no pain or clinically significant signs at either elbow, but some symptoms, clinically significant signs or specific diagnosis at another site. The worse elbow was used for each subject. The numbers of subjects in each elbow medical diagnosis category were as follows:

Category 1 - no diagnoses and no symptoms or clinically significant signs at any of the sites - N=443;
Category 2 - no elbow diagnosis and no pain or clinically significant signs at either elbow, but some symptoms, clinically significant signs or specific diagnosis at another site - N=1234;
Category 3 - non-specific elbow signs only - N=229;
Category 4 – non-specific elbow pain - N=158;
Category 5 – specific elbow diagnosis – N=76.

For those analyses that included the postal questionnaire asymptomatic subjects the baseline 'control' category was used in both the cluster classification and the medically based categorisation.

7.3.4 <u>The wrist/hand</u>

Fourteen wrist/hand clusters were identified, based on location and nature of disease involvement. Of the 2104 subjects included in the cluster analysis, 1323 (63%) had symmetrical profiles, 449 (21%) had unilateral wrist/hand disorders (one wrist/hand was in the 'normal' cluster), 59 (3%) had only one wrist/hand included in the analysis and 273 (13%) had wrists/hands in different non-'normal' clusters (Tables 46a, 46b and 46c). The more severe wrist/hand disease involvement was used as the subject classification in those subjects with unilateral wrist/hand profiles or missing data. Of the 273 subjects with bilateral but not symmetrical wrist/hand profiles, 39 had one wrist/hand in Cluster 14, and the other wrist/hand in a non-'normal' cluster with fewer reported symptoms or signs (Cluster 2 – Cluster 12), and were categorised with the more severe profile. Similarly, 9 subjects had one wrist/hand in Cluster 13, and the other wrist/hand in a non-'normal' cluster with fewer reported symptoms or signs (Cluster 2 – Cluster 12), and were categorised again with the more severe profile. Three subjects had one wrist/hand in the 'Thumb signs' cluster and the other wrist/hand in the 'Radial wrist and thumb' cluster and were categorised as an individual with the latter more severe profile. The remaining 222 subjects were categorised according to the divisions identified in the hierarchical cluster analysis (see Figure 37a), into a 'Mixed signs' group (both wrists/hands in two of Clusters 2, 3 or 4, N=42), into a 'Mixed musculoskeletal' group (both wrists/hands in two of Clusters 5, 6 or 7, N=16), and into a 'Mixed numbness or tingling' group (both wrists/hands in two of Clusters 8, 9, 10, 11 or 12, N=29). All remaining subjects (N=135) were classified as 'Mixed – other', and consisted of those subjects with wrists/hands in two non-'normal' clusters that were not in the same overriding cluster defined above. This last group of individuals was therefore highly heterogeneous compared to the other subject-based wrist/hand cluster categories.

As for the other sites, subjects in the 'normal' group were divided into those who had normal profiles at every site, and those who had normal profiles at both wrists/hands, but a non-normal profile at other sites. Thus 19 exclusive and exhaustive person-based wrist/hand cluster classifications were investigated for their validity:

1 – Normal at all sites (N=291)
2 – Normal at both wrists/hands, but abnormal at another site (N=555)
3 – Positive Phalen's test only (N=97)
4 – Thumb signs (N=55)
5 – Heberden's nodes (N=199)
6 – Radial wrist and thumb symptoms and signs (N=71)
7 – Finger joint symptoms and signs (N=43)
8 – Wrist symptoms and signs (N=70)
9 – Numbness and/or tingling in the palmar fingers (N=81)
10 - Numbness and/or tingling throughout the hand except for the thumb (N=97)
11 - Numbness and/or tingling throughout the hand (N=133)
12 - Numbness and/or tingling throughout the hand alongside signs (N=20)
13 - Numbness and/or tingling in the palm or dorsum (N=71)
14 - All (N=25)

15 - Numbness and/or tingling and radial wrist and thumb involvement (N=74)
16 – Bilateral – mixed numbness and/or tingling profiles (N=29)
17 – Bilateral – mixed musculoskeletal profiles (N=16)
18 – Bilateral – mixed signs (N=42)
19 – Bilateral – mixed other (N=135).

Five medically derived wrist/hand diagnoses were used in the study. Tenosynovitis and de Quervain's disease of the wrist were not differentiated well at examination (32% of de Quervain's disease diagnoses were made alongside a diagnosis of tenosynovitis) and these two diagnoses were combined into an 'any tenosynovitis' diagnosis for the validity analyses. Although the two sites of OA were distinguished by the examination, an overall diagnosis of 'any OA' was investigated. This was partly because the underlying disease process is the same at the two sites, and the disability, healthcare utilisation, risk factors and psychological profiles associated with them might be expected to be similar, and partly in order to keep the analyses simple where possible. Carpal tunnel syndrome was also investigated as a separate diagnosis. Multiple diagnoses within an individual (whether on the same, or different sides) were investigated as separate categories.

The remaining subjects were categorised according to whether they had non-specific wrist/hand pain, non-specific numbness and/or tingling, or both (regardless of physical signs). Subjects with neither diagnoses nor symptoms were classified into three final mutually exclusive groups: those with clinically significant signs (i.e. signs which contributed to one of the five diagnoses); those with no symptoms or clinically significant signs at the wrist/hand, but some pain, clinically significant signs or specific diagnoses at another site; and those with no symptoms or clinically significant signs anywhere in the neck or upper limb (as before). The numbers of subjects in each person-based wrist/hand category were thus:

Category 1 - no diagnoses and no symptoms or clinically significant signs at any of the sites - N=443;
Category 2 - no wrist/hand diagnosis and no pain or clinically significant signs at either wrist/hand, but some symptoms, clinically significant signs or specific diagnosis at another site - N=314;
Category 3 - non-specific wrist/hand signs only - N=460;
Category 4 - non-specific wrist/hand numbness and/or tingling - N=395;
Category 5 - non-specific wrist/hand pain – N=143;

Category 6 - non-specific numbness and/or tingling and pain - N=93;
Category 7 - any tenosynovitis - N=37;
Category 8 - any OA – N=147;
Category 9 – carpal tunnel syndrome - N=28;
Category 10 - any tenosynovitis plus any OA - N=40;
Category 11 – carpal tunnel syndrome plus any OA – N=3;
Category 12 – carpal tunnel syndrome plus any tenosynovitis plus any OA – N=1.

Clearly, Categories 11 and 12 had too few subjects to yield meaningful results in the statistical models of risk for each outcome, but they were included so that the observed prevalence of each outcome could be presented.

7.4 Construct validity of the neck clusters and medical diagnosis

Severe disability due to neck pain was reported by 4.6% of subjects who underwent the clinical interview. A graded relationship between each neck disorder category and disability was observed for both classification systems, with the medical classifications yielding more extreme odds ratios (Table 48). Restriction of neck movement alone was associated with a small increase in odds ratio, although this was not significantly different from unity (odds ratio (OR) 1.5, 95% confidence interval (95%CI) 0.5 - 4.0). It should be noted that these two classification systems led to very similar neck disorder categories (the two categories characterised by pain in each system contained the same 579 subjects) driven predominantly by neck pain status. Having a restricted range of neck movement in addition to neck pain or cervical spondylosis gave rise to notably elevated odds ratios (OR 13.8, 95%CI 5.2 – 36.7 and OR 20.0, 95%CI 8.4 - 47.7 respectively) and it appeared that the medically driven classification system (restricted range of neck movement was age and sex adjusted) identified more severe disease involvement.

Table 48: Association of disability due to neck pain with neck disorders

Classification system	N	% reporting disability	Odds ratio (95% Confidence Interval)
Neck clusters (data-driven)			
Normal all sites	290	1.7	1
Normal neck	480	1.0	0.6 (0.2 – 1.9)
Restricted ROM*	781	2.7	1.5 (0.5 – 4.0)
Pain	345	4.9	2.6 (0.9 – 7.2)
Pain & restricted ROM	234	21.4	13.8 (5.2 – 36.7)
Neck diagnoses (medically-driven)			
No symptoms or signs all sites	441	1.6	1
No neck pain	1110	2.2	1.1 (0.5 – 2.7)
Non-specific neck pain	472	7.8	4.3 (1.9 – 9.8)
Cervical spondylosis	107	28.0	20.0 (8.4 – 47.7)

*ROM=range of movement. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems (data-driven and medically-driven) were performed separately.

A similar pattern of association was seen between the two classification systems and seeing a doctor and receiving prescribed medication for neck pain (Table 49). In these two outcomes, however, restricted range of movement alone was also associated with a statistically significant increase in odds ratio, indicating that an excess of these subjects had experienced neck pain in the recent past (which caused them to seek and receive medical treatment) compared to the baseline neck disorder category. This may indicate that part of the natural history of neck pain that comes to the attention of health services is its resolution accompanied by residual restricted range of movement. The odds ratios for those subjects with no neck pain, but some disease involvement elsewhere in the upper limb were raised for both of these outcomes and in both classification systems, indicating again that there was some excess of recently resolved neck pain among subjects with other upper limb complaints.

Occupational mechanical neck activity (holding the neck bent or twisted) was common (49.9% in all postal responders). The odds ratios should therefore be viewed in this light: while they remain a valid measure of association, they cannot be seen as accurate estimates of relative risk. All subjects who were called forward for examination had statistically significantly higher odds ratios for mechanical neck use compared to the postal asymptomatic subjects (Table 50), and among these, subjects with neck pain with restricted movement in the data-driven classification scheme, and subjects with cervical spondylosis in the medically driven classification system had the highest odds ratios (OR 3.3, 95%CI 2.3 – 4.9 and OR 3.6, 95%CI 2.1 – 6.4 respectively).

Table 49: Association of healthcare use due to neck pain with neck disorders

Classification system	N	% seeing doctor	Odds ratio (95% Confidence Interval)
Neck clusters (data-driven)			
Normal all sites	290	2.4	1
Normal neck	480	5.2	2.3 (1.0 – 5.3)
Restricted ROM*	781	9.9	4.6 (2.1 – 10.3)
Pain	345	31.3	18.7 (8.5 – 41.1)
Pain & restricted ROM	234	52.2	47.1 (21.0 – 105.8)
Neck diagnoses (medically-driven)			
No symptoms or signs all sites	441	5.2	1
No neck pain	1110	7.8	1.5 (0.9 – 2.4)
Non-specific neck pain	472	34.1	9.1 (5.7 – 14.6)
Cervical spondylosis	107	64.5	32.2 (18.0 – 57.6)
		% treated**	
Neck clusters (data-driven)			
Normal all sites	290	1.0	1
Normal neck	480	3.1	3.0 (0.9 – 10.6)
Restricted ROM*	781	7.0	7.1 (2.2 – 23.2)
Pain	343	17.2	19.5 (6.0 – 63.2)
Pain & restricted ROM	234	43.2	71.4 (21.9 – 232.8)
Neck diagnoses (medically-driven)			
No symptoms or signs all sites	441	2.5	1
No neck pain	1110	5.6	2.1 (1.1 – 4.0)
Non-specific neck pain	470	21.9	10.2 (5.3 – 19.3)
Cervical spondylosis	107	53.3	41.7 (20.4 – 85.1)

*ROM=range of movement. **prescribed medication. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table 50: Association of occupational mechanical neck activity with neck disorders

Classification system	N	% exposed	Odds ratio (95% Confidence Interval)
Neck clusters (data-driven)			
Control	2126	43.1	1
Normal all sites	246	52.9	1.5 (1.1 – 1.9)
Normal neck	387	55.0	1.7 (1.3 – 2.1)
Restricted ROM*	539	57.7	1.9 (1.6 – 2.3)
Pain	279	66.3	2.7 (2.1 – 3.6)
Pain & restricted ROM	137	70.1	3.3 (2.3 – 4.9)
Neck diagnoses (medically-driven)			
Control	2126	43.1	1
No symptoms or signs all sites	356	54.8	1.6 (1.3 – 2.0)
No neck pain	816	56.3	1.8 (1.5 – 2.1)
Non-specific neck pain	355	66.8	2.8 (2.2 – 3.6)
Cervical spondylosis	61	72.1	3.6 (2.1 – 6.4)

*ROM=range of movement. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Associations between psychosocial exposures (work control, support and demand) and the two neck classification systems were inconsistent (Table 51). Poor work support was statistically significantly associated with the disease categories characterised by neck pain in both systems, whereas poor work control was not associated strongly with any of the disease categories (the odds ratio furthest from unity was seen for non-specific neck pain OR 1.4, 95%CI 1.1 – 1.9). High work demand was most strongly associated with neck pain with restricted range of movement in the data-driven classification scheme (OR 2.5 95%CI 1.7 – 3.7) and with non-specific neck pain in the medically driven classification system (OR 1.7, 95%CI 1.3 – 2.3).

A clear graded relationship was observed between individual psychological factors (vitality and mental health as defined by the SF-36 questionnaire) and both neck classification systems (Table 52). The odds ratios were more extreme for poor vitality than for poor mental health, but very similar in the two systems (OR for vitality in the 'Pain & restricted ROM ' group 6.0, 95%CI 4.5 – 8.0; OR for poor vitality in 'Cervical spondylosis' 5.9, 95%CI 3.9 – 8.9). Notice that the control group is identical in the two analyses.

It was evident that both classification systems discriminated between neck disorders with varying levels of disability, healthcare use, mechanical occupational activity and psychological factors. Since the two systems had almost identical bases, it was to be expected that their validity would be similar. Although a restricted range of neck movement alone was associated with increased healthcare use and disability, this was likely to be via an increased occurrence of resolved neck pain among these subjects.

Little distinction was seen in either classification system with respect to levels of work control, support or demand. This may be because such psychosocial variables do not vary between distinct neck disorders, because the study measurements of these variables were inadequate, or that the classification systems did not show construct validity for these measures.

Table 51: Association of psychosocial variables with neck disorders

Classification system	N	% poor work control	Odds ratio (95% Confidence Interval)
Neck clusters (data-driven)			
Control	2126	17.5	1
Normal all sites	246	15.5	0.9 (0.6 – 1.3)
Normal neck	387	22.2	1.3 (1.0 – 1.7)
Restricted ROM*	539	20.6	1.1 (0.9 – 1.5)
Pain	279	23.7	1.4 (1.0 – 1.9)
Pain & restricted ROM	137	26.3	1.5 (1.0 – 2.3)
Neck diagnoses (medically-driven)			
Control	2126	17.5	1
No symptoms or signs all sites	356	16.0	0.9 (0.7 – 1.2)
No neck pain	816	21.8	1.2 (1.0 – 1.5)
Non-specific neck pain	355	24.8	1.4 (1.1 – 1.9)
Cervical spondylosis	61	23.0	1.3 (0.7 – 2.4)
	N	**% poor work support**	**Odds ratio (95% Confidence Interval**
Neck clusters (data-driven)			
Control	2126	10.1	1
Normal all sites	246	10.6	1.2 (0.8 – 1.8)
Normal neck	387	10.6	1.1 (0.8 – 1.6)
Restricted ROM*	539	13.5	1.2 (0.9 – 1.7)
Pain	279	15.8	2.0 (1.4 – 2.8)
Pain & restricted ROM	137	23.4	2.9 (1.8 – 4.5)
Neck diagnoses (medically-driven)			
Control	2126	10.1	1
No symptoms or signs all sites	356	11.5	1.2 (0.8 – 1.7)
No neck pain	816	12.1	1.2 (0.9 – 1.5)
Non-specific neck pain	355	17.2	2.1 (1.5 – 2.8)
Cervical spondylosis	61	24.6	3.6 (1.9 – 6.8)
		% high work demand	
Neck clusters (data-driven)			
Control	2126	17.2	1
Normal all sites	246	12.6	0.7 (0.5 – 1.1)
Normal neck	387	19.9	1.3 (1.0 – 1.7)
Restricted ROM*	539	19.1	1.3 (1.0 – 1.6)
Pain	279	19.4	1.4 (1.0 – 1.9)
Pain & restricted ROM	137	29.2	2.5 (1.7 – 3.7)
Neck diagnoses (medically-driven)			
Control	2126	17.2	1
No symptoms or signs all sites	356	12.6	0.7 (0.5 – 1.0)
No neck pain	816	20.3	1.4 (1.1 – 1.7)
Non-specific neck pain	355	23.1	1.7 (1.3 – 2.3)
Cervical spondylosis	61	19.7	1.4 (0.8 – 2.8)

*ROM=range of movement. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table 52: Association of psychological variables with neck disorders

Classification system	N	% poor vitality	Odds ratio (95% Confidence Interval)
Neck clusters (data-driven)			
Control	2701	20.0	1
Normal all sites	291	25.1	1.2 (0.9 – 1.6)
Normal neck	484	27.1	1.4 (1.2 – 1.8)
Restricted ROM*	787	33.0	2.3 (1.9 – 2.8)
Pain	345	47.0	3.4 (2.7 – 4.3)
Pain & restricted ROM	234	56.0	6.0 (4.5 – 8.0)
Neck diagnoses (medically-driven)			
Control	2701	20.0	1
No symptoms or signs all sites	443	27.1	1.4 (1.1 – 1.8)
No neck pain	1119	30.7	1.9 (1.6 – 2.2)
Non-specific neck pain	472	48.7	3.9 (3.1 – 4.8)
Cervical spondylosis	107	58.9	5.9 (3.9 – 8.9)
		% poor mental health	
Neck clusters (data-driven)			
Control	2701	26.3	1
Normal all sites	291	34.0	1.3 (1.0 – 1.7)
Normal neck	484	33.9	1.4 (1.1 – 1.7)
Restricted ROM*	787	33.7	1.6 (1.4 – 2.0)
Pain	345	44.6	2.2 (1.7 – 2.7)
Pain & restricted ROM	234	48.3	3.1 (2.4 – 4.2)
Neck diagnoses (medically-driven)			
Control	2701	26.3	1
No symptoms or signs all sites	443	33.0	1.3 (1.1 – 1.7)
No neck pain	1119	34.1	1.6 (1.3 – 1.8)
Non-specific neck pain	472	44.9	2.3 (1.9 – 2.9)
Cervical spondylosis	107	51.4	3.2 (2.2 – 4.9)

*ROM=range of movement. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

7.5 Construct validity of the shoulder clusters and medical diagnoses

Severe disability due to shoulder pain was reported by 7.2% of subjects invited to the clinical interview. Statistically significantly elevated odds ratios were seen for shoulder disability in the 'Signs' and 'Pain' data-driven classifications (OR 4.3, 95%CI 1.8 – 10.3, OR 4.4, 95%CI 1.9 – 10.6 respectively) (Table 53). The 'Pain and signs' category had a markedly higher association (OR 22.4, 95%CI 9.7 – 51.5). The medically driven 'Non-specific signs' profile had a higher odds ratio than the 'Non-specific shoulder pain' (OR 4.9, 95%CI 2.4 – 9.9, OR 0.5, 95%CI 0.1 – 4.2 respectively), but having a specific shoulder diagnosis had the highest (OR 12.2, 95%CI 6.3 – 23.5). Note that although the characteristics of the data-driven and medically driven classifications seem concordant, they contain very different numbers of subjects and are not synonymous.

Table 53: Association of disability due to shoulder pain with shoulder disorders

Classification system	**N**	**% reporting disability**	**Odds ratio (95% Confidence Interval)**
Shoulder clusters (data-driven)			
Normal all sites	288	2.4	1
Normal shoulders	1208	1.9	0.7 (0.3 – 1.6)
Signs	194	12.4	4.3 (1.8 – 10.3)
Pain	232	11.2	4.4 (1.9 – 10.6)
Pain & signs	178	40.5	22.4 (9.7 – 51.5)
Shoulder diagnoses (medically-driven)			
No symptoms or signs all sites	438	2.5	1
Normal shoulders	995	1.4	0.5 (0.2 – 1.1)
Non-specific shoulder signs only	249	13.7	4.9 (2.4 – 9.9)
Non-specific shoulder pain	67	1.5	0.5 (0.1 – 4.2)
Any shoulder diagnosis	351	26.2	12.2 (6.3 – 23.5)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

The patterns of association between seeing a doctor and receiving prescribed medication or an injection for shoulder pain with each shoulder disorder were similar to those for disability (Appendix VI, Table A15). For these two outcomes, however, non-specific shoulder pain had higher odds ratios than those for non-specific shoulder signs only, and thus the pattern of association was similar between the two classification systems.

Holding the hands above the shoulders for an hour or longer at work each day was associated more with shoulder pain than physical signs in both classification systems (Table 54). Having non-specific shoulder signs alone, or a complaint in another part of the neck or upper limb, or being called forward for examination all had similar odds ratios, which were elevated but not generally statistically significant. In contrast, carrying weights (a common activity) was associated more with shoulder signs than pain in both classification systems, although having both in the data-driven classification scheme had the highest odds ratio of all (OR 2.7, 95%CI 1.7 – 4.1).

No clear relationship was seen between the shoulder disorders and psychosocial work exposures in either classification scheme (Appendix VI, Table A16).

Table 54: Association of occupational mechanical shoulder activity with shoulder disorders

Classification system	**N**	**% holding hands above shoulders**	**Odds ratio (95% Confidence Interval)**
Shoulder clusters (data-driven)			
Control	2126	7.2	1
Normal all sites	246	9.8	1.5 (0.9 – 2.4)
Normal shoulders	914	10.0	1.6 (1.2 – 2.2)
Signs	138	8.0	1.7 (0.9 – 3.3)
Pain	173	16.8	2.8 (1.8 – 4.5)
Pain & signs	109	18.4	4.0 (2.3 – 6.9)
Shoulder diagnoses (medically-driven)			
Control	2126	7.2	1
No symptoms or signs all sites	356	9.8	1.5 (1.0 – 2.3)
Normal shoulders	764	9.8	1.6 (1.2 – 2.2)
Non-specific shoulder signs only	170	9.4	1.9 (1.1 – 3.3)
Non-specific shoulder pain	54	18.5	3.7 (1.8 – 7.9)
Any shoulder diagnosis	236	16.5	3.0 (2.0 – 4.4)
		% carrying weights	
Shoulder clusters (data-driven)			
Control	2126	30.2	1
Normal all sites	246	34.6	1.3 (1.0 – 1.8)
Normal shoulders	914	36.4	1.6 (1.3 – 1.9)
Signs	138	37.0	2.3 (1.5 – 3.3)
Pain	173	43.9	2.1 (1.5 – 3.0)
Pain & signs	109	44.0	2.7 (1.7 – 4.1)
Shoulder diagnoses (medically-driven)			
Control	2126	30.2	1
No symptoms or signs all sites	356	33.4	1.3 (1.0 – 1.6)
Normal shoulders	764	36.9	1.7 (1.4 – 2.0)
Non-specific shoulder signs only	170	38.8	2.3 (1.6 – 3.3)
Non-specific shoulder pain	54	40.7	2.1 (1.1 – 3.8)
Any shoulder diagnosis	236	44.1	2.3 (1.7 – 3.1)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

The 'Pain and signs' shoulder cluster had the highest odds ratio for poor vitality (OR 5.0, 95%CI 3.4 – 6.9) (Table 55), and this was higher than the odds ratio seen for poor vitality in the 'any shoulder diagnosis' category, as already seen for the disability and healthcare utilisation outcomes. Shoulder disorders characterised by physical signs but no symptoms also had high odds ratios in both the data-driven and the medically based classification schemes (OR 2.9, 95%CI 2.1 - 4.0 and OR 3.5, 95%CI 2.7 – 4.7 respectively). The associations between mental health and the shoulder disorders were similar to those already presented for vitality (Appendix VI, Table A17).

Table 55: Association of vitality with shoulder disorders

Classification system	N	% poor vitality	Odds ratio (95% Confidence Interval)
Shoulder clusters (data-driven)			
Control	2701	20.0	1
Normal all sites	291	25.1	1.2 (0.9 – 1.6)
Normal shoulders	1216	33.3	2.1 (1.8 – 2.4)
Signs	196	41.3	2.9 (2.1 – 4.0)
Pain	233	37.3	2.6 (1.9 – 3.5)
Pain & signs	178	52.3	5.0 (3.4 – 6.9)
Shoulder diagnoses (medically-driven)			
Control	2701	20.0	1
No symptoms or signs all sites	443	27.1	1.4 (1.1 – 1.8)
Normal shoulders	1000	32.3	2.0 (1.7 – 2.3)
Non-specific shoulder signs only	251	45.0	3.5 (2.7 – 4.7)
Non-specific shoulder pain	67	37.3	2.4 (1.5 – 4.1)
Any shoulder diagnosis	353	44.8	3.6 (2.9 – 4.6)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

The two shoulder classification systems were in general agreement in their associations with disability, healthcare utilisation, occupational activities and psychological profile. The most severe data-driven diagnostic category had higher odds ratios with these outcomes than the medical diagnoses and it was clear that the former category was the more stringent (demonstrated by the fewer numbers of subjects eligible to be classified in this group). The more severe disorders in both classification systems had markedly higher odds ratios compared to the other groups in these outcomes, except for holding the hands above shoulder height for an hour or longer at work, which was associated most strongly with non-specific shoulder pain in the medically based classification system. Having a poor psychological state was associated with physical signs more than it was with pain alone, indicating that this clinical profile has its own distinct risk factors and/or sequelae.

7.6 Construct validity of the elbow clusters and medical diagnoses

Disability due to elbow pain was associated most strongly with the data-driven elbow clusters characterised by pain only (at the lateral, medial or posterior elbow, OR 36.9, 95%CI 4.3 – 315.0), lateral symptoms and signs (OR 40.3, 95%CI 5.0 – 326.3) and mixed bilateral profiles (OR 40.5, 95%CI 4.3 – 380.0) (Table 56). All non-normal elbow profiles had elevated odds ratios compared to the baseline profile, although those with lateral disease involvement were associated with a higher occurrence of disability that those with medial or posterior involvement. Having a specific elbow diagnosis had the highest odds ratio, with 21% of these subjects reporting severe disability.

Table 56: Association of disability due to elbow pain with elbow disorders

Classification system	**N**	**% reporting disability**	**Odds ratio (95% Confidence Interval)**
Elbow clusters (data-driven)			
Normal all sites	290	0.3	1
Normal elbows	1384	0.4	0.9 (0.1 – 7.7)
Pain	55	12.7	36.9 (4.3 – 315.0)
Medial signs	64	3.1	5.8 (0.5 – 67.2)
Posterior/ Antecubital Fossa	53	3.8	8.7 (0.8 – 100.5)
Medial symptoms and signs	44	6.8	15.7 (1.5 – 160.3)
Lateral signs	134	5.2	11.5 (1.4 – 97.4)
Lateral symptoms and signs	82	14.6	40.3 (5.0 – 326.3)
Bilateral – mixed profiles	26	19.2	40.5 (4.3 – 380.0)
Elbow diagnoses (medically-driven)			
No symptoms or signs all sites	443	0.2	1
Normal elbows	1227	0.6	2.0 (0.2 – 16.2)
Non-specific elbow signs only	229	4.4	15.6 (2.0 – 124.4)
Non-specific elbow pain	158	7.0	29.0 (3.7 – 230.3)
Any elbow diagnosis	75	21.3	98.1 (12.5 – 770.1)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

The relationship between seeing a doctor, and receiving prescribed medication or an injection with the two classification systems were similar to those seen for disability (Appendix VII, Table A18), although the mixed bilateral profiles had much higher odds ratios for both of these outcomes than any of the other data-driven profiles.

Bending and straightening the elbow repeatedly was a common occupational activity (44.7% reported it in the postal questionnaire). Again, those data-driven profiles that were characterised by pain, except for medial symptoms and signs, had the highest odds ratios for this measure, along with the profiles characterised by lateral elbow signs and mixed bilateral disease involvement (Table 57). Those profiles with non-specific elbow pain, non-specific elbow signs and disease involvement elsewhere in the neck or upper limb according to the medically derived classification system had similar associations with occupational elbow activity (statistically significantly higher than unity), and those subjects with a specific elbow diagnosis had the highest odds ratio within the medically derived disease categories (OR 3.8, 95%CI 2.1 – 6.9).

Table 57: Association of occupational mechanical elbow activity with elbow disorders

Classification system	N	% bending elbow repeatedly	Odds ratio (95% Confidence Interval)
Elbow clusters (data-driven)			
Control	2126	39.5	1
Normal all sites	246	39.0	1.0 (0.8 – 1.3)
Normal elbows	1018	51.0	1.7 (1.5 – 2.0)
Pain	38	68.4	3.2 (1.6 – 6.5)
Medial signs	45	53.3	1.8 (1.0 – 3.3)
Posterior/ Antecubital Fossa	38	60.5	2.4 (1.2 – 4.6)
Medial symptoms and signs	30	53.3	1.8 (0.9 – 3.8)
Lateral signs	93	63.4	2.7 (1.6 – 4.2)
Lateral symptoms and signs	62	69.4	2.6 (2.0 – 6.1)
Bilateral – mixed profiles	20	80.0	6.9 (2.3 – 20.8)
Elbow diagnoses (medically-driven)			
Control	2126	39.5	1
No symptoms or signs all sites	356	42.4	1.2 (0.9 – 1.5)
Normal elbows	912	51.2	1.7 (1.5 – 2.0)
Non-specific elbow signs only	158	60.8	2.4 (1.7 – 3.4)
Non-specific elbow pain	111	64.0	2.7 (1.8 – 4.1)
Any elbow diagnosis	53	69.8	3.8 (2.1 – 6.9)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Mixed associations were seen between the two sets of disease categories and psychosocial work variables (Appendix VII, Table A19), with little distinction between categories for poor work control, and only the data-driven profile of mixed bilateral elbow disease involvement showing a marked increase in odds ratio for work support (OR 6.5, 95%CI 2.5 – 16.9) and high work demand (OR 3.4, 95%CI 1.3 – 8.9).

The data-driven clusters characterised by pain ('Pain', 'Posterior/ Antecubital Fossa', Medial symptoms and signs', 'Lateral symptoms and signs', and 'Bilateral mixed') had the highest odds ratios for reporting poor vitality, followed by the two profiles characterised by signs only (Table 58). Those subjects with disease involvement in another part of the neck or upper limb also had a statistically significantly increased odds ratio for poor vitality compared to the controls (OR 2.2, 95%CI 1.9 – 2.6). In the medically driven categories of elbow disease, non-specific pain had the highest odds ratio for poor vitality (OR 3.6, 95%CI 2.6 – 5.1), followed by a specific elbow diagnosis (OR 3.1, 95%CI 1.9 – 5.0). Even subjects who had no clinically significant symptoms or signs, but attended the physical examination had a significantly higher occurrence of reported poor vitality than the controls (OR 1.4, 95%CI 1.1 – 1.8). In contrast, the patterns of association between the elbow disease categories and mental health were highest in the profiles characterised by signs only and pain only in both the data-driven and the medically driven classification schemes

(Appendix VII, Table A20). The association was particularly high for the data-driven 'Medial signs only' profile (OR 3.2, 95%CI 1.9 – 5.4).

Table 58: Association of vitality with elbow disorders

Classification system	N	% poor vitality	Odds ratio (95% Confidence Interval)
Elbow clusters (data-driven)			
Control	2701	20.0	1
Normal all sites	291	25.1	1.2 (0.9 – 1.6)
Normal elbows	1390	35.2	2.2 (1.9 – 2.6)
Pain	55	43.6	3.2 (1.9 – 5.6)
Medial signs	64	37.5	2.7 (1.6 – 4.6)
Posterior/ Antecubital Fossa	53	43.4	3.5 (2.0 – 6.1)
Medial symptoms and signs	45	51.1	4.5 (2.5 – 8.3)
Lateral signs	134	39.6	2.8 (2.0 – 4.1)
Lateral symptoms and signs	82	40.2	3.0 (1.9 – 4.7)
Bilateral – mixed profiles	26	50.0	6.1 (2.6 – 14.5)
Elbow diagnoses (medically-driven)			
Control	2701	20.0	1
No symptoms or signs all sites	443	27.1	1.4 (1.1 – 1.8)
Normal elbows	1234	36.1	2.3 (2.0 – 2.7)
Non-specific elbow signs only	229	38.0	2.7 (2.0 – 3.6)
Non-specific elbow pain	158	44.9	3.6 (2.6 – 5.1)
Any elbow diagnosis	76	40.8	3.1 (1.9 – 5.0)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

The two elbow classification schemes were in broad agreement in their associations with the outcomes of disability, healthcare use, occupational factors and psychological profile, suggesting that the presence of pain and pain accompanied by physical signs had particularly high odds ratios for these measures. The data-driven classification scheme also indicated that lateral involvement was more closely associated with disability and occupational mechanical elbow activity than other sites of elbow involvement, whilst medial elbow disease involvement was more closely associated with poor vitality and mental health. The data-driven classification scheme identified three profiles not investigated by the medically driven one: pain only, posterior or antecubital fossa involvement and mixed bilateral involvement. The 'Pain only' profile had some of the highest odds ratios for the outcome measures investigated, similar to other profiles involving both symptoms and signs. The 'posterior or antecubital fossa' profile was a heterogeneous group in the original limb-based cluster analysis, but again it had comparatively high odds ratios for healthcare use, occupational mechanical activity and vitality. The mixed bilateral profile had universally high odds ratios, which were higher than those for all other categories in both classification schemes for the outcomes of seeing a doctor, mechanical occupational activity, work support, work demand and vitality. Having a specific elbow diagnosis according to the medically derived classification scheme had higher odds ratios than either of the data-driven classifications of 'Lateral symptoms and signs' or 'Medial symptoms and signs' for disability, healthcare use, work demand and occupational mechanical elbow activity, but not for work control, work support, vitality or mental health.

7.7 Construct validity of the wrist/hand clusters and medical diagnoses

Substantial disability due to wrist/hand symptoms (pain or numbness or tingling) was reported by 8.2% of subjects attending the clinical interview, and was reported more commonly because of wrist/hand pain (7.4%) than because of numbness and/or tingling (2.1%). The data-driven classifications of 'numbness and/or tingling plus physical signs' and 'All' (a group characterised by multiple symptoms and signs at most of the fingers, thumb, hand and wrist) had the highest odds ratios and prevalences of reported disability, with their lower confidence limits of the odds ratios being above 9.0 (Table 59). Odds ratios were generally higher in the groups characterised by pain (other than the aforementioned 'numbness and/or tingling plus physical signs' group), although elevated odds ratios were seen in most groups compared to the baseline category. Having pain and physical signs located at the radial wrist and thumb was associated with more frequent reporting of disability than having pain and physical signs at the wrist or at finger joints alone. Among the medically derived classifications, those groups characterised by a diagnosis of OA and/or tenosynovitis as well as those with non-specific wrist/hand pain and numbness and/or tingling had the highest odds ratios for reported disability (lower confidence limits for the odds ratios were all above 6.0). In both classification systems, having symptoms or signs in another part of the upper limb or neck was not associated with more frequent reporting of disability due to wrist/hand symptoms.

Seeing a doctor because of wrist/hand symptoms was also more prevalent and had higher odds ratios in the data-driven groups characterised by pain, and the highest odds ratios were seen in the 'numbness and/or tingling plus physical signs' (OR 12.5, 95%CI 4.6 – 34.2) and 'All' (OR 13.9, 95%CI 5.5 – 35.1) categories (Appendix VIII, Table A21). In contrast, seeing a doctor was most common in the medically derived group with a diagnosis of carpal tunnel syndrome alone (prevalence 50.0%, OR 18.3, 95%CI 7.8 - 43.0), followed by the groups with tenosynovitis alone and non-specific numbness and/or tingling and pain.

Table 59: Association of disability due to wrist/hand symptoms with wrist/hand disorders

Classification system	N	% reporting disability	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Normal all sites	291	1.7	1
Normal wrists/hands	551	2.2	1.2 (0.4 – 3.4)
Phalen's test positive	96	8.3	4.8 (1.5 – 15.1)
Thumb signs	55	16.4	8.3 (2.6 – 26.7)
Heberden's nodes	196	3.0	1.3 (0.4 – 4.3)
Radial wrist & thumb involvement	71	23.9	13.2 (4.5 – 38.3)
Finger joint symptoms and signs	43	11.6	5.6 (1.5 – 20.7)
Wrist symptoms and signs	70	11.4	6.7 (2.1 – 21.4)
N/T* in the palmar fingers	80	2.5	1.4 (0.3 – 7.3)
N/T* all except thumb	97	9.3	5.6 (1.8 – 17.4)
N/T* all	132	6.8	3.4 (1.1 – 10.6)
N/T* all plus signs	20	35.0	34.4 (9.3 – 127.2)
N/T* palm or dorsum	71	11.3	7.2 (2.3 – 23.1)
All	25	48.0	35.2 (10.5 – 118.0)
N/T* and radial wrist and thumb	74	27.0	16.0 (5.6 – 45.6)
Bilateral – N/T*	27	6.9	3.6 (0.7 – 19.7)
Bilateral – musculoskeletal	16	25.0	14.3 (3.3 – 61.4)
Bilateral – signs	42	9.5	3.9 (1.0 – 15.8)
Bilateral – mixed	135	19.3	10.1 (3.7 – 27.8)
Wrist/hand diagnoses (medically-driven)			
No symptoms or signs all sites	442	1.4	1
Normal wrists/hands	311	1.6	1.2 (0.4 – 3.9)
Non-specific wrist/hand signs	456	6.5	4.2 (1.7 – 10.2)
Non-specific wrist/hand pain	391	7.6	5.5 (2.3 – 13.5)
Non-specific wrist/hand N/T*	143	14.0	11.1 (4.3 – 28.3)
Non-spec. wrist/hand pain & N/T*	93	22.6	18.8 (7.3 – 48.8)
Tenosynovitis	37	24.3	20.3 (6.7 – 61.7)
OA	147	22.5	16.4 (6.6 – 40.9)
Carpal tunnel syndrome	28	7.1	5.7 (1.1 – 30.1)
Tenosynovitis & OA	40	37.5	29.9 (10.5 – 85.4)
Carpal tunnel syndrome & OA	3	66.7	122.7 (8.6 – 1757.1)
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Receiving prescribed treatment (medication, injection or operation) was most common in the 'All' (52%), 'Bilateral - musculoskeletal' (37.5%), 'Numbness and/or tingling and radial wrist and thumb' (36.5%), and 'numbness and/or tingling plus physical signs' (35%) data-driven groups and in the medically derived diagnosis groups involving tenosynovitis (Appendix VIII, Table A22). Receiving prescribed treatment was more commonly associated with groups characterised by pain in the data-driven classification system, although non-specific numbness and/or tingling had a higher odds ratio than non-specific wrist/hand pain (OR 9.6, 95%CI 4.5 – 20.6 and OR 2.7, 95%CI 1.3 – 5.8 respectively) in the medically derived categories.

Table 60: Association of occupational mechanical wrist/hand activity with wrist/hand disorders

Classification system	N	% performing repetitive movements of wrist or fingers	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2126	27.1	1
Normal all sites	246	26.4	1.0 (0.8 – 1.4)
Normal wrists/hands	443	34.3	1.4 (1.1 – 1.8)
Phalen's test positive	84	23.8	0.9 (0.5 – 1.6)
Thumb signs	36	33.3	1.3 (0.7 – 2.7)
Heberden's nodes	123	31.7	1.2 (0.8 – 1.8)
Radial wrist & thumb involvement	44	43.2	2.0 (1.1 – 3.8)
Finger joint symptoms and signs	27	48.2	2.7 (1.2 – 5.8)
Wrist symptoms and signs	54	35.2	1.5 (0.9 – 2.7)
N/T* in the palmar fingers	62	40.3	1.7 (1.0 – 2.9)
N/T* all except thumb	77	28.6	1.1 (0.7 – 1.8)
N/T* all	101	38.6	1.8 (1.2 – 2.7)
N/T* all plus signs	13	38.5	1.5 (0.5 – 4.7)
N/T* palm or dorsum	51	49.0	2.6 (1.5 – 4.7)
All	8	62.5	5.3 (1.3 – 22.6)
N/T* and radial wrist and thumb	46	47.8	2.5 (1.4 – 4.6)
Bilateral – N/T*	22	40.9	2.1 (0.9 – 5.1)
Bilateral – musculoskeletal	11	27.3	1.3 (0.3 – 5.0)
Bilateral – signs	23	47.8	2.4 (1.0 – 5.5)
Bilateral – mixed	101	44.6	2.2 (1.5 – 3.3)
Wrist/hand diagnoses (medically-driven)			
Control	2126	27.1	1
No symptoms or signs all sites	356	29.5	1.2 (0.9 – 1.5)
Normal wrists/hands	263	32.7	1.3 (1.0 – 1.7)
Non-specific wrist/hand signs	321	31.8	1.3 (1.0 – 1.7)
Non-specific wrist/hand pain	305	38.0	1.7 (1.0 – 2.8)
Non-specific wrist/hand N/T*	106	43.4	2.2 (1.5 – 3.3)
Non-spec. wrist/hand pain & N/T*	63	38.1	1.7 (1.0 – 2.8)
Tenosynovitis	24	33.3	1.5 (0.6 – 3.5)
OA	94	45.7	2.3 (1.5 – 3.5)
Carpal tunnel syndrome	19	63.2	3.9 (1.5 – 10.2)
Tenosynovitis & OA	19	42.1	2.1 (0.8 – 5.3)
Carpal tunnel syndrome & OA	1	0.0	-
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

A mixture of statistically significant associations was seen between the data-driven wrist/hand disease categories and frequent use of repetitive movements of the wrist or fingers at work (Table 60). The 'All' profile had the highest odds ratio (OR 5.3, 95%CI 1.3 – 22.6), followed by the numbness and tingling in the palm or dorsum profile (OR 2.6, 95%CI 1.5 – 4.7) and other moderately elevated odds ratios were seen in groups characterised by both musculoskeletal and sensorineural disease involvement. Those subjects with disease involvement in the neck or upper limb, but not at the wrist/hand also had an elevated odds ratio (OR 1.4, 95%CI 1.1 – 1.7). In the medically derived categories, those subjects with a diagnosis of carpal tunnel syndrome only had the highest odds ratio (OR 3.9, 95%CI 1.5 – 10.2), followed by those with a diagnosis of OA only, non-specific numbness and/or tingling, tenosynovitis plus OA, non-specific wrist pain, and non-specific wrist pain and numbness or tingling.

Extensive keyboard use at work was only statistically significantly associated with those subjects characterised by no wrist/hand disease involvement, but some involvement at another site in the neck or upper limb only in both classification systems (Appendix VIII, Table A23). Other moderately elevated odds ratios included those for numbness and tingling and radial wrist and thumb involvement (OR 1.7, 95%CI 0.9 – 3.4), bilateral signs (OR 1.7, 95%CI 0.7 – 4.5), finger joint symptoms and signs (OR 1.6, 95%CI 0.7 – 3.8) and tenosynovitis plus OA (OR 2.1, 95%CI 0.8 - 5.6).

Psychosocial factors showed an inconsistent pattern of association with the wrist/hand disorders (Appendix VIII, Tables A24, A25 and A26). Poor work control was associated most strongly with numbness and/or tingling in the palmar fingers (OR 2.2, 95%CI 1.3 – 3.9) of the data-driven profiles, and with carpal tunnel syndrome of the medically derived diagnoses (OR 3.3, 95%CI 1.3 – 8.3). Other profiles characterised by non-specific pain in the medically derived diagnoses also had slightly elevated odds ratios, as did some of the data-driven profiles involving sensorineural disturbance (numbness and/or tingling in the palm or dorsum, and numbness and/or tingling throughout the hand). Poor work support was associated most strongly with a mixed bilateral numbness and/or tingling profile (OR 3.8, 95%CI 1.4 – 10.2) and the 'All' profile (OR 3.1, 95%CI 0.6 – 15.4) in the data-driven wrist/hand disorders, and with non-specific numbness and/or tingling and pain (OR 2.1, 95%CI 1.1 – 4.2) and OA plus tenosynovitis (OR 2.1, 95%CI 0.6 – 7.7) in the medically derived diagnoses. High work demand was associated most strongly with the 'All' profile of the data-driven system (OR 4.6, 95%CI 1.1 – 19.7), and with carpal tunnel syndrome only (OR 5.7, 95%CI 2.2 – 14.3) in the medically derived diagnoses.

Table 61: Association of vitality with wrist/hand disorders

Classification system	N	% poor vitality	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2701	20.0	1
Normal all sites	291	25.1	1.2 (0.9 – 1.6)
Normal wrists/hands	555	36.2	2.3 (1.9 – 2.9)
Phalen's test positive	97	42.3	2.7 (1.8 – 4.1)
Thumb signs	55	29.1	1.8 (1.0 – 3.2)
Heberden's nodes	199	31.7	2.1 (1.5 – 3.0)
Radial wrist & thumb involvement	71	23.9	1.5 (0.9 – 2.7)
Finger joint symptoms and signs	43	41.9	1.5 (0.9 – 2.7)
Wrist symptoms and signs	70	24.3	3.0 (1.6 – 5.6)
N/T* in the palmar fingers	81	30.9	1.2 (0.7 – 2.1)
N/T* all except thumb	97	46.4	2.0 (1.2 – 3.2)
N/T* all	133	36.1	3.5 (2.3 – 5.4)
N/T* all plus signs	20	35.0	2.3 (1.6 – 3.4)
N/T* palm or dorsum	71	46.5	2.4 (0.9 – 6.2)
All	25	56.0	3.6 (2.2 – 5.8)
N/T* and radial wrist and thumb	74	44.6	4.8 (2.1 – 10.7)
Bilateral – N/T*	29	27.6	1.4 (0.6 – 3.2)
Bilateral – musculoskeletal	16	56.3	5.5 (1.9 – 15.6)
Bilateral – signs	42	35.7	2.6 (1.3 – 4.9)
Bilateral – mixed	135	38.5	2.7 (1.8 – 3.8)
Wrist/hand diagnoses (medically-driven)			
Control	2701	20.0	1
No symptoms or signs all sites	443	27.1	1.4 (1.1 – 1.8)
Normal wrists/hands	314	38.2	2.5 (2.0 – 3.2)
Non-specific wrist/hand signs	460	35.4	2.3 (1.9 – 2.9)
Non-specific wrist/hand pain	395	36.0	2.3 (1.8 – 2.9)
Non-specific wrist/hand N/T*	143	30.1	1.7 (1.2 – 2.5)
Non-spec. wrist/hand pain & N/T*	93	52.7	4.8 (3.1 – 7.4)
Tenosynovitis	37	37.8	2.5 (1.2 – 4.9)
OA	147	41.5	3.3 (2.3 – 4.7)
Carpal tunnel syndrome	28	14.3	0.8 (0.3 – 2.2)
Tenosynovitis & OA	40	45.0	3.5 (1.8 – 6.6)
Carpal tunnel syndrome & OA	3	0.0	-
All three diagnoses	1	100.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

A poor psychological profile was reported in most wrist/hand disorders compared to the baseline controls in both classification schemes for vitality (Table 61) and mental health (Appendix VIII, Table A27). Odds ratios for poor vitality were particularly high in the data-driven bilateral mixed musculoskeletal profile (OR 5.5, 95%CI 1.9 – 15.6), and numbness and/or tingling and radial wrist and thumb involvement (4.8, 95%CI 2.1 – 10.7), but were not statistically significant for the unilateral or bilateral symmetrical musculoskeletal profiles of radial wrist & thumb involvement or finger joint symptoms and signs. The odds ratio for poor vitality among the medically derived diagnoses was highest in the non-specific numbness and/or tingling and pain profile (OR 4.8, 95%CI 3.1 – 7.4), and this was statistically significantly higher than the odds ratios for either the non-specific numbness and/or tingling profile (OR 2.3, 95%CI 1.8 – 2.9) and the non-specific wrist/hand pain profile (OR 1.7, 95%CI 1.2 – 2.5). In contrast, those subjects with a diagnosis of carpal tunnel syndrome only had the lowest odds ratio (OR 0.8, 95%CI 0.3 – 2.2). The 'All', 'Numbness and/or tingling plus signs' and 'Thumb signs' had the highest odds ratios for poor mental health among the data-driven profiles, whilst the pattern of association between the medically derived diagnoses and mental health were similar to those seen for vitality.

The two classification schemes of wrist/hand disorders gave complex information about the association of these conditions with disability, healthcare use, occupational factors and psychological profiles. Whilst the two schemes distinguished broadly between characteristics of musculoskeletal and sensorineural conditions and some of the data-driven profiles tallied well with the medically derived diagnoses (Section 6.4), none of them was identical to a medically derived diagnosis, and this was reflected in the relationships explored in this chapter.

Reported disability, seeing a doctor and use of prescribed treatment each had similar relationships with the wrist/hand conditions, as might be expected. Both classification systems suggested that disorders characterised by pain were more strongly associated with these outcomes than numbness and/or tingling or physical signs. The exception to this was the data-driven disorder characterised by numbness and tingling with accompanying physical signs (sensorineural and thumb signs), which was associated with high occurrences of these morbidity outcomes. While having a diagnosis of carpal tunnel syndrome was statistically significantly associated with all three measures, the odds ratios were not as high for disability and treatment as those for the diagnoses of tenosynovitis or OA; the latter two diagnoses had similar associations with the morbidity measures investigated. There was some suggestion from the data-driven classifications that different sites of pain within the wrist/hand were associated with different levels of morbidity, and that greater disease involvement was associated with greater morbidity.

In contrast to the morbidity measures, occupational activities did not distinguish different wrist/hand conditions well, and were not strongly associated with tenosynovitis, or any of the tenosynovitis-like data-driven profiles, which might have been expected. Keyboard use was more strongly associated with conditions at other parts of the upper limb or neck rather than any of the wrist/hand disease profiles.

Carpal tunnel syndrome, however, was identified as a condition associated with extensive repetitive wrist or finger movements.

Investigation of psychological factors in relation to the wrist/hand disorders suggested that both musculoskeletal and sensorineural symptoms (and signs) were associated with poor vitality and mental health, although carpal tunnel syndrome was not.

Although a coherent pattern of association was not seen in the analyses presented, some of the wrist/hand profiles clearly had distinct risk factors or outcomes. In particular, the 'numbness and/or tingling with physical signs' data-driven profile was associated with high morbidity and poor mental health, and the diagnosis of carpal tunnel syndrome had a set of associations with the outcome measures that was dissimilar to other profiles of numbness and/or tingling, or the diagnoses of tenosynovitis or OA.

7.8 Summary

The limb-based clusters identified in Chapter 6 along with the medically derived diagnoses were re-categorised into two person-based classification schemes at each of the neck, shoulders, elbow and wrist/hand. Their validity was investigated in terms of disability, healthcare use, putative mechanical and psychosocial risk factors and individual psychological profile.

The two classification systems of neck disorders were in strong agreement in their associations with disability, healthcare utilisation, occupational activities and psychological profile, and this was to be expected because the two systems were similar. Generally, a graded relationship was seen of increasing severity of neck disorder associated with more frequent disability, healthcare utilisation, mechanical neck activity and poor psychological state.

Similarly, the two shoulder classification systems were in general agreement in their associations with the outcome measures investigated, and the more severe disorders in both classification systems had markedly higher odds ratios compared to the other disorders. Occupational holding of the hands above shoulder height was associated with pain regardless of physical signs. Occupational carrying of weights and poor psychological state was associated with physical signs more than it was with pain alone.

The presence of pain and pain accompanied by physical signs had high odds ratios for the outcome measures investigated in both of the elbow classification schemes. There was a distinction between different sites of involvement in terms of their associations, and an indication that the three data-driven profiles not corresponding to medical diagnoses had their own discrete patterns of association that involved significantly elevated odds ratios (some of the highest odds ratios observed for the elbow disorders).

In contrast to the other three sites, the two classification schemes of wrist/hand disorders did not have clear patterns of association with disability, healthcare use, occupational factors and psychological profiles. Pain appeared to be more strongly associated with reported morbidity than numbness and/or tingling (and tenosynovitis and OA had higher odds ratios than carpal tunnel syndrome), except for the data-driven profile characterised by numbness and tingling with accompanying physical signs. There was also some suggestion that different sites of wrist/hand disease involvement gave rise to different associations with morbidity. Occupational activities did not distinguish between different wrist/hand disorders other than to identify an association between repetitive wrist or finger movements and carpal tunnel syndrome.

CHAPTER 8: DISCUSSION

8 Discussion

Soft-tissue musculoskeletal disorders of the neck and upper limb are common and can cause substantial disability and time lost from work. They comprise a heterogeneous group of conditions, some of which have clearly defined pathology and natural history and others whose disease process is currently unclear. Diagnoses used for these conditions may be based on underlying pathology (e.g. tendonitis, indicating inflammation of a tendon), assumed cause (e.g. repetitive strain injury) or clinical description (e.g. painful wrist) and these different types of diagnosis may therefore overlap. Research into these disorders has been impeded by the lack of an agreed classification system and diagnostic criteria, and establishing these has been particularly problematic in the classification of shoulder disorders and the diagnosis of carpal tunnel syndrome in general or working populations [30,56,58,60,63,66,125,126].

A large population-based study of musculoskeletal disorders of the neck and upper limbs has been performed in Southampton. The first stage of the study was a postal questionnaire sent to subjects aged 25-64 registered at one of two GP practices in different areas of the city. 6038 (62%) subjects responded, answering questions about occupational activities, leisure time activities, physical and psychological health. A surprisingly large proportion of these subjects (52%) reported pain or numbness or tingling in their neck or upper limbs in the previous week, and many reported multiple sites of complaints and long-term symptoms (Chapter 5). All of the subjects reporting symptoms in their baseline questionnaire were invited to attend a physical examination, and 1960 (62%) agreed. A further 185 asymptomatic subjects were also examined. The study employed recently proposed diagnostic criteria (the HSE classifications) for eight common musculoskeletal complaints [67] (shoulder capsulitis, rotator cuff tendonitis, bicipital tendonitis, lateral epicondylitis, medial epicondylitis, de Quervain's disease, tenosynovitis and carpal tunnel syndrome) and additionally diagnosed four other conditions (acromio-clavicular joint disorder, subacromial bursitis, olecranon bursitis and cervical spondylosis). The validated physical examination schedule [71] was used to evaluate subjects' necks and upper limbs, and the diagnoses were applied to the observations from the physical examinations.

An alternative approach to the classification and diagnosis of musculoskeletal disorders in the neck and arms is to abandon *a priori* clinical assumptions about the range of disorders that exist, their underlying pathology and associated symptoms and signs, and investigate instead the way in which symptoms and signs actually coexist within individuals. This is essentially a pattern recognition exercise which may serve to validate the disease categories and diagnoses proposed in the medical literature by replicating them, or find other profiles of disease which would need validating. Underpinning the data-driven classification approach is the thinking that symptoms and signs which cluster together within individuals indicate clinical syndromes, and that a clinical syndrome may be indicative of an underlying disease process. Thus, the symptom-sign constellations identified by data-driven methods have some validity purely because they exist in the population, and in sufficient numbers to be detected. What the constellations mean, and how useful they are then needs to be examined by characterising them and comparing them with putative risk factors, associated disability, response to treatment and outcome.

Many methods exist for data-driven pattern recognition, but only a few of them have been rigorously investigated to test their performance, most frequently variations of cluster analysis. Often, these methods are applied to identify unknown patterns, and their performance cannot be tested against a known correct answer. It is, therefore, important to use methods appropriate for the data, and where possible, to use those methods that are known to perform well. Confidence in the results of a cluster analysis will depend on the methodology used, but also on the quality of the data being analysed. Thus, the repeatability of the physical examination was investigated in a subset of the study sample (Chapter 4) [118] and was found to be satisfactory, although it was poorer than that seen in a hospital setting [71].

This thesis aimed to use data-driven cluster analysis techniques to contribute a fresh approach to the classification of neck and upper limb musculoskeletal disorders in the general population. Furthermore, the thesis aimed to evaluate the findings from this work alongside recent proposals for classification and diagnoses derived from mainstream medical rationale.

The objectives of the thesis were:

1) To classify and characterise each of the neck, shoulder, elbow and wrist/hand symptom-sign profiles amongst a working-age population from the UK using cluster analysis techniques.

2) To compare the resulting classifications with the HSE classifications, once these have been refined.
3) To validate both classification systems by associated disability, healthcare utilisation and risk factors.

The work undertaken to address these aims was:
1 a) A wide selection of cluster analysis methods were appraised, and the most suitable combination of approaches was chosen prior to the commencement of analysis (Chapter 3).
1 b) All data were checked and cleaned following entry onto computer (Section 2.3).
1 c) Cluster analysis techniques were employed to identify the predominant symptom-sign complexes of each of the neck, shoulder, elbow and wrist/hand, using all 2145 physical examinations (Chapter 6). The stability of the solutions was tested by performing the analysis first in one subset of the data (in subjects recruited from the first 'Hill Lane' GP practice) and then repeating it in a second subset (in subjects recruited from the second 'Bitterne' GP practice) and comparing the results. For those physical examinations which incorporated only binary observations (elbow and wrist/hand) two cluster analysis techniques (Ward's method and average linkage) were employed, and the results were compared.
1 d) The symptom-sign profiles were characterised and the distinguishing features of each cluster described (Chapter 6).
2 a) The diagnostic criteria for cervical spondylosis were revised, as were those for carpal tunnel syndrome. Additional diagnoses of symptomatic distal interphalangeal joint (DIP) osteoarthritis (OA) and thumb base OA were used (Chapter 6, Sections 6.1.4 and 6.4.4).
2 b) The symptom-sign profiles were compared with the medically-based classifications for each of the neck, shoulder, elbow and wrist/hand (Chapter 6).
3 a) The medically based categories' and symptom-sign profiles' associated disability, healthcare use, and relation to putative mechanical and psycho-social occupational and psychological risk factors were compared using logistic regression (Chapter 7).

8.1 Main Findings

Neck

Cluster analysis at the neck identified four main symptom-sign profiles based on the 7-item physical examination (1 reported symptom of neck pain, and 6 recorded ranges of movement, Appendix I) (Section 6.1.3). These profiles were stable when the cluster analysis was performed separately in the subjects recruited from the two different general practice registers (Section 6.1.2). The neck profiles were primarily characterised by the presence or absence of pain. These two main profiles were each subdivided into those subjects with a normal range of neck movement, and those whose neck movements (no particular direction was isolated) were restricted. There was some evidence, however, that the subjects displayed a continuum across the spectrum of neck movements irrespective of neck pain, rather than demonstrating the distinct 'normal range of movement' and 'abnormal range of movement' dichotomy indicated by the cluster analysis. This finding was suggested by the large numbers of inter-cluster moves of subjects within the two overriding clusters in the refining stage of the cluster analysis following the initial hierarchical analysis.

It is important that any conclusions from the findings of a cluster analysis take into account the way in which this form of pattern recognition works. It is crucial to realise that cluster analysis will always, by definition, find patterns in data, even if there are none to be found except for those due to random variation (Section 3.6). It is for this reason that the choice of methodology and the final assessment of the clusters identified are so important. Two of the most widely used and tested methods of cluster analysis (hierarchical and partitioning) were employed in a two-stage process to exploit the benefits of each, and the high-performing [114] technique of Ward's method was used to search for tightly homogeneous clusters (Section 3.5). The variables used in the cluster analysis were clearly dependent on the information collected at the physical examination, and these were weighted in an attempt to assign all variables equal importance in the analysis (Section 3.3). It is this last aspect which was likely to have the greatest impact in the neck cluster analysis. Mathematically, it was to be expected that the presence or absence of pain would be the primary distinguishing feature in this analysis since it was the only binary variable, and therefore always took an extreme value (0 or 1) between pairs of subjects. The numerical range of movement differences between subjects only took on an extreme value in a minority of occasions, when the range of movement was exactly the same between a pair of subjects, or when the difference between a pair was the maximum amongst all pairs.

The full spectrum of movement was seen in both the pain-positive and the pain-negative subjects, and lower ranges of movement were seen predominantly with older age. Although some of the limitation in range of movement may have been due to specific neck disorders, much of it may have been due to the increased prevalence of degenerative changes seen at the neck in older age [20].

Criteria for the diagnosis of cervical spondylosis were the presence of neck pain with accompanying restriction of movement. The cut-points for restricted ranges of movement were not originally based on a general population, and there was evidence from a validation study of this physical examination and diagnoses [71] (Section 1.2.2) that they were too high. As a result, 85 of the 88 subjects in the study met the criteria for restricted neck movement, and as a result, all subjects who reported neck pain were diagnosed with cervical spondylosis. Hospital-based rheumatologists gave a diagnosis of cervical spondylosis much less frequently than the newly proposed criteria did.

Investigations to try to establish new cut-offs for restricted neck movement were based on this study sample [123] and suggested that a) neck movement in all directions was inversely correlated with age, b) there was little difference between the sexes in neck movement, and c) there was a small but statistically significant reduction in neck ranges of movement in association with neck pain, even after accounting for any age or sex effects.

In order to maximise the possibility of detecting true cervical spondylosis it was decided to employ the strict criterion of a restricted range of movement defined as movement less than two standard deviations below the mean in any direction, within each 10-year age and sex stratum.

Comparison of the two classification systems (Table 32, Section 6.1.4) showed that cervical spondylosis was a less prevalent condition than pain with accompanying low range of movement, although virtually all cases of cervical spondylosis were found in this most severe cluster. It was hardly surprising that these two classifications were so similar when they were based on the two same domains: pain and range of neck movement. It is of note that neither classification system identified different patterns of neck restriction, and it is likely that in order to distinguish cervical spondylosis from other mechanical neck disorders (if such a distinction is useful) other information such as history may be required.

Validation of the two neck classification systems again demonstrated little difference between them (Sections 7.3.1 and 7.4). Both discriminated between varying levels of severity of disease involvement in terms of associated disability, healthcare use, mechanical neck activity and psychological factors.

Shoulder

Cluster analysis at the shoulder (Section 6.2) also identified four main symptom-sign profiles based on the 31-item physical examination (7 reported symptoms of shoulder pain and 24 recorded physical signs). These clusters were less mathematically stable than those identified at the neck (Section 6.2.2): Analysis on the shoulders from subjects recruited by the first GP practice suggested five main shoulder profiles, including one indicative of AC joint involvement. In the overall analysis and that for the limbs from the second (larger) GP practice only four profiles were identified, and none were suggestive of involvement of any particular site (Section 6.2.3). Rather, they were characterised by 1) an absence of pain or physical signs, 2) a variety of physical signs, mostly tenderness or positive AC joint stress test, 3) Pain at different sites with or without some physical signs, and 4) multiple sites of pain, tenderness and other physical signs. Thus the cluster analysis distinguished severity of disease involvement rather than distinct pathology. Interestingly, although the physical examination and diagnoses were designed to distinguish between five shoulder conditions, in practice this did not happen (Table 37, Section 6.2.4). Shoulders were frequently diagnosed with two or more conditions, and rotator cuff tendonitis (RT) in particular was nearly always diagnosed alongside shoulder capsulitis (SC). Similarly, the less prevalent disorders of bicipital tendonitis, AC joint disorder and subacromial bursitis were diagnosed more frequently with either RT or SC than alone. This may have occurred for a number of reasons: firstly, it may be that the physical examination was not detailed enough to distinguish the different disorders. However, signs and symptoms included in the examination were taken from the HSE criteria and were therefore based on expert experience and literature reviews. Secondly, it may be that the examination, whilst suitable in a secondary care setting, is not suitable for the general population, where the presence of physical signs may be uncertain and the report of pain more subject to inaccuracies if it was mild or not debilitating. However, the validation study of the examination and diagnoses highlighted this same problem of multiple diagnoses at the shoulders even though subjects were attending a rheumatology clinic after referral from their GP [71], (Section 1.2.2). Thirdly, it may be that whilst the different shoulder disorders can be distinguished by other investigations such as arthroscopy, they cannot by clinical examination alone. This has been suggested by a number of authors [40,41], but refuted by others [127]. A cluster analysis of shoulder signs and symptoms from patients presenting to general practice in the Netherlands had similar findings to this Southampton study, and suggested that detailed classifications of shoulder disorders were unsuitable in the primary care setting [61]. Their analysis was based on past history of shoulder complaints as well as physical examination, and thus no distinction could be made between different shoulder disorders even though the subjects may have had more severe disease compared to those in the Southampton study, and had more clinical information available. It may be, therefore, that clinicians in practice only give one diagnosis because they believe only one condition is present, even if clinically there is evidence for more than one condition. Finally, since the shoulder joint is complex and a disorder in one part may lead to secondary conditions in others, clinicians may diagnose what they consider to be the single primary condition even though others exist: this has been documented for the condition of subacromial bursitis, which is thought by some to be part of the RT process [58].

As seen in the neck cluster analysis, there was a lack of discrimination between the ranges of shoulder movement seen in the pain-positive and pain-negative shoulder clusters, and other authors have reported this [128]. This may indicate that the diagnostic criteria for shoulder capsulitis (based on pain and a restricted range of shoulder movement) are not specific enough, and would tend to make this diagnosis a blunt indication of shoulder disease severity rather than one indicating specific pathology. The relative importance of the ranges of movement in the present cluster analysis compared to other signs or symptoms will have been diminished in the same way as the neck ranges of movement were in the neck cluster analysis.

Validation of the two shoulder classifications had to be modified because of the extensive overlap of the medical diagnoses, and was based on subject comparisons rather than on limb-based ones (Section 7.3.2). As seen in the neck analyses, the two classification systems had similar findings with the most severe categories of disease involvement being associated with higher levels of disability and healthcare utilisation (Section 7.5 and Appendix VI). Exposure of the shoulder to occupational holding of the hands above shoulder height was associated with shoulder pain regardless of whether there were accompanying physical signs, and such a relationship gives weight to a shoulder classification proposal made by Kuorinka and Viikari-Juntura [53] which includes a category of 'temporary symptoms of overuse' characterised by fatigue, stiffness or soreness following strenuous exercise such as work exposure, but not by pathological changes. Psychological factors were associated more strongly with physical signs alone than with pain alone, a finding which contrasts with the hypothesis that the origin of non-specific pain has a strong psychological component, whilst pain arising from underlying pathology (diagnosed with presenting pain plus accompanying physical signs in a clinical setting) has other pathogeneses. The most severe data-driven shoulder profile had higher odds ratios for these validation measures compared to the medical diagnoses. This may again be an indication that the criteria for shoulder capsulitis in particular were not specific enough.

Elbow

Cluster analysis at the elbow (Section 6.3) identified seven main symptom-sign profiles based on the 17-item physical examination (6 reported symptoms of elbow pain and 11 recorded physical signs). Three of these clusters were highly mathematically robust and were identified by two clustering methods and in both subsets of the data as well as in the whole population. These three clusters were characterised by 1) an absence of pain or physical signs, 2) symptoms and accompanying signs over the medial epicondyle, and 3) symptoms and accompanying signs over the lateral epicondyle. The remaining four clusters were identified by Ward's method only, and were characterised by 1) pain alone at either the lateral or medial epicondyle or posterior elbow, 2) tenderness over the medial epicondyle and some pain on resisted wrist flexion over the medial epicondyle, 3) tenderness over the lateral epicondyle and some pain on resisted wrist extension over the lateral epicondyle, and 4) tenderness over the posterior elbow or antecubital fossa, with some pain over the posterior elbow or antecubital fossa and some swelling over the posterior joint (Section 6.3.3). The cluster analysis demonstrated clearly that symptoms and signs at the same location on the elbow tend to co-exist, and that a range of disease severity was present in the population. Diagnoses at the elbow (lateral and medial epicondylitis and olecranon bursitis) were distinct from each other and tallied well with three of the clusters identified (Table 42, Section 6.3.4). Classification and diagnosis of disorders at the elbow is less controversial than that at other sites, and this cluster analysis supports these well-recognised diagnoses [20,129]. These findings also indicate some validity of the cluster analysis methods employed in that they were able to identify the anticipated conventional physical profiles. Cluster analysis did, however, find a heterogeneous cluster characterised by pain alone, and the significance of this profile needs to be established.

Validation of the elbow classification systems was again subject-based rather than limb-based (Section 7.3.3) and showed broad agreement between the two classification systems (Section 7.3 and appendix VII). The presence of pain was again the distinguishing feature of the subject profiles most strongly associated with disability and healthcare use, and those profiles which led to clinical diagnoses had the strongest associations of all. It was of note that the heterogeneous 'pain' cluster had some of the highest odds ratios, on a par with profiles involving both symptoms and signs. Disease involvement at the lateral epicondyle was more strongly associated with reported disability and occupational mechanical elbow exposure whilst that at the medial epicondyle was more strongly associated with a poor psychological profile. Having any specific diagnosis was associated with consistently high odds ratios.

These findings gave encouraging evidence that the data-driven clusters represent profiles that have important differences, and that they are therefore distinctions worth making, even though a number of them did not represent recognised clinical conditions.

Wrist/hand

A two-phase approach was taken to perform the cluster analysis at the wrist/hand (Section 6.4). The first stage was a cluster analysis on the Katz hand diagram data alone (Appendix IV), to reduce the 30 binary variables denoting numbness or tingling in 30 regions of the hand down to 7 binary variables denoting the main patterns of numbness or tingling throughout each hand. This was essentially a smoothing process, and was performed so that the detailed record of sensorineural symptoms would not overwhelm the cluster analysis of the whole wrist/hand examination. The second stage of the cluster analysis proceeded as for the elbow, using two techniques on the 57-item physical examination data (9 symptoms of pain, 7 symptoms of sensorineural disturbance, 35 musculoskeletal signs and 6 signs of sensorineural disturbance). Cluster analysis at the wrist/hand identified three overriding clusters made up of four, three and seven main clusters respectively (Section 6.4.3). Ten of these main clusters were identified in both datasets by Ward's method. The alternative cluster analysis technique did not produce meaningful clusters. Thus there were some mathematically stable wrist/hand examination profiles, but there was considerably more heterogeneity among the profiles than was seen at other locations. This was in part due to the greater number of examination items contributing to the analysis.

The three overriding clusters were identified as having 1) no symptoms of any kind, 2) pain with or without musculoskeletal signs but no sensorineural symptoms, and 3) sensorineural symptoms with or without signs or pain. The four main asymptomatic clusters were characterised by 1) no symptoms and very few signs, 2) a positive Phalen's test only, 3) thumb base tenderness, some Heberden's nodes and pain on resisted thumb extension, and 4) Heberden's nodes only (Tables A9, A10, Appendix V). The latter two clusters suggested asymptomatic OA at the thumb base, DIP joints or both; the first one identified those subjects with resolved wrist/hand conditions, or conditions at other parts of the upper limb or neck. Those wrists/hands with only a positive Phalen's test were of particular interest: Phalen's test is an indicator of carpal tunnel syndrome, although it has been suggested that it is not a very sensitive or specific one [65,130]. This cluster of 194 wrists/hands seems to confirm this lack of specificity since none of them had any sensorineural symptoms and few had other sensorineural signs. (Phalen's test was performed by holding the wrists in full flexion for a timed 1-minute duration. It seems unlikely that this widely used technique [131] caused such a large number of false positives.)

The three main clusters with pain but no numbness or tingling were characterised by 1) thumb base pain and signs, with some radial wrist pain and signs, 2) Finger joint pain, tenderness and swelling and some Heberden's nodes, and 3) a variety of sites of wrist pain with some wrist tenderness or swelling (Tables A9, A10, A11, A12, Appendix V). Diagnoses of thumb base OA and DIP OA were made (Section 6.4.4) as well as the original diagnoses of tenosynovitis, De Quervain's disease and carpal tunnel syndrome. Cluster 1 above (Cluster 5 in Section 6.4.3, 6.4.4 and Appendix V) contained wrists with diagnoses of thumb base OA (71.4%), de Quervain's disease (8.6%), tenosynovitis (7.8%) and DIP OA (7.8%). Cluster 2 above contained wrists with predominantly DIP OA (57.8%) or no diagnosis (39.2%), and Cluster 3 above contained wrists with primarily tenosynovitis (15.1%) or no diagnosis (82.5%). Thus clinically defined musculoskeletal and articular conditions were distinguished to some extent by the cluster analysis methods, although de Quervain's disease was mostly seen in the same cluster as thumb base OA.

Five of the remaining seven clusters were characterised by patterns of sensorineural disturbance (Tables A11, A12, A13 and A14, Appendix V). None of these clusters was indicative of either median or ulnar nerve compression, although closer inspection of the data suggested that a small subset of wrist/hands had symptoms in a median nerve distribution [124], and these were associated with higher prevalences of positive Phalen's and Tinel's tests, but not neck pain. These wrists/hands fell into a variety of different clusters, including all of the aforementioned five (Table 47b and 47c, Section 6.4.4).

The final two clusters were characterised by multiple sites of pain, numbness and /or tingling and signs (Tables A13 and A14, Appendix V). Both clusters contained wrist/hands with multiple diagnoses (66.7% and 11.4%), single diagnoses (25.6% and 57.2%) and no diagnoses (7.7% and 31.3%) (Table 47c, Section 6.4.4). These clusters identified the small number of hand/wrists with multiple disease involvement, either predominantly in the wrist and thumb base, or in the thumb base and finger joints.

Validation of the wrist/hand classifications was substantially more complicated than that at the other three sites due to the larger number of clusters and diagnoses, and the need to create new categories for subjects with bilateral disease involvement (Section 7.3.4). As previously seen at the other sites, disability and healthcare use were most strongly associated with profiles characterised by pain in both classifications (Section 7.7 and Appendix VIII). Numbness and/or tingling was not as strongly associated with these measures, except for the profile of numbness and /or tingling with accompanying physical signs. Occupational exposures to keyboard use and repetitive wrist or finger movements appeared to be more common in those subjects whose musculoskeletal complaints were at sites other than the wrist/hand, although the 'All' profile with multiple sites of pain, numbness and/ or tingling and signs, carpal tunnel syndrome and OA showed elevated odds ratios for repeated wrist or finger movements. An association between carpal tunnel syndrome and such mechanical exposure has been reported in other studies [132]. Subjects with more severe or bilateral musculoskeletal conditions had poorer vitality and mental health, although both musculoskletal and sensorineural symptoms showed some association with these factors. Carpal tunnel syndrome had the lowest levels of poor vitality and poor mental health of all the profiles investigated, a finding that is not reported in the literature.

These analyses yielded complex findings and a coherent pattern of association was not obvious. Of all the sites investigated, the two classification systems for the wrist/hand were the most divergent, and comparing them was therefore difficult. These analyses may have been more affected than the others by the comparative heterogeneity of many of the profiles (both data-driven and medical) (Appendix V). However, there were indications that some of the wrist/hand profiles had distinct risk factors or outcomes.

8.2 Limitations of the main study and of the methods used

Study design

The first aim of this thesis was to use cluster analysis to identify the different symptom-sign profiles existing amongst working-age UK subjects. The study design, however, biased the population called forward for physical examination so that the majority of subjects examined had recent musuloskeletal symptoms in the neck or upper limbs. Only 185 examinations were of asymptomatic subjects. It is possible, therefore, that some common profiles in the population were not represented in this study at all. In all four cluster analyses (neck, shoulder, elbow and wrist/hand), profiles were detected of abnormal physical signs without any symptoms. These may result from: resolving conditions; early stages of the disease; mild cases of disorders; or steady state abnormalities (for example, the wrist/hand profile of Heberden's nodes only). However, they would have been detected in our study only if they were the result of a resolving condition, or if the subject had symptoms at another site in their neck or upper limb (which may well occur more frequently than would be expected by chance if such disorders tend to be correlated).

Additionally, it may be that the reported disability and healthcare use associated with the asymptomatic profiles that were detected was unrepresentative of, and more prevalent than that in the general population of asymptomatic profiles. This was because subjects found to have asymptomatic profiles at examination were more likely to have had recently resolved symptoms than such subjects in the general population, and these symptoms could have caused them disability or to seek healthcare in the past year. Only disability or healthcare use due to pain or numbness and tingling at baseline or examination was investigated, which might underestimate the true burden of disability and healthcare use, particularly in subjects without symptoms but with signs such as restricted range of movement. These associations, therefore, should be viewed with caution. This problem did not arise in the same way with the occupational and psychological risk factors examined in the validation analysis because these data were collected at baseline. However, any misclassification of 'signs only' profiles as control in these analyses due to the categorisation of all non-examined subjects as controls would tend to bias the odds ratios towards unity.

Whilst these limitations of the study introduce some uncertainty as to the prevalence and associations of the clusters characterised by physical signs alone, it is the physical profiles that do involve symptoms that are of more interest in relation to the further aims of this thesis. It was for this reason that subjects were selectively sampled from respondents reporting recent musculoskeletal symptoms on the neck or upper limbs, and it is of note that the medical diagnoses proposed by the HSE under investigation all included symptoms as part of their diagnostic criteria. Studies of musculoskeletal conditions frequently focus on symptoms alone or investigate physical signs only in the presence of symptoms. The only common musculoskeletal conditions that have been studied epidemiologically even though they have asymptomatic profiles are early-stage Dupuytren's contracture [133], and Heberden's nodes, which are an indicator of osteoarthritis, but are not necessarily painful.

Southampton examination proforma

The physical examination was designed to distinguish the HSE-defined disorders, and incorporated extra information in order to diagnose cervical spondylosis, AC joint disorder, subacromial bursitis, olecranon bursitis, thumb base OA, DIP OA and Dupuytren's contracture. It may be that there are other common conditions which the physical examination was not designed to detect, which would obscure the current cluster analysis findings. Possible candidates might be tension neck syndrome, thoracic outlet syndrome (the HSE did not discuss this diagnosis since the condition is considered rare in the UK), ulnar nerve entrapment and nerve entrapment at the elbow (which could be detected by the SEP) and trigger finger, all of which have been included in other classification systems [45,50-52]. Additionally, there may have been significant omissions from the physical examination which meant that the planned diagnoses could not be made, and this might be the case for the shoulder diagnoses, as already discussed.

The repeatability of the examination was investigated (Chapter 4) and found to be satisfactory, although it was worse than in a hospital setting [71,118]. The diagnosis of tenosynovitis had poor repeatability, and this was due to the low repeatability in observing pain on resisted radial or ulnar deviation. The range in reliability of the physical examination items was not accounted for in the cluster analysis. This could have been achieved by weighting observations according to their repeatability. Such a scheme has a sound mathematical basis but would have made the cluster analysis results harder to interpret: medical conditions not identified in the data-driven categories might truly be expressed in a number of heterogeneous physical profiles, or might simply be diagnosed using less reliable physical signs, making them less influential in the analysis. Furthermore, the inter-subject repeatability of symptom reporting is immeasurable, and it would be difficult to know how best to weight these components of the examination.

<u>Cluster analysis</u>

As has been stated already, cluster analysis (and indeed any statistical analysis) is highly dependent on the information included in it. These analyses used all of the physical examination findings but no other data. It may be that more informative clusters could have been identified by incorporating the previous history of symptoms, age, or other variables into the cluster analysis. One cluster analysis of shoulder disorders based on physical profile and history found that past history of symptoms was a useful distinguishing feature, but still did not lead to clusters that indicated recognised shoulder pathology [61]. It was of note that only the HSE diagnoses at the shoulder used past history of pain in their criteria, and shoulder capsulitis, which is generally understood to have three distinct phases (the last of which is characterised by stiffness but not pain), required current pain for a positive diagnosis. The revised diagnostic criteria for cervical spondylosis used age and sex information, which led to a more specific diagnosis than the cluster analysis provided. The main reasons for restricting the cluster analysis to the physical examination data only were to keep other variables (age, sex, history of symptoms) available to make inter-cluster comparisons, and so investigate the utility of the physical examination alone.

Cluster analysis is designed so that it will always find patterns in data, and that is why careful consideration of the methods used at each stage of the analysis (as described in Chapter 3) is so important [87]. However, the analyses at the elbow and wrist/hand showed that the use of a less efficient technique (group average linkage, compared to Ward's method [103]) gave less meaningful results (with a few very large heterogeneous clusters and some very small clusters comprising only a few limbs), although it was able to identify three highly robust elbow clusters. Characterisation of the clusters, their replicability, heterogeneity and their size all contribute to the assessment of the legitimacy of a cluster. For the analyses in this thesis, relatively large clusters were considered (the smallest comprising 39 limbs in the 'All' cluster of wrist/hand profiles). Clusters in the neck, shoulder and elbow analyses were replicable in the two data subsets, as were the majority of the clusters in the wrist/hand analysis. The least homogeneous clusters were wrist/hand profiles, even after refinement of the cluster by k-means partition. This could have been anticipated because the wrist/hand examination had so many observations, and thus there was more potential for variation even in wrists/hands with the same underlying pathology. This issue would have been augmented in the cluster analyses that used squared Euclidean distance, because mathematically profiles could be similar, based on a large proportion of shared absences of traits, but clinically look heterogeneous because of a few differences in the presence of traits. The wrist/hand heterogeneity may also have occurred because of the preliminary cluster analysis performed on the Katz hand diagram data: this smoothing process may have masked some important differences in the distribution of numbness and tingling throughout the hand, and led to wrist/hands being clustered together even though their profiles with respect to their numbness and tingling as well as some of the other examination findings were different. This may be why the diagnosis of carpal tunnel syndrome was evident in so many of the wrist/hand clusters. It was important to employ the preliminary cluster analysis, otherwise patterns of numbness and tingling would have dominated the analysis. As it was, five of the fourteen clusters identified were characterised primarily by different distributions of sensorineural symptoms. Characterisation of the clusters demonstrated that the findings were clinically plausible, even when they did not reflect recognised musculoskeletal conditions.

The cluster analyses performed on the shoulder, elbow and wrist/hand examination profiles included the right and left limbs from each subject, and no mathematical account was taken of the fact that the profiles of pairs of limbs might not be independent of each other. It was also evident that the clusters identified at these sites were predominantly seen bilaterally, or in subjects who only had one limb profile with symptoms or signs. Few subjects had two different 'non-normal' profiles on their right and left limbs. It is therefore possible that some of the non-normal profiles observed were identified because they frequently occur bilaterally and were thus represented twice in the cluster analysis for a single individual, rather than another profile that might have a similar prevalence in individuals, but which occurs unilaterally.

A small proportion of neck or limb profiles were not used in the cluster analyses because they contained missing data. The majority of the missing data was due to physical examination items being omitted at random, but a small number of profiles were incomplete because an examination item was too painful to perform. These latter profiles might represent the most extreme disease involvement in this population sample, but were not represented in the cluster analyses. It is possible that the clusters identified do not include profiles of significant disease involvement that were nevertheless experienced in this population. An analysis that imputed extreme values for the missing data that was not missing at random, and then included these additional profiles in a cluster analysis would investigate the impact of this small group of profiles.

Validation of the clusters

A severe limitation of the validation of the clusters was the lack of limb-specific construct validity measures. Thus analyses were subject-based rather than limb-based, which may have obscured some associations. Assumptions might have been reasonable for the side of some of the occupational exposures that are usually performed in the dominant hand, but this accounted for only a couple of factors. Disability and healthcare use could be attributed to one side when there were unilateral conditions, although this information covered the previous year, and was not necessarily attributable to the current condition or side. Such assumptions seemed to be at least as crude as employing subject-based analyses, and possibly harder to interpret. Psychological and psychosocial factors of necessity had to be considered by person rather than by limb. This issue was less important in the shoulder analyses, where there were only four clusters and it was straightforward to assign appropriate subject-based categories. The elbow profiles were highly symmetrical, with only 33 (1.5%) subjects of the total 2140 having two non-'normal' data-driven elbow profiles, and the majority of these were placed in their own 'mixed bilateral elbow profiles' subject category. The wrist/hand profiles were less symmetrical, with 273 subjects (13%) having two non-'normal' data-driven profiles. 222 of these subjects had to be placed into heterogeneous summary groups in order to keep the analysis as simple as possible.

These analyses were all performed with each classification scheme (data-driven or medically based) as a single predictor taking on multiple levels, using a category of no symptoms and no signs as the baseline comparator. As has already been discussed, such a baseline may have been inappropriate because of the particular validity measure being explored (for example disability due to pain, Section 8.2), but also may not have been the comparison of most interest. Validation of the classification systems aimed to explore whether different categories of disease differed from each other in their associations with morbidity, putative risk factors or outcome rather than whether they differed in these associations from a baseline category of no disease. Clearly, the odds ratios can and were compared informally between disease categories, but direct comparisons may have been more useful and appropriate. Furthermore, such comparisons may have led to tighter confidence intervals since the small numbers of positive outcomes in the baseline category would have no longer affected the analysis.

Odds ratios were used as the measurement of association in the validation analyses, even when the outcomes were common (such as observed with the mechanical occupational exposures). Whilst these are valid measures of association, they may have inflated the estimate of the relative risk in a way that prevalence rate ratios would not have done. In these data, however, prevalence rate ratios would have been impossible to calculate for some analyses due to convergence problems in the model.

In all subject-based categories (data-driven and clinical) an attempt was made to differentiate between those subjects who reported symptoms only, those with signs only and those with both symptoms and signs. In the clinical categories, subjects with non-specific symptoms and signs were grouped with subjects who had non-specific symptoms only, reflecting clinical assessment in practice. Thus a mathematical hierarchy was assumed that the associations between symptoms and the variables used to test validity overrode those between clinical signs and the variables used to test validity. The subject-based clinical categories made no distinction between different shoulder and elbow diagnoses: this was unavoidable in the former situation because of the large overlap of diagnoses, and because of the low prevalences of diagnoses in the latter situation.

The variables used to assess validity were all self-reported measures, and may therefore have been affected by subjects' perception of their musculoskeletal conditions and their likely causes. Measures of disability and GP consultation will have reflected subjects' beliefs about effective treatment for certain conditions, and may have been influenced by subjects' occupations and lifestyles (if subjects were constantly using their disordered limb, they may have been more likely to report disability and seek treatment). More interestingly, associations with prescribed treatment may reflect GPs' beliefs about treatment, which they might be more ready to administer when presented with 'barn-door' cases of musculoskeletal conditions, leading to the high odds ratios seen between specific diagnoses and prescribed treatment in this study. It should be noted, however, that other non-specific profiles also had elevated odds ratios for this outcome. The analyses of these three measures in relation to the elbow classifications led to very wide confidence intervals due to the low prevalences in the baseline category. Mathematically, an alternative reference category would have been more appropriate, but retaining the baseline kept identical reference groups (varying only because of missing data) across the analyses at all sites.

Associations between physical occupational exposures and categories of disease were only explored at the anatomical sites deemed most likely to be relevant (for example exposure to activities involving twisting or bending the neck was explored in relation to the neck disorder classification systems only), based on previously reported associations and the probable biomechanical effect of the activity. This subset of analyses may have failed to present important relationships and thus may have been misleading in suggesting that these associations were site-specific. It may be that certain occupational activities have associations with disease involvement at other sites because of the strong interrelation of the neck and upper limbs, because the activities are directly associated with disease involvement at specific sites not considered in these analyses, or because subjects with disease involvement tend to report these physical activities in the workplace more readily.

Questions regarding psychosocial factors at baseline were structured according to a model of work strain involving elements of control, support and demand in the workplace. This study showed no clear relationships between job control, support or demand with disease categories at any site, and it may be that a combination of exposures would have shown some discrimination between the categories where the individual factors did not.

The obvious omissions from the validity analyses were the lack of any measures of long-term outcome or response to treatment. These data were outside the scope of the main cross-sectional study and this thesis, but a follow-up of around 400 of the subjects 18 months later will allow these investigations to be made. Response to treatment will be harder to assess in this extension to the study due to the limited data available relating to specific treatments used.

Interrelation of the neck, shoulders, elbows and wrist/hands

Investigation of the physical profiles was made at each of the neck, shoulder, elbow and wrist/hand in isolation. This was desirable in the original cluster analyses, which aimed simply to describe physical profiles existing in the general population, but clearly is a poor representation of the way in which disorders might impact on surrounding tissues and the use of nearby joints, opposite limbs and the upper body overall. Further examination of these clusters considering their inter-relation with clinical findings at other sites might help to clarify the nature of some of the profiles, in particular identifying cases of referred pain, sensorineural disturbance in the wrist/hand due to nerve compression higher up the body, and widespread pain or fibromyalgia. The validity analyses made a basic attempt to consider the relationship between different parts of the upper limb and neck by distinguishing those subjects with no abnormalities at examination from those with normal local profiles but abnormalities at other sites. Even this minimal adjustment confirmed the interrelation of different sites, with the mechanical occupations of elbow bending and repetitive movements of the wrist or fingers being statistically significantly associated with abnormalities at sites other than the elbow and wrist respectively. The association between extensive keyboard use and abnormalities at sites other than the wrist (the only statistically significantly elevated odds ratios seen) suggest that this activity was an indicator of neck, shoulder or elbow mechanical exposure more than it was of the wrist/hand. It was also clear that the subjects called forward to the second stage of the study were different in their mechanical exposures and psychological profiles from those who were not called forward, even when their physical examination was normal at all sites (the majority of these subjects had resolved symptoms).

It may be that investigation of construct validity of symptom-sign profiles which model site interaction more completely would reveal disease processes more clearly. However, the potential increase in understanding would be counter-balanced by the complexity of such modelling and the probable need for more data points, and is beyond the scope of this thesis.

8.3 Implications and future work

This thesis has adopted a novel approach to the problem of diagnosing and classifying musculoskeletal conditions of the neck and upper limb. Despite the danger of the self-fulfilling nature of cluster analysis and the simplistic validation models used, findings suggest that data-driven classification techniques yield clinically plausible categories of disease involvement, and moreover are able to identify distinct physical profiles indicative of recognised underlying pathology.

Cluster analysis results were in good agreement with medically based diagnoses at the neck, shoulder (both classification schemes agreed in their lack of power to distinguish different shoulder pathologies) and elbow. Thus, independent validity for current medical thinking on neck and elbow musculoskeletal conditions has been provided. Musculoskeletal and articular disorders at the wrist/hand were identified and distinguished from sensorineural conditions. A variety of different sensorineural profiles was seen, none of which tallied with the recognised disorders of median or ulnar nerve compression, and it is here, along with musculoskeletal disorders at the shoulder, that data-driven findings and medical understanding appear to diverge.

Validation of the clusters and medical diagnoses by investigation of associations with disability, healthcare use, occupational and psychological factors demonstrated that both the clusters and medical diagnoses made useful distinctions between disorders with different risk factors and natural history as well as physical profiles, which need further investigation. The main implication of these findings is the importance of the physical examination. Currently epidemiological research into musculoskeletal disorders is split into two strands, one of which investigates symptoms alone. Such work has clear advantages in that it can be conducted on a large scale relatively quickly and cheaply, and arguably focuses on the problem which causes the community a large burden and which clinicians are called upon to treat. The second strand of research focuses on pathology, indicated by a complex of symptoms and signs and therefore incorporates a physical examination. This work has suffered from a lack of standardised diagnoses and is therefore difficult to interpret and generalise (much of it has also been conducted in specialised occupational settings). This thesis suggests, however, that such work is worth pursuing because valid categorisations of physical symptom-sign profiles can be made, and that these may lead to more specific understanding of musculoskeletal conditions and possibly more effective prevention and treatment as a result.

In particular, findings of the neck and elbow examination suggest that current medically derived approaches to classification and diagnosis are satisfactory: thus clinicians can continue to use them as helpful labels to direct treatment and management, and researchers can continue to investigate these conditions to further our knowledge of their aetiology, natural history, effective treatment and outcome. It is of note, however, that a large number of elbows had symptoms and signs of medial or lateral epicondylitis, but still did not meet the criteria for a medical diagnosis. Further research into the natural history and treatment efficacy of disease in such cases might confirm whether milder, early or resolving forms of epicondylitis were being identified, and whether they benefit from treatment.

Findings at the shoulder and of sensorineural disorders in the wrist/hand suggest that current medically derived approaches to classification and diagnosis are more controversial and require further investigation. Although physical examination of the shoulder suggested that underlying pathology cannot be identified, this finding may be restricted to less severe cases, which may have been the majority of cases in this study, or be due to an inadequacy in the physical examination, or in the diagnostic criteria (the cut-points for restricted range of movement could be a problem in the case definition of shoulder capsulitis, for example). MRI, ultrasound, arthroscopy and more detailed clinical examination of shoulders in general populations, general practice and secondary care settings could shed light on this diagnostic problem. This thesis suggests that clinical examination of the painful shoulder in a primary care setting is currently unlikely to produce a single reliable diagnosis, and that general practitioners might be advised to concentrate on treatment and management directed by shoulder pain rather than by a particular underlying pathology.

The classification and diagnosis of sensorineural disorders in the wrist/hand is more complicated because, although pathology such as carpal tunnel syndrome is well described and defined, no gold standard tests are available to diagnose them, even in secondary care settings (nerve conduction tests are regarded as the gold standard in predicting response to surgical carpal tunnel release, but not for the diagnosis itself). Focused studies in the different settings already mentioned which gather enough information to identify some of the different underlying processes of sensorineural symptoms may help to form classifications and diagnostic criteria.

Alternatively studies may need to concentrate on natural history, response to treatment and risk factors in order to make progress in this area.

At the heart of epidemiological research is the gain of knowledge to effectively treat, as well as prevent disease. In the area of musculoskeletal disorders, treatments are often in the form of pain relief, anti-inflammatory medication, manipulation or physiotherapy and rest, with surgical intervention being considered for persistent and resistant cases. The response to and efficacy of these treatments, particularly of physiotherapy, manipulation and rest are poorly described [134] and this is compounded in conditions seen to be self-limiting and self-resolving. Investigation of the various treatment options in these conditions by observational studies and clinical trials following the establishment of valid classification systems and diagnoses, as progressed in this thesis, is the next stage.

Further work leading directly from this thesis is the follow-up of these subjects, and investigation of the data-driven and medically based categories' long-term prognosis. Already a subset of these subjects has been contacted 18 months after their physical examination with another questionnaire and examination. Further plans to contact the entire cohort are being considered.

8.4 Summary of the principal findings of this thesis

- Cluster analysis of physical examination observations at the neck, shoulder, elbow and wrist/hand produced robust symptom-sign profiles representing clinically plausible categories of disease involvement.
- Data-driven profiles at the elbow and neck tallied well with diagnoses based on clinical understanding.
- Data-driven and clinically based profiles of shoulder disease involvement identified different levels of severity but could not distinguish underlying pathology.
- Data-driven symptom-sign profiles at the wrist/hand separated musculoskeletal and articular disorders from sensorineural ones and were more heterogeneous than the clusters produced at other sites.
- The medically derived diagnostic criteria for carpal tunnel syndrome needed refining from those originally proposed. None of the sensorineural data-driven

profiles tallied with this diagnosis or appeared to be indicative of this condition.

- Validation of both the data-driven clusters and the medically derived diagnoses confirmed that important aetiological as well as physical differences were distinguished in the two classification schemes.

Papers and Presentations from the thesis

Papers:

1. Walker-Bone K, Byng P, Linaker C, Reading I, Coggon D, Palmer KT, Cooper C. Reliability of the Southampton examination schedule for the diagnosis of upper limb disorders in the general population *Ann Rheum Dis* 2002; Vol. 61, No. 12: 1103-1106

2. Reading I, Walker-Bone K, Palmer KT, Cooper C, Coggon D. Anatomical Distribution of Sensory Symptoms in the Hand and Their Relation to Neck Pain, Psychosocial Variables, and Occupational Activities. *American Journal of Epidemiology* 2003; Vol. 157, No. 6: 524 – 530

Abstracts:

Poster presentation at the Annual General Meeting of the British Society for Rheumatology, Brighton, 2000

1. Reading IC, Walker-Bone K, Palmer K, Coggon D, Cooper C. Classification Algorithms for Musculoskeletal Disorders of the Upper Limb and Neck. *Rheumatology* 2000; Vol. 39: 74, 133S

Oral poster presentation at the Annual General Meeting of the British Society for Rheumatology, Edinburgh, 2001

2. Reading IC, Walker-Bone K, Palmer K, Coggon D, Cooper C. Classification Algorithms for Musculoskeletal Disorders of the Hand and Wrist. *Rheumatology* 2001; Vol. 40: 132, 375S

Poster presentations at the International Scientific Conference on the Prevention of Work-Related Musculoskeletal Disorders, Amsterdam, 2001

3. Reading IC, Walker-Bone K, Palmer K, Coggon D, Cooper C. Classification Algorithms for Musculoskeletal Disorders of the Hand and Wrist. *Proceedings of*

the Fourth International Scientific Conference on the Prevention of Work-Related Musculoskeletal Disorders 2001;34S

4. Reading IC, Walker-Bone K, Palmer K, Coggon D, Cooper C. Classification Algorithms for Musculoskeletal Disorders of the Shoulder and Elbow. *Proceedings of the Fourth International Scientific Conference on the Prevention of Work-Related Musculoskeletal Disorders* 2001;191S

Poster presentation at the Annual General Meeting of the American College of Rheumatology, San Francisco, 2001

5. Reading IC, Walker-Bone K, Palmer K, Coggon D, Cooper C. Classification of Musculoskeletal Disorders of the Upper Limb and Neck: A Novel Approach. *Arthritis Rheum* 2001; Vol. 44:S118

Poster presentation at the Annual General Meeting of the British Society for Rheumatology, Brighton, 2002

6. Reading IC, Walker-Bone K, Palmer K, Coggon D, Cooper C. Diagnosis of Cervical Disorders in Epidemiological Studies: Utility of Cervical Ranges of Movement. *Rheumatology* 2002; Vol. 41: 45, 67S

Oral poster presentation at the Annual General Meeting of the British Society for Rheumatology, Manchester, 2003

7. Reading IC, Walker-Bone KE, Palmer KT, Cooper C, Coggon D. Do Hand Symptoms Predict Carpal Tunnel Syndrome: Results of a Population-based Survey. *Rheumatology* 2003; Vol. 42: 37, 46S

Oral presentation at the Annual General Meeting of the British Society for Rheumatology, Manchester, 2003

8. Reading IC, Walker-Bone KE, Palmer KT, Coggon D, Cooper C. Does Physical Examination Assist in the Evaluation of Disorders of the Neck and Upper Limb?. *Rheumatology* 2003; Vol. 42: 15, OP46

Appendices

Appendix I: The Southampton Examination Proforma and instructions

SERIAL NO: [ ][ ] [ ][ ][ ]

MRC

MEDICAL RESEARCH COUNCIL

and

ARC

ARTHRITIS AND RHEUMATISM COUNCIL
FOR RESEARCH

Community Survey of Musculoskeletal Complaints:

Examination Proforma

EXAMINATION PROFORMA

Interviewer: ☐

Date: ☐☐ ☐☐ ☐☐
day month year

Height cms ☐☐☐ *Weight* kg ☐☐☐.☐

NECK

Range of movement (°)?

		Active Movement
Rotation	right side	☐☐☐
	left side	☐☐☐
Flexion		☐☐☐
Extension		☐☐☐
Lateral flexion	right side	☐☐☐
	left side	☐☐☐

SHOULDERS

Left Side

1 History: Where is the pain located?

	Yes
No pain	☐
Deltoid area	☐
Anterior shoulder	☐
Acromioclavicular joint	☐
Subacromial bursa	☐
Diffuse	☐
Elsewhere?	☐

(describe) ______________________

2 Palpation: Where is it maximally tender?

	Yes
No tenderness	☐
	☐
	☐
	☐
	☐
	☐
	☐

(describe) ______________________

3 Pain on resisted movement?

	No	Yes
Elbow flexion	☐	☐
Forearm supination	☐	☐
External rotation	☐	☐
Internal rotation	☐	☐
Abduction	☐	☐

	No	Yes	Range of movement (°)?	
Painful arc?	☐	☐	☐☐☐ (started)	☐☐☐ (stopped)

4 Stress test, acromioclavicular joint

	No	Yes
Acromioclavicular joint pain on adduction?	☐	☐

5 Range of movement (°)?

	Active Movement	*Passive Movement*
Abduction	☐☐☐ °	☐☐☐ °
Forward flexion	☐☐☐ °	☐☐☐ °
Extension	☐☐☐ °	☐☐☐ °
External rotation	☐☐☐ °	☐☐☐ °
Internal rotation	☐☐☐ °	☐☐☐ °

SHOULDERS

Right Side

1 History: Where is the pain located?

	Yes
No pain	☐
Deltoid area	☐
Anterior shoulder	☐
Acromioclavicular joint	☐
Subacromial bursa	☐
Diffuse	☐
Elsewhere?	☐

(describe) ____________________

2 Palpation: Where is it maximally tender?

	Yes
No tenderness	☐
	☐
	☐
	☐
	☐
	☐
	☐

(describe) ____________________

3 Pain on resisted movement?

	No	Yes
Elbow flexion	☐	☐
Forearm supination	☐	☐
External rotation	☐	☐
Internal rotation	☐	☐
Abduction	☐	☐

	No	Yes	Range of movement (°)?	
Painful arc?	☐	☐	☐☐☐ (started)	☐☐☐ (stopped)

4 Stress test, acromioclavicular joint

	No	Yes
Acromioclavicular joint pain on adduction?	☐	☐

5 Range of movement (°)?

	Active Movement	*Passive Movement*
Abduction	☐☐☐°	☐☐☐°
Forward flexion	☐☐☐°	☐☐☐°
Extension	☐☐☐°	☐☐☐°
External rotation	☐☐☐°	☐☐☐°
Internal rotation	☐☐☐°	☐☐☐°

ELBOWS

Left Side

1 History: Where is the pain located?	Yes	2 Palpation: Where is it maximally tender?	Yes
No pain	☐	No tenderness	☐
Lateral elbow	☐		☐
Medial elbow	☐		☐
Posterior elbow	☐		☐
Other	☐		☐
(describe) ____________		*(describe)* ____________	

Other observations/procedures:

	No	Yes	Crepitus? Yes
Pain lateral elbow on resisted wrist extension?	☐	☐	☐
Pain medial elbow on resisted wrist flexion?	☐	☐	☐
Swelling over posterior elbow joint?	☐	☐	

ELBOWS

Right Side

1 History: Where is the pain located?	Yes	2 Palpation: Where is it maximally tender?	Yes
No pain	☐	No tenderness	☐
Lateral elbow	☐		☐
Medial elbow	☐		☐
Posterior elbow	☐		☐
Other	☐		☐
(describe) ______________________		*(describe)* ______________________	

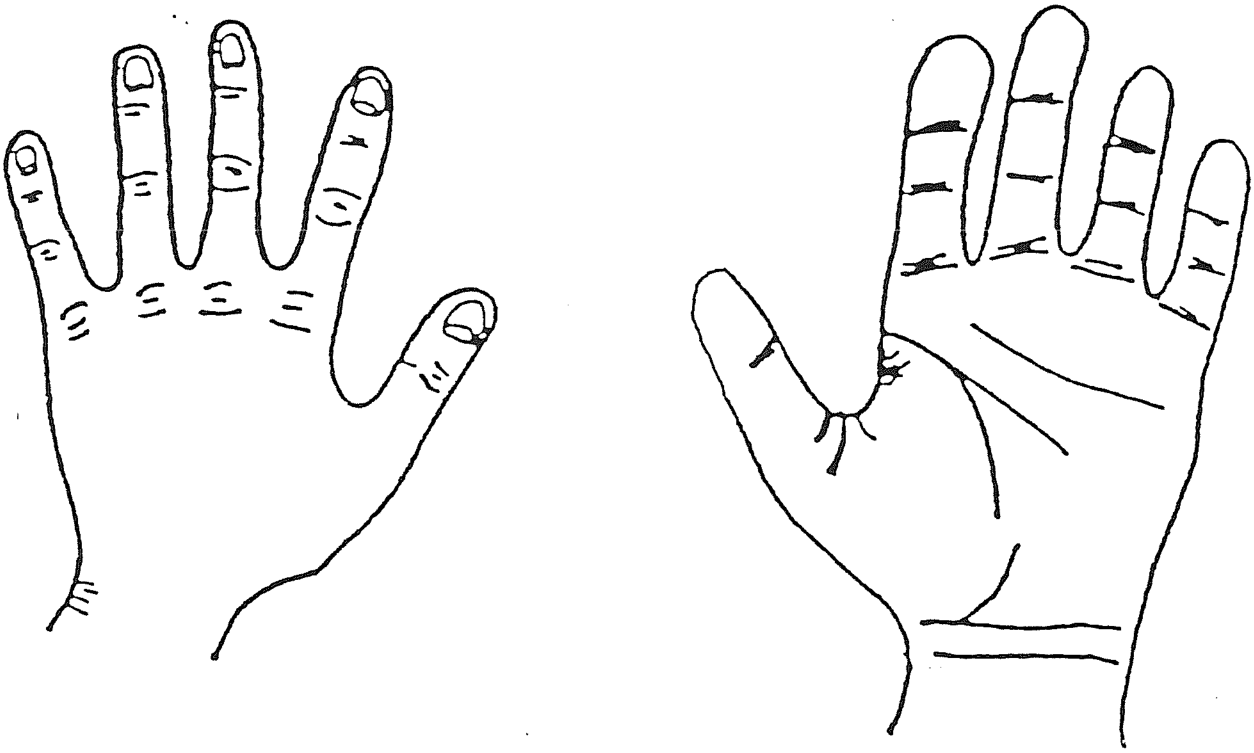
Other observations/procedures:

	No	Yes	Crepitus? Yes
Pain lateral elbow on resisted wrist extension?	☐	☐	☐
Pain medial elbow on resisted wrist flexion?	☐	☐	☐
Swelling over posterior elbow joint?	☐	☐	

KATZ HAND DIAGRAM

Left Side

If the subject has indicated tingling or numbness in the hand(s)/arm(s) in the past 7 days (question 30), indicate where it (they) occurred by shading the affected parts on the diagram below.

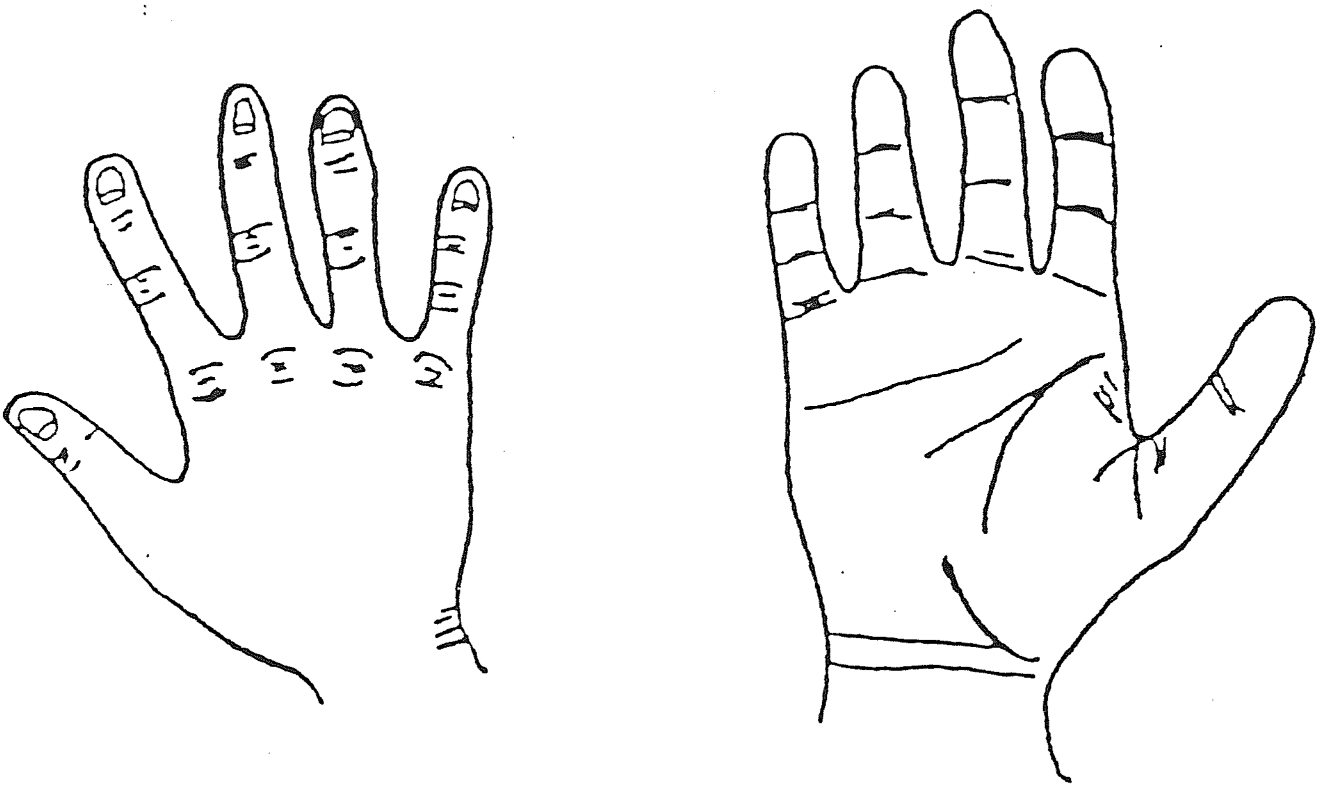


Diagnosis: classical ☐ probable ☐ possible ☐ unlikely ☐

KATZ HAND DIAGRAM

Right Side

If the subject has indicated tingling or numbness in the hand(s)/arm(s) in the past 7 days (question 30), indicate where it (they) occurred by shading the affected parts on the diagram below.



Diagnosis: classical ☐ probable ☐ possible ☐ unlikely ☐

FOREARMS AND HANDS

Left Side

1 History: location of pain (on movement) / Palpation: maximum tenderness? / Swelling?

	History: location of pain (on movement) Yes	Palpation: maximum tenderness? Yes	Swelling? Yes
dorsal forearm	☐	☐	☐
palmar forearm	☐	☐	☐
dorsal wrist	☐	☐	☐
palmar wrist	☐	☐	☐
radial wrist	☐	☐	☐
medial wrist	☐	☐	☐
other	☐	☐	☐

(describe) ______________________ *(describe)* ______________________ *(describe)* ______________________

2 Pain on resisted movement

	No	Yes	Crepitus? Yes
radial wrist	☐	☐	☐
medial wrist	☐	☐	☐
finger extension	☐	☐	☐
finger flexion	☐	☐	☐

3 Hand examination

		No	Yes		No	Yes
Muscle wasting	thenar eminence	☐	☐	hypothenar eminence	☐	☐
Dupuytren's contracture		☐	☐			
Heberden's nodes		☐	☐			

Light touch:

	normal	*abnormal*
thumb	☐	☐
index finger	☐	☐
little finger	☐	☐

Thumb base:

	No	Yes
Pain?	☐	☐
Tenderness?	☐	☐

	No	Yes
Positive Phalen's test?	☐	☐
Positive Tinel's test?	☐	☐
Weakness of thumb abduction	☐	☐
Pain on resisted left thumb extension?	☐	☐
Positive Finkelstein test?	☐	☐

	No	Yes
thumb opposition	☐	☐

FOREARMS AND HANDS

Right Side

1 History: location of pain (on movement)	Yes	Palpation: maximum tenderness? Yes	Swelling? Yes
dorsal forearm	☐	☐	☐
palmar forearm	☐	☐	☐
dorsal wrist	☐	☐	☐
palmar wrist	☐	☐	☐
radial wrist	☐	☐	☐
medial wrist	☐	☐	☐
other	☐	☐	☐

(describe) ______________________ *(describe)* ______________________ *(describe)* ______________________

2 Pain on resisted movement	No	Yes	Crepitus? Yes
radial wrist	☐	☐	☐
medial wrist	☐	☐	☐
finger extension	☐	☐	☐
finger flexion	☐	☐	☐

3 Hand examination

		No	Yes		No	Yes
Muscle wasting	thenar eminence	☐	☐	hypothenar eminence	☐	☐
Dupuytren's contracture		☐	☐			
Heberden's nodes		☐	☐			

Light touch:

	normal	*abnormal*
thumb	☐	☐
index finger	☐	☐
little finger	☐	☐

Thumb base:

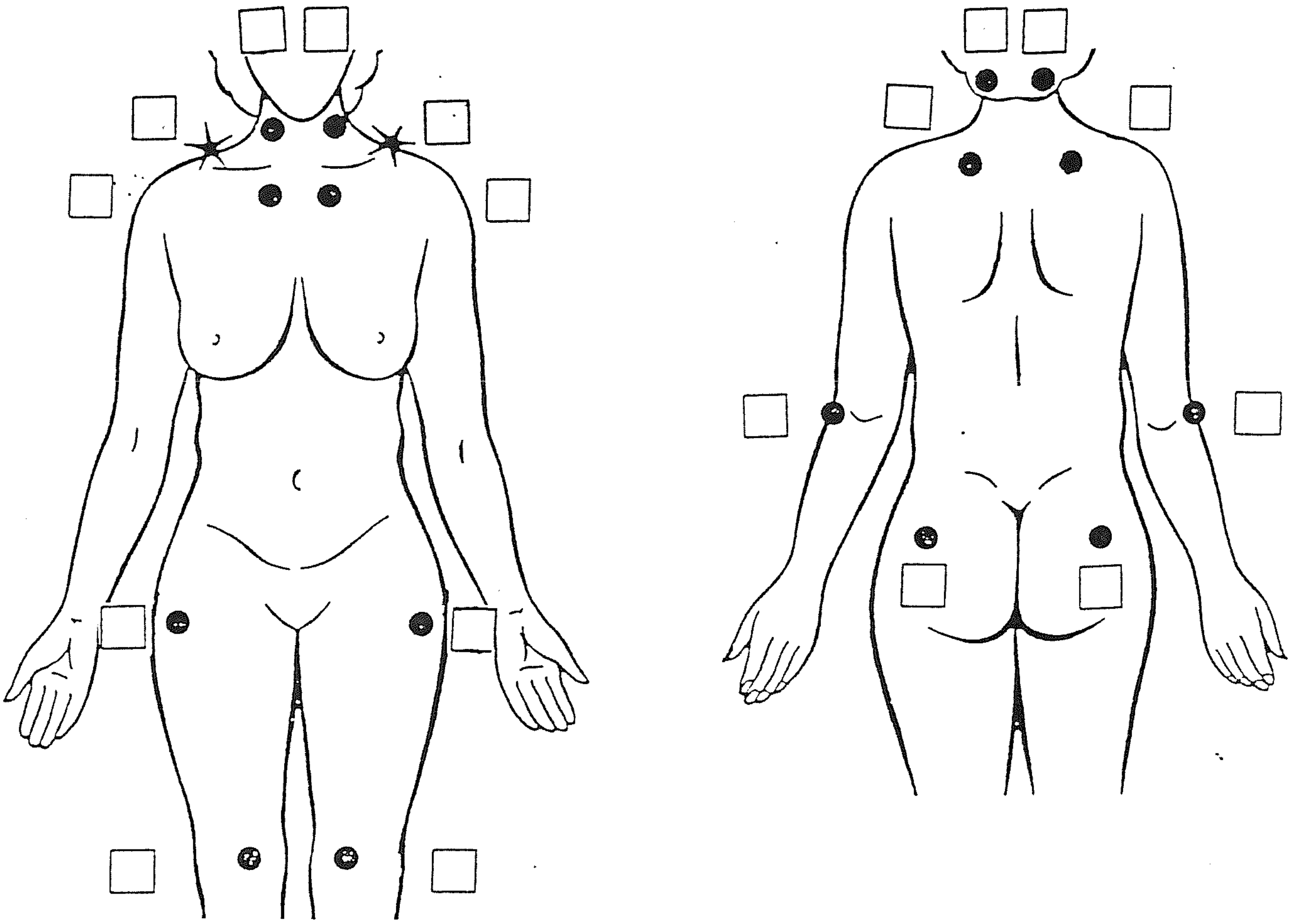
	No	Yes
Pain?	☐	☐
Tenderness?	☐	☐

	No	Yes
Positive Phalen's test?	☐	☐
Positive Tinel's test?	☐	☐
Weakness of thumb abduction	☐	☐
Pain on resisted right thumb extension?	☐	☐
Positive Finkelstein test?	☐	☐

	No	Yes
thumb opposition	☐	☐

Fibromyalgia tender spots

(Tick those that are tender)



Electroneurometry

	R			L		
Latency (milliseconds)						
Sensory	☐	☐	☐	☐	☐	☐
Motor	☐	☐	☐	☐	☐	☐

This completes the Examination Proforma. Thank the subject for their assistance.

Clinical assessment schedule for upper limb and neck disorders:

instructions for the examiner

Page 1:

Weight - is measured using electronic portable scales.

Height - is measured using a portable stadiometer.

Neck - Range of neck movements - *the examiner explains to the subject that she wishes to measure a number of neck movements.*

Rotation

The goniometer is placed on the shoulders and the examiner requests the subject to turn the neck "as far as he/she can go without discomfort".

Flexion/extension

The plurimeter is placed on top of the head in the midline with the dial facing towards the subject's side. After adjusting the head position so that the plurimeter reads zero, the subject is requested firstly to flex and then extend the neck "as far as he/she can go without discomfort", and the maximum position of active movement is recorded.

Lateral flexion

The plurimeter is placed on top of the head in the midline with the dial facing in the direction the patient is looking. After adjusting the head position so that the plurimeter reads zero, the subject is requested to flex the neck firstly to one side, then to return to the neutral position, and then to flex the neck to the opposite side, and the maximum position of active movement is recorded.

Pages 2/3:

Location of shoulder pain (in those with pain lasting a day or more in the previous 7 days)

The examiner asks subject to point to the site of the pain.

1. The shoulder area is defined as the area bounded by a vertical line taken from the midclavicular line (anteriorly) or the mid-scapular line (posteriorly) and a horizontal line taken from the inferior border of the axilla, as illustrated in Figures 1-3. Pain outside this area is not classified as shoulder pain.

2. Where a subject has pain within the shoulder area, the location is coded by reference to the specified areas in Figure 3.
3. More than one site can be volunteered.
4. 'Diffuse' pain is said to exist if the pain is located outside the shaded areas (whether or not within them).

Location of shoulder tenderness – *the examiner asks to 'feel whether there are any tender spots', and systematically palpates the shoulder area, so that all of the areas specified in Figure 3 are included. In palpation sufficient pressure is applied to induce blanching in the nail of the examining finger.*

1. Tenderness is denoted by the presence of complaint, facial grimace, flinch or withdrawal.
2. She records the area of *maximum* tenderness, the location being coded by reference to the areas specified in Figure 3.
3. If *equally tender*, more than one site can be recorded.

Pain on resisted movement - *the examiner explains that she wants the subject to perform a number of movements which she will attempt to resist. The subject is then encouraged to perform (in order) the movements of elbow flexion; forearm supination; and external rotation, internal rotation and abduction of the shoulder while the examiner resists each movement. The subject is asked whether pain is induced by any of these movements, and if so, where it is felt.*

1. The site of the pain is determined with reference to the shoulder illustration (Figure 3).
2. Pain during resisted *elbow flexion* or *forearm supination* is only counted as positive if felt over the anterior shoulder area (see Figure 3).
3. Pain during resisted *internal/external rotation* or *abduction* of the shoulder is only counted if felt over the deltoid area (see Figure 3).

Painful arc - *With the arms at the side and the palms facing outwards, the subject is asked to abduct the shoulder. If the movement incites deltoid pain which diminishes during the arc of movement, the subject is asked to pause at the points when the pain begins and ends. Using the pleurimeter, these points are measured in a single movement.*

1. If no 'deltoid area' pain is experienced during the first attempt at shoulder abduction, or pain is felt but does not diminish by full abduction, then a painful arc is deemed to be absent.

Stress test, acromioclavicular joint - *a fingertip is applied to the acromioclavicular joint on the side under examination; the examiner then asks the subject to adduct the arm across the chest as far as possible, and during this movement applies downward pressure to the acromioclavicular joint. The subject is asked to indicate whether pain arises.*

1. Tenderness is denoted by the presence of complaint, facial grimace, flinch or withdrawal, and is only present if it occurs over the site of finger pressure.

Range of shoulder movements - *The examiner explains to the subject that she wishes to measure several shoulder movements.*

Abduction - The plurimeter is positioned on the humerus and the examiner ensures that it reads zero. With the elbow fully extended, the subject is asked to abduct the shoulder as far as possible without assistance. This position is measured (active abduction), and then the examiner assists the subject in abduction "to see how far movement can go", and the maximum range of passive movement is measured.

Forward flexion

The plurimeter is positioned on the humerus and the examiner ensures that it reads zero. With the elbow fully extended, the subject is asked to flex the shoulder as far as possible without assistance. This position is measured (active flexion), and then the examiner assists the subject in flexion "to see how far movement can go", and the maximum range of passive movement is measured.

Extension

The plurimeter is positioned on the humerus and the examiner ensures that it reads zero. With the elbow fully extended, the subject is asked to extend the shoulder as far as possible without assistance. This position is measured (active extension), and then the examiner assists the subject in extension "to see how far movement can go", and the maximum range of passive movement is measured.

External rotation

With the elbow flexed at 90^{o} and by the subject's side (neutral shoulder abduction), the subject is asked to rotate the forearm away from his/her chest keeping the elbow in the same *position. A universal goniometer is placed below the olecranon and used to measure the angle of rotation.*

Internal rotation

With the elbow flexed at 90° and by the subject's side (neutral shoulder abduction), the subject is asked to rotate the shoulder towards his/her anterior chest and then behind his/her back. If the subject can perform the full movement, a value of 110° is recorded. If the subject can only perform the manoeuvre so far as the anterior chest (and not behind the back), a value of 90° is recorded.

<u>Pages 4/5</u>

Location of elbow pain (in those with pain lasting a day or more in the previous 7 days). *The examiner asks the subject to point to the site of the pain.*

1. The elbow area is defined as an area defined superiorly and inferiorly by horizontal lines 5 cm above and below the epicondyles, when the elbow rests in the neutral position under the influence of gravity (Figures 1 and 2).
2. Pain outside this area is not classified as elbow pain.
3. Where a subject has pain in the 'elbow' region, the location is coded by reference to the specified areas on accompanying elbow illustrations (Figure 4), as belonging to the medial, lateral or posterior elbow, or other elbow sites.
4. More than one site can be volunteered.

Location of elbow tenderness - *the examiner asks to 'feel whether there are any tender spots', and systematically palpates all the areas in Figure 4. In palpation sufficient pressure is applied to induce blanching in the nail of the examining finger.*

1. Tenderness is denoted by the presence of complaint, facial grimace, flinch or withdrawal.
2. She records the area of *maximum* tenderness, the location being coded by reference to the specified areas in Figure 4.
3. If *equally tender*, more than one site can be recorded.

Pain on resisted wrist movements - *the examiner extends the subject's elbow as fully as possible and asks the subject first to extend and then flex their wrist, while she attempts to prevent the movement. She enquires whether resisted movement elicits pain.*

1. Pain on resisted wrist extension is only counted as present if it is experienced over the lateral epicondyle (see Figure 4).
2. Pain on resisted wrist flexion is only counted as present if it is experienced over the medial epicondyle (see Figure 4).

Pages 6/7

Katz hand diagrams - *if the subject has experienced numbness or tingling in the previous seven days, the examiner asks the subject to identify where symptoms have occurred. The examiner shades in the appropriate area on the diagram and asks the subject to confirm the areas shaded.*

1. The subject is encouraged to offer his own response, but in the case of digit involvement care is taken to check whether or not symptoms are present (or also present) in the index finger, thumb and little finger of each hand.
2. If symptoms are present in the arm, the position and orientation of the arm in the hand diagrams is demonstrated by the examiner as a check on understanding and accuracy.
3. Responses are coded according to the criteria proposed by Katz *et al* 1990[17], as indicated in the table below.

Rating system for hand diagrams[17]

Classic	Tingling, numbness, or decreased sensation with or without pain in at least two of digits 1, 2 or 3. Palm and dorsum of the hand excluded; wrist pain or radiation proximal to the wrist allowed
Probable	Same as for classic, except palmar symptoms allowed unless confined solely to ulnar aspect
Possible	Tingling, numbness, decreased sensation and/or pain in at least one of digits 1, 2 or 3
Unlikely	No symptoms in digits 1, 2 or 3

Pages 8/9

Location of forearm and wrist pain (in those with pain lasting a day or more in the previous 7 days). *the examiner asks subject to point to the site of the pain.*

1. The forearm area is defined as an area bounded proximally by a horizontal line 5 cm below the epicondyles and distally by the proximal palmar crease. The wrist area is defined as

the area bounded by the base of the first and fifth metacarpals distally and the proximal palmar crease proximally (see Figures 1, 2 and 5). The dorsal, palmar, radial and medial aspects of forearm and wrist are defined in relation to the anatomical position.

2. Pain outside these areas is not classified as forearm or wrist pain.

Location of forearm and wrist tenderness - *the examiner asks to 'feel whether there are any tender spots'. She inspects for swelling over the dorsal and palmar wrist. She then asks the subject to place both palms together and to rest the forearms on top of the table fully extended, and inspects for swelling around the radial styloid and extending proximally for 8 - 10 cm above the thumb base. She then systematically palpates the forearm/wrist so that all of the areas specified in Figures 1 and 2 (forearm and wrist) and Figure 5 (wrist) are included in the examination. In palpation sufficient pressure is applied to induce blanching in the nail of the examining finger.*

1. Tenderness is denoted by the presence of complaint, facial grimace, flinch or withdrawal.
2. Swelling is counted as present over the radial wrist if there is swelling in the area illustrated by diagram (Figure 5).
3. The examiner records the area of *maximum* tenderness, (although more than one site can be volunteered if they are equally tender).
4. She also records any palpable crepitus over the dorsal and palmar wrist.

Pain on resisted finger movements - *The examiner explains that she wants the subject to perform a number of movements which she will attempt to overcome. Finger extension and finger flexion are demonstrated. The examiner attempts to overcome each of these movements using a similar one on her own part (e.g. testing the strength of finger extension with the examiner's own fingers in extension). Inquiry is made about pain on resisted movement.*

1. Pain on resisted finger extension is only counted as present if it is experienced over the dorsal surface of the forearm.
2. Pain on resisted finger flexion is only counted as present if it is experienced over the anterior aspect of the forearm.

Hand examination - *the examiner asks the subject to place his open hand on the table, palm uppermost. She inspects the thenar and hypothenar eminences (see Figure 5) for wasting; inspects and palpates the palm for evidence of Dupuytren's contracture; and then inspects the fingers for Heberden's nodes.*

1. Wasting is denoted by flattening or concavity of the area, rather than the normal convexity. The two palms should be compared for asymmetry. Only definite wasting should be recorded.
2. Dupuytren's contracture causes a hard, thickened nodule or tract, visible and palpable in the palm overlying the tendon of the ring and/or little finger at the distal palmar crease. The ring/little finger may be permanently flexed. Count only a definite visible and palpable palmar swelling following the line of the flexor tendon.
3. Heberden's nodes are bulbous visible and palpable swellings of the lateral aspects of the distal interphalangeal joints (see Figure 5).

Light touch - *the examiner then asks the subject to close his eyes, keeping his palms uppermost. She explains that she intends to touch the fingers and thumbs gently to test for normal feeling. The subject is asked to say 'Yes' each time contact is felt. The distal parts of the thumb, index finger and little finger of each hand are touched lightly, and a response sought. The procedure is repeated two or three times to gauge the consistency of response; and once more to compare responses at symmetrical positions on the other side.*

1. The result is counted as abnormal if the subject indicates that feeling is diminished or absent; or that it is clearly less than at the same position on the other hand.
2. The result is only counted as abnormal if responses appear to be consistent when the procedure is repeated.

Phalen's test - The *examiner asks the subject to rest both elbows on the table, with forearms vertical, and allow the hands to assume a posture of marked wrist flexion, maintained for a minute. The subject is asked to describe any discomfort, pain, pins and needles or numbness.*

1. The test is counted as positive if pain, pins and needles or numbness occur in one or more of the following sites: the thumb, index finger, middle finger or medial palmar surface of either hand.

Tinel's test - *The examiner asks the subject to extend the wrist, with the volar wrist surface uppermost. She then percusses each wrist briskly three times with a tendon hammer over the flexor retinaculum, just radial to the palmaris longus tendon at the distal palmar crease. The subject is asked to describe any discomfort, pain, pins and needles or numbness.*

1. The test is counted as positive if pain, pins and needles or numbness occur in one or more of the following sites: the thumb, index finger, middle finger or medial palmar surface of either hand.

Weakness of thumb abduction and opposition - *the examiner asks the subject to place the dorsum of each hand in turn flat on a surface, then to abduct the thumb so that it points to the ceiling and to maintain its position against resistance. The examiner applies a counterforce with her own thumb in an attempt to defeat thumb abduction. The examiner then asks the subject to oppose the tips of his thumb and little finger for each hand in turn, and applies a counterforce with her own thumb and index finger in an attempt to defeat thumb opposition.*

1. Weakness is deemed to be present if the subject's attempts to maintain the position are readily and easily defeated. (Check the subject has understood the instruction).

Resisted extension of the thumb – *The examiner asks the subject to extend the thumb. She resists the movement and asks the subject whether this causes pain.*

1. Pain on resisted thumb extension is only counted as present if it is felt over the extensor tendons around the radial styloid process.

Finkelstein's test - *The examiner asks the subject to place the thumb in the palm of his hand with the fingers flexed over it. She then imparts passive ulnar deviation at the wrist and asks whether this causes pain.*

1. The test is counted as positive if pain is reproduced over the distal radius and radial side of the wrist.

<u>Page 10</u>

Fibromyalgia tender spot assessment - *the examiner will palpate and enquire about tenderness in each of the areas proposed in the American College of Rheumatology's classification criteria for fibromyalgia*[14]. *She records the findings on the diagram in the* *proforma. In palpation sufficient pressure is applied to induce blanching in the nail of the examining finger.*

1. Tenderness is denoted by the presence of complaint, facial grimace, flinch or withdrawal. (N.B. a feeling of *pressure* experienced by the subject at each site is *not* sufficient to be classed as a positive reaction).

Measurements of median nerve conduction velocity - *Measurement of median nerve conduction velocities will be undertaken using the Nerve Pace S-200 Nerve Conduction Monitor (Electroneuronometer), Neutron Medical, Lawrenceville, New Jersey USA. The subject sits with his arm resting on a flat surface and the elbow flexed at 30^0. The thenar eminence, wrist area and dorsum of the hand are cleansed with alcohol and three electrodes applied to the subject's hand. A ground electrode is placed on the dorsum of the hand. The active electrode is placed over the belly of the abductor pollicis brevis muscle and the reference electrode on the radial/volar aspect of the thumb. Conductive gel is applied to the stimulating probe, and the distal portion of the probe placed over the path of the median nerve 3 cm proximal to the distal wrist crease. A low intensity stimulus is delivered to the skin over the median nerve. Measurements are made bilaterally of distal sensory and motor latencies, and the values (in milliseconds) recorded.*

Figure 1: Anatomical landmarks in the upper limb: anterior view

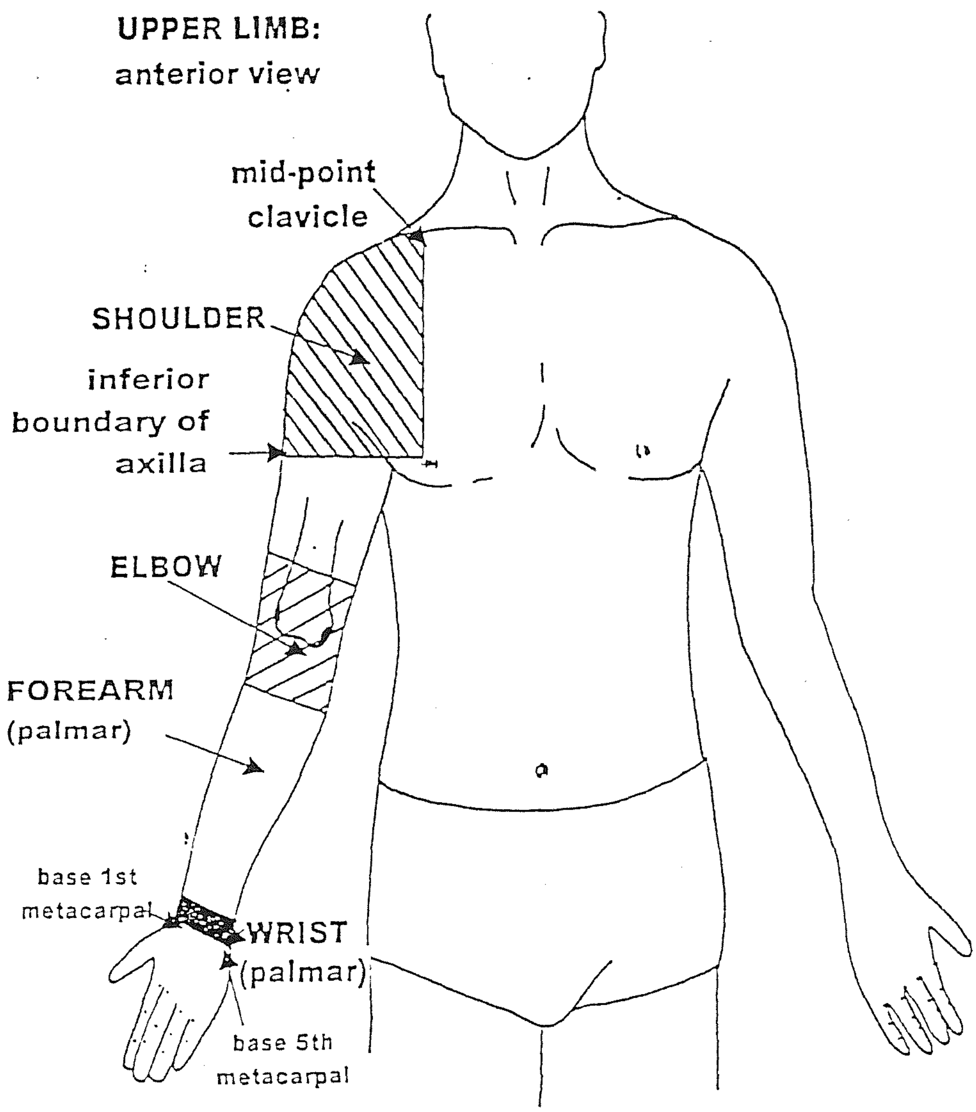


Figure 2: Anatomical landmarks in the upper limb: posterior view

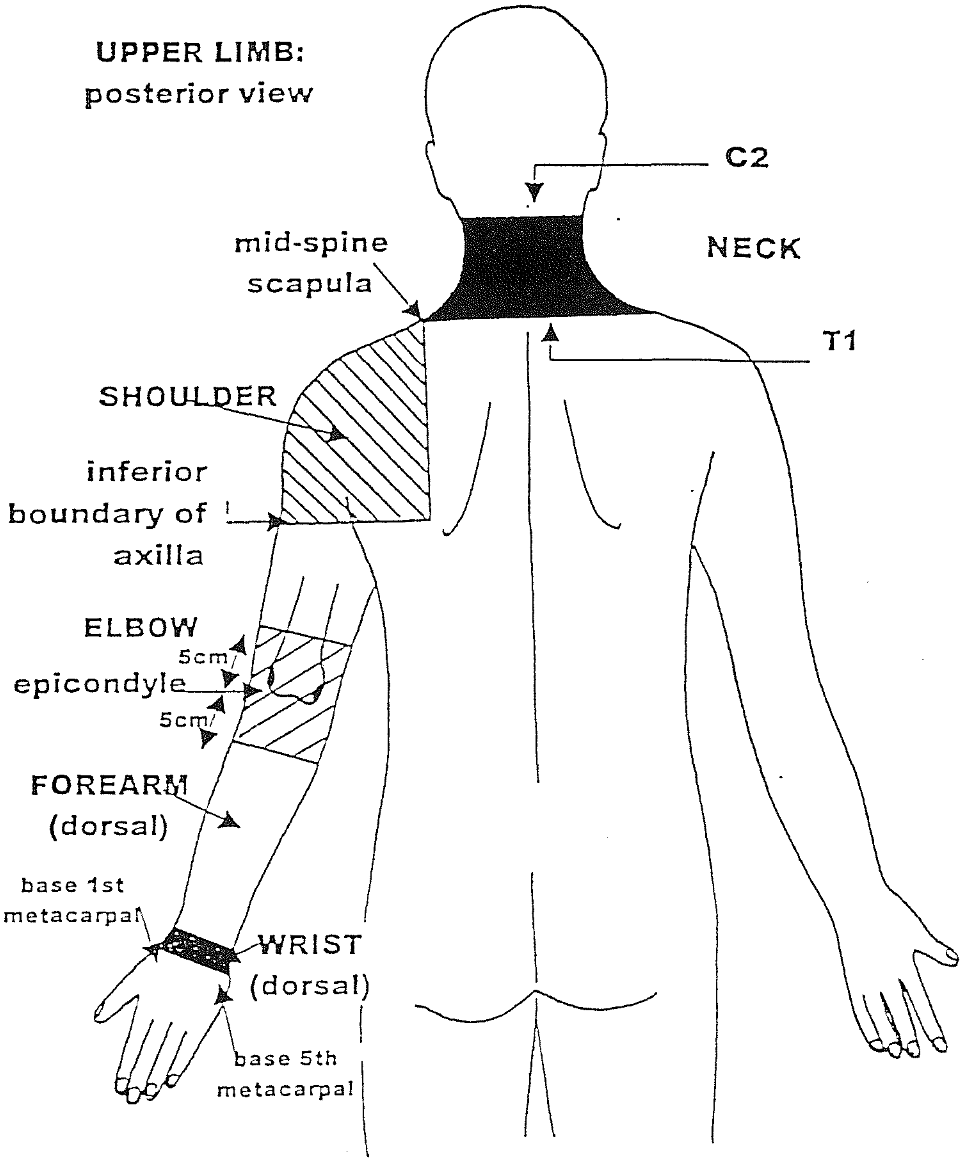
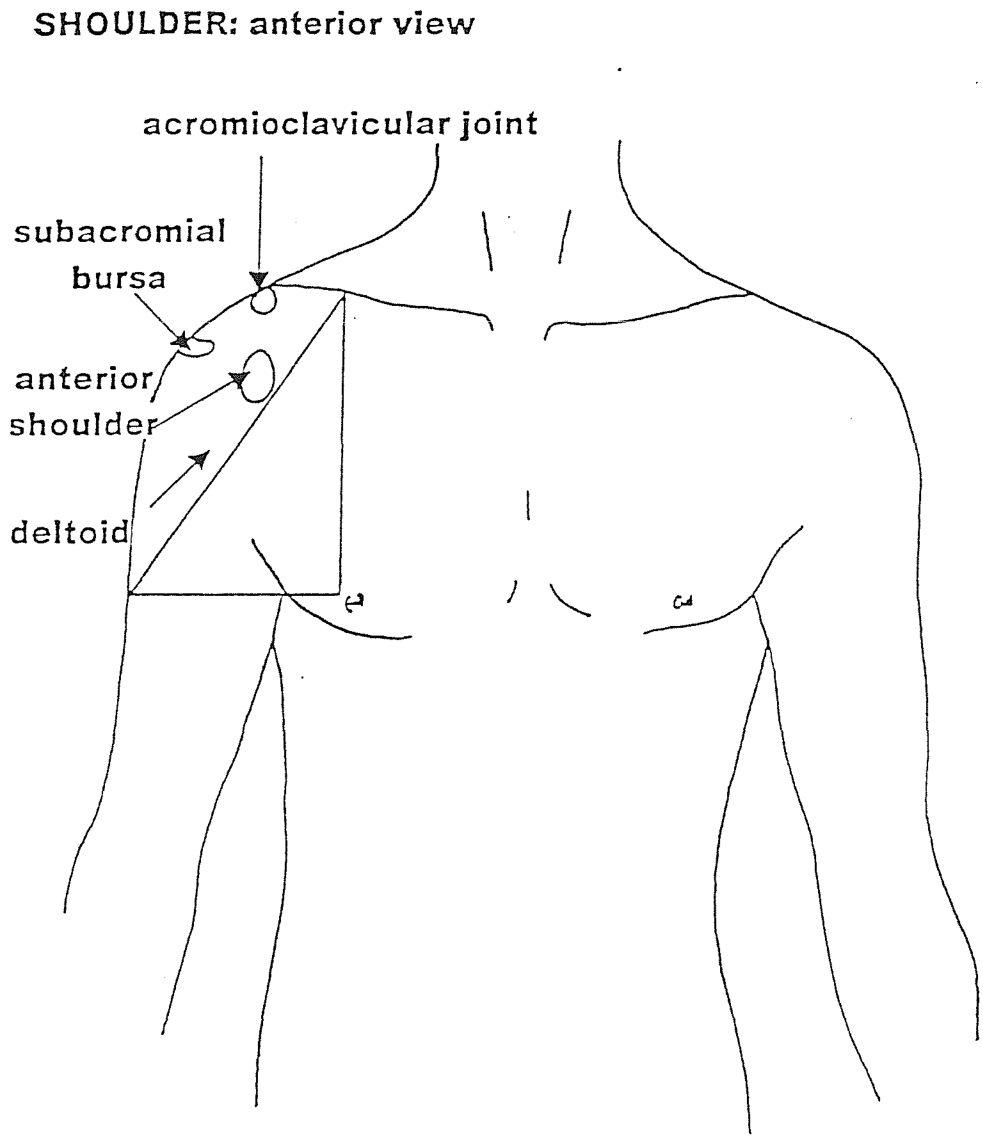


Figure 3: Anatomical landmarks in the shoulder



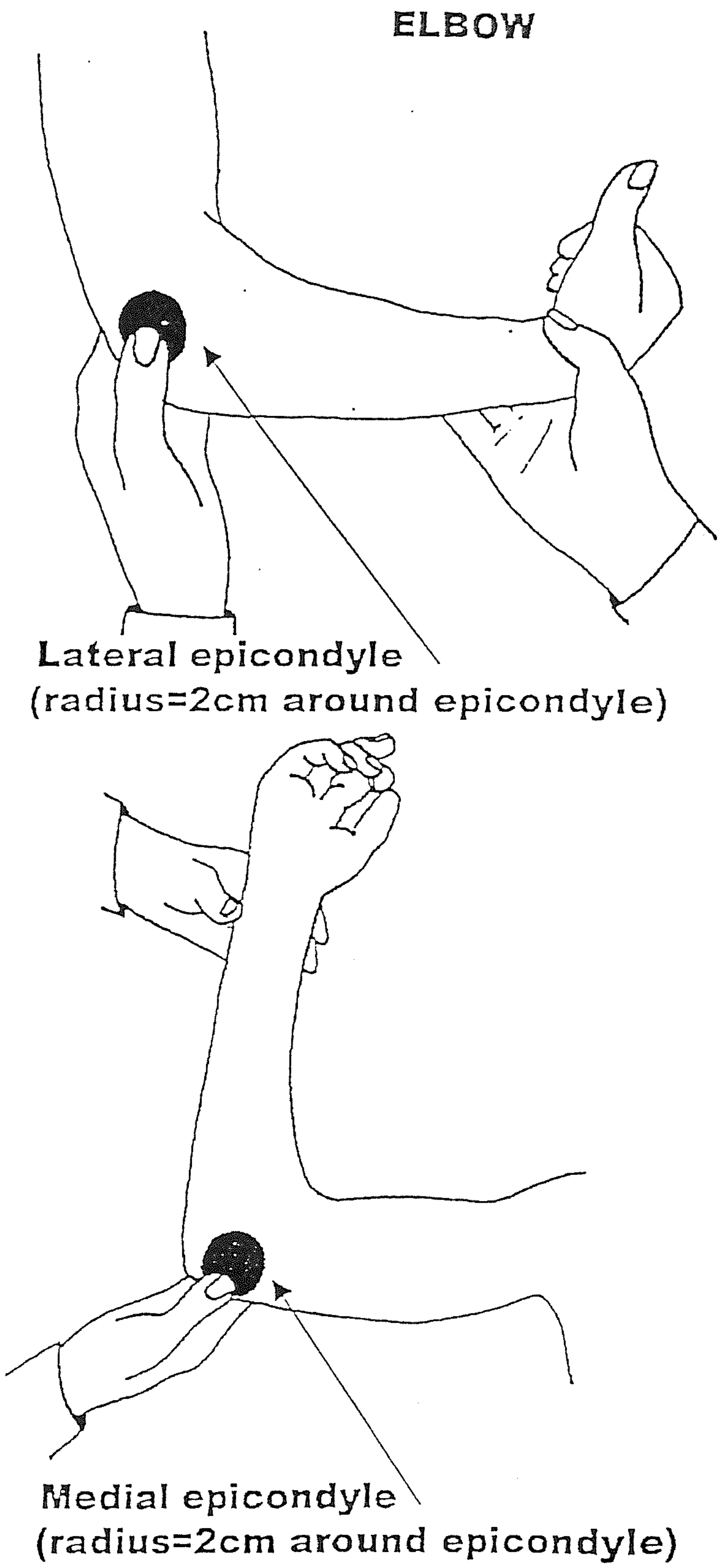


Figure 4: Anatomical landmarks at the elbow

Appendix II: The nurse interview

APPENDIX E
[Nurse administered]

MRC

MEDICAL RESEARCH COUNCIL

and

ARC

ARTHRITIS AND RHEUMATISM COUNCIL
FOR RESEARCH

Follow-up Questionnaire on Musculoskeletal Complaints

Subject's name: ______________________________ *Serial No:* ☐☐☐☐☐

Interviewer: ☐ *Date*: ☐☐ ☐☐ ☐☐

day month year

Is the pain you reported in the questionnaire in the same place today?

No ☐ Yes ☐ N/A ☐

Have you developed a new pain since completing the questionnaire?

No ☐ Yes ☐

ECTION ONE: NECK PAIN

) Earlier complainers

(These questions apply to those who complained of neck pain in the postal survey. For others, go to question 6, page 3).

1 When you filled in our earlier questionnaire, you told us that you had had pain in your neck. If you added up all the days on which you have had neck pain over the past 12 months, how many days would that make altogether?

less than 7 days ☐ 7 - 13 days ☐ 14 - 27 days ☐ 28 days or more ☐ don't know ☐

2 During the past 12 months have you:

consulted a doctor about your neck pain?	No ☐	Yes ☐
consulted a physiotherapist about your neck pain?	No ☐	Yes ☐
consulted a chiropractor or osteopath about your neck pain?	No ☐	Yes ☐
taken a prescribed medicine for your neck pain?	No ☐	Yes ☐
taken a non-prescribed medicine for your neck pain?	No ☐	Yes ☐
had physiotherapy or manipulation for your neck pain?	No ☐	Yes ☐

3 During the past 12 months have you taken time off work because of neck pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

If *YES*, how many days have you taken off work altogether over the whole 12 months because of neck pain? ☐☐☐ days

4 During the past 12 months have you changed what you do at work because of neck pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

5 In the past 12 months have any of the following been made more difficult or impossible by neck pain?

	No difficulty	*Difficult but not impossible*	*Impossible*
Sleeping	☐	☐	☐
Driving	☐	☐	☐
Carrying bags (eg shopping)	☐	☐	☐
Getting dressed	☐	☐	☐

) All Subjects

(These questions apply to everyone, whether pain positive or pain negative in the postal survey)

6 During the past 7 days have you had pain lasting a day or longer in your neck? No ☐ Yes ☐

ECTION 2: SHOULDER PAIN

) Earlier complainers

(These questions apply to those who complained of shoulder pain in the postal survey. For others, go to question 12, page 4).

7 When you filled in our earlier questionnaire, you told us that you had had pain in your shoulder(s). If you added up all the days on which you have had shoulder pain over the past 12 months, how many days would that make altogether?

less than 7 days ☐ 7 - 13 days ☐ 14 - 27 days ☐ 28 days or more ☐ don't know ☐

8 During the past 12 months have you:

consulted a doctor about your shoulder pain?	No ☐	Yes ☐
consulted a physiotherapist about your shoulder pain?	No ☐	Yes ☐
consulted a chiropractor or osteopath about your shoulder pain?	No ☐	Yes ☐
taken a prescribed medicine for your shoulder pain?	No ☐	Yes ☐
taken a non-prescribed medicine for your shoulder pain?	No ☐	Yes ☐
had physiotherapy or manipulation for your shoulder pain?	No ☐	Yes ☐
had an injection in your shoulder(s) to relieve the pain?	No ☐	Yes ☐

9 During the <u>past 12 months</u> have you taken time off work because of shoulder pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

If *YES*, how many days have you taken off work altogether over the whole 12 months because of shoulder pain? ☐☐☐ days

10 During the <u>past 12 months</u> have you changed what you do at work because of shoulder pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

11 In the <u>past 12 months</u> have any of the following been made more difficult or impossible by shoulder pain?

	No difficulty	*Difficult but not impossible*	*Impossible*
Sleeping	☐	☐	☐
Driving	☐	☐	☐
Carrying bags (eg shopping)	☐	☐	☐
Getting dressed	☐	☐	☐
Opening doors	☐	☐	☐
Getting things down from high shelves	☐	☐	☐
Fastening your clothing (eg buttons, shoelaces, zip, bra)	☐	☐	☐
Heavy jobs around the house	☐	☐	☐
Moving your arm(s) or hand(s)	☐	☐	☐

All Subjects

(These questions apply to everyone, whether pain positive or pain negative in the postal survey)

12 During the <u>past 7 days</u> have you had pain lasting <u>a day or longer</u> in one or both shoulder(s)?

Right shoulder	No ☐	Yes ☐
Left shoulder	No ☐	Yes ☐

SECTION THREE: ELBOW PAIN

A) Earlier complainers

(These questions apply to those who complained of elbow pain in the postal survey. For others, go to question 18, page 5).

13 When you filled in our earlier questionnaire, you told us that you had had pain in your elbow(s). If you added up all the days on which you have had elbow pain over the past 12 months, how many days would that make altogether?

less than 7 days ☐ 7 - 13 days ☐ 14 - 27 days ☐ 28 days or more ☐ don't know ☐

14 During the past 12 months have you:

consulted a doctor about your elbow pain?	No ☐	Yes ☐
consulted a physiotherapist about your elbow pain?	No ☐	Yes ☐
consulted a chiropractor or osteopath about your elbow pain?	No ☐	Yes ☐
taken a prescribed medicine for your elbow pain?	No ☐	Yes ☐
taken a non-prescribed medicine for your elbow pain?	No ☐	Yes ☐
had physiotherapy or manipulation for your elbow pain?	No ☐	Yes ☐
had an injection in your elbow(s) to relieve the pain?	No ☐	Yes ☐

15 During the past 12 months have you taken time off work because of elbow pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

If *YES*, how many days have you taken off work altogether over the whole 12 months because of elbow pain?

☐☐☐ days

16 During the past 12 months have you changed what you do at work because of elbow pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

17 In the past 12 months have any of the following been made more difficult or impossible by elbow pain?

	No difficulty	*Difficult but not impossible*	*Impossible*
Sleeping	☐	☐	☐
Driving	☐	☐	☐
Carrying bags (eg shopping)	☐	☐	☐
Getting dressed	☐	☐	☐

3) All Subjects

(These questions apply to everyone, whether pain positive or pain negative in the postal survey)

18 During the past 7 days have you had pain lasting a day or longer in one or both elbow(s)?

Right elbow	No ☐	Yes ☐
Left elbow	No ☐	Yes ☐

ECTION FOUR: WRIST/HAND PAIN

) Earlier complainers

(These questions apply to those who complained of wrist/hand pain in the postal survey. For others, go to question 24, page 6).

19 When you filled in our earlier questionnaire, you told us that you had had pain in your wrist/hand(s). If you added up all the days on which you have had wrist/hand pain over the past 12 months, how many days would that make altogether?

less than 7 days ☐ 7 - 13 days ☐ 14 - 27 days ☐ 28 days or more ☐ don't know ☐

20 During the past 12 months have you:

consulted a doctor about your wrist/hand pain?	No ☐	Yes ☐
consulted a physiotherapist about your wrist/hand pain?	No ☐	Yes ☐
consulted a chiropractor or osteopath about your wrist/hand pain?	No ☐	Yes ☐
taken a prescribed medicine for your wrist/hand pain?	No ☐	Yes ☐
taken a non-prescribed medicine for your wrist/hand pain?	No ☐	Yes ☐
had physiotherapy or manipulation for your wrist/hand pain?	No ☐	Yes ☐
had an injection in your wrist/hand(s) to relieve the pain?	No ☐	Yes ☐

21 During the past 12 months have you taken time off work because of wrist/hand pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

If *YES*, how many days have you taken off work altogether over the whole 12 months because of wrist/hand pain? ☐☐☐ days

22 During the past 12 months have you changed what you do at work because of wrist/hand pain?

No ☐ Yes ☐ Not worked in past 12 months ☐

23 In the past 12 months have any of the following been made more difficult or impossible by wrist/hand pain?

	No difficulty	*Difficult but not impossible*	*Impossible*
Sleeping	☐	☐	☐
Driving	☐	☐	☐
Carrying bags (eg shopping)	☐	☐	☐
Getting dressed	☐	☐	☐
Writing	☐	☐	☐
Undoing lids on bottles or jars	☐	☐	☐

3) All Subjects

(These questions apply to everyone, whether pain positive or pain negative in the postal survey)

24 During the past 7 days have you had pain lasting a day or longer in one or both wrist/hand(s)?

Right wrist/hand No ☐ Yes ☐

Left wrist/hand No ☐ Yes ☐

ECTION FIVE: NUMBNESS AND TINGLING

) Earlier complainers

(These questions apply to those who complained of numbness or tingling in their hands or arms that lasted at least three minutes in the postal survey. For others, go to question 30 below).

25 When you filled in our earlier questionnaire, you told us that you had had numbness or tingling in your fingers, thumbs or arms. If you added up all the days on which you have had such numbness or tingling over the past 12 months, how many days would that make altogether?

less than 7 days ☐ 7 - 13 days ☐ 14 - 27 days ☐ 28 days or more ☐ don't know ☐

26 During the past 12 months have you:

consulted a doctor about the numbness or tingling?	No ☐	Yes ☐
consulted a physiotherapist about the numbness or tingling?	No ☐	Yes ☐
consulted a chiropractor or osteopath about the numbness or tingling?	No ☐	Yes ☐
taken a prescribed medicine for the numbness or tingling?	No ☐	Yes ☐
taken a non-prescribed medicine for the numbness or tingling?	No ☐	Yes ☐
had physiotherapy or manipulation for the numbness or tingling?	No ☐	Yes ☐
had an injection in your wrist or arm for the numbness or tingling?	No ☐	Yes ☐
had an operation on your wrist to relieve the numbness or tingling?	No ☐	Yes ☐

27 During the past 12 months have you taken time off work because of the numbness or tingling?

No ☐ Yes ☐ Not worked in past 12 months ☐

If *YES*, how many days have you taken off work altogether over the whole 12 months because of the numbness or tingling? ☐☐☐ days

28 During the past 12 months have you changed what you do at work because of the numbness or tingling?

No ☐ Yes ☐ Not worked in past 12 months ☐

29 In the past 12 months have any of the following been more difficult or impossible by the numbness or tingling?

	No difficulty	*Difficult but not impossible*	*Impossible*
Sleeping	☐	☐	☐
Carrying bags (eg shopping)	☐	☐	☐
Doing up buttons or zips	☐	☐	☐

30 During the past 7 days have you had **numbness** (a lack of feeling) that lasted at least three minutes in your:

fingers or thumbs No ☐ Yes ☐

other parts of your hands No ☐ Yes ☐

other parts of your arm No ☐ Yes ☐

AND during the past 7 days have you had **tingling** (pins and needles) that lasted at least three minutes in your:

fingers or thumb No ☐ Yes ☐

other parts of your hands No ☐ Yes ☐

other parts of your arm No ☐ Yes ☐

31 **IF SO**, during the past 7 days, did your symptoms of numbness or tingling disturb your sleep? No ☐ Yes ☐

Appendix III: The survey questionnaire

SERIAL NO: 3622

MRC

MEDICAL RESEARCH COUNCIL

and

ARC

ARTHRITIS AND RHEUMATISM COUNCIL FOR RESEARCH

Community Survey of Musculoskeletal Complaints

The answers given on this form are confidential.

Replies will be seen by MRC staff ONLY.

RHEUMATIC QUESTIONNAIRE

SECTION ONE: ABOUT YOURSELF

1 Please fill in your name: ______________________________

date of birth: ☐☐ ☐☐ ☐☐
day month year

and your sex: male ☐ female ☐

2 How would you best describe your racial origin?

European ☐ India/Pakistan ☐ Afro-Caribbean ☐
South East Asia ☐ other ☐ ______________________________

3 What is your height? ft ☐☐ ins ☐☐ **or** cms ☐☐☐

and your weight? st ☐☐ lbs ☐☐ **or** kg ☐☐

4 Have you ever smoked regularly (ie at least once a day for a month or longer?)

no ☐ yes ☐

If *YES*, do you still smoke regularly? no ☐ yes ☐

5 Are you right or left handed? ☐ right ☐ left ☐ able to use both hands equally

6 Do you currently have a paid job? no ☐ yes ☐

If *NO*, go to question 13 page 4. If *YES*, continue.

SECTION TWO: MAIN JOB

7 What is your main job? Occupation: ______________________

AND in what industry do you carry out this occupation?
(eg farming, shipyard, car factory, shoe factory, hospital, insurance office)

Industry: ______________________

8 How long have you done this job? ☐☐ months or ☐☐ years

9 Does an average working day in the job involve any of the following?
*(Please tick **no** or **yes** for each question)*

		No	*Yes*
a)	Use of a keyboard or typewriter for longer than one hour in total?	☐	☐
b)	Other tasks involving repeated movements of the wrist or fingers for longer than one hour in total (eg using a screwdriver or soldering iron)?	☐	☐
c)	Use of a keyboard or typewriter for longer than four hours in total?	☐	☐
d)	Other tasks involving repeated movements of the wrist or fingers for longer than four hours in total?	☐	☐
e)	Repeated bending and straightening of your elbow for longer than one hour in total?	☐	☐
f)	Working for longer than one hour in total with a powered tool that makes your hand(s) or arm(s) vibrate (eg chain saw, pneumatic drill)?	☐	☐
g)	Working for longer than one hour in total with your hand held above shoulder height?	☐	☐
h)	Carrying weights on one shoulder?	☐	☐
i)	Lifting or carrying weights of 5Kg (10lbs) or more in one hand (eg a tool bag or heavy briefcase)?	☐	☐
j)	Working for longer than two hours in total with your neck bent forward?	☐	☐
k)	Working for longer than half an hour in total with your neck twisted (eg when looking to one side)?	☐	☐
l)	Piecework in which you are paid according to the number of articles or tasks you or your team make or finish in the day?	☐	☐
m)	A target number of articles or tasks that you or your team are expected to make or finish in the day?	☐	☐
n)	Payment of a bonus if you make or finish more than an agreed number of articles/tasks in the day?	☐	☐
o)	Working to tight deadlines?	☐	☐

		Often	*Sometimes*	*Seldom*	*Never*	*N/A*
10	When you have difficulties in your work, how often do you get help and support from your colleagues?	☐	☐	☐	☐	☐
	When you have difficulties in your work, how often do you get help and support from your immediate superior?	☐	☐	☐	☐	☐

		Often	*Sometimes*	*Seldom*	*Never/almost never*
11	Do you have a choice in deciding <u>how</u> you do your work?	☐	☐	☐	☐
	Do you have a choice in deciding <u>what</u> you do at work?	☐	☐	☐	☐
	Do you have a choice in deciding your work timetable and breaks?	☐	☐	☐	☐

		Very satisfied	*Satisfied*	*Dissatisfied*	*Very Dissatisfied*
12	How satisfied have you been with your job as a whole, taking everything into consideration?	☐	☐	☐	☐

		No	*Yes*
13	Have you ever changed from a job because of a problem (or problems) with your neck, arm, shoulder, elbow, wrist or hand?	☐	☐

If ***YES***, what was the first job that you had to leave and what was the problem?

Job ______________________________

Problem(s) ______________________________

____

____

SECTION THREE: ABOUT YOUR HOBBIES

14 Did you play any of the following SPORTS during the past 12 months?

Sport	*No*	*Yes*	*If YES, please indicate number of times played over past 12 months* ***(If you cannot give the exact figure, please give your best estimate)***
tennis	☐	☐	☐☐☐ times
squash	☐	☐	☐☐☐ times
badminton	☐	☐	☐☐☐ times
any other racquet sport *(describe)* ________	☐	☐	☐☐☐ times
hockey	☐	☐	☐☐☐ times
golf	☐	☐	☐☐☐ times
cricket	☐	☐	☐☐☐ times
swimming	☐	☐	☐☐☐ times
football	☐	☐	☐☐☐ times
rugby	☐	☐	☐☐☐ times
aerobics/keep fit	☐	☐	☐☐☐ times
weight training	☐	☐	☐☐☐ times

15 In the past 12 months did you undertake any of the following DIY or CRAFT ACTIVITIES for <u>more than 20 hours in total</u> in your spare time?

Activity	*No*	*Yes*
digging/shovelling	☐	☐
house or fence painting	☐	☐
cutting/sawing (by hand)	☐	☐
drilling (by hand)	☐	☐
sewing (by hand)	☐	☐
typing/computer use	☐	☐
other hobby or DIY pursuit involving use of shoulder, arm or hand muscles	☐	☐

describe *1* ____________________

2 ____________________

SECTION FOUR: ABOUT YOUR HEALTH

16 Have you ever been told by a doctor that you suffered from:

Diabetes?	no ☐	yes ☐
Rheumatoid arthritis?	no ☐	yes ☐

17 Have you ever broken any of the following bones?

	No, never	*Yes, within past 12 months*	*Yes, but more than 12 months ago*
collar bone	☐	☐	☐
bone of the upper arm	☐	☐	☐
bone(s) of the forearm (wrist)	☐	☐	☐
bone(s) of the finger, thumb or hand	☐	☐	☐

18 ***(For women only)*** Have you taken any of the following medicines during the past 12 months?

contraceptive pill (birth control pill)	no ☐	yes ☐
hormone replacement therapy	no ☐	yes ☐

19 *These questions concern **pain** in the <u>past 7 days</u>. Answer the questions using the tick boxes - 1 tick for each question.*

*During the **past 7 days**, have you had pain lasting a **day or longer** in your:*	*If **YES**, did the pain make it difficult for you to carry out your normal activities (eg job, housework, hobbies) during the **past 7 days**?*	***AND** how long ago did the pain <u>first</u> begin?*
Neck No ☐ Yes ☐	No difficulty ☐ Difficult but not impossible ☐ Some activities impossible ☐	In the past week ☐ Longer than a week, but within the past month ☐ Longer than a month, but within the past year ☐ Longer than a year ☐
Shoulder(s) No ☐ Yes ☐ right shoulder ☐ left shoulder ☐ both shoulders	No difficulty ☐ Difficult but not impossible ☐ Some activities impossible ☐	In the past week ☐ Longer than a week, but within the past month ☐ Longer than a month, but within the past year ☐ Longer than a year ☐
Elbow(s) No ☐ Yes ☐ right elbow ☐ left elbow ☐ both elbows	No difficulty ☐ Difficult but not impossible ☐ Some activities impossible ☐	In the past week ☐ Longer than a week, but within the past month ☐ Longer than a month, but within the past year ☐ Longer than a year ☐
Wrist(s)/hand(s) No ☐ Yes ☐ right wrist/hand ☐ left wrist/hand ☐ both wrists/hands	No difficulty ☐ Difficult but not impossible ☐ Some activities impossible ☐	In the past week ☐ Longer than a week, but within the past month ☐ Longer than a month, but within the past year ☐ Longer than a year ☐

20 *These questions concern* ***numbness and tingling*** *in the past 7 days. Answer the questions using the tick boxes - 1 tick for each question. Please answer these questions even if you have never had problems of this sort in these parts of your body.*

In the past 7 days have you had numbness or tingling that lasted ***at least three minutes*** *in your:*	*Numbness (lack of feeling)?*	*Tingling (pins and needles)?*
Fingers/thumbs?	no ☐ yes ☐	no ☐ yes ☐
Other parts of the hand(s)?	no ☐ yes ☐	no ☐ yes ☐
Other part(s) of the arm(s)?	no ☐ yes ☐	no ☐ yes ☐

21 These questions are about how you feel and how things have been with you **during the past 4 weeks.**

For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the **past 4 weeks**:

(circle one number on each line)

		All of the time	*Most of the time*	*A good bit of the time*	*Some of the time*	*A little of the time*	*None of the time*
a)	did you feel full of life?	1	2	3	4	5	6
b)	have you been a very nervous person?	1	2	3	4	5	6
c)	have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
d)	have you felt calm and peaceful?	1	2	3	4	5	6
e)	did you have a lot of energy?	1	2	3	4	5	6
f)	have you felt downhearted and low?	1	2	3	4	5	6
g)	did you feel worn out?	1	2	3	4	5	6
h)	have you been a happy person?	1	2	3	4	5	6
i)	did you feel tired?	1	2	3	4	5	6

You have finished. Thank you for completing this form. Please could you now post it back in the envelope provided? We are grateful for your help.

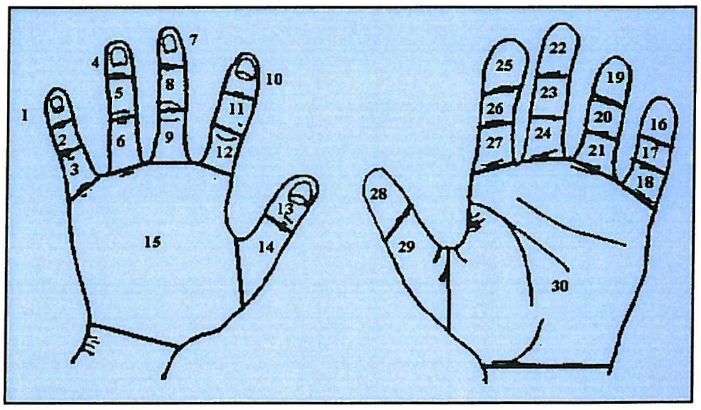
Appendix IV: Katz hand diagram cluster analysis

A preliminary cluster analysis was performed on the Katz hand diagram data alone, the results of which were included in the cluster analysis of all the hand and wrist examination data.

Thirty regions of the hand were individually coded as 1='shaded' or 0='not shaded' from the completed Katz hand diagrams (Figure A1). Subjects who had no numbness or tingling at the physical examination would therefore have all regions coded as 0='not shaded'.

Of the 4290 hands examined, six (three pairs) had missing Katz hand diagrams (the subjects reported symptoms of numbness or tingling in the clinical interview, but did not shade in the Katz diagram) and were excluded from the analysis. Amongst the Hill Lane data there were 1402 complete Katz hand diagrams, and these were analysed first.

Figure A1: Regions of the hand used in the Katz diagram cluster analysis



An extra variable was generated to denote no numbness or tingling anywhere in the hand, and was coded 1=true, 0=false. This was done to create a variable with a non-zero value in those hands with no numbness or tingling so that they would be retained in an analysis based on Jaccard's coefficient of community. Since this distance measure ignores the agreement of the absence of a trait (cell *d* in Table 18), any pairs of hands both with no numbness or tingling would have a Jaccard similarity of 0/0, and would be dropped from the analysis. Including the extra non-zero above-mentioned variable, such pairs would have a Jaccard similarity of 1/1=1, indicating perfect agreement.

The 31 binary variables were analysed in two ways, as per the elbow cluster analysis (Section 6.3), yielding the dendrograms presented in Figures A2 and A3. Ward's hierarchical method on squared Euclidean distance identified eight main clusters which were grouped into three overriding clusters. The fusion values for this hierarchical model (Figure A4, note the breaks in the y-axis) indicate the strong separation between the three overriding clusters, followed by less distinct separation between the next five cluster divisions. There is little difference in fusion values beyond those for the first eight divisions.

The group average method on Jaccard's coefficient of community identified two main clusters: one of hands with no numbness or tingling (N=1048), the leftmost vertical black line, and one of hands with some numbness or tingling (N=333), the heterogeneous cluster third from the left (Figure A3). A further five clusters were identified, four of which contained one hand, and one which contained 17 hands, and was characterised by numbness or tingling in the palm, dorsum or thumb only.

It was decided to continue analysis with the Ward's method solution which gave a clearer group structure. The k-means procedure moved 38 (2.7%) hands into another cluster, 31 of which were from Cluster 2 (coloured red in Figure A2) into Clusters 1, 3, 4 or 5.

Figures A2 and A3: Dendrograms of a) Ward's and b) group average linkage hierarchical clustering on Hill Lane data Katz hand diagram observations

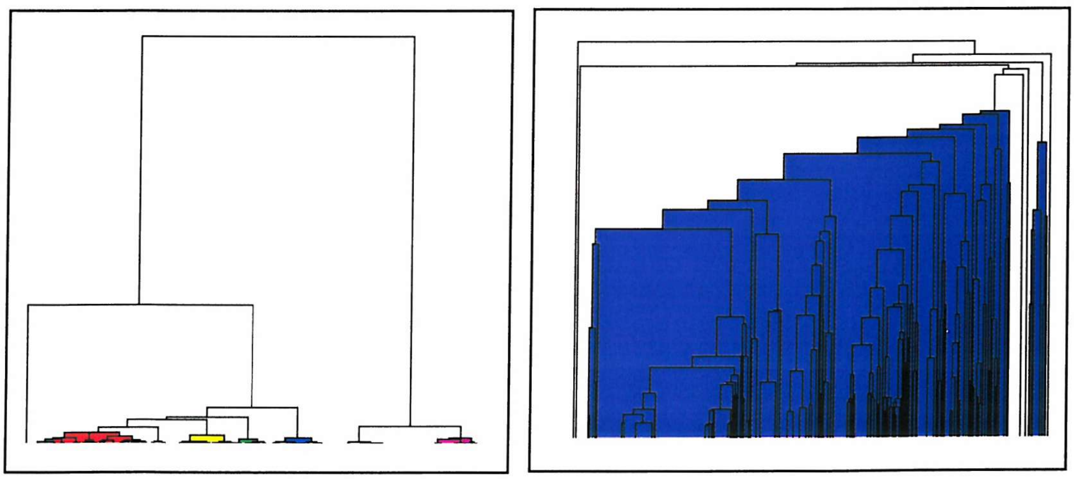
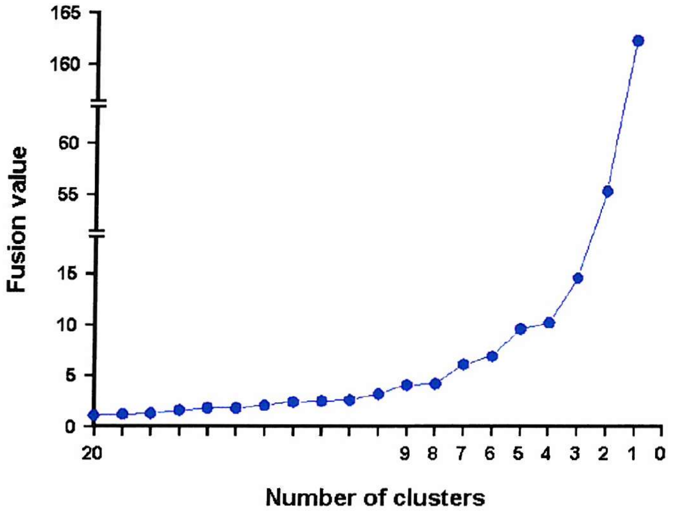


Figure A4: Fusion values for successive Katz diagram cluster joins



<u>*Characterisation of the Katz hand clusters (Hill Lane data only)*</u>

Cluster 1 (N=1065) contained hands with mostly no numbness or tingling (N=1048), or symptoms at only 1 or 2 sites, mostly in the palm or dorsum (Table A1).

Table A1: Patterns of numbness and tingling in the first four Hill Lane Katz diagram clusters

Characteristics	**Cluster**			
	1: "Normal"	**2: Mixed**	**3: Dorsal aspect**	**4: Distal**
N	**1065**	**65**	**25**	**57**
Colour in dendrogram	Black	Red	Orange	Yellow
No numbness/ tingling	*98.4*	*0*	*0*	*0*
Digit involvement:				
Little finger	0.4	22	68	93
Ring finger	0.2	52	80	100
Middle finger	0.3	66	100	100
Index finger	0.1	60	84	100
Thumb	0.2	43	56	46
Aspect of involvement:				
Dorsal	0.8	69	100	100
Palmar	1.4	98	4	100
Location of involvement:				
Distal phalanges	0.3	100	100	100
Middle phalanges	0.1	65	100	42
Proximal phalanges	0.4	62	84	13
Palm/Dorsum	1.1	23	48	14
Total number of symptomatic regions (out of 30):				
Median (range)	0 (0 – 4)	6 (3 – 17)	9 (6 – 17)	10 (6 – 19)

Cluster 2 (N=65) comprised a heterogeneous group of hands, as indicated in Figure A2 with a small amount of numbness and tingling present at a variety of sites. Hands in this cluster had a median of 6 regions of positive symptoms. Cluster 3 (N=25) was characterised by hands with numbness or tingling in the dorsal aspect only; they had a median of 9 regions involved, mostly in the finger, although 48% had numbness or tingling in the dorsum. Cluster 4 (N=57) identified hands with symptoms in at least three digits (ring, middle and index fingers), on both the palmar and dorsal aspects, predominantly in the distal regions.

Cluster 5 (N=25) was the smallest cluster along with Cluster 3, and contained hands with numbness or tingling predominantly in the whole of the little and ring fingers, with some involvement of the palm or dorsum (Table A2). Cluster 6 (N=43) comprised hands with symptoms in the palmar aspect only, seen predominantly in the ring, middle and index fingers, but also at the little finger and thumb. 40% of these hands had numbness or tingling in the palm.

Table A2: Patterns of numbness and tingling in the second four Hill Lane Katz diagram clusters

Characteristics	**Cluster**			
	5: Little/ ring finger	**6: Palmar aspect**	**7: All not thumb**	**8: All sites**
N	**25**	**43**	**50**	**72**
Colour in dendrogram	Green	Blue	Cyan	Pink
No numbness/ tingling	*0*	*0*	*0*	*0*
Digit involvement:				
Little finger	100	65	94	96
Ring finger	76	81	100	100
Middle finger	12	100	100	100
Index finger	8	95	100	100
Thumb	8	44	0	100
Aspect of involvement:				
Dorsal	96	2	100	100
Palmar	100	100	100	100
Location of involvement:				
Distal phalanges	100	95	100	100
Middle phalanges	100	98	100	100
Proximal phalanges	100	100	100	100
Palm/Dorsum	76	40	54	75
Total number of symptomatic regions (out of 30):				
Median (range)	14 (6 – 20)	12 (6 – 15)	24 (18 – 26)	30 (22 – 30)

Clusters 2, 3, 4, 5 and 6 formed the second of the three overriding clusters, and were therefore more similar to each other than Clusters 1, 7 or 8. Cluster 7 (N=50) identified hands with symptoms at all sites except the thumb, and Cluster 8 (N=72) was the second largest cluster, characterised by symptoms at all sites.

Replication of the Katz hand clusters (Bitterne data only)

Katz hand examination data from the Bitterne practice (2882 hands with complete information) were clustered using the two methods previously employed on the Hill Lane data. As was seen in the Hill Lane analysis, Ward's method on squared Euclidean distance produced a more clearly structured solution of six clusters (Figures A5 and A7, note the break in the y-axis) than the group average method on Jaccard's coefficient of community (Figure A6).

Figures A5 and A6: Dendrograms of a) Ward's and b) group average linkage hierarchical clustering on Bitterne data Katz hand diagram observations

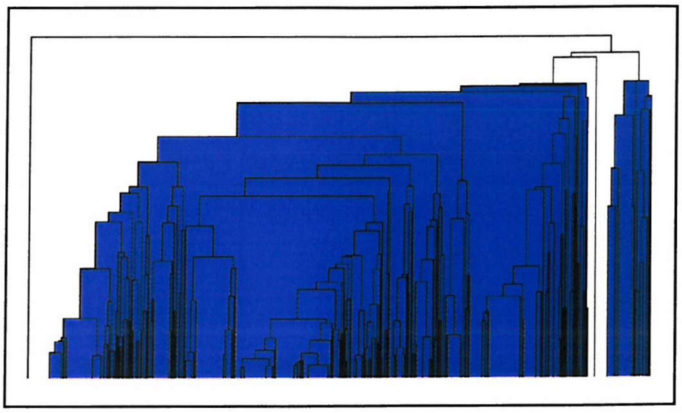
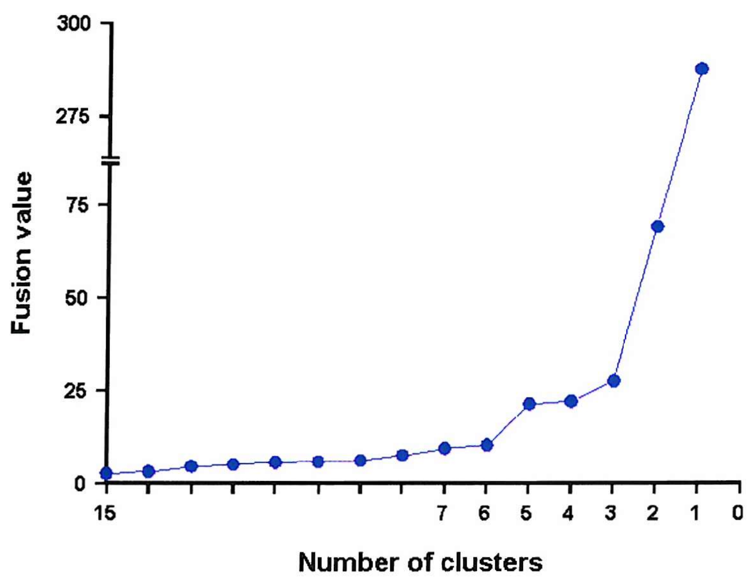


Figure A7: Fusion values for successive Katz diagram cluster joins



The group average method produced three main clusters (of sizes 2254, 580 and 47 hands) and one cluster containing one hand only. After k-means refinement, which made 450 (16%) inter-cluster moves, only the largest original cluster retained its original characteristics, suggesting that the solution was mathematically unstable. The Ward's method solution was refined using the k-means procedure which made 43 (1.5%) inter-cluster moves.

The same Katz diagram data were analysed using the k-means procedure, with the eight Hill Lane cluster centroids as the seed points. The two cluster solutions for the Bitterne data were compared using the kappa statistic and Rand index (Table A3).

The observed agreement between the clusters (numbers in bold) was 93%, whilst the expected was 64%, yielding a κ of 0.79. The agreement between each pair of hands as to whether they were placed in the same cluster or in different clusters from each other was 96.5% according to the Rand index, and 94.7% according to the Jaccard statistic. These values were adjusted to 84.2% and 43.8% respectively when the 2258 hands in the 'Normal' cluster according to both clustering algorithms had been removed. These indices suggested a reasonable level of agreement and therefore replicability of the clusters. In particular, the 'Normal', 'Little/ ring finger', 'palmar aspect', and 'all sites' clusters were identified in both the Hill Lane and Bitterne datasets (*coloured black, green, blue and purple in* Figures A2 and A5). The characteristics of the 'mixed', 'dorsal aspect', and 'distal' clusters from the Hill Lane analysis were lost in the Bitterne data, but two further clusters (coloured dark red and dark yellow in Figure A5) were identified that combined 197 (97%) of the hands from those three original clusters.

Table A3: Comparison of the Bitterne Katz diagram clusters obtained by two different methods

K-means clustering using Hill Lane seed points:	**Hierarchical (Ward's method) clustering plus k-means:**						
Kappa statistic							
Clusters:	1	2	4	5	3	6	N
Colour in dendrogram	Black	Dark Red	Dark Yellow	Green	Blue	Purple	
1: 'Normal'	**2258**	50	0	0	0	0	2308
2: Mixed	0	**121**	31	0	4	0	156
3: Dorsal aspect	0	0	**25**	1	0	0	26
4: Distal	0	6	14	1	0	0	21
5: Little/ ring finger	0	33	0	**47**	0	0	80
6: Palmar aspect	0	1	0	0	**99**	0	100
7: All not thumb	0	0	16	2	1	50	69
8: All sites	0	0	6	0	0	**116**	122
N	2258	211	92	51	104	166	2882
			κ=0.79				

Rand index/Jaccard statistic

Pairs of hands:	In the same cluster	In different clusters	N
In the same cluster	2,572,006	120,734	2,692,740
In different clusters	22,814	1,435,967	1,458,781
N	2,597,820	1,556,701	4,151,521
	Rand index=96.5% , Jaccard statistic=94.7%		

Rand index/Jaccard statistic*

Pairs of hands:	In the same cluster	In different clusters	N
In the same cluster	23,853	7,834	31,687
In different clusters	22,814	139,875	162,689
N	46,667	147,709	194,376
	Rand index=84.2% , Jaccard statistic=43.8%		

*Removing the 2258 hands in the 'Normal' clusters according to both clustering algorithms

Cluster analysis of the Katz hand examination data from the whole population

The Katz hand diagrams from the whole population (N=4284) were clustered using the same methods as for the Hill Lane Katz diagram data. Ten main clusters were yielded by Ward's method on squared Euclidean distance (Figures A8 and A10, note the breaks in the y-axis) compared with 3 main clusters and 4 minor ones by the group average algorithm on Jaccard's coefficient of community (Figure A9). As the solution giving the clearer hierarchical structure, and being more mathematically stable, the Ward's method solution was used for further exploration.

Figure A8: Dendrogram of Ward's hierarchical clustering on both Hill Lane and Bitterne data Katz hand diagram observations

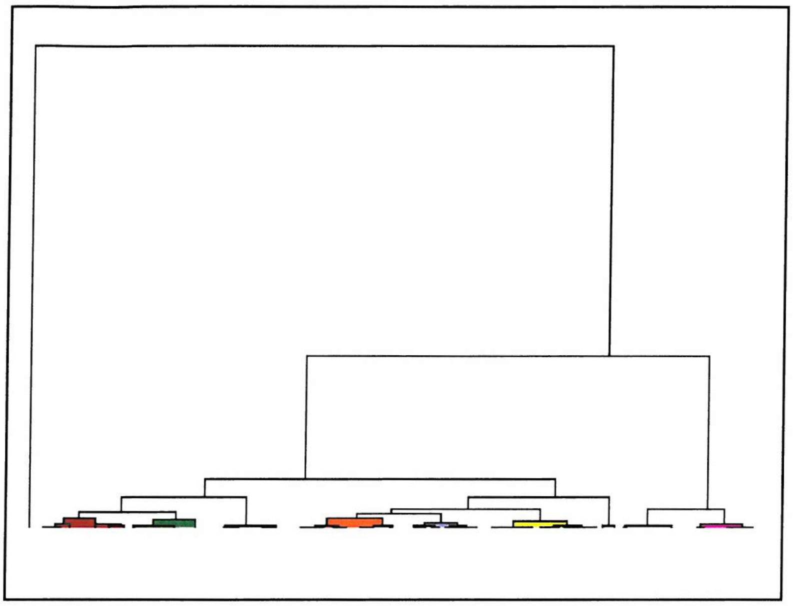


Figure A9: Dendrogram of group average linkage hierarchical clustering on both Hill Lane and Bitterne data Katz hand diagram observations

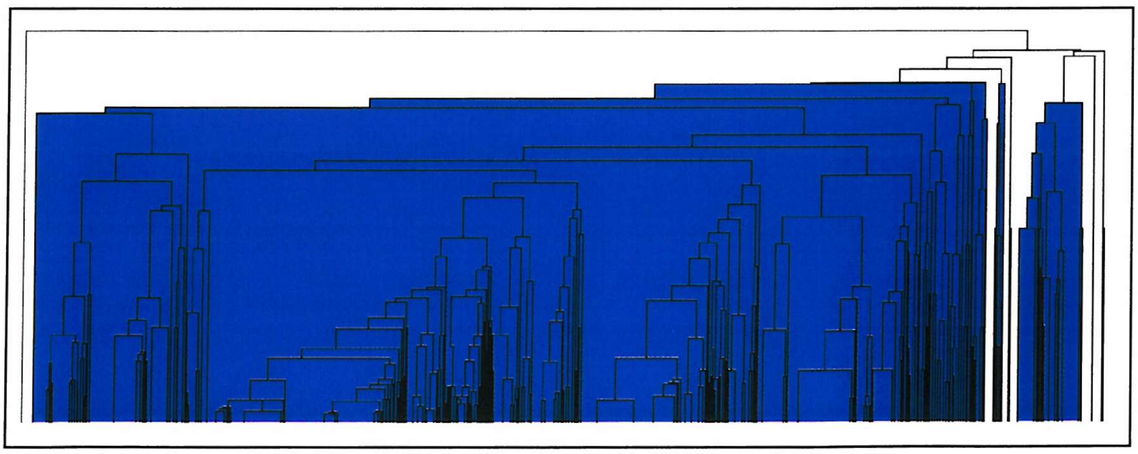
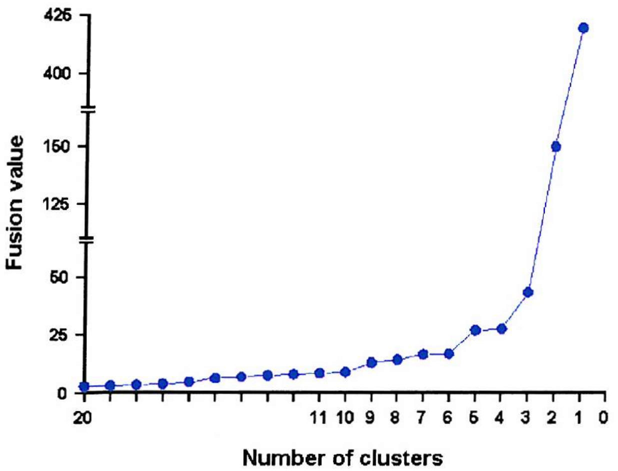


Figure A10: Fusion values for successive Katz diagram cluster joins



The k-means procedure made 119 (2.8%) inter-cluster moves, mostly involving Clusters 2, 4 and 5. (By contrast, the k-means procedure made 722 moves to refine the cluster solution from the group average algorithm.) As indicated by the colours of the clusters in Figures A2, A5 and A8, characteristics identified in Hill Lane only (orange, yellow, cyan and pink clusters), Bitterne only (dark red cluster) and both Hill Lane and Bitterne individually (black, blue and green clusters) were identified in this cluster solution. Two clusters (dark green and lilac in Figure A8) not previously identified were also seen in this analysis.

The characteristics of these clusters are presented in Tables A4 and A5. Cluster 1 (N=3307) was the largest cluster and most separate from the others. It comprised hands with minimal numbness or tingling. Cluster 2 (N=150) identified hands with numbness or tingling predominantly on the palmar aspect of the middle and index fingers, and mostly distally. Cluster 3 (N=139) formed a new cluster not seen in the Hill Lane or Bitterne analyses and was characterised by a few regions of numbness or tingling only, mostly in the palm or dorsum or the little finger. Hands with numbness or tingling on the majority of the palmar aspect were grouped together in Cluster 4 (N=122), whilst those with symptoms predominantly on the dorsal aspect (especially the ring and middle fingers) were identified in Cluster 5 (N=71).

Table A4: Patterns of numbness and tingling in the first five final Katz diagram clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Middle/ index distally**	**3: Palm / dorsum**	**4: Palmar aspect**	**5: Dorsal aspect**
N	**3307**	**150**	**139**	**122**	**71**
Colour in dendrogram	Black	Dark Red	Dark Green	Blue	Orange
No numbness/ tingling	*99.85*	*0*	*0*	*0*	*0*
Digit involvement:					
Little finger	0	50	50	79	54
Ring finger	0	69	28	98	100
Middle finger	0.06	91	4	99	75
Index finger	0.03	93	6	92	75
Thumb	0.09	32	17	52	31
Aspect of involvement:					
Dorsal	0.09	13	59	2	100
Palmar	0.06	100	85	100	48
Location of involvement:					
Distal phalanges	0.09	100	65	98	96
Middle phalanges	0.06	33	52	100	100
Proximal phalanges	0.09	23	68	100	94
Palm/Dorsum	0	7	76	56	55
Total number of symptomatic regions (out of 30):					
Median (range)	0 (0 – 4)	4 (1 – 12)	4 (1 – 9)	13 (7 – 15)	14 (6 – 21)

Cluster 6 (N=56) comprised hands with numbness or tingling primarily in the middle or index fingers or thumb on both the palmar and dorsal aspects. Numbness or tingling reported predominantly in the distal phalanges was identified in 79 hands (Cluster 7). Cluster 8 (N=64) was characterised by symptoms in the little and ring fingers with substantial palm or dorsum involvement. Clusters 9 (N=103) and 10 (N=193) were clearly separate from the other clusters (Figure A8), and were characterised by numbness or tingling throughout all fingers either without (Cluster 9) or with (Cluster 10) additional symptoms in the thumb.

Table A5: Patterns of numbness and tingling in the second five final Katz diagram clusters

Characteristics	**Cluster**				
	6: Middle/ index/ thumb	**7: Distal phalanges**	**8: Little/ ring fingers**	**9: All not thumb**	**10: All sites**
N	**56**	**79**	**64**	**103**	**193**
Colour in dendrogram	Lilac	Yellow	Green	Cyan	Pink
No numbness/ tingling	0	0	0	0	0
Digit involvement:					
Little finger	0	97	100	100	96
Ring finger	4	100	100	99	100
Middle finger	77	100	20	99	99
Index finger	87	99	3	100	100
Thumb	62	43	3	2	100
Aspect of involvement:					
Dorsal	100	100	100	100	100
Palmar	89	100	100	100	100
Location of involvement:					
Distal phalanges	100	100	100	98	100
Middle phalanges	100	51	100	100	100
Proximal phalanges	95	15	100	100	100
Palm/Dorsum	43	16	81	53	80
Total number of symptomatic regions (out of 30):					
Median (range)	12 (6 – 21)	11 (6 – 20)	14 (11 – 20)	24 (16 – 26)	30 (22 – 30)

The demographic characteristics of these clusters, their history of numbness, tingling and neck pain at baseline and their laterality were explored (Tables A6 and A7). A high proportion of women were seen in Clusters 5, 6 and 9 compared to the asymptomatic group (Cluster 1), whilst a higher proportion of men was seen in Cluster 3 compared to Cluster 1. Lower proportions of those in the youngest age band (25 – 34 years) were seen in Clusters 2, 4, 5, 7 and 10 compared to Cluster 1, whilst a higher proportion of 45 – 54 year olds were seen in Clusters 2, 4, 9 and 10, and a higher proportion of 55 – 64 year olds were seen in Clusters 4, 5, 6 and 8. An equally high proportion of subjects had reported numbness or tingling at baseline in the symptomatic clusters, (note that the baseline report did not distinguish between left and right sides), and neck pain was reported similarly across all of the clusters. Symptom profiles were frequently bilateral, with 83% of subjects having both their hands assigned to the same cluster, and a further 15% having unilateral symptoms only.

Table A6: Demographic and baseline characteristics of hands in the Katz clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Middle/ index distally**	**3: Palm / dorsum**	**4: Palmar aspect**	**5: Dorsal aspect**
N	**3307**	**150**	**139**	**122**	**71**
Colour in dendrogram	Black	Dark Red	Dark Green	Blue	Orange
Demographic characteristics:					
% female	60	55	50	65	72
% 25 – 34 yrs	16	5	19	6	7
% 35 – 44 yrs	24	20	19	20	34
% 45 – 54 yrs	30	45	30	37	21
% 55 – 64 yrs	30	30	32	38	38
Report of numbness or tingling in the fingers or hand or neck pain at baseline*:					
% numbness/tingling present	31	94	85	88	87
% neck pain present	43	54	52	44	52
Laterality – Number (%) of hands whose pair was:					
In the same cluster	2990 (90)	94 (63)	60 (43)	70 (57)	34 (48)
In the 'Normal' cluster	-	44 (29)	67 (48)	41 (34)	32 (45)
In a different cluster (not 'Normal')	317 (10)	12 (8)	12 (9)	11 (9)	5 (7)
Not in the analysis	-	-	-	-	-

* Numbness or tingling at baseline was not reported separately for the right and left side

Table A7: Demographic and baseline characteristics of hands in the Katz clusters

Characteristics	**Cluster**				
	6: Middle/ index/ thumb	**7: Distal phalanges**	**8: Little/ ring fingers**	**9: All not thumb**	**10: All sites**
N	**56**	**79**	**64**	**103**	**193**
Colour in dendrogram	Lilac	Yellow	Green	Cyan	Pink
Demographic characteristics:					
% female	68	62	45	71	66
% 25 – 34 yrs	13	6	13	12	8
% 35 – 44 yrs	16	27	28	13	20
% 45 – 54 yrs	36	38	23	41	41
% 55 – 64 yrs	36	29	36	35	31
Report of numbness or tingling in the fingers or hand or neck pain at baseline*:					
% numbness/tingling present	89	82	89	88	91
% neck pain present	45	46	53	49	47
Laterality – Number (%) of hands whose pair was:					
In the same cluster	24 (43)	62 (78)	32 (50)	72 (70)	138 (72)
In the 'Normal' cluster	24 (43)	13 (16)	29 (45)	22 (21)	45 (23)
In a different cluster (not 'Normal')	8 (14)	4 (5)	3 (5)	9 (9)	10 (5)
Not in the analysis	-	-	-	-	-

* Numbness or tingling at baseline was not reported separately for the right and left side

Patterns of numbness and tingling that would be expected with median nerve compression were most frequently seen in Clusters 2, 4 and 6, whilst those more closely associated with ulnar nerve compression were seen in Cluster 8. Numbness or tingling on the dorsal aspect of the hand only, such as that seen in Cluster 5 might be associated with radial nerve compression, although this condition is very rare. Conditions other than nerve entrapments may lead to symptoms of paraesthesia in the hand: Raynaud's phenomenon, cervical spondylosis and other neck disorders. Raynaud's phenomenon is associated with paraesthesia in the distal and middle phalanges, seen most commonly in Clusters 2 and 7.

It was decided to use the 10-cluster solution from the Katz diagram cluster analysis in the wrist/hand examination cluster analysis, rather than the 4-point Katz grading originally recorded. The 10-cluster solution clearly provided a more informative classification of the patterns of numbness and tingling presenting in the study, having more distinct categories, and being symptom-based rather than being pathology (median nerve compression)-based. The cluster solution was converted into seven binary variables, denoting numbness or tingling at the following sites: 1) little or ring fingers, 2) middle or index fingers, 3) thumb, 4) palm or dorsum, 5) palmar side, 6) dorsal side, 7) middle or proximal phalanges. Individuals in each of the 10 clusters were assigned values to these seven variables as shown in Table A8. Over 50% of the hands in each cluster had to have each characteristic to be assigned a value of one rather than zero.

Table A8: The seven binary variables created to denote the 10-point Katz cluster solution

Cluster	Variable						
	Little/ ring finger	Middle/ index finger	Thumb	Palm/ dorsum	Palmar aspect	Dorsal aspect	Middle/ proximal phalanges
1: "Normal"	0	0	0	0	0	0	0
2: Middle/ index distally	1	1	0	0	1	0	1
3: Palm/ dorsum	0	0	0	1	1	1	0
4: Palmar aspect	1	1	1	1	1	0	1
5: Dorsal aspect	1	1	0	1	0	1	1
6: Middle/ index/ thumb	0	1	1	0	1	1	1
7: Distal phalanges	1	1	0	0	1	1	0
8: Little/ ring fingers	1	0	0	1	1	1	1
9: All not thumb	1	1	0	1	1	1	1
10: All sites	1	1	1	1	1	1	1

Appendix V: Prevalence of symptoms and signs in the final wrist/hand clusters

Table A9: Prevalence of musculoskeletal symptoms and signs in the first five final wrist/hand clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Phalen**	**3: Thumb signs**	**4: Heberden**	**5: Radial wrist & thumb**
N	**2115**	**194**	**123**	**428**	**140**
Colour in dendrogram	Red	Orange	Light Blue	Yellow	Light Green
Any pain:	*0*	*0*	*0*	*1*	*100*
Dorsal forearm				0	4
Palmar forearm				0	4
Dorsal wrist				0	11
Palmar wrist				0	6
Radial wrist				0	**27**
Medial wrist				0	5
Thumb base				0	**97**
Finger joints				0	19
Other (mostly hand)				1	6
Any tenderness:	*3*	*15*	*100*	*11*	*79*
Dorsal forearm	0.3	3	7	2	4
Palmar forearm	0.1	1	2	0.2	2
Dorsal wrist	0.8	2	11	3	6
Palmar wrist	0.1	1	4	0.5	2
Radial wrist	0.8	3	7	1	**22**
Medial wrist	0.4	1	4	1	2
Thumb base	0	6	**100**	0	**74**
Finger joints	0.6	4	9	6	9
Other (mainly hand)	0.5	1	2	0.2	4
Any swelling:	*4*	*7*	*33*	*21*	*37*
Dorsal forearm	0.5	1	5	1	1
Palmar forearm	0.1	1	1	1	0
Dorsal wrist	0.6	1	7	2	4
Palmar wrist	0.2	1	6	1	1
Radial wrist	0.3	1	2	1	5
Medial wrist	0.6	0	0	1	2
Thumb base	0.1	0	11	1	**20**
Finger joints	2	4	15	15	12
Other (mainly hand)	0.6	2	1	1	0
Pain on resisted movement:	*3*	*6*	*25*	*7*	*50*
Radial wrist	0.7	1	6	1	**15**
Medial wrist	0.4	1	4	1	4
Finger extension	0.6	1	2	1	5
Finger flexion	0.4	2	2	1	2
Thumb extension	2	2	**24**	4	**40**
Other hand signs:					
Dupuytren's contracture	2	2	10	9	7
Heberden's nodes	0	16	**47**	**100**	**35**
Finkelstein's test positive	1	3	7	2	**14**
Thenar muscle wasting	0.1	1	4	1	6
Hypothenar muscle wasting	0.2	1	1	0.2	1
Weakness of thumb abduction	0.7	1	6	3	**11**
Weakness of thumb opposition	1.0	3	4	2	**11**

Table A10: Prevalence of sensorineural symptoms and signs in the first five final wrist/hand clusters

Characteristics	**Cluster**				
	1: "Normal"	**2: Phalen**	**3: Thumb signs**	**4: Heberden**	**5: Radial wrist & thumb**
N	**2115**	**194**	**123**	**428**	**140**
Colour in dendrogram	Red	Orange	Light Blue	Yellow	Light Green
Any numbness/ tingling (according to Katz hand cluster definition):					
Little/ ring finger	0	0	0	0	0
Middle/ index finger	0	0	0	0	0
Thumb	0	0	0	0	0
Palm/ Dorsum	0	0	0	0	0
Palmar Aspect	0	0	0	0	0
Dorsal Aspect	0	0	0	0	0
Proximal/ Middle phalanges	0	0	0	0	0
Sensorineural signs:					
Abnormal sensation:					
thumb	0.4	1	0	1	1
index finger	0.6	2	1	2	1
little finger	0.6	2	2	2	1
Phalen's test positive	0	**100**	3	0	21
Tinel's test positive	0.7	6	2	1	9
Sleep disturbed because of numbness/tingling	0	1	0	0	0

Table A11: Prevalence of musculoskeletal symptoms and signs in the second five final wrist/hand clusters

Characteristics	**Cluster**				
	6: Joint	**7: Wrist**	**8: NT – palmar fingers**	**9: NT – all not thumb**	**10: NT – all**
N	**102**	**126**	**173**	**179**	**268**
Colour in dendrogram	Dark green	Blue	Cyan	Pink	Dark red
Any pain:	*100*	*100*	*13*	*9*	*8*
Dorsal forearm	1	8	1	1	1
Palmar forearm	1	5	1	0	0.4
Dorsal wrist	8	**49**	3	3	2
Palmar wrist	3	**30**	1	2	3
Radial wrist	6	**28**	2	2	1
Medial wrist	8	**25**	2	2	0.4
Thumb base	10	2	3	1	0
Finger joints	**98**	1	4	2	0.4
Other (mostly hand)	3	17	1	1	1
Any tenderness:	*68*	*51*	*21*	*22*	*18*
Dorsal forearm	2	6	3	4	4
Palmar forearm	1	4	5	2	1
Dorsal wrist	3	**21**	4	6	3
Palmar wrist	2	7	0	2	1
Radial wrist	5	8	3	2	2
Medial wrist	2	11	1	3	1
Thumb base	10	3	9	6	7
Finger joints	**62**	2	5	6	3
Other (mainly hand)	3	2	0	2	0.4
Any swelling:	*53*	*21*	*12*	*7*	*6*
Dorsal forearm	1	2	1	1	1
Palmar forearm	0	2	0	0	0
Dorsal wrist	6	10	2	1	1
Palmar wrist	1	2	1	0	0.4
Radial wrist	1	7	1	1	1
Medial wrist	1	8	2	0	0.4
Thumb base	2	1	1	1	0.4
Finger joints	**52**	3	6	5	4
Other (mainly hand)	0	1	0	0	0
Pain on resisted movement:	*26*	*29*	*13*	*12*	*7*
Radial wrist	4	13	2	3	1
Medial wrist	3	12	5	4	2
Finger extension	10	7	1	1	2
Finger flexion	10	6	5	0	1
Thumb extension	11	7	6	7	4
Other hand signs:					
Dupuytren's contracture	7	2	3	2	4
Heberden's nodes	**58**	17	21	21	27
Finkelstein's test positive	6	6	2	4	3
Thenar muscle wasting	2	1	0	2	1
Hypothenar muscle wasting	0	1	1	0	2
Weakness of thumb abduction	2	2	1	1	3
Weakness of thumb opposition	10	5	1	4	4

Table A12: Prevalence of sensorineural symptoms and signs in the second five final wrist/hand clusters

Characteristics	**Cluster**				
	6: Joint	**7: Wrist**	**8: NT – palmar fingers**	**9: NT – all not thumb**	**10: NT – all**
N	**102**	**126**	**173**	**179**	**268**
Colour in dendrogram	Dark green	Blue	Cyan	Pink	Dark red
Any numbness/ tingling (according to Katz hand cluster definition):					
Little/ ring finger	0	0	**100**	**100**	**85**
Middle/ index finger	0	0	**100**	**73**	**100**
Thumb	0	0	**0**	**0**	**100**
Palm/ Dorsum	2	0	**0**	**100**	**85**
Palmar Aspect	2	0	**100**	**69**	**100**
Dorsal Aspect	2	0	**34**	**100**	**64**
Proximal/ Middle phalanges	0	0	**66**	**100**	**100**
Sensorineural signs:					
Abnormal sensation:					
thumb	0	1	1	2	1
index finger	1	2	3	2	3
little finger	1	2	2	7	1
Phalen's test positive	12	10	**32**	**38**	**41**
Tinel's test positive	3	3	8	6	10
Sleep disturbed because of numbness/tingling	0	0	**19**	**31**	**39**

Table A13: Prevalence of musculoskeletal symptoms and signs in the last four final wrist/hand clusters

Characteristics	**Cluster**			
	11: NT – all plus signs	**12: NT – hand only**	**13: All**	**14: Thumb and NT**
N	**48**	**118**	**39**	**96**
Colour in dendrogram	Dark Grey	Black	Brown	Light Grey
Any pain:	*25*	*20*	*92*	*99*
Dorsal forearm	0	0	13	5
Palmar forearm	0	1	18	3
Dorsal wrist	8	6	**77**	8
Palmar wrist	10	4	**54**	7
Radial wrist	4	4	**72**	**21**
Medial wrist	0	3	**69**	**15**
Thumb base	6	6	**62**	**66**
Finger joints	10	3	26	**39**
Other (mostly hand)	2	4	5	10
Any tenderness:	*29*	*21*	*100*	*87*
Dorsal forearm	0	3	13	5
Palmar forearm	4	1	13	2
Dorsal wrist	4	4	**46**	11
Palmar wrist	6	2	**26**	2
Radial wrist	2	2	**51**	10
Medial wrist	0	1	**62**	3
Thumb base	19	7	**77**	**60**
Finger joints	13	5	28	**34**
Other (mainly hand)	2	2	10	4
Any swelling:	*15*	*9*	*67*	*44*
Dorsal forearm	4	1	8	3
Palmar forearm	2	0	3	0
Dorsal wrist	2	3	**21**	5
Palmar wrist	4	0	**15**	0
Radial wrist	0	0	**18**	7
Medial wrist	0	1	**23**	5
Thumb base	4	1	3	5
Finger joints	13	4	**28**	**31**
Other (mainly hand)	2	1	0	0
Pain on resisted movement:	*19*	*14*	*95*	*60*
Radial wrist	2	3	**56**	17
Medial wrist	0	2	**72**	14
Finger extension	4	5	**69**	11
Finger flexion	6	5	**41**	9
Thumb extension	10	5	**77**	**47**
Other hand signs:				
Dupuytren's contracture	8	5	8	9
Heberden's nodes	29	17	41	42
Finkelstein's test positive	4	5	**64**	11
Thenar muscle wasting	2	1	21	7
Hypothenar muscle wasting	4	3	13	1
Weakness of thumb abduction	10	0	**36**	16
Weakness of thumb opposition	15	4	**54**	18

Table A14: Prevalence of sensorineural symptoms and signs in the last four final wrist/hand clusters

Characteristics	**Cluster**			
	11: NT – all plus signs	**12: NT – hand only**	**13: All**	**14: Thumb and NT**
N	**48**	**118**	**39**	**96**
Colour in dendrogram	Dark Grey	Black	Brown	Light Grey
Any numbness/ tingling (according to Katz hand cluster definition):				
Little/ ring finger	**96**	**0**	**74**	**91**
Middle/ index finger	**98**	**0**	**77**	**91**
Thumb	**50**	**0**	**26**	**48**
Palm/ Dorsum	**54**	**100**	**67**	**71**
Palmar Aspect	**98**	**100**	**82**	**94**
Dorsal Aspect	**58**	**100**	**74**	**75**
Proximal/ Middle phalanges	**81**	**0**	**82**	**89**
Sensorineural signs:				
Abnormal sensation:				
thumb	**92**	8	15	4
Index finger	**100**	8	23	4
little finger	**81**	15	15	1
Phalen's test positive	**56**	8	**51**	**46**
Tinel's test positive	**29**	3	**26**	11
Sleep disturbed because of numbness/tingling	**38**	21	**62**	**51**

Appendix VI: Construct validity of the shoulder clusters and medical diagnoses

Table A15: Association of healthcare use due to shoulder pain with shoulder disorders

Classification system	N	% seeing doctor	Odds ratio (95% Confidence Interval)
Shoulder clusters (data-driven)			
Normal all sites	288	3.5	1
Normal shoulders	1207	4.4	1.2 (0.6 – 2.5)
Signs	194	15.0	4.6 (2.2 – 9.9)
Pain	230	40.4	18.1 (9.0 – 36.1)
Pain & signs	178	48.3	24.5 (12.0 – 50.0)
Shoulder diagnoses (medically-driven)			
No symptoms or signs all sites	438	3.9	1
Normal shoulders	994	3.6	0.9 (0.5 – 1.6)
Non-specific shoulder signs only	249	14.5	4.0 (2.2 – 7.4)
Non-specific shoulder pain	67	28.4	9.6 (4.7 – 19.8)
Any shoulder diagnosis	349	46.7	20.8 (12.2 – 35.7)
		% treated*	
Shoulder clusters (data-driven)			
Normal all sites	288	2.1	1
Normal shoulders	1206	2.7	1.1 (0.4 – 2.6)
Signs	194	12.4	5.3 (2.1 – 13.4)
Pain	230	27.0	14.3 (6.0 – 34.1)
Pain & signs	177	45.8	30.5 (12.7 – 73.2)
Shoulder diagnoses (medically-driven)			
No symptoms or signs all sites	438	2.3	1
Normal shoulders	994	2.5	1.0 (0.5 – 2.1)
Non-specific shoulder signs only	249	10.8	4.3 (2.0 – 9.1)
Non-specific shoulder pain	66	15.2	7.1 (2.8 – 17.8)
Any shoulder diagnosis	348	38.2	22.4 (11.4 – 43.7)

**prescribed medication or injection.* Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A16: Association of psychosocial variables with shoulder disorders

Classification system	N	% poor work control	Odds ratio (95% Confidence Interval)
Shoulder clusters (data-driven)			
Control	2126	17.5	1
Normal all sites	246	15.5	0.9 (0.6 – 1.3)
Normal shoulders	914	21.7	1.2 (1.0 – 1.5)
Signs	138	18.8	1.0 (0.6 – 1.5)
Pain	173	25.4	1.5 (1.1 – 2.2)
Pain & signs	109	25.7	1.5 (0.9 – 2.3)
Shoulder diagnoses (medically-driven)			
Control	2126	17.5	1
No symptoms or signs all sites	356	16.0	0.9 (0.7 – 1.2)
Normal shoulders	764	21.6	1.2 (1.0 – 1.5)
Non-specific shoulder signs only	170	23.5	1.3 (0.9 – 1.9)
Non-specific shoulder pain	54	18.5	1.0 (0.5 – 2.0)
Any shoulder diagnosis	236	26.3	1.6 (1.1 – 2.1)
		% poor work support	
Shoulder clusters (data-driven)			
Control	2126	10.1	1
Normal all sites	246	10.6	1.2 (0.8 – 1.8)
Normal shoulders	914	13.6	1.4 (1.1 – 1.8)
Signs	138	15.2	1.7 (1.0 – 2.8)
Pain	173	15.6	1.6 (1.0 – 2.5)
Pain & signs	109	14.7	1.6 (0.9 – 2.9)
Shoulder diagnoses (medically-driven)			
Control	2126	10.1	1
No symptoms or signs all sites	356	11.5	1.2 (0.8 – 1.7)
Normal shoulders	764	13.7	1.4 (1.1 – 1.9)
Non-specific shoulder signs only	170	14.7	1.6 (1.0 – 2.5)
Non-specific shoulder pain	54	11.1	1.1 (0.4 – 2.6)
Any shoulder diagnosis	236	15.7	1.7 (1.1 – 2.5)
		% high work demand	
Shoulder clusters (data-driven)			
Control	2126	17.2	1
Normal all sites	246	12.6	0.7 (0.5 – 1.1)
Normal shoulders	914	19.7	1.3 (1.1 – 1.6)
Signs	138	23.2	1.9 (1.2 – 2.8)
Pain	173	22.0	1.5 (1.0 – 2.2)
Pain & signs	109	20.2	1.5 (0.9 – 2.4)
Shoulder diagnoses (medically-driven)			
Control	2126	17.2	1
No symptoms or signs all sites	356	12.6	0.7 (0.5 – 1.0)
Normal shoulders	764	22.1	1.5 (1.2 – 1.9)
Non-specific shoulder signs only	170	17.1	1.2 (0.8 – 1.9)
Non-specific shoulder pain	54	14.8	0.9 (0.4 – 2.0)
Any shoulder diagnosis	236	22.0	1.5 (1.1 – 2.1)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A17: Association of mental health with shoulder disorders

Classification system	N	% poor mental health	Odds ratio (95% Confidence Interval)
Shoulder clusters (data-driven)			
Control	2701	26.3	1
Normal all sites	291	34.0	1.3 (1.0 – 1.7)
Normal shoulders	1216	34.8	1.6 (1.3 – 1.8)
Signs	196	45.4	2.5 (1.9 – 3.4)
Pain	233	36.1	1.7 (1.3 – 2.2)
Pain & signs	178	48.3	3.1 (2.3 – 4.3)
Shoulder diagnoses (medically-driven)			
Control	2701	26.3	1
No symptoms or signs all sites	443	33.0	1.3 (1.1 – 1.7)
Normal shoulders	1000	34.8	1.5 (1.3 – 1.8)
Non-specific shoulder signs only	251	44.6	2.6 (1.9 – 3.4)
Non-specific shoulder pain	67	47.8	2.5 (1.6 – 4.2)
Any shoulder diagnosis	353	40.5	2.1 (1.7 – 2.7)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Appendix VII: Construct validity of the elbow clusters and medical diagnoses

Table A18: Association of healthcare use due to elbow pain with elbow disorders

Classification system	N	% seeing doctor	Odds ratio (95% Confidence Interval)
Elbow clusters (data-driven)			
Normal all sites	290	1.7	1
Normal elbows	1384	2.1	1.0 (0.4 – 2.6)
Pain	55	29.1	20.1 (6.9 – 58.5)
Medial signs	64	4.7	2.2 (0.5 – 9.7)
Posterior/ Antecubital Fossa	53	26.4	17.3 (5.8 – 51.5)
Medial symptoms and signs	44	29.6	18.7 (6.1 – 56.9)
Lateral signs	133	16.5	9.2 (3.4 – 25.2)
Lateral symptoms and signs	82	34.2	22.4 (8.1 – 61.5)
Bilateral – mixed profiles	26	57.7	62.6 (18.8 – 208.2)
Elbow diagnoses (medically-driven)			
No symptoms or signs all sites	443	2.0	1
Normal elbows	1227	1.9	0.8 (0.4 – 1.8)
Non-specific elbow signs only	228	14.5	7.2 (3.3 – 15.4)
Non-specific elbow pain	158	31.7	19.7 (9.3 – 41.5)
Any elbow diagnosis	75	40.0	26.4 (11.7 – 59.9)
		% treated*	
Elbow clusters (data-driven)			
Normal all sites	290	1.0	1
Normal elbows	1384	1.5	1.0 (0.3 – 3.4)
Pain	55	25.5	26.3 (7.2 – 96.6)
Medial signs	64	4.7	3.0 (0.6 – 15.6)
Posterior/ Antecubital Fossa	53	22.6	20.7 (5.5 – 77.7)
Medial symptoms and signs	44	25.0	22.3 (5.8 – 85.3)
Lateral signs	133	13.5	10.6 (3.0 – 37.3)
Lateral symptoms and signs	82	30.5	30.0 (8.6 – 104.3)
Bilateral – mixed profiles	26	46.2	51.8 (12.8 – 209.6)
Elbow diagnoses (medically-driven)			
No symptoms or signs all sites	443	0.9	1
Normal elbows	1227	1.2	1.1 (0.4 – 3.4)
Non-specific elbow signs only	228	12.3	12.1 (4.2 – 35.3)
Non-specific elbow pain	158	25.3	31.4 (10.9 – 90.0)
Any elbow diagnosis	75	41.3	60.6 (20.2 – 181.6)

*prescribed medication or injection. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A19: Association of psychosocial variables with elbow disorders

Classification system	N	% poor work control	Odds ratio (95% Confidence Interval)
Elbow clusters (data-driven)			
Control	2126	17.5	1
Normal all sites	246	15.5	0.9 (0.6 – 1.3)
Normal elbows	1018	21.4	1.2 (1.0 – 1.5)
Pain	38	21.1	1.3 (0.6 – 2.8)
Medial signs	45	28.9	1.7 (0.9 – 3.4)
Posterior/ Antecubital Fossa	38	15.8	0.8 (0.3 – 2.0)
Medial symptoms and signs	30	13.3	0.7 (0.2 – 2.0)
Lateral signs	93	26.9	1.6 (1.0 – 2.6)
Lateral symptoms and signs	62	32.3	2.2 (1.3 – 3.9)
Bilateral – mixed profiles	20	30.0	1.7 (0.7 – 4.6)
Elbow diagnoses (medically-driven)			
Control	2126	17.5	1
No symptoms or signs all sites	356	16.0	0.9 (0.7 – 1.2)
Normal elbows	912	22.0	1.2 (1.0 – 1.5)
Non-specific elbow signs only	158	26.0	1.6 (1.1 – 2.3)
Non-specific elbow pain	111	21.6	1.3 (0.8 – 2.0)
Any elbow diagnosis	53	28.3	1.8 (1.0 – 3.3)
		% poor work support	
Elbow clusters (data-driven)			
Control	2126	10.1	1
Normal all sites	246	10.6	1.2 (0.8 – 1.8)
Normal elbows	1018	12.4	1.3 (1.0 – 1.6)
Pain	38	21.1	2.5 (1.1 – 5.8)
Medial signs	45	17.8	1.9 (0.8 – 4.2)
Posterior/ Antecubital Fossa	38	18.4	2.0 (0.8 – 4.7)
Medial symptoms and signs	30	6.7	0.5 (0.1 – 2.2)
Lateral signs	93	16.1	1.6 (0.9 – 2.8)
Lateral symptoms and signs	62	25.8	2.8 (1.5 – 5.1)
Bilateral – mixed profiles	20	40.0	6.5 (2.5 – 16.9)
Elbow diagnoses (medically-driven)			
Control	2126	10.1	1
No symptoms or signs all sites	356	11.5	1.2 (0.8 – 1.7)
Normal elbows	912	12.4	1.3 (1.0 – 1.7)
Non-specific elbow signs only	158	15.8	1.5 (1.0 – 2.4)
Non-specific elbow pain	111	22.5	2.5 (1.5 – 4.0)
Any elbow diagnosis	53	22.6	2.6 (0.8 – 1.7)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A19 (continued): Association of psychosocial variables with elbow disorders

Classification system	N	% high work demand	Odds ratio (95% Confidence Interval)
Elbow clusters (data-driven)			
Control	2126	17.2	1
Normal all sites	246	12.6	0.7 (0.5 – 1.0)
Normal elbows	1018	19.4	1.3 (1.1 – 1.6)
Pain	38	23.7	1.5 (0.7 – 3.2)
Medial signs	45	20.0	1.4 (0.7 – 3.0)
Posterior/ Antecubital Fossa	38	26.3	1.9 (0.9 – 3.9)
Medial symptoms and signs	30	23.3	1.6 (0.7 – 3.9)
Lateral signs	93	21.5	1.5 (0.9 – 2.4)
Lateral symptoms and signs	62	24.2	1.6 (0.9 – 2.9)
Bilateral – mixed profiles	20	35.0	3.4 (1.3 – 8.9)
Elbow diagnoses (medically-driven)			
Control	2126	17.2	1
No symptoms or signs all sites	356	12.6	0.7 (0.5 – 1.0)
Normal elbows	912	20.0	1.4 (1.1 – 1.7)
Non-specific elbow signs only	158	22.8	1.6 (1.0 – 2.3)
Non-specific elbow pain	111	25.2	1.7 (1.1 – 2.7)
Any elbow diagnosis	53	26.4	2.0 (1.0 – 3.7)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A20: Association of psychological variables with elbow disorders

Classification system	N	% poor mental health	Odds ratio (95% Confidence Interval)
Elbow clusters (data-driven)			
Control	2701	26.3	1
Normal all sites	291	34.0	1.3 (1.0 – 1.7)
Normal elbows	1390	35.9	1.6 (1.4 – 1.9)
Pain	55	45.5	2.4 (1.4 – 4.2)
Medial signs	64	50.0	3.2 (1.9 – 5.4)
Posterior/ Antecubital Fossa	53	39.6	2.1 (1.2 – 3.7)
Medial symptoms and signs	45	37.8	1.9 (1.0 – 3.5)
Lateral signs	134	44.8	2.4 (1.7 – 3.5)
Lateral symptoms and signs	82	41.5	2.1 (1.3 – 3.4)
Bilateral – mixed profiles	26	26.9	1.1 (0.5 – 2.8)
Elbow diagnoses (medically-driven)			
Control	2701	26.3	1
No symptoms or signs all sites	443	33.0	1.3 (1.1 – 1.7)
Normal elbows	1234	36.6	1.7 (1.5 – 2.0)
Non-specific elbow signs only	229	43.7	2.4 (1.8 – 3.2)
Non-specific elbow pain	158	43.7	2.4 (1.7 – 3.3)
Any elbow diagnosis	76	35.5	1.7 (1.0 – 2.7)

Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Appendix VIII: Construct validity of the wrist/hand clusters and medical diagnoses

Table A21: Association of seeing a doctor due to wrist/hand symptoms with wrist/hand disorders

Classification system	N	% seeing doctor	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Normal all sites	291	6.5	1
Normal wrists/hands	551	8.8	1.4 (0.8 – 2.4)
Phalen's test positive	96	7.2	1.1 (0.4 – 2.7)
Thumb signs	55	27.3	5.1 (2.4 – 11.0)
Heberden's nodes	196	10.6	1.6 (0.8 – 3.2)
Radial wrist & thumb involvement	71	35.2	7.5 (3.8 – 15.0)
Finger joint symptoms and signs	43	30.2	6.0 (2.7 – 13.4)
Wrist symptoms and signs	70	30.0	6.1 (3.0 – 12.1)
N/T* in the palmar fingers	80	21.0	3.9 (1.9 – 8.0)
N/T* all except thumb	97	23.7	4.5 (2.3 – 8.7)
N/T* all	132	30.1	6.0 (3.3 – 11.0)
N/T* all plus signs	20	45.0	12.5 (4.6 – 34.2)
N/T* palm or dorsum	71	25.4	5.0 (2.5 – 10.2)
All	25	52.0	13.9 (5.5 – 35.1)
N/T* and radial wrist and thumb	74	40.5	9.4 (4.8 – 18.4)
Bilateral – N/T*	27	34.5	7.3 (2.9 – 17.9)
Bilateral – musculoskeletal	16	43.8	10.4 (3.5 – 31.1)
Bilateral – signs	42	23.8	4.2 (1.8 – 10.1)
Bilateral – mixed	135	33.3	6.9 (3.8 – 12.6)
Wrist/hand diagnoses (medically-driven)			
No symptoms or signs all sites	442	5.4	1
Normal wrists/hands	311	8.3	1.6 (0.9 – 2.8)
Non-specific wrist/hand signs	456	13.7	2.6 (1.6 – 4.3)
Non-specific wrist/hand pain	391	24.1	5.4 (3.4 – 8.7)
Non-specific wrist/hand N/T*	143	31.5	7.9 (4.6 – 13.6)
Non-spec. wrist/hand pain & N/T*	93	40.9	11.7 (6.5 – 21.1)
Tenosynovitis	37	48.7	15.9 (7.4 – 34.3)
OA	147	34.7	8.7 (5.0 – 15.0)
Carpal tunnel syndrome	28	50.0	18.3 (7.8 – 43.0)
Tenosynovitis & OA	40	40.0	10.2 (4.8 – 22.1)
Carpal tunnel syndrome & OA	3	66.7	33.3 (2.9 – 386.8)
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A22: Association of receiving treatment due to wrist/hand symptoms with wrist/hand disorders

Classification system	N	% treated**	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Normal all sites	291	2.8	1
Normal wrists/hands	551	4.2	1.5 (0.6 – 3.3)
Phalen's test positive	96	3.1	1.1 (0.3 – 4.2)
Thumb signs	55	12.7	4.5 (1.5 – 13.2)
Heberden's nodes	196	5.5	1.7 (0.7 – 4.5)
Radial wrist & thumb involvement	71	23.9	9.2 (3.7 – 23.1)
Finger joint symptoms and signs	43	16.3	5.9 (2.0 – 17.5)
Wrist symptoms and signs	70	20.0	8.4 (3.3 – 21.0)
N/T* in the palmar fingers	80	7.4	2.8 (0.9 – 8.4)
N/T* all except thumb	97	7.2	2.7 (0.9 – 7.7)
N/T* all	132	8.3	2.9 (1.1 – 7.4)
N/T* all plus signs	20	35.0	19.4 (6.0 – 62.6)
N/T* palm or dorsum	71	5.6	2.1 (0.6 – 7.1)
All	25	52.0	31.0 (10.5 – 91.1)
N/T* and radial wrist and thumb	74	36.5	17.5 (7.4 – 41.7)
Bilateral – N/T*	27	10.3	3.7 (0.9 – 14.8)
Bilateral – musculoskeletal	16	37.5	18.2 (5.2 – 63.5)
Bilateral – signs	42	19.1	6.7 (2.3 – 19.5)
Bilateral – mixed	135	17.0	6.1 (2.6 – 14.4)
Wrist/hand diagnoses (medically-driven)			
No symptoms or signs all sites	442	2.3	1
Normal wrists/hands	311	3.5	1.6 (0.7 – 3.7)
Non-specific wrist/hand signs	456	7.0	2.8 (1.4 – 5.9)
Non-specific wrist/hand pain	391	6.3	2.7 (1.3 – 5.8)
Non-specific wrist/hand N/T*	143	18.9	9.6 (4.5 – 20.6)
Non-spec. wrist/hand pain & N/T*	93	21.5	10.8 (4.8 – 24.1)
Tenosynovitis	37	46.0	33.4 (13.5 – 82.8)
OA	147	26.5	13.0 (6.2 – 27.4)
Carpal tunnel syndrome	28	21.4	11.7 (3.8 – 35.6)
Tenosynovitis & OA	40	40.0	22.3 (9.0 – 55.3)
Carpal tunnel syndrome & OA	3	66.7	71.6 (5.8 – 888.9)
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. **medication, injection or operation. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A23: Association of occupational mechanical wrist/hand activity with wrist/hand disorders

Classification system	N	% reporting keyboard use	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2126	21.8	1
Normal all sites	246	25.2	1.1 (0.8 – 1.5)
Normal wrists/hands	443	26.2	1.3 (1.0 – 1.7)
Phalen's test positive	84	25.0	1.1 (0.7 – 1.9)
Thumb signs	36	19.4	1.0 (0.4 – 2.4)
Heberden's nodes	123	14.6	0.8 (0.5 – 1.4)
Radial wrist & thumb involvement	44	18.2	1.0 (0.5 – 2.2)
Finger joint symptoms and signs	27	29.6	1.6 (0.7 – 3.8)
Wrist symptoms and signs	54	20.4	0.9 (0.5 – 1.8)
N/T* in the palmar fingers	62	11.3	0.5 (0.2 – 1.2)
N/T* all except thumb	77	19.5	0.9 (0.5 – 1.6)
N/T* all	101	25.7	1.4 (0.9 – 2.2)
N/T* all plus signs	13	7.7	0.3 (0.04 – 2.6)
N/T* palm or dorsum	51	29.4	1.6 (0.8 – 2.9)
All	8	12.5	0.6 (0.1 – 4.8)
N/T* and radial wrist and thumb	46	28.3	1.7 (0.9 – 3.4)
Bilateral – N/T*	22	9.1	0.4 (0.1 – 1.6)
Bilateral – musculoskeletal	11	27.3	1.3 (0.3 – 4.8)
Bilateral – signs	23	26.1	1.7 (0.7 – 4.5)
Bilateral – mixed	101	15.8	0.8 (0.5 – 1.4)
Wrist/hand diagnoses (medically-driven)			
Control	2126	21.8	1
No symptoms or signs all sites	356	25.6	1.2 (0.9 – 1.5)
Normal wrists/hands	263	27.0	1.4 (1.0 – 1.9)
Non-specific wrist/hand signs	321	20.6	1.0 (0.8 – 1.4)
Non-specific wrist/hand pain	305	19.7	1.0 (0.7 – 1.3)
Non-specific wrist/hand N/T*	106	21.7	1.0 (0.6 – 1.6)
Non-spec. wrist/hand pain & N/T*	63	19.1	1.0 (0.5 – 1.8)
Tenosynovitis	24	20.8	1.0 (0.4 – 2.7)
OA	94	21.3	1.2 (0.7 – 2.1)
Carpal tunnel syndrome	19	10.5	0.6 (0.1 – 2.6)
Tenosynovitis & OA	19	31.6	2.1 (0.8 – 5.6)
Carpal tunnel syndrome & OA	1	0.0	-
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A24: Association of poor work control with wrist/hand disorders

Classification system	N	% poor work control	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2126	17.4	1
Normal all sites	246	15.5	0.9 (0.6 – 1.3)
Normal wrists/hands	443	20.8	1.2 (1.0 – 1.6)
Phalen's test positive	84	21.4	1.2 (0.7 – 2.1)
Thumb signs	36	19.4	1.0 (0.4 – 2.3)
Heberden's nodes	123	20.3	1.1 (0.7 – 1.7)
Radial wrist & thumb involvement	44	15.9	0.8 (0.3 – 1.7)
Finger joint symptoms and signs	27	33.3	2.1 (0.9 – 4.8)
Wrist symptoms and signs	54	13.0	0.7 (0.3 – 1.5)
N/T* in the palmar fingers	62	32.3	2.2 (1.3 – 3.9)
N/T* all except thumb	77	23.4	1.4 (0.8 – 2.3)
N/T* all	101	30.7	1.9 (1.2 – 3.0)
N/T* all plus signs	13	7.7	0.4 (0.1 – 3.1)
N/T* palm or dorsum	51	29.4	2.0 (1.1 – 3.6)
All	8	25.0	1.2 (0.2 – 6.3)
N/T* and radial wrist and thumb	46	19.6	1.0 (0.5 – 2.1)
Bilateral – N/T*	22	18.2	1.0 (0.3 – 2.9)
Bilateral – musculoskeletal	11	9.1	0.4 (0.1 – 3.1)
Bilateral – signs	23	8.7	0.4 (0.1 – 1.6)
Bilateral – mixed	101	26.7	1.5 (1.0 – 2.4)
Wrist/hand diagnoses (medically-driven)			
Control	2126	17.5	1
No symptoms or signs all sites	356	16.0	0.9 (0.7 – 1.2)
Normal wrists/hands	263	20.5	1.2 (0.9 – 1.7)
Non-specific wrist/hand signs	321	21.5	1.2 (0.9 – 1.6)
Non-specific wrist/hand pain	305	25.6	1.5 (1.2 – 2.0)
Non-specific wrist/hand N/T*	106	17.0	0.9 (0.6 – 1.6)
Non-spec. wrist/hand pain & N/T*	63	28.6	1.8 (1.0 – 3.1)
Tenosynovitis	24	16.7	0.8 (0.3 – 2.4)
OA	94	24.5	1.3 (0.8 – 2.1)
Carpal tunnel syndrome	19	42.1	3.3 (1.3 – 8.3)
Tenosynovitis & OA	19	21.1	1.1 (0.4 – 3.3)
Carpal tunnel syndrome & OA	1	0.0	-
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A25: Association of poor work support with wrist/hand disorders

Classification system	N	% poor work support	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2126	10.1	1
Normal all sites	246	10.6	1.2 (0.8 – 1.8)
Normal wrists/hands	443	15.6	1.6 (1.2 – 2.2)
Phalen's test positive	84	10.7	1.1 (0.6 – 2.3)
Thumb signs	36	5.6	0.6 (0.1 – 2.5)
Heberden's nodes	123	11.4	1.1 (0.6 – 2.0)
Radial wrist & thumb involvement	44	15.9	2.1 (0.9 – 5.0)
Finger joint symptoms and signs	27	14.8	1.9 (0.6 – 5.8)
Wrist symptoms and signs	54	9.3	0.9 (0.3 – 2.2)
N/T* in the palmar fingers	62	17.7	1.7 (0.9 – 3.4)
N/T* all except thumb	77	7.8	0.8 (0.3 – 1.8)
N/T* all	101	13.9	1.5 (0.8 – 2.8)
N/T* all plus signs	13	15.4	1.2 (0.3 – 5.6)
N/T* palm or dorsum	51	15.7	1.7 (0.8 – 3.8)
All	8	25.0	3.1 (0.6 – 15.4)
N/T* and radial wrist and thumb	46	19.6	2.4 (1.1 – 5.2)
Bilateral – N/T*	22	27.3	3.8 (1.4 – 10.2)
Bilateral – musculoskeletal	11	18.2	2.3 (0.5 – 10.9)
Bilateral – signs	23	21.7	2.9 (1.0 – 8.4)
Bilateral – mixed	101	11.9	1.1 (0.6 – 2.1)
Wrist/hand diagnoses (medically-driven)			
Control	2126	10.1	1
No symptoms or signs all sites	356	11.5	1.2 (0.8 – 1.7)
Normal wrists/hands	263	16.7	1.7 (1.2 – 2.5)
Non-specific wrist/hand signs	321	12.2	1.3 (0.9 – 1.9)
Non-specific wrist/hand pain	305	14.1	1.5 (1.0 – 2.1)
Non-specific wrist/hand N/T*	106	11.3	1.2 (0.6 – 2.2)
Non-spec. wrist/hand pain & N/T*	63	19.1	2.1 (1.1 – 4.2)
Tenosynovitis	24	8.3	0.8 (0.2 – 3.3)
OA	94	16.0	1.8 (1.0 – 3.3)
Carpal tunnel syndrome	19	10.5	0.8 (0.2 – 3.4)
Tenosynovitis & OA	19	15.8	2.1 (0.6 – 7.7)
Carpal tunnel syndrome & OA	1	0.0	-
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A26: Association of high work demand with wrist/hand disorders

Classification system	N	% high work demand	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2126	17.2	1
Normal all sites	246	12.6	0.7 (0.5 – 1.0)
Normal wrists/hands	443	19.2	1.2 (0.9 – 1.6)
Phalen's test positive	84	20.2	1.4 (0.8 – 2.4)
Thumb signs	36	5.6	0.3 (0.1– 1.4)
Heberden's nodes	123	15.5	1.1 (0.7 – 1.8)
Radial wrist & thumb involvement	44	25.0	2.1 (1.0 – 4.3)
Finger joint symptoms and signs	27	25.9	2.0 (0.8 – 4.9)
Wrist symptoms and signs	54	22.2	1.5 (0.8 – 2.9)
N/T* in the palmar fingers	62	22.6	1.5 (0.8 – 2.8)
N/T* all except thumb	77	23.4	1.6 (0.9 – 2.8)
N/T* all	101	19.8	1.4 (0.8 – 2.4)
N/T* all plus signs	13	30.8	2.1 (0.6 – 7.0)
N/T* palm or dorsum	51	21.6	1.4 (0.7 – 2.7)
All	8	37.5	4.6 (1.1 – 19.7)
N/T* and radial wrist and thumb	46	28.3	2.4 (1.2 – 4.6)
Bilateral – N/T*	22	27.3	2.4 (0.9 – 6.2)
Bilateral – musculoskeletal	11	0.0	-
Bilateral – signs	23	13.0	0.9 (0.3 – 3.1)
Bilateral – mixed	101	24.8	2.0 (1.3 – 3.3)
Wrist/hand diagnoses (medically-driven)			
Control	2126	17.2	1
No symptoms or signs all sites	356	12.6	0.7 (0.5 – 1.0)
Normal wrists/hands	263	21.3	1.4 (1.0 – 1.9)
Non-specific wrist/hand signs	321	16.5	1.1 (0.8 – 1.6)
Non-specific wrist/hand pain	305	20.0	1.4 (1.0 – 1.9)
Non-specific wrist/hand N/T*	106	23.6	1.6 (1.0 – 2.6)
Non-spec. wrist/hand pain & N/T*	63	22.2	1.6 (0.9 – 3.0)
Tenosynovitis	24	16.7	1.2 (0.4 – 3.7)
OA	94	29.8	2.8 (1.7 – 4.5)
Carpal tunnel syndrome	19	52.6	5.7 (2.2 – 14.3)
Tenosynovitis & OA	19	21.1	1.9 (0.6 – 5.8)
Carpal tunnel syndrome & OA	1	100.0	-
All three diagnoses	1	0.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

Table A27: Association of mental health with wrist/hand disorders

Classification system	N	% poor mental health	Odds ratio (95% Confidence Interval)
Wrist/hand clusters (data-driven)			
Control	2701	26.3	1
Normal all sites	291	34.0	1.3 (1.0 – 1.7)
Normal wrists/hands	555	36.4	1.6 (1.3 – 2.0)
Phalen's test positive	97	41.2	1.9 (1.2 – 2.8)
Thumb signs	55	49.1	3.0 (1.7 – 5.1)
Heberden's nodes	199	28.6	1.3 (1.0 – 1.9)
Radial wrist & thumb involvement	71	32.4	1.6 (1.0 – 2.8)
Finger joint symptoms and signs	43	27.9	1.2 (0.6 – 2.3)
Wrist symptoms and signs	70	41.4	1.9 (1.2 – 3.2)
N/T* in the palmar fingers	81	32.1	1.4 (0.9 – 2.3)
N/T* all except thumb	97	43.3	2.2 (1.4 – 3.3)
N/T* all	133	45.1	2.4 (1.7 – 3.4)
N/T* all plus signs	20	50.0	3.4 (1.4 – 8.6)
N/T* palm or dorsum	71	35.2	1.6 (1.0 – 2.6)
All	25	64.0	5.4 (2.3 – 12.7)
N/T* and radial wrist and thumb	74	46.0	2.7 (1.7 – 4.3)
Bilateral – N/T*	29	34.5	1.4 (0.6 – 3.1)
Bilateral – musculoskeletal	16	37.5	1.7 (0.6 – 4.8)
Bilateral – signs	42	40.5	2.4 (1.2 – 4.5)
Bilateral – mixed	135	34.1	1.6 (1.1 – 2.4)
Wrist/hand diagnoses (medically-driven)			
Control	2701	26.3	1
No symptoms or signs all sites	443	33.0	1.3 (1.1 – 1.7)
Normal wrists/hands	314	35.7	1.5 (1.2 – 2.0)
Non-specific wrist/hand signs	460	38.0	1.9 (1.5 – 2.3)
Non-specific wrist/hand pain	395	38.5	1.8 (1.5 – 2.3)
Non-specific wrist/hand N/T*	143	39.2	1.8 (1.3 – 2.6)
Non-spec. wrist/hand pain & N/T*	93	46.2	2.7 (1.8 – 4.2)
Tenosynovitis	37	35.1	1.6 (0.8 – 3.1)
OA	147	40.1	2.2 (1.5 – 3.1)
Carpal tunnel syndrome	28	25.0	1.1 (0.5 – 2.7)
Tenosynovitis & OA	40	42.5	2.4 (1.3 – 4.7)
Carpal tunnel syndrome & OA	3	0.0	-
All three diagnoses	1	100.0	-

*N/T=numbness and/or tingling. Analyses were adjusted for sex and age in four strata, and associations for the two classification systems were performed separately.

References

Reference List

1. Croft P. Soft tissue rheumatism. In: Silman A, Hochberg MC, eds. Epidemiology of the Rheumatic Diseases. London: Oxford University Press,1993: 375-421.
2. Lawrence JS. Disk degeneration: its frequency and relationship to symptoms. *Annals of the Rheumatic Diseases* 1969; **28**: 121-137.
3. Cunningham LS, Kelsey JL. Epidemiology of musculoskeletal impairments and their associated disability. *American Journal of Public Health* 1984; **74**: 574-579.
4. Brattberg G, Thorslund M, Wikman A. The prevalence of pain in a general population. The results of a postal survey in a county of Sweden. *Pain* 1989; **37**: 215-222.
5. Makela M, Heliovaara M, Sievers K, Knekt P, Maatela J, Aromaa A. Musculoskeletal disorders as determinants of disability in Finns ages 30 years or more. *Journal of Clinical Epidemiology* 1993; **46**: 549-559.
6. Cote P, Cassidy JD, Carroll L. The Saskatchewan health and back pain survey - the prevalence of neck pain and related disability in Saskatchewan adults. *Spine* 1998; **23**: 1689-1698.
7. Allander E. Prevalence, incidence, and remission rates of some common rheumatic diseases or syndromes. *Scandinavian Journal of Rheumatology* 1974; **3**: 145-153.
8. Chard MD, Hazleman BL. Shoulder disorders in the elderly: a hospital study. *Annals of the Rheumatic Diseases* 1987; **46**: 684-687.
9. Pope DP, Croft PR, Pritchard CM, Silman AJ. Prevalence of shoulder pain in the community: The influence of case definition. *Annals of the Rheumatic Diseases* 1997; **56**: 308-312.
10. Macfarlane GJ, Hunt IM, Silman AJ. Role of mechanical and psychosocial factors in the onset of forearm pain: prospective population based study. *British Medical Journal* 2000; **321**: 676-679.
11. McCormick A, Fleming D, Charlton J. Morbidity statistics from general practice. 4th national study 1991-2.Series MB5 No.3. London: HMSO. 1991
12. Chard MD, Hazleman R, Hazleman BL, King RH, Reiss BB. Shoulder disorders in the elderly: a community survey. *Arthritis & Rheumatism* 1991; **34**: 766-769.

13. Chakravarty KK, Webley M. Disorders of the shoulder: an often unrecognised cause of disability in elderly people. *British Medical Journal* 1990; **300**: 848-849.
14. Bogduk N. The neck. *Bailliere's best practice and research.Clinical rheumatology* 1999; **13**: 261-285.
15. Van der Heijden GJMG. Shoulder disorders: a state-of-the-art review. *Bailliere's best practice and research.Clinical rheumatology* 1999; **13**: 287-309.
16. Helliwell PS. The elbow, forearm, wrist and hand. *Bailliere's best practice and research.Clinical rheumatology* 1999; **13**: 311-328.
17. Barnsley L. The spine: neck pain. In: Klippel JH, Dieppe PA, eds. Rheumatology. London: CV Mosby,1998:
18. Barnsley L, Lord S, Bogduk N. The pathophysiology of whiplash. In: Malanga GA, ed. Cervical Flexion-Extension/Whiplash Injuries. Philadelphia: Hanley & Belfus,1998: 209-242.
19. Currey HLF, Hull S. Rheumatology for General Practitioners. Oxford: Oxford University Press 1987.
20. Klippel JH, Dieppe PA, (eds). Rheumatology. London: CV Mosby 1994.
21. Urwin M, Symmons D, Allison T, et al. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence at different anatomical sites, and the relation to social deprivation. *Annals of the Rheumatic Diseases* 1998; **57**: 649-655.
22. Takala J, Sievers K, Klaukka T. Rheumatic symptoms in the middle-aged population in south-western Finland. *Scandinavian Journal of Rheumatology* 1982; **suppl 47**: 15-29.
23. Bergenudd H, Lindgärde F, Nilsson B, Petersson CJ. shoulder pain in middle age a study of prevalence and relation to occupational work load and psychosocial factors. *Clinical Orthopaedics & Related Research* 1988; **231**: 234-238.
24. Palmer KT, Coggon D, Kellingray S, Griffin MJ, Cooper C. A national survey of occupational exposure to hand-transmitted vibration and associated health effects. *Proceedings from PREMUS-ISEOH, abstracts* 1998; 263.
25. Early PF. Population studies in Dupuytren's contracture. *Journal of Bone & Joint Surgery* 1962; **44B**: 602-613.
26. Mikkelsen OA. The prevalence of Dupuytren's disease in Norway. *Acta Chirurgica Scandinavica* 1972; **138**: 695-700.

27. Gordon S. Dupuytren's contracture: the significance of various factors in its aetiology. *Annals of Surgery* 1954; **140**: 683-687.
28. Hueston JD. The incidence of Dupuytren's contracture. *Medical Journal of Australia* 1960; **47**: 999-1000.
29. Hagberg M, Silverstein B, Wells R, et al. Work Related Musculoskeletal Disorders (WMSDs): a reference book for prevention. London: Taylor & Francis Ltd. 1995.
30. Homan MM, Franzblau A, Werner RA, Albers JW, Armstrong TJ, Bromberg MB . Agreement between symptom surveys, physical examination procedures and electrodiagnostic findings for the carpal tunnel syndrome. *Scandinavian Journal of Work, Environment & Health* 1999; **25**: 115-124.
31. De Krom MCTF, Knipschild PG, Kester ADM, Thijs CT, Boekkooi PF, Spaans F . Carpal tunnel syndrome: Prevalence in the general population. *Journal of Clinical Epidemiology* 1992; **45**: 373-376.
32. Wolfe F, Smyth HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis & Rheumatism* 1990; **33**: 160-172.
33. Croft P, Schollum J, Silman A. Population study of tender point counts and pain as evidence of fibromyalgia. *British Medical Journal* 1994; **309**: 696-699.
34. Croft P, Burt J, Schollum J, Thomas E, Macfarlane G, Silman A. More pain, more tender points: Is fibromyalgia just one end of a continuous spectrum? *Annals of the Rheumatic Diseases* 1996; **55**: 482-485.
35. Ide CW. Again, again....repetitive strain. *Rheumatology in Practice* 1989; 24-27.
36. Helliwell PS. A review of diagnostic criteria for work-related upper limb disorders (unpublished report). 1996;
37. McBeth J, Silman AJ. Unraveling the association between chronic widespread pain and psychological distress: An epidemiological approach. *Journal of Psychosomatic Research* 1999; **47**: 109-114.
38. Dudley F, Fry J. Practical problems in rheumatology. London: Martin Dunitz Ltd. 1983.
39. Hodgekin K. Towards earlier diagnosis. A guide to primary care. Edinburgh: Churchill Livingstone 1985.
40. Croft P. Diagnosing regional pain: the view from primary care. *Bailliere's best practice and research.Clinical rheumatology* 1999; **13**: 231-242.
41. Bamji AN, Erhardt CC, Price TR, Williams PL. The painful shoulder: Can consultants agree? *British Journal of Rheumatology* 1996; **35**: 1172-1174.

42. Buchbinder R, Goel V, Bombardier C, Hogg-Johnson S. Classification systems of soft tissue disorders of the neck and upper limb: Do they satisfy methodological guidelines? *Journal of Clinical Epidemiology* 1996; **49**: 141-149.
43. Davis TRC. Diagnostic criteria for upper limb disorders in epidemiological studies. *Journal of Hand Surgery* 1998; **23B**: 567-569.
44. Cooper C, Baker PD. Upper limb disorders. *Occupational Medicine* 1996; **46**: 435-437.
45. Waris P, Kuorinka I, Kurppa K. Epidemiologic screening of occupational neck and upper limb disorders. Methods and criteria. *Scandinavian Journal of Work, Environment & Health* 1979; **5**: 25-38.
46. Kuorinka I, Koskinen P. Occupational rheumatic diseases and upper limb strain in manual jobs in a light mechanical industry. *Scandinavian Journal of Work, Environment & Health* 1979; **5**: 39-47.
47. Luopajarvi T, Kuorinka I, Virolainen M, Holmberg M. Prevalence of tenosynovitis and other injuries of the upper extremities in repetitive work. *Scandinavian Journal of Work, Environment & Health* 1979; **5**: 48-55.
48. Kurppa K, Viikari-Juntura E, Kuomas E, Huuskonen M, Kivi P. Incidence of tenosynovitis or peritendinitis and epicondylitis in a meat-processing factory. *Scandinavian Journal of Work, Environment & Health* 1991; **17**: 32-37.
49. Viikari-Juntura E, Kurppa K, Kuosma E, et al. Prevalence of epicondylitis and elbow pain in the meat-processing industry . *Scandinavian Journal of Work, Environment & Health* 1991; **17**: 38-45.
50. Viikari-Juntura E. Neck and upper limb disorders among slaughterhouse workers. An epidemiologic and clinical study. *Scandinavian Journal of Work, Environment & Health* 1983; **9**: 283-290.
51. Silverstein BA. The prevalence of upper extremity cumulative disorders in industry (Thesis). The University of Michigan, Occupational Health and Safety 1985.
52. McCormack RRJr, Inman RD, Wells A, Berntsen C, Imbus HR. Prevalence of tendinitis and related disorders of the upper extremity in a manufacturing workforce. *Journal of Rheumatology* 1990; **17**: 958-964.
53. Kuorinka I, Viikari-Juntura E. Prevalence of neck and upper limb disorders (NLD) and work load in different occupational groups. Problems in classification and diagnosis. *Journal of Human Ergology* 1982; **11**: 65-72.

54. Norregaard J, Jacobsen S, Kristensen JH. A narrative review on classification of pain conditions of the upper extremities. *Scand J Rehab Med* 1999; **31**: 153-164.
55. Turner-Stokes L. Clinical differential diagnosis of shoulder pain. *British Journal of Hospital Medicine* 1996; **56**: 73-77.
56. Larson HM, O'Connor FG, Nirschl RP. Shoulder pain: The role of diagnostic injections. *American Family Physician* 1996; **53**: 1637-1643.
57. Hawkins RJ, Abrams JS. Impingement syndrome in the absence of rotator cuff tear (stages 1 and 2). *Orthopedic Clinics of North America* 1987; **18**: 373-382.
58. Bland JH, Merrit JA, Boushey DR. The painful shoulder. *Seminars in Arthritis and Rheumatisim* 1977; **7**: 21-47.
59. Oxner KG. Magnetic resonance imaging of the musculoskeletal system: Part 6. The shoulder . *Clinical Orthopaedics & Related Research* 1997; **334**: 354-373.
60. Booth REJr, Marvel JPJr. Differential diagnosis of shoulder pain . *Orthopedic Clinics of North America* 1975; **6**: 353-379.
61. Winters JC, Groenier KH, Sobel JS, Arendzen HH, Meyboom-de Jongh B. Classification of shoulder complaints in general practice by means of cluster analysis. *Archives of Physical Medicine & Rehabilitation* 1997; **78**: 1369-1374.
62. Winters JC, Sobel JS. Shoulder complaints in general practice. A simple approach. *Nederlands Tijdschrift voor Geneeskunde* 1998; **142**: 1241-1243.
63. Katz JN, Larson MG, Fossel AH, Liang MH. Validation of a surveillance case definition of carpal tunnel syndrome. *American Journal of Public Health* 1991; **81**: 189-193.
64. Katz JN, Larson MG, Sabra A, et al. The carpal tunnel syndrome: diagnostic utility of the history and physical examination findings. *Annals of Internal Medicine* 1990; **112**: 321-327.
65. De Krom MCTF, Knipschild PG, Kester ADM, Spaans F. Efficacy of provocative tests for diagnosis of carpal tunnel syndrome. *Lancet* 1990; **335**: 393-395.
66. Rempel D, Evanoff B, Amadio PC, et al. Consensus criteria for the classification of carpal tunnel syndrome in epidemiologic studies. *American Journal of Public Health* 1998; **88**: 1447-1451.

67. Harrington JM, Carter JT, Birrell L, Gompertz D. Surveillance case definitions for work related upper limb pain syndromes. *Occupational & Environmental Medicine* 1998; **55**: 264-271.
68. Helliwell PS. Diagnostic criteria for work-related upper limb disorders. *British Journal of Rheumatology* 1997; **35**: 1195-1196.
69. Palmer K, Coggon D, Cooper C, Doherty M. Work related upper limb disorders: Getting down to specifics. *Annals of the Rheumatic Diseases* 1998; **57**: 445-446.
70. Katz JN, Stirrat CR. A self-administered hand diagram for the diagnosis of carpal tunnel syndrome. *Journal of Hand Surgery* 1990; **15A**: 360-363.
71. Palmer K, Walker-Bone K, Linaker C, et al. The Southampton examination schedule for the diagnosis of musculoskeletal disorders of the upper limb. *Annals of the Rheumatic Diseases* 2000; **59**: 5-11.
72. Croft P, Pope D, Boswell R, Rigby A, Silman A. Observer variability in measuring elevation and external rotation of the shoulder. *British Journal of Rheumatology* 1994; **33**: 942-946.
73. Campbell L, Pannett B, Egger P, Cooper C , Coggon D. Validity of a questionnaire for assessing occupational activities. *American Journal of Industrial Medicine* 1997; **31**: 422-426.
74. Pope DP, Silman AJ, Cherry NM, Pritchard C, Macfarlane GJ. Validity of a self-completed questionnaire measuring the physical demands of work. *Scandinavian Journal of Work, Environment & Health* 1998; **24**: 376-385.
75. Kuorinka I, Jonsson B, Kilbom A, et al. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Applied Ergonomics* 1987; **18**: 233-237.
76. Palmer K, Smith G, Kellingray S, Cooper C. Repeatability and validity of an upper limb and neck discomfort questionnaire: the utility of the standardized Nordic questionnaire. *Occupational Medicine* 1999; **49**: 171-175.
77. Baron S, Hales T, Hurrell J. Evaluation of symptom surveys for occupational musculoskeletal disorders. *American Journal of Industrial Medicine* 1996; **29**: 609-617.
78. Dickinson CE, Campion K, Foster AF, Newman SJ, O'Rourke AMT, Thomas PG. Questionnaire development: an examination of the Nordic musculoskeletal questionnaire. *Applied Ergonomics* 1992; **23**: 197-201.
79. Ohlsson K, Attewell RG, Johnsson B, Ahlm A, Skerfving S. An assessment of neck and upper extremity disorders by questionnaire and clinical examination. *Ergonomics* 1994; **37**: 891-897.

80. Ware J. The MOS 36-item short-form health survey. *Medical Care* 1992; **30**: 473-483.
81. Green G, Buchbinder R, Forbes A, Bellamy N. A standardized protocol for measurement of range of movement of the shoulder using the Plurimeter-V Inclinometer and assessment of its intrarater and interrater reliability. *Arthritis Care & Research* 1998; **11**: 43-52.
82. Office of Population Censuses and Surveys. Standard Occupational Classification 1990. London: HMSO. 1990
83. SPSS Inc. SPSS for Windows. 1993;
84. StataCorp. Stata Statistical software. 1999;
85. Wishart D. ClustanGraphics. 1999;
86. Krzanowski WJ. Principles of multivariate analysis: a user's perspective. Oxford: Clarendon 1988.
87. Milligan GW, Cooper MC. Methodology review: clustering methods. *Applied Psychological Measurement* 1987; **11**: 329-354.
88. Fowlkes EB, Gnanadesikan R, Kettenring JR. Variable selection in clustering. *Journal of Classification* 1988; **5**: 205-228.
89. Blackburn R. Replicated personality disorder clusters among mentally disordered offenders and their relation to dimensions of personality. *Journal of Personality Disorders* 1996; **10**: 68-81.
90. Milligan GW, Cooper MC. A study of standardization of variables in cluster analysis. *Journal of Classification* 1988; **5**: 181-204.
91. Fleiss JL, Zubin J. On the methods and theory of clustering. *Multivariate Behavioural Research* 1969; **4**: 235-250.
92. Hubert L, Arabie P. Comparing partitions. *Journal of Classification* 1985; **2**: 193-218.
93. Hannappel S, Piepho B. Cluster analysis of environmental data which is not interval scaled but categorical. *Chemosphere* 1996; **33**: 335-342.
94. Gower JC. A general coefficient of similarity and some of its properties. *Biometrics* 1971; **27**: 857-874.
95. De Soete G. Optimal variable weighting for ultrametric and additive tree clustering. *Quality and Quantity* 1986; **20**: 169-180.
96. Milligan GW. A validation study of variable weighting algorithm for cluster analysis. *Journal of Classification* 1989; **6**: 53-71.
97. Punj G, Steward DW. Cluster analysis in market research: review and suggestions for application. *Journal of Marketing Research* 1983; **20**: 134-148.

98. Sokal RR, Rohlf FJ. The comparison of dendrograms by objective methods. *Taxon* 1962; **11**: 33-40.
99. Macnaughton-Smith P, Williams WT, Dale MB, Mockett LG. Dissimilarity analysis: a new technique of hierarchical sub-division. *Nature* 1964; **202**: 1034-1035.
100. Gordon AD. A review of hierarchical classification. *Journal of the Royal Statistical Society Series A* 1987; **150**: 119-137.
101. Everitt BS. Unresolved problems in cluster analysis. *Biometrics* 1979; **35**: 169-181.
102. Milligan GW. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika* 1980; **45**: 325-342.
103. Ward JHJr. Hierarchical grouping to optimize an objective function. *American Statistical Association Journal* 1963; 236-244.
104. Milligan GW. Ultrametric hierarchical clustering algorithms. *Psychometrika* 1979; **44**: 343-346.
105. Pruzansky S, Tversky A, Carroll JD. Spatial versus tree representations of proximity data. *Psychometrika* 1982; **47**: 3-24.
106. Shepard RN. Multidimentional scaling, tree-fitting, and clustering. *Science* 1980; **210**: 390-398.
107. Selim SZ, Ismail MA. k-means-type algorithms: a generalised convergence theorem and characterization of local optimality. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1984; **PAMI-6**: 81-86.
108. Cormack RM. A review of classification. *Journal of the Royal Statistical Society Series A* 1971; **134**: 321-367.
109. Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 1985; **50**: 159-179.
110. Dubes RC. how many clusters are best? - an experiment. *Pattern Recognition* 1987; **20**: 645-663.
111. Ellis TE, Rudd MD, Rajab MH. Cluster analysis of MCMI scores of suicidal psychiatric patients: four personality profiles. *Journal of Clinical Psychology* 1996; **52**: 411-422.
112. Taylor E, Everitt B, Thorley G, Schachar R, Rutter M, Wieselberg M. Conduct disorder and hyperactivity: II A cluster analytic approach to the identification of a behavioural syndrome. *British Journal of Psychiatry* 1986; **149**: 768-777.
113. Milligan GW. A Monte Carlo study of thirty internal criterion measures for cluster analysis. *Psychometrika* 1981; **46**: 187-199.

114. Breckenridge JN. Replicating cluster analysis: method, consistency, and validity. *Multivariate Behavioural Research* 1989; **24**: 147-161.
115. McIntyre RM, Blashfield RK. A nearest-centroid technique for evaluating the minimum-variance clustering procedure. *Multivariate Behavioural Research* 1980; **2**: 225-238.
116. Lapointe F-J, Legendre P. A classification of pure malt scotch whiskies. *Applied Statistics - Journal of the Royal Statistical Society Series C* 1994; **43**: 237-257.
117. Dollfus S, Everitt B, Petit M. Identification de sous-groupes de patients schizophrènes par une analyse typologique. *Ann.Méd.-Psychol.* 1995; **153**: 517-521.
118. Walker-Bone K, Byng P, Linaker C, et al. Reliability of the Southampton examination schedule for the diagnosis of upper limb disorders in the general population. *Annals of the Rheumatic Diseases* 2002; **61**: 1103-1106.
119. Marx RG, Bombardier C, Wright JG. What do we know about the reliability and validity of physical examination tests used to examine the upper extremity? *Journal of Hand Surgery* 1999; **24A**: 185-193.
120. Liesdek C, Van Der Windt DAWM, Koes BW, Bouter LM. Soft-tissue disorders of the shoulder. A study of inter-observer agreement between general practitioners and physiotherapists and an overview of physiotherapeutic treatment. *Physiotherapy* 1997; **83**: 12-17.
121. Pellecchia GL, Paolino J, Connell J. Intertester reliability of the Cyriax evaluation in assessing patients with shoulder pain. *Journal of Orthopaedic & Sports Physical Therapy* 1996; **23**: 34-38.
122. Pope D, Croft P. Surveys using general practice registers: Who are the non-responders? *Journal of Public Health Medicine* 1996; **18**: 6-12.
123. Reading I, Walker-Bone K, Palmer K, Coggon D, Cooper C. Diagnosis of cervical disorders in epidemiological studies: utility of cervical ranges of movement. *Rheumatology* 2002; **41**: 45
124. Reading I, Walker-Bone K, Palmer KT, Cooper C, Coggon D. Anatomical Distribution of Sensory Symptoms in the Hand and Their Relation to Neck Pain, Psychosocial Variables and Occupational Activities. *American Journal of Epidemiology* 2003; **157**: 524-530.
125. Ferry S, Silman AJ, Pritchard T, Keenan J, Croft P. The association between different patterns of hand symptoms and objective evidence of median nerve compression: A community-based survey. *Arthritis & Rheumatism* 1998; **41**: 720-724.

126. Dunne CA, Thompson PW, Cole J, et al. Carpal tunnel syndrome: evaluation of a new method of assessing median nerve conduction at the wrist. *Annals of the Rheumatic Diseases* 1996; **55**: 396-398.

127. Cyriax J. Textbook of orthopaedic medicine: diagnosis of soft tissue lesions. London: Bailliere Tindall 1982.

128. Pope DP, Croft PR, Pritchard CM, Macfarlane GJ, Silman AJ. The frequency of restricted range of movement in individuals with self-reported shoulder pain: Results from a population-based survey. *British Journal of Rheumatology* 1996; **35**: 1137-1141.

129. Hamilton PG. The prevalence of humeral epicondylitis: a survey in general practice. *Journal of the Royal College of General Practitioners* 1986; 464-465.

130. Massy-Westropp N, Grimmer K, Bain G. A systematic review of the clinical diagnostic tests for carpal tunnel syndrome. *Journal of Hand Surgery* 2000; **25**: 120-127.

131. Phalen GS. Spontaneous compression of the median nerve at the wrist. *JAMA* 1951; **145**: 1128-1133.

132. Bernard BP, (ed). Musculoskeletal disorders (MSDs) and workplace factors. Cincinnati (OH): US department of Health and Human Services 1997.

133. Gudmundsson KG, Arngrimsson R, Jonsson T. Eighteen years follow-up study of the clinical manifestations and progression of Dupuytren's disease. *Scandinavian Journal of Rheumatology* 2001; **30**: 31-34.

134. Green S, Buchbinder R, Glazier R, Forbes A. Systematic review of randomised controlled trials of interventions for painful shoulder: selection criteria, outcome assessment, and efficacy. *British Medical Journal* 1998; **316**: 354-360.