ISSN 1324-5627

Australasian Musculoskeletal Medicine

- The Physiology of Deep, Somatic Pain
- The Interobserver Reliability of Thoracic Spinal Examination
- Autologous Chondrocyte Implantation: Is it the answer?
- Cervical Pain and Regional Relationships
- Knowledge of Musculoskeletal Medicine at Undergraduate & Postgraduate Levels
- Reliability and Validity Studies of Diagnostic Procedures in M/M Medicine

Vol. 7 No. 1 May 2002 **Registered by Australia Post Publication No. NAW 6362**

AAMM Website: www.musmed.com

Australian Association of Musculoskeletal Medicine Office Bearers 1999 - 2002

President
Dr Steve Jensen MB BS, Grad Dip Musc Med
5 Stanlake St, Footscray Vic 3011
Ph: +61 3 93185233
Fax: +61 3 93186630

Vice President
Dr Scott Masters MB BS, FRACGP, FAFMM
Caloundra Medical Centre
39 Minchinton St, Caloundra, Q 4551
Ph: +61 7 54911144
Fax: +61 7 54911253

Honorary Secretary
Dr Michael Yelland MB BS, FRACGP, FAFMM, Grad Dip Musc Med
64 Wirraway Pde, Inala, Q 4077
Ph: +61 7 32755444
Fax: +61 7 32789987

Honorary Treasurer
Dr Derek Davey MB BS, D Obst, RCOG, Grad Dip Occ H, Grad Dip Musc Med, FAFMM
29 Craigie Rd, Newtown Vic 3220
Ph: +61 3 52223677
Fax: +61 3 52213104

Committee Members
Dr Robert Gassin MB BS, FAFMM, Grad Dip Musc Med
Cranbourne North, Vic 3977
Ph: +61 3 59956999
Fax: +61 3 59956700

Dr Geoff Harding MB BS, FAFMM, Grad Dip Musc Med (immediate past president)
Sandgate, Q
Ph: +61 7 32695522
Fax: +61 7 38693288

Dr Michael Oei MB BS, Dip Phys Med, M Med Phys Med, Cert Man Med
Darlinghurst, NSW
Ph: +61 2 8302 1111
Fax: +61 2 8302 1195

Dr Des Schimeld MB BS, Grad Dip Musc Med, BSc(Hons)
Dulwich, SA 5065
Ph: +61 8 83612769

Dr Phil Watson BSc, MB, ChB, FRACGP, Dip Obst, Dip MusMed, CIME
Sunnybank, Q 4109
Ph: +61 7 33457117
Fax: +61 7 3216 9052

New Zealand Association of Musculoskeletal Medicine Office Bearers 2002 - 2003

President
Dr Steve Bentley
2 Hadley Close Dunedin
Ph: +64 3 467 9616
Fax: +64 3 474 1986
stevelf@cs.co.nz

Secretary
Dr Gary Collinson
4 Lynbrooke Ave Blockhouse Bay, Auckland
Ph; +64 9 627 1024
Fax: +64 9 672 1181
gajisuda@xtra.co.nz

Treasurer
Dr Clemens Franzmayr
Longstaffs Rd RD 4, Christchurch
Ph: + 64 3 352 9357

Past President
Dr James Watt
2 Pupuke Rd Takapuna, Auckland
Ph: +64 9 486 3146
jameswatt@clear.net.nz

Academic Co-ordinator
Dr Jim Borowczyk
256 Papanui Rd Merivale, Christchurch
Ph: +64 3 355 0342
Fax: + 64 3 355 7071
jimbo@clear.net.nz

Committee Members
Dr Peter McKenzie, Nelson
Ph: +64 3 366 1536

Dr Peter Catt, New Plymouth
Ph: +64 6 753 2939

Dr Keith Laubscher, Hamilton
Ph: +64 9 524 0585

Dr John MacVicar, Christchurch
Ph: +64 3 366 1536

Dr Giresh Kanji, Wellington

Dr David Jacks, Havelock North
Ph: +64 6 877 7555

AAMM Website: www.musmed.com

Contents

Editorial 3

From AAMM President 4

From NZAMM President 5

The Physiology of Deep, Somatic Pain 6

The Interobserver Reliability of Thoracic Spinal Examination 16

Quiz 22

Autologous Chondrocyte Implantation: Is it the answer? ... 23

Cervical Pain and Regional Relationships 25

Knowledge of Musculoskeletal Medicine at Undergraduate and Postgraduate Levels 28

Reliability and Validity Studies of Diagnostic Procedures in Musculoskeletal/Manual Medicine 33

The Effects of a Pragmatic Approach to Low Back Pain 44

Difficulties in Management: Chronic Patients 47

Practice Tip: "Pseudo" Gout not Pseudo-Gout 49

Journal Abstracts 49

FIMM Report 55

Letters to editor 56

Answers to Quiz 57

Educational Activities 58

Australasian Musculoskeletal Medicine is published by the Australian Association of Musculoskeletal Medicine for medical practitioners interested in the aetiology and management of musculoskeletal disorders. Opinions expressed are those of the authors and not necessarily those of the editor or the Association. Editorial comment may reflect the opinions of the editor alone. Contributions on any relevant topic are welcome for submission to the editor, Dr Scott Masters, Caloundra Sports Medicine Centre, Suite 4 "Trinity House", 39 Minchinton St., Caloundra Qld 4551. Phone: + 61 7 5491 1144(w), +61 7 5443 9475. Fax: + 61 7 5491 1253(w). Email: scotty1@ozemail.com.au.

Editorial

Dr Scott Masters

The search for truth often starts with much hope and optimistic expectation. The starry-eyed glimmer in explorers' eyes can fade quickly as the enormity of their task becomes apparent. One such exploration was that of the Search for Extra-Terrestrial Intelligence (SETI) which began in earnest in the 1970s. Speculation abounded as to the message content from the superior intelligence that would be out there in deep space. The enthusiasts never doubted that their multimillion dollar scanning sweep of the skies would divulge to them a plethora of information. Technological miracles could well abound from this knowledge, moving human culture along hitherto unthought of pathways. Or so was thought.

One of the original investigators spent the first 10 years of the search faithfully manning his station, intently listening for the space sounds that intelligent civilisations must be broadcasting. After 10 years his grand accumulation of data was one signal, a signal that was unidentifiable. After being swept up in the initial euphoria of SETI, this researcher and others had to face the reality of the enormous complexity of their chosen task. What wavelength would ETs be likely to communicate on? Would it be the 21 cm line, i.e., the natural frequency of vibrating atoms of hydrogen gas, the simplest abundant element of the cosmos? Would it be in a form unrecognisable to mankind? Or would aliens simply ignore earthlings as inferior philistines not worth the trouble?

Among all this doubt, SETI has continued over 30 years, still without one successful hit. The search has enormous potential but is plagued by unanswerable questions resulting in scientists making informed guesses as to best options for search routes. Similarly, scientists face dilemmas now as to the future direction for stem cell research. Although their dilemmas are different, they both are pointers to the problems that lay ahead for souls who reach out to discover the truth about the universe in which they reside.

Now that the first publication has surfaced from the Australian Musculoskeletal Medicine Initiative (see p.8 of this journal), it will be interesting to watch the response from parties with vested interests in back pain management. Criticism will abound as to the design of the study, the reason behind the differences in treatment results and the relevance of the study to different communities. Some will see the Initiative as trying to be all-conquering, when in fact it just seeks to take one step along the pathway to truth. Let the debate begin! I only hope the Initiative results in more testing of evidence-based guidelines, particularly in chronic musculoskeletal pain.

I've just taken the first tentative steps into arranging some primary care research into the management of frozen shoulder syndrome, currently an evidence-based remote zone. Just agreeing on methodology is enough of a headache, but then to find funding and recruitment on top makes me wonder what type of Pandora's box I've opened. Perhaps I should have stuck to the straighter, less convoluted path of guru-based medicine. Or even the economically sound six-minute medicine. But then I'm reminded of the scientist's pin-up boy, Mr Spock, who told us that "in truth there is beauty". May the truth reveal itself at warp speed!

As such, the emphasis on primary care research seems appalling. Up to the 80s it was virtually nonexistent and in the following 20 years has just started to find its feet. Now, there is sponsorship for GPs to do primary care research and receive some financial support. The Primary Care Novice Research Fellowships are now available through funding from the Commonwealth Department of Health and Ageing. I can vouch for their availability in Queensland and would urge readers to check through their local divisions as to funding through their own states. The odyssey towards truth continues.

Scott Masters

From the AAMM President, Dr Steve Jensen

An open letter to all GP Divisions

Dear GP Division

I am writing to offer you the teaching services and expertise of the Australian Association of Musculoskeletal Medicine. We have a wealth of teaching experience and expertise within the organisation willing and able to educate your members in matters musculoskeletal.

Generally speaking, members of AAMM belong to this organisation because they recognised the deficiency of musculoskeletal teaching at both undergraduate and postgraduate levels, and sought to further their knowledge in this area.

But why did they need to?

Did you realise that somewhere in the vicinity of 15-20% of all GP consultations are for musculoskeletal problems? And that back pain alone accounts for approximately 5% of all consultations?

And this figure of 15-20% does not include musculoskeletal problems that may masquerade as visceral disease. For example, it has been shown in a general population study,[1] using the criteria outlined by the International Headache Society (IHS),[2] that the incidence of cervicogenic headache (CGH) is 18% of all chronic headaches. Yet how many of your members realise that the diagnostic criteria for CGH include autonomic phenomenon such as nausea, vomiting and photophobia and phonophobia, and thus include it in the differential diagnosis of patients presenting to their practice with headache? And the current evidence base has shown that manual therapy to the cervical spine can be efficacious for CGH and certain types of migraine.[3]

Thus, if we include the likes of CGH and musculoskeletal causes of chest pain,[4] I would suggest that musculoskeletal problems presenting to general practice would be much higher than the published 15-20% figure.

Yet with the deficiency in teaching in this vast area of medicine, I would humbly suggest that most of your members could be better equipped to diagnose and manage these problems. Thus I believe that your members owe it to themselves and their patients to upgrade their skills in this area.

It has been shown that musculoskeletal practitioners, using evidence based guidelines for acute back pain[5] used less diagnostic imaging, and used less analgesics and NSAIDs, resulting in a much lower cost to the patient and thus the community as a whole. Furthermore, far fewer patients ended up with chronic pain, and patient satisfaction was very high.

Generally speaking, the basic practical skills required in musculoskeletal medicine are relatively easily learnt. We believe we can at least teach your members to better recognise and diagnose musculoskeletal problems, including those masking as visceral disease such as CGH and musculoskeletal causes of chest pain. If they wish, we can teach them simple and proven management strategies, such as explanation and reassurance, and home rehabilitation[6,7] which forms the basis of musculoskeletal management generally, and the evidenced based guidelines for management of low back pain.

For those wishing to develop further skills, we can also teach them simple and safe manual therapy techniques to further enhance their overall clinical skills. In Australasia, for those who wish to develop a special interest in musculoskeletal medicine, the option exists for further study through locally held workshops and even postgraduate diploma courses. We have a major forum each year, our annual scientific meeting, which almost invariably runs practical hands-on workshops for beginners in the field. The 2002 conference is in Melbourne, October 17-20, and will again run such a workshop. Your members are most welcome to attend.

I hereby cordially invite you to avail your organisation of our educational services, so that your members may advance their clinical skills in this area with the resultant improved and cost-effective patient care, and heightened patient and practitioner satisfaction.

All you need to do is contact me, and I can arrange an appropriate AAMM educator in your state to speak to your members at one of your divisional meetings.

I look forward to hearing from you.

Yours faithfully

Steve Jensen, MB BS FAFMM
President, Australian Association of Musculoskeletal Medicine

References

1. Nilsson N. Prevalence of cervicogenic headache in a random population sample of 20-59 year olds. *Spine* 1995; 20: 1884-88.
2. Sjaastad O, Fredericksen TA, Pfaffernrath V. Cervicogenic headache: diagnostic criteria. *Headache* 1998; 38: 442-45.
3. Gassin R, Masters S. Spinal manual therapy - the evidence. *Aust Musculoskeletal Med* 2001; 6:1: 26-31.
4. Jensen S. Musculoskeletal causes of chest pain. *Aust Fam Physician* 2001; 30: 834-39.
5. McGuirk B, King W, Govind J, Lowry J, Bogduk N. Safety, efficacy, and cost-effectiveness of evidenced-based guidelines for the management of acute low back pain in primary care. *Spine* 2001; 23: 2615-22.
6. Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered: a randomised controlled trial. *Spine* 1995; 20: 473-77.
7. Indahl A, Haldorsen EH, Holm, et al. Five-year follow-up study of a controlled clinical trial using light mobilisation and an informative approach to low back pain. *Spine* 1998; 23: 2625-30.

From the NZAMM President, Dr James Watt

My term as president in New Zealand has come to an end with our group making substantial headway in our aims of furthering the recognition of musculoskeletal medicine as a specialty.

We have been fortunate in having a close relationship with our Australian colleagues, as the workload has been very heavy and borne by a small number of us. The need now is to develop a training program that will provide learning opportunities for our successors.

The field of musculoskeletal medicine is broad and has many niches each with its advocates. However, the patient and the referring doctor need to have their problems dealt with and the best care possible provided. The presenting symptom is almost always pain, often associated with impairment or disability. Thus a clear understanding of the mechanism of pain is essential to the practice of musculoskeletal medicine.

Management of the pain and the disability then depends on the skills available to the patient from the particular physician. The database contained within the medical literature provides information that helps in the choice of management options. However, it should be remembered that the literature offers epidemiological data that are not always relevant to the individual standing before them. Frequently the quality of studies is poor. With poor quality of data, all the cases that should be excluded become "polluting cases", which will cause a trend towards the normal unmodified outcome. Thus the reported effect of any treatment, or indeed any examination, will be nullified.

Musculoskeletal medicine by its nature is a branch of medicine where the outcomes depend on the skills of the person applying them. It differs from fields such as drug therapy where a certain amount of a chemical will be administered. In this example the variable will be the metabolism of a specific amount of that chemical within the individual. Almost invariably, treatment other than prescribing drugs relies on the manual dexterity of the therapist. Thus we need to provide training programs which address these needs. Accuracy is imperative, whether it is in the placement of a needle for injection or acupuncture, or in the application of a force to change the biomechanics of a joint complex. Those of us who have gained skills in a particular field need to have both a well-designed teaching program, and sufficient time to supervise the student in order to ascertain that knowledge or a particular skill has been acquired.

Case series are still relevant in the field. Each operator is able to collect data and report what they find. If one person using a particular treatment shows better results than his or her colleagues, the treatment and its application warrant closer inspection. The treatment used should then be disseminated and incorporated within the armamentarium.

Training in practical skills should borrow heavily from the surgical model as procedural technique is of the essence in most cases we manage. When we refer on, we need a close understanding of both the type and quality of work done by therapists to whom we refer.

Training in communication is important and cannot be counted on having occurred during the undergraduate years. Indahl, among others, has shown that a very powerful therapeutic tool is the restoration of confidence and the removal of perception of nursing a disease. Pain does not necessarily equate with serious damage. Abnormality demonstrated on investigation is not unusual. Many doctors know this but most patients do not. Such abnormality is sometimes used to cause patients to become dependent and thus is a useful business tool (if outcome is judged by income rather than the patients' wellbeing). Correction of this misapprehension should be one of the first issues addressed in any musculoskeletal consultation.

Restoration of the patient's sense of wellbeing is one of the most powerful things we can offer and one where the epidemiological data are very important tools. We should all be proficient in this art.

Appropriate application of the pharmacological armoury is also a critically important skill.

We have come from diverse fields and hold a wide variety of skills and knowledge. We have a communal interest in the field of musculoskeletal medicine and should each ask ourselves how best we can further its development. We must continue to pool our knowledge and share the burden of development by offering all we can into the association's pool, rather than sitting back and expecting it to offer something to us.

The Physiology of Deep, Somatic Pain

by Nikolai Bogduk, Newcastle Bone and Joint Institute

Contemporary knowledge about pain physiology is dominated by cutaneous pain, neuroma pain and neuropathic pain. The reason for this is understandable. The skin provides a target that can be stimulated in a controlled manner using a variety of stimuli – touch, pin-prick, heat, and applications of chemicals, both in experimental animals and in human volunteers. Cutaneous pain can be studied without invading the organism. Neuromas and nerve injuries can be induced at selected and desired sites and provide a known and isolatable source of nociception. Phenomena such as cutaneous hyperalgesia and receptive fields can be readily mapped because they are distributed across only a two-dimensional surface.

The irony is that epidemiologically, cutaneous pain, neuroma, and neuropathic pain are relatively uncommon. Far more common is deep, somatic pain, otherwise referred to as musculoskeletal pain for the reason that, to the patient, the pain seems to arise in muscles, bones or joints; it is felt deeply and definitely not in the skin.

For something as common as musculoskeletal pain, knowledge of its physiology is meagre compared to that of cutaneous pain. Although research into cutaneous pain has been critical in elucidating nociceptive pathways and control mechanisms, and although these principles might be applied to musculoskeletal pain, unless they have been explicitly demonstrated to apply, the possibility remains that different and distinctive processes might apply to musculoskeletal pain.

Certain obvious differences are immediately evident. Skin is exteroceptive, designed to respond to external physical stimuli such as heat and touch. Teleologically, there is no reason for deep somatic structures to be heat nociceptive in the same way as skin. It can be construed that the purpose of cutaneous nociception is to avoid or escape external, threatening stimuli; deep somatic pain cannot be escaped. Since deep tissues lack touch transduction, there is no reason to expect they exhibit Aβ allodynia.

The Legacy

The history of research into musculoskeletal pain can be depicted graphically in three time lines (Figure 1). The earliest studies can be classified as clinical experimental studies, in which pain phenomena were studied in normal, human volunteers. These were then followed by anatomical studies, which pursued the histological substrates of deep, somatic pain. The youngest style of research has been animal experiments in which nociception from musculoskeletal tissues, as opposed to skin, has been studied. Each of these streams of research commenced at various times during the twentieth century, and has continued into the present time.

Another dimension of musculoskeletal pain research has been the target structure. Clinical studies have focused largely on pain stemming from the

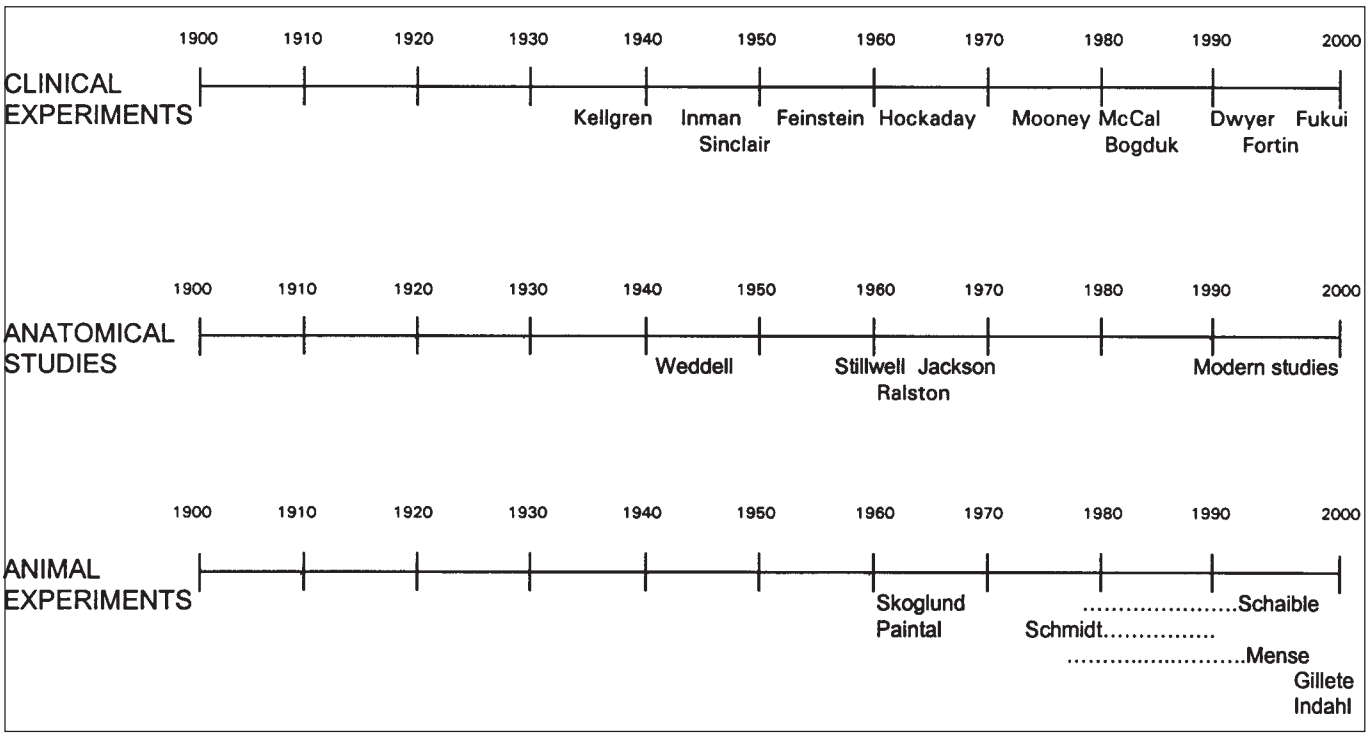


Figure 1. Time lines indicated the occasion, by principal author, of cardinal studies on the mechanisms of deep, somatic pain, in the categories of clinical experiments, anatomical studies in humans, and animal experiments.

joints and muscles of the vertebral column, largely perhaps because of all the musculoskeletal pains, spinal pain has remained the most poorly understood or rather the one least able to be ascribed, conveniently and dismissively, to "arthritis". Meanwhile, animal experiments have focussed on the knee joint because this joint is the most accessible joint whose behaviour can be controlled and studied in perfect isolation. To a lesser extent, animal experiments have used the ankle joint, again ostensibly because it can be isolated and controlled.

Clinical Experiments

Much of our present understanding of the phenomenology of musculoskeletal pain can be traced to work of Kellgren in the late 1930s. In an effort to understand musculoskeletal pain in patients he explored how deep somatic pain might be elicited in normal volunteers, where it was perceived, what it felt like, and what other features were associated with it.

Kellgren's first study[1] was on referred pain from muscle. He demonstrated that noxious stimulation of muscle, with injections of hypertonic saline, produced pain that was diffuse and perceived remote from the site of stimulation. Moreover, in the limbs, muscle pain tended to be perceived towards the joint upon which the muscle acted. Stimulation of axial and paraxial muscles produced pain anteriorly in the trunk or abdomen or into the upper or lower limb.

Kellgren's most lasting and penetrating contribution, however, was in the study of spinal referred pain.

In an era when disc prolapse had just been discovered and spinal pain was ascribed to nerve root compression, Kellgren[2] ventured a competing paradigm. He showed that noxious stimulation of the interspinous ligaments, by injection of hypertonic saline, could produce referred pain in remote areas.[2] Stimulation of thoracic ligaments produced pain in the posterior and anterior chest wall. Stimulation of cervical and lumbar ligaments produced pain in the respective limbs.

Kellgren's experiment was not intended to demonstrate that interspinous ligaments were the source of back pain and neck pain. Rather, they established several principles:

1. Spinal pain could arise from noxious stimulation of intrinsic structures of the vertebral column.
2. Such stimulation produced referred pain in the trunk and limbs.
3. Referred pain could be produced by mechanisms other than nerve root irritation.
4. This referred pain was not neuralgic in nature, in that it was not shooting, burning or stabbing in quality, and not associated with numbness or paraesthesiae in the skin; rather, it was dull and aching in quality, diffuse and hard to localise in distribution, and perceived deeply, in which respects it resembled the complaints of many patients.
5. In order to distinguish this type of referred pain from pain caused by nerve root irritation or pain arising from viscera, it could be referred to as somatic referred pain. That term specified that the source of pain was in the somatic tissues of the body as opposed to viscera or nerves.
6. Somatic referred pain followed a segmental distribution that was not dermatomal in nature (Figure 2). Stimulation of successively lower spinal segments produced pain in successively more caudal regions of the body wall or limbs, but these regions did not correspond to the known dermatomes of the body. Kellgren believed this pattern to reflect a segmental pattern of innervation of deep tissues.

Kellgren's report and interpretation were not well accepted, because they ran contrary to prevailing wisdom that

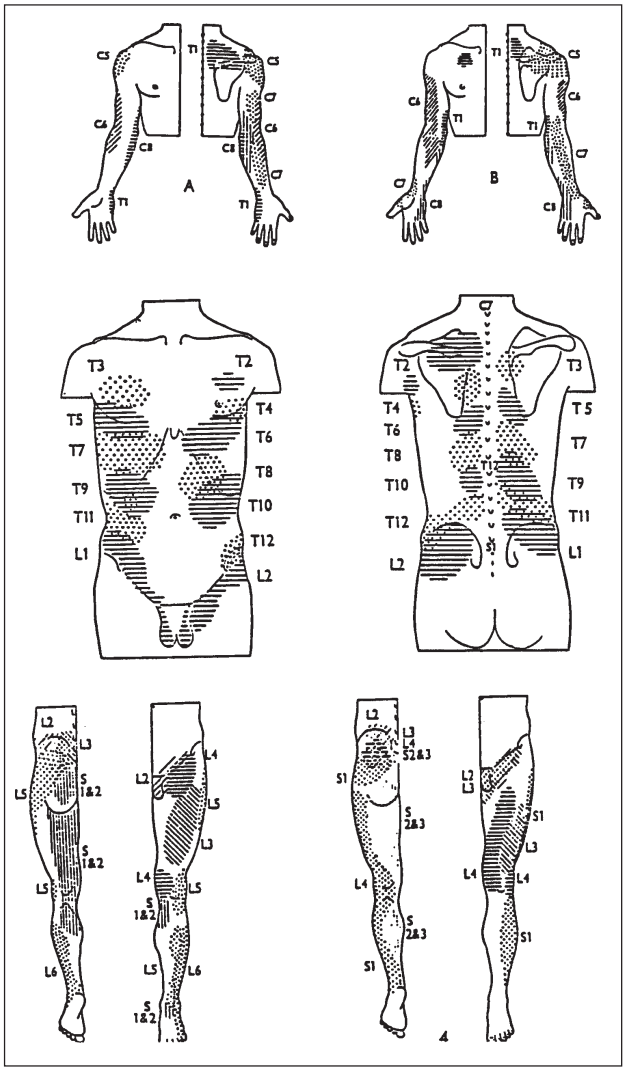


Figure 2. Selections from the maps of Kellgren[2] showing the distribution of referred pain following the noxious stimulation of interspinous ligaments in normal volunteers, at the segments indicated.

The Physiology of Deep, Somatic Pain

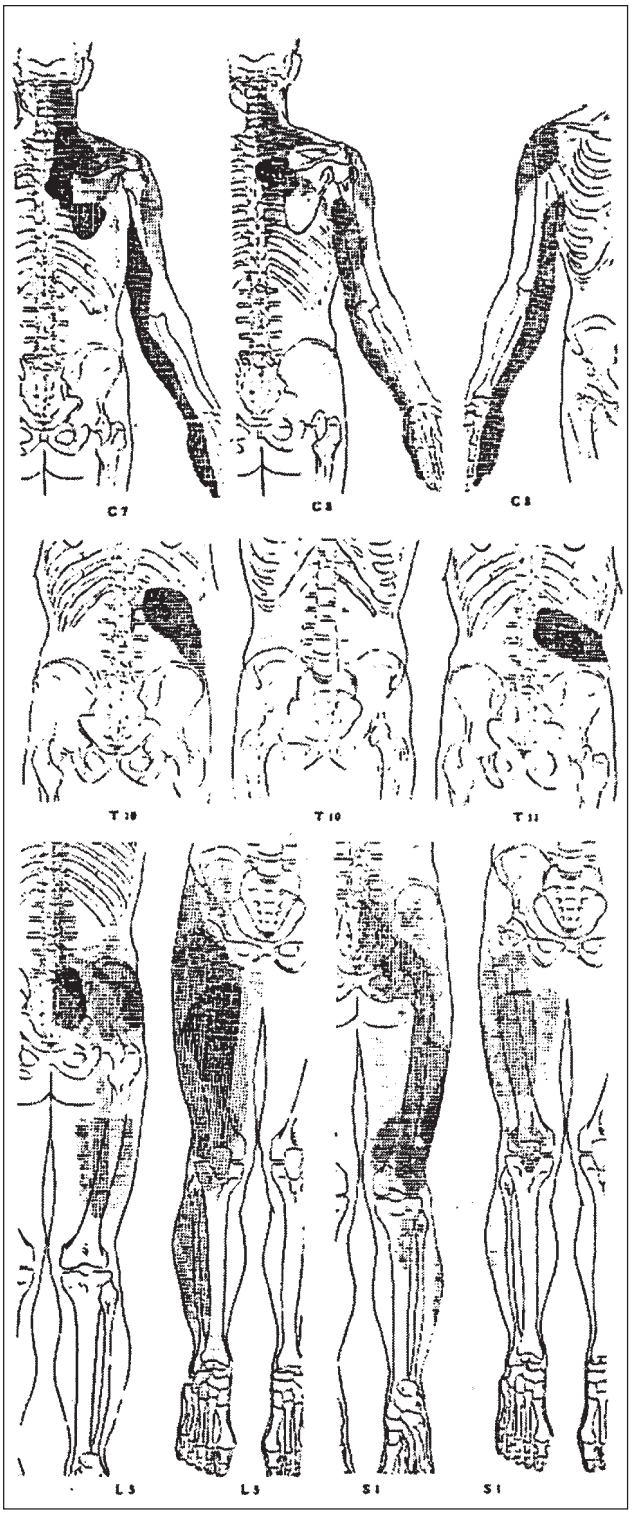


Figure 3. A selection from the maps of Feinstein et al[7] showing the distribution of referred pain following the noxious stimulation of interspinous tissues in normal volunteers, at the segments indicated.

referred pain must be causal by root irritation. Sinclair et al[3] tried to reproduce Kellgren's experiment and failed to produce referred pain to the limbs. They argued against his interpretations and submitted that his injections must have inadvertently stimulated nerve roots. In a contemporary essay on referred pain Sinclair and associates[4] argued that referred pain was due to axonal branching in the periphery, and involved antidromic propagation of impulses to the referred zone, which then triggered pain in that zone, which was then propagated orthodromically back along the same nerve.

However, Kellgren's observations were subsequently reproduced by Hockaday and Whitty[5] and by Whitty and Willison,[6] although the frequency and extent of referred pain to the limbs that they encountered was not as dramatic as that reported by Kellgren. Full corroboration was provided by Feinstein et al[7] who published maps of referred pain that resembled those of Kellgren in extent but not in exact location (Figure 3).

In a short but inordinately influential paper Inman and Saunders[8] firmly consolidated the concept of deep, somatic referred pain. The paper presented little information on methods beyond stating that deep somatic tissues – periosteum, ligaments, bone, joints and muscles, throughout the body were noxiously stimulated by scratching with a needle, drilling with a wire, or by injections of formic acid or 6% saline; it presented no quantitative data; but it assertively declared profound results. The sensitivity of deep somatic tissues was ranked in the order – periosteum > ligament > joint capsule > tendon > fascia > muscle. Most influentially, the paper depicted maps of the dermatomes, the myotomes, and the sclerotomes of the body, in order to contrast their patterns. Dermatomes are the regions of skin innervated by individual spinal nerves, and myotomes are the regions of muscle innervated by a given spinal nerve. Sclerotomes were presented as the regions of bones, joints and ligaments purportedly innervated by the same spinal cord segment. The latter were declared to be the basis for somatic referred pain, and have been repeatedly quoted in the literature since. This paper was influential because it declared an attractive concept but its influence was inordinate because the maps of sclerotomes that it provided were idealised and not based on published quantitative data. The consistency of patterns of referred pain was not stipulated.

In 1950, Kellgren left the spine, and together with Samuel[9] studied the knee joint. In normal volunteers they explored the sensitivity of different structures in the knee with a needle introduced through anaesthetised skin; in patients undergoing arthrotomy they studied the sensitivity of synovium; but in a dramatic experiment they opened the knee of Samuel in order to explore the sensitivity of the synovial membrane across its entire extent. They found the fibrous structures: ligaments and capsule to be nociceptive to me-

chanical and chemical stimulation, but the synovial membrane was largely insensitive to pin-prick, crushing with forceps, and chemical stimulation, except on a few occasions in isolated areas near the upper border of the patella and towards the sides of the joint.

The tradition of Kellgren was resurrected after 1976 when investigators ventured to determine referred pain patterns from specific structures that might be more likely sources of spinal pain than the interspinous ligaments.

Using injections of hypertonic saline, Mooney and Robertson[10] showed in normal volunteers that the lower lumbar zygapophysial joints could be sources of low back pain and referred pain in the lower limbs. They complemented their study with observations of relief of similar patterns of pain in patients following anaesthetisation of the lumbar zygapophysial joints.

This work was corroborated by McCall et al[11] who confirmed that local and referred pain could be evoked in normal volunteers by stimulating the lumbar zygapophysial joints, but the patterns of referred pain that these investigators encountered were not as extensive as those reported by Mooney and Robertson.[10] Moreover, McCall et al demonstrated that the areas of referral from upper lumbar joints overlapped those from lower joints. Consequently, the location of referred pain could not be used to identify the segmental location of a painful joint.

In their experiments, Hockaday and Whitty[5] and Feinstein et al[7] had noted that somatic referred pain could be associated with muscle spasm in the zone of referred pain; and Mooney and Robertson[10] mentioned that referred pain from the lumbar zygapophysial joints was associated with activity in the hamstring muscles, which they demonstrated by EMG. This phenomenon was explored by Bogduk[12] who reproduced Kellgren's experiments but also showed that referred pain from the lower lumbar interspinous ligaments and muscles was accompanied by involuntary activity in the multifidus muscles, tensor fasciae latae and gluteus medius. This activity started shortly after the onset of pain and dissipated as the pain eased over the next few minutes.

The work of Mooney and Robertson[10] was reproduced in the neck by Dwyer et al[13] who stimulated the cervical zygapophysial joints in normal volunteers by distending the joint with injections of contrast medium. The evoked pain was perceived in distinctive, segmental locations (Figure4). In a companion study, the same workers showed how these maps could be used to guide diagnostic investigations.[14]

The sacroiliac joint was the next target in the study of spinal pain. Fortin et al [15] stimulated the sacroiliac joints of normal volunteers with injections of contrast medium and found that the induced pain was perceived over the sacral and gluteal region.

Most recently a pair of Japanese studies[16,17] revisited the lumbar and the cervical zygapophysial joints. The investigators used injections of contrast medium to stimulate individual joints, and electrical stimulation of the nerves that supplied them. They provided quantitative data on the incidence of pain in selected regions following stimulation of a given joint or a given nerve, which was the first time that such data were provided in the history of study of spinal pain. However, the data did not alter the thrust of the conclusion of previous studies. In the lumbar spine, referred pain from the zygapophysial joints could occur in the buttock or lower limbs but not in a reliably distinctive segmental pattern. In the cervical spine, the patterns were more consistent and distinctive, as shown by Dwyer et al[13] (Figure 4).

Work not yet published, but under peer review, will show that the pain patterns reported for the cervical zygapophysial joints (Figure 4) is not specific for zygapophysial joints, for the cervical intervertebral discs exhibit essentially identical pain patterns. Referred pain maps, therefore, are not indicative of the structure that is the actual source of pain, but they do indicate the likely segmental location of the structure. The common factor is neurology. Referred pain maps indicate the segmental innervation of the source of pain but not the responsible structure. Thus, for example, all structures innervated by C5,6 refer pain to the C5-6 area (Figure 4), be they a disc or zygapophysial joint.

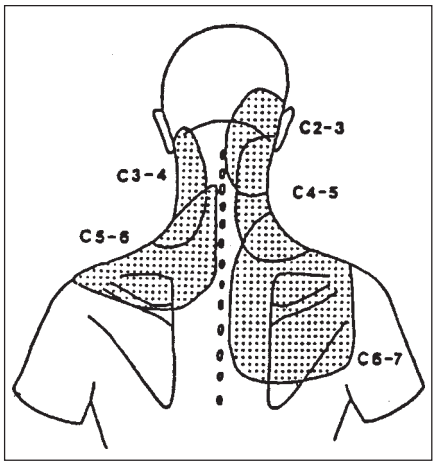


Figure 4. The distribution of referred pain following the stimulation of cervical zygapophysial joints in normal volunteers, at the segmental levels indicated. Based on Dwyer et al.[13]

Summary

Clinical experimental studies in normal volunteers have shown that:

- deep somatic structures are nociceptive;
- in rank order, the sensitivity of structures is periosteum > ligament > joint capsule > tendon > fascia > muscle;
- pain from deep, somatic structures is referred to remote sites;
- pain from muscle tends to be referred to the joint on which the muscle acts;
- pain from spinal structures is referred in a quasi-segmental fashion, at thoracic levels to regions of

The Physiology of Deep, Somatic Pain

thoracic and abdominal walls, at cervical and lumbar levels into the respective limb girdles and limbs.

- Of specific structures of the spine, those that are nociceptive and capable of producing referred pain are:
  the interspinous ligaments
  the paraspinal muscles
  the zygapophysial joints
  the sacroiliac joint.

Comment

The concept of "sclerotomes" was invented to provide an explanation of the patterns of deep, somatic referred pain. It implies that deep tissues are innervated in a segmental fashion analogous to dermatomes and myotomes, and therefore, referred pain is perceived in deep tissues with the same segmental innervation as the source of pain.

Although this concept is attractive as a helpful explanation of referred pain, there is no explicit evidence for it. Dermatomes and myotomes are valid anatomical entities. Their segmental innervation can be demonstrated by anatomical and physiological means. Dermatomes were mapped by studying the zones of eruption of the vesicles of herpes zoster, which constitute a physical tracer of segmental nerves; and by studying the zones of numbness after dorsal rhizotomies. Myotomes were established by mapping zones of weakness after segmental nerve injury, and by mapping EMG activity evoked by electrical stimulation of segmental nerves. Were the experiments ethically feasible, dermatomes and myotomes could be determined by introducing tracer substances into segmental nerves.

There is no equivalent evidence about sclerotomes. No-one has traced segmental nerves to deep tissues using anatomical or physiological means. Sclerotomes lack a physical substrate. Maps of sclerotomes have been based exclusively on the subjective descriptions of patterns of referred pain in individuals undergoing experimental noxious stimulation of deep, somatic tissue.

Although sclerotomes may, indeed, reflect deep segmental innervation that has yet to be demonstrated, another interpretation is that they simply represent perceptual patterns, in which case they are determined more by connections with the central nervous system than by peripheral patterns of innervation. This contrasting interpretation does not invalidate the concept or utility of pain maps but it does challenge the propriety of regarding a sclerotome as an anatomical substrate for deep somatic pain. Rather than a physical entity it may be a psychophysical entity.

Anatomical Studies

Towards the end of the 19th century and in the early 20th century, anatomists had studied the innervation of various tissues: the epithelia of skin and cornea, teeth, mucous and serous membranes, and blood vessels. The pursuit of the anatomical substrate of deep, somatic pain in humans began in 1940.

Weddell and Harpman[18] studied the sensations evoked from deep fascia, tendons and periosteum, and correlated these with the structure of nerve endings found in these tissues. Some 20 years later these studies were complemented by those of Stillwell,[19] on tendons and aponeuroses, and by Ralston et al[20] who studied human fasciae, tendons, ligaments, periosteum, joint capsules and synovium.

Deep tissues were found to be innervated by three types of nerve endings: free nerve endings, complex, unencapsulated receptors, and encapsulated receptors. Fasciae, joint capsules and ligaments typically exhibited all three types of endings. Tendons contained mainly free nerve endings and relatively simple unencapsulated endings and small encapsulated endings. Periosteum exhibited all three types of endings, which were particularly abundant near the sites of attachment of muscles, tendons or ligaments. In synovial membrane, only free nerve endings were detected. The anatomists ascribed a nociceptive function to the free nerve endings. To the unencapsulated endings they ascribed a proprioceptive function. The encapsulated endings they considered to be pressure transducers.

Receptors in spinal tissues were first systematically studied by Jackson et al[21] in 1966, who established that the ligaments and joints of the spine were innervated in a manner like those of the appendicular skeleton. More recent studies, using immunohistochemical and other advanced staining techniques, have confirmed and elaborated these findings.[22-25]

Whereas it was accepted that the spinal ligaments, muscle and synovial joints received a nociceptive innervation, the innervation of intervertebral discs remained controversial until 1980. The earliest studies found nerve endings in the outer most fibres of the anulus fibrosus but subsequent studies failed to confirm this. Malinsky's study[26] in 1959 was definitive and was later corroborated by others.[27-29] The outer third of the anulus fibrosus is consistently innervated from birth. Accordingly the intervertebral discs join the other deep, somatic tissues as having an innervation.

Animal Experiments

Animal experiments on deep, somatic pain lagged substantially behind clinical experiments. Consequently more was known sooner about the phenomenology of somatic pain and somatic referred pain than about its physiological mechanisms. The reasons for this lag are multiple. Foremost is probably fashion. For many reasons neurophysiologists focused their attention on the operation of nerves, synapses, muscle spindles and neu-

romuscular effectors. Leadership and expertise developed in these domains; and young scientists were more likely to pursue a career in one of these established disciplines than to enter a field that lacked leadership. Pain research became possible and attractive when individuals trained in other domains of neurophysiology turned their attention to pain. As a result, the concerted study of pain physiology did not commence until the 1970s. A second factor was technology. Pain is mediated by small diameter peripheral affected fibres and by small neurons in the spinal cord. These could not be studied until devices were developed that provided access to small neurons. Thirdly, once the incentive arose to study pain and once the necessary technology to do so became available, it was convenient and pragmatic to study cutaneous pain first. The study of deep, somatic pain followed.

Articular Nociception

The earliest electrophysiological studies of the cat knee joint, by Gardner[30] in 1950, and by Skoglund[31] in 1960, described the effect of stimulating articular nerves on reflexes evoked from the joint. The first concerted efforts to study joint nociception electrophysiologically were undertaken by Schaible and Schmidt, who performed preliminary work in the late 1970s, and published their first comprehensive study in 1983.[32,33] They showed that group III and group IV afferents could be activated by mechanical and by noxious stimuli.

Work undertaken since that time has conveniently been reviewed by Schaible and Grubb.[34] The following summarises the cardinal features of contemporary knowledge of articular nociception:

- The fibrous tissues of joints – periosteum, capsules, menisci and ligaments are well endowed with nerve endings.
- Earlier studies provided conflicting results concerning the innervation of synovium, but immunohistochemical studies have confirmed the presence of nerve fibres in this tissue.
- The articular nerves of joints consist of myelinated and unmyelinated fibres, the proportion differs in different nerves but the majority of fibres (ca 80%) are unmyelinated.
- A minority of myelinated fibres are group III fibres with free nerve endings.
- Half of the unmyelinated fibres are group IV fibres with free nerve endings.
- Fibres with free nerve endings exhibit a beaded structure suggestive of multiple transducer sites.
- Group III and group IV fibres exhibit a variety of response characteristics.
- Some are low threshold mechanoreceptors with respect to innocuous movements but also in a graded fashion to increasingly noxious strains of the joint.
- Some are weak low threshold mechanoreceptors that respond to innocuous movements but exhibit a graded response only to noxious stimuli.
- Some are high threshold mechanoreceptors that respond to extremes of movement and also in a graded fashion to noxious strains.
- Some fibres respond only to noxious pressures applied to the joint capsule.
- Some fibres respond only to chemical stimuli.
- Other fibres are silent under normal conditions but become sensitive in inflamed joints.
- Glutamate and substance P are the cardinal neurotransmitters of primary afferents from joints.
- Joint afferents project to lamina I, laminae V and VI, and the dorsal part of lamina VII of the dorsal horn.
- Spinal cord neurones responding to articular stimulation are located in laminae I, and IV-VIII of the dorsal horn.
- Joint afferents activate interneurones, motor neurones, and cells of the spinocerebellar and spinothalamic tracts.
- Second order neurones in the dorsal horn, responsive to joint afferents consist of nociceptive-specific and wide-dynamic-range neurones.
- The receptive fields of second order neurones innervated by joint afferents are under tonic descending inhibitory control.

Inflammation

Articular nerves not only mediate nociception from inflamed joints, they also contribute to the inflammation. In regard to the latter, the role of articular nerves can be regarded as nocifensive in that the nerves act to promote repair of an ostensibly damaged joint.

Inflammation affects articular afferents in different ways. Group II afferents are minimally affected, if at all.[34] In contrast, many but not all Group III and Group IV afferents are either activated or sensitised.[34] Low threshold mechanoreceptors respond more strongly; high threshold mechanoreceptors respond at lower thresholds; and silent nociceptors become active.

Articular afferents are sensitised by serotonin, PGE_2, PGI_2 and bradykinin.[34] Bradykinin is the most potent mediator and acts initially on B_2 receptors but subsequently on B_1 receptors which become upregulated in inflamed joints. The prostaglandins facilitate the effect of bradykinin. All these mediators are released from damaged tissue cells or inflammatory cells.

The inflammatory process is promoted by substance P, neurokinin A and CGRP, which are released from the peripheral terminals of articular nociceptors.[34] Substance P increases vascular permeability, and CGRP causes vasodilatation. These effects

The Physiology of Deep, Somatic Pain

are compounded and reinforced by noradrenaline and neuropeptide Y, that are released from sympathetic nerve terminals.[34] Collectively these processes constitute neurogenic inflammation, that is, sensory and sympathetic nerves contribute to mechanisms of inflammation.

Afferent input from inflamed joints also affects second-order neurones that subtend those joints.[34] The threshold for activation of nociceptive-specific neurones is lowered. The response of wide-dynamics-range neurones is increased. Receptive fields enlarge. More cells exhibit ongoing discharge; and descending inhibition is increased. This probably reflects a continued input from the inflamed areas.

Muscle Nociception

The first neurophysiological study of nociception from muscle was that of Paintal,[35] in 1960, who demonstrated that muscle nociception was mediated, at least in part, by group III afferents. He showed that these afferents could be activated by pressing or squeezing muscle fibres, and that they were activated by injections of hypertonic saline, like those used by Kellgren[2] to elicit pain from muscle in humans.

However, the study of muscle nociception remained relatively dormant for some 20 years. It was resurrected by Mense and Schmidt,[36] and sustained by Mense,[37,38] in the mid 1970s. These investigators showed that Group IV afferents from muscle could be activated by bradykinin, potassium, serotonin, and histamine. Progress since that time has been summarised in a review by Mense.[39]

The cardinal features of muscle nociception are:

- Muscles are innervated by nerves which differ in their composition, from muscle to muscle, but a typical profile would be 40% myelinated and 60% unmyelinated fibres.
- Of the myelinated fibres, 60% are motor.
- Of the unmyelinated fibres, about half are sympathetic efferents.
- Of the myelinated sensory fibres, 50% are Ia, Ib and Group II afferents; 20% are Group III afferents.
- About half of the unmyelinated fibres are Group IV afferents.
- Of the Group III and Group IV afferents, at least 40% are nociceptive.
- Most of the nociceptive afferents are activated by squeezing the muscle or by chemical stimulation with bradykinin, serotonin or potassium ions.
- Some nociceptive afferents are silent under normal conditions and become active only in damaged or inflamed muscles.
- Some nociceptive afferents exhibit ongoing activity in undistributed resting muscles.
- Muscle nociceptive afferents project to second-order neurones in lamina I and lamina V of the dorsal horn which project to the thalamus and hence to the cortex. (These pathways are not necessarily exclusive to muscle afferents, for they may involve convergence with cutaneous and other deep afferents.)
- Second-order neurones are subject to tonic descending inhibition.
- The peripheral terminals of muscle nociceptive afferents release substance P and CGRP.
- Muscle nociceptive afferents are sensitised by bradykinin, prostaglandins and serotonin; they are desensitised by LTD_4.
- Muscle pain is induced by trauma, inflammation or ischaemia of a muscle; each of these processes seems to involve the activation of muscle nociceptors by bradykinin, prostaglandins or potassium.
- There are no experimentally validated explanations of chronic muscle pain in the absence of inflammation.

Referred Pain

A mechanism for somatic referred pain has been demonstrated in multiple animal experiments. Both articular and muscle afferents exhibit convergence.[34,39] They synapse on second-order neurones that also receive an input from other deep somatic tissues and from skin. With respect to spinal referred pain, animal studies have revealed hyperconvergent neurones – ones that respond to stimulation of muscles, joints and intervertebral discs of the lumbar spine as well as the lower limbs.[40]

Accordingly, referred pain can be explained on the basis of perceptual ambiguity. When a dorsal horn neurone is stimulated by one of its convergent afferents, pain is evoked but the neurone does not convey information to the brain as to which of its afferents was the source of pain. At best, the cortex deduces that the source lies in one or other or all of the structures subtended by the activated neurone. The experience becomes one of pain throughout all of the structures rather than from a single, specific source.

Intriguing are recent studies of spinal pain mechanisms.[41] Noxious stimulation of intervertebral discs evokes reflex muscle activity in the paraspinal muscles. Distension of the zygapophysial joints inhibits this activity. These observations indicate that therapeutic interventions directed at one element in a vertebral motion segment can influence the effects of nociception arising from other elements in the same segment.

Muscle Spasm

Deep somatic referred pain can be associated with involuntary activity in muscles. How consistent this phenomenon is has not been determined, nor has the distribution of such activity been mapped for particular sites of noxious stimulation. Nevertheless, it

seems that muscle activity can occur in muscles adjacent to and remote from a site of noxious stimulation. In the case of lumbar spinal pain, activity can occur in paraspinal muscles, and in muscles of the lower limb girdle and lower limb. Such phenomena have been observed in human experiments[5,7,10,12] and reproduced in animal experiments.[39.42]

The teleological purpose of such activity has not been explained. It cannot be ascribed to guarding, for that does not explain activity in remote sites. Activity in gluteal and hamstring muscles does not serve to guard the lumbar spine in the same way as spasm of the abdominal muscles might be perceived to protect underlying viscera. There is no evidence to suggest that it is more than an epiphenomenon of local and referred pain.

Vexatious is the issue of whether this muscle activity is a secondary source of pain. There is no evidence that it is. More particularly, there is no evidence in support of a pain-muscle spasm-pain cycle, and some evidence against this concept. Although a popular concept in some clinical circles, it has not been substantiated experimentally, and recent reviews have dismissed it as invalid.[39,42]

Incomplete Explanations

Clinical experiments have shown that the joints and muscles can be sources of local and referred pain. Complementary animal studies have shown that joints and muscles have a nociceptive innervation that under normal conditions can be activated by excessive strains or pressure, or by chemical insults. Comprehensive models are available for pain produced by inflamed joints or muscles. What remain unexplained are the mechanisms of chronic pain from joints or muscles that are not inflamed.

Whereas injury and inflammation are adequate explanations for acute muscle pain, there is no satisfying or compelling model for chronic pain stemming from muscle, and no evidence, clinical or experimental, that it occurs. Even ischaemia in tonically active muscle has been challenged as an explanation of acute, let alone chronic, pain from muscle.[39]

With respect to joints, inflammation is clearly an acceptable explanation for the mechanism of pain in rheumatoid arthritis, in which features of inflammation are clinically obvious. However, the same does not pertain to osteoarthrosis. In that condition, inflammation is not consistently present. It is manifest in sudden "flares" or effusions,[43] but at other times the degree of inflammation varies between patients at different phases of the disease.[44.45] Although prominent in some cases of osteoarthrosis, inflammation is low grade or absent in others.[44,45]

Among the proffered, alternative explanations are capsular contracture and intraosseous venous hypertension.[45] Capsular contracture is a valid explanation for joint stiffness, and even of pain at the limits of available movement, but it does not explain pain at rest. On the other hand, intraosseous venous hypertension could.

The model proposed that as subchondral sclerosis occurs in osteoarthrosis, venous channels become obstructed, causing distension of veins proximal to the obstruction.[46] Stretch of the adventitia of these veins becomes the mechanism of nociception.

Testing this theory is difficult for it requires puncture of the putatively painful bone and manometric study of its intraosseous veins. Such studies of this nature that have been conducted reveal trends in favour of the model but insufficient differences to discriminate consistently between normal and painful joints.[46]

Another emerging but unexplored concept is that of subchondral bone pain. Histological studies have demonstrated nerves in the subchondral bone of synovial joints[47-49] and in the end plates of intervertebral discs.[50] This invites the proposition that if, as a result of age, injury or disease, the subchondral bone is weakened, it might undergo excessive strain under compression loading, which activates the subchondral nerves. Such a process would explain pain in weight-bearing joints that is relieved by rest, and may gain favour as an explanation of lumbar discogenic pain.[50]

These various concepts have one thing in common. They place the nociception of joint pain not in the fibrous tissues or synovium of the joint but within its bones. Proof or refutation of either the concept of intraosseous venous hypertension or of subchondral bone pain awaits the next technological advance in the study of deep, somatic pain: the ability to study the neurophysiology of nerves inside, or innervating, bones.

References

1. Kellgren JH. Observations on referred pain arising form muscle. *Clin Sci* 1938; 3: 175-90.
2. Kellgren JH. On the distribution of pain arising from deep somatic structures with charts of segmental pain areas. *Clin Sci* 1939; 4: 35-46.
3. Sinclair DC, Feindel WH, Weddell G, Falconer MA. The intervertebral ligaments as a source of segmental pain. *J Bone Joint Surg* 1948; 30B: 515-21.
4. Sinclair DC, Weddell G, Feindel WH. Referred pain and associated phenomena. *Brain* 1948; 71: 184-211.
5. Hockaday JM, Whitty CWM. Patterns of referred pain in the normal subject. *Brain* 1967; 90: 481-96.
6. Whitty CWM, Willison RG. Some aspects of referred pain. *Lancet* 1958; 2: 226-31.
7. Feinstein B, Langton JNK, Jameson RM, Schiller F. Experiments on pain referred from deep somatic tissues. *J Bone Joint Surg* 1954; 35A: 981-87.
8. Inman VT, Saunders JBD. Referred pain from skeletal structure. *J Nerv*

The Physiology of Deep, Somatic Pain

Ment Dis 1944; 99: 660-67.
9. Kellgren JH, Samuel EP. The sensitivity and innervation of the articular capsule. *J Bone Joint Surg* 1950; 32B: 84-92.
10. Mooney V, Robertson J. The facet syndrome. *Clin Orthop* 1976; 115: 149-56.
11. McCall IW, Park WM, O'Brien JP. Induced pain referred from posterior lumbar elements in normal subjects. *Spine* 1979; 4: 441-46.
12. Bogduk N. Lumbar dorsal ramus syndrome. *Med J Aust* 1980; 2: 537-41.
13. Dwyer A, Aprill C, Bogduk N. Cervical zygapophyseal joint pain patterns I: a study in normal volunteers. *Spine* 1990; 15: 453-57.
14. Aprill C, Dwyer A, Bogduk N. Cervical zygapophyseal joint pain patterns II: a clinical evaluation. *Spine* 1990; 15: 458-61.
15. Fortin JD, Dwyer AP, West S, Pier J. Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique: Part I: asymptomatic volunteers. *Spine* 1994; 19:1475-82.
16. Fukui S, Ohseto K, Shiotani M, Ohno K, Karasawa H, Nagaauma Y, Yuda Y. Referred pain distribution of the cervical zygapophyseal joints and cervical dorsal rami. *Pain* 1996; 68: 79-83.
17. Fukui S, Ohseto K, Shiotani M, Ohno K, Karasawa H, Nagaauma Y. Distribution of referred pain from the lumbar zygapophyseal joints and dorsal rami. *Clin J* 1997; 13: 303-307.
18. Weddell G, Harpman JA. The neurohistological basis for the sensation of pain provoked from deep fascia, tendon, and periosteum. *J Neurol Psychiat* 1940; 3: 319-28.
19. Stillwell DL. The innervation of tendons and aponeuroses. *Am J Anat* 1957; 100: 289-318.
20. Ralston HJ, Miller MR, Kasahara M. Nerve endings in human fasciae, tendons, ligaments, periosteum, and joint synovial membrane. *Anat Rec* 1960; 136:137-47.
21. Jackson HC, Winklemann RK, Bickel WH. Nerve endings in the human lumbar spinal column and related structures. *J Bone Joint Surg* 1966; 48A: 1272-81.
22. Ashton I K, Ashton B A, Gibson S J, et al. Morphological basis for back pain: the demonstration of nerve fibres and neuropeptides in the lumbar facet joint capsule but not in ligamentum flavum. *J Orthop Res* 1992; 10: 72-78.
23. Jiang H, Russell G, Raso J, et al. The nature and distribution of the innervation of human supraspinal and interspinal ligaments. *Spine* 1995; 20: 869-76.
24. Rhalmi S, Yahia L, Newman N, Isler M. Immunohistochemical study of nerves in lumbar spine ligaments. *Spine* 1993; 18: 264-67.
25. Yahia L H, Newman N, Rivard C H. Neurohistology of lumbar spine ligaments. *Acta Orthop Scandinav* 1988; 59: 508-12.
26. Malinsky J. The ontogenetic development of nerve terminations in the intervertebral discs of man. *Acta Anat* 1959; 38: 96-113.
27. Yoshizawa H, O'Brien JP, Thomas-Smith W, Trumper M. The neuropathology of intervertebral discs removed for low-back pain. *J Path* 1980; 132: 95-104.
28. Ashton IK, Walsh DA, Polak JM, Eisenstein SM. Substance P in intervertebral discs: binding sites on vascular endothelium of the human annulus fibrosus. *Acta Orthop Scandinav* 1994; 65: 635-39.
29. Roberts S, Eisenstein SM, Menage J, et al. Mechanoreceptors in intervertebral discs: morphology, distribution, and neuropeptides. *Spine* 1995; 20: 2645-51.
30. Gardner E. Reflex muscle responses to stimulation of articular nerves in the cat. *Am J Physiol* 1950; 161:133.
31. Skoglund S. Anatomical and physiological studies of knee joint innervation in the cat. *Acta Physiol Scand* 1956; Sup 124: 1-101.
32. Schaible HG, Schmidt RF. Activation of group III and IV sensory units in medial articular nerve by local mechanical stimulation of knee joint. *J Neurophysiol* 1983; 49: 35-44.
33. Schaible HG, Schmidt RF. Responses of fine medial articular nerve afferents to passive movements of knee joint. *J Neurophysiol* 1983; 49: 1118-26.
34. Schaible HG, Grubb BD. Afferent and spinal mechanisms of joint pain. *Pain* 1993; 55: 5-54.
35. Paintal AS. Functional analysis of group III afferent fibres of mammalian muscles. *J Physiol* 1960; 152: 250-70.
36. Mense S, Schmidt RF. Activation of group IV afferent units from muscle by algesic agents. *Brain Res* 1974; 72: 305-10.
37. Franz M, Mense S. Muscle receptors with group IV afferent fibres responding to application of bradykinin. *Brain Res* 1975; 92: 369-83.
38. Fock S, Mense S. Excitatory effects of 5-hydroxytryptamine, histamine and potassium ions on muscular group IV afferent units: a comparison with bradykinin. *Brain Res* 1976; 105: 459-69.
39. Mense S. Nociception from skeletal muscle in relation to clinical muscle pain. *Pain* 1993; 54: 241-89.
40. Gillette RG, Kramis RC, Roberts WJ. Characterization of spinal somatosensory neurons having receptive fields in lumbar tissues of cats. *Pain* 1993; 54: 85-98.
41. Indahl A, Kaigle AM, Reikeras O, Holm SH. Interaction between the porcine lumbar intervertebral disc, zygapophysial joints, and paraspinal muscles. *Spine* 1997; 22: 2834-40.
42. Simons D, Mense DG. Understanding and measurement of muscle tone as related to clinical muscle pain. *Pain* 1998; 75: 1-17.
43. Peyron J. Inflammation in osteoarthrosis (OA): review of its role in clinical picture, disease progress, subsets, and pathophysiology. *Sem Arthritis Rheum* 1981; 11(Supp1): 115-

16.
44. Creamer P, Hunt M, Dieppe P. Pain mechanisms in osteoarthritis of the knee: effect of intraarticular anesthetic. *J Rheumatol* 1996; 23: 1031-36.
45. Dieppe P, Lim K. Osteoarthritis and related disorders. Clinical features and diagnostic problems. In: Klippel JH, Dieppe PA (eds). *Rheumatology*, 2nd edn. London: Mosby, 1998, pp 8.3.1-8.3.16.
46. Arnoldi CC. Vascular aspects of degenerative joint disorders: a synthesis. *Acta Orthop Scand* 1994; Supp 261: 1-82.
47. Beaman DN, Graziano GP, Glover RA, et al. Substance P innervation of lumbar spine facet joints. *Spine* 1993;18: 1044-49.
48. Badalamente MA, Cherney SB. Periosteal and vascular innervation of the human patella in degenerative joint disease. *Sem Arthritis Rheum* 1989; 18 (Supp 2); 61-66.
49. Wojtys EM, Beaman DN, Glover RA, Janda D. Innervation of the human knee joint by substance P fibres. *Arthroscopy* 1990; 6: 254-63.
50.Brown MF, Hukkanen MVJ, McCarthy ID, et al. Sensory and sympathetic innervation of the vertebral endplate in patients with degenerative disc disease. *J Bone Joint Surg* 1997; 79B: 147-53.

Discussion Topics

What criteria would have to be satisfied for you to credit that a patient's pain is arising from:

a muscle?
a ligament?
a joint?
a bone?

What clinical evidence (a) is available, (b) might be pursued, that is consistent with the proposition that the pain of osteoarthrosis is due to intraosseous venous hypertension? Does any of this evidence prove the mechanism? What would you consider to be the most definitive data that should be obtained?

Explain why resurfacing an osteoarthritic knee so promptly relieves pain.

The Interobserver Reliability of Thoracic Spinal Examination

Dr Michael J Yelland MB BS, FRACGP, FAFMM, Grad Dip Musculoskeletal Medicine – Senior Lecturer in General Practice, University of Queensland, Brisbane, Australia
Professor Paul P Glasziou MB BS, PhD, FAFPHM – Professor of Evidence Based Health Care, University of Queensland, Brisbane, Australia
Ms Jennifer Purdie BA, M Med Sci – Manager, Statistical Services, Queensland Police Service, Brisbane, Australia

Abstract

The interobserver reliability of thoracic spinal examination was tested in a general practice referral clinic on 97 subjects with any combination of pain in the back in the back, chest, or abdomen or either with no pain. Subjects were examined by a general practitioner and a physiotherapist without knowledge of their history. Kappa coefficients indicated agreement that was generally fair for kyphosis, scoliosis, movement restriction, pain with gross active movements and pain with overpressure at end range. Agreement was moderate for point tenderness, moderate to substantial for regional tenderness and slight for maximal tenderness within a region. Spearman's coefficients for pressure threshold readings for these sites of maximal tenderness showed substantial agreement. Based on examination findings alone, agreement about intervertebral dysfunction was moderate for its presence, fair for its side and exact region and slight for its exact level.

The results question the utility of some thoracic spinal signs and demonstrate that the cost of anatomical precision in assessment is decreased reliability.

Introduction

In the assessment of spinal pain, where investigations rarely provide definitive results, physical examination plays a central role in refining the provisional diagnosis made on history. First, it continues the diagnostic process of excluding red-flag conditions. Then it focuses on detecting areas of intervertebral dysfunction (defined as "reversible, benign, painful, segmental vertebral dysfunction of mechanical and reflex origin")[1] and defining them as precisely as possible in the hope that precision diagnosis will lead to precision treatment, and hence to better results. However, for physical examination to be useful, its components must have both good reliability and validity. There are very few data on either of these properties of physical examination of the thoracic spine. In particular, the reliability of tests presented in standard textbooks on the thoracic spine,[2,3] including inspection for deformities, gross movement testing for pain and restriction and palpation of structures for tenderness, have not been rigorously tested. Whilst pain maps have been drawn, based on irritant injections of thoracic spinal segments in normal volunteers,[4,5] there are no validity studies comparing the thoracic spinal signs of intervertebral dysfunction with a criterion standard such as selective anaesthetisation of putative vertebral structures. This study aims to address the gap in the literature on interobserver reliability by testing a number of signs commonly used in the examination of the thoracic spine and to estimate the level of precision which is possible in examination and diagnosis before reliability is lost.

Methods

Setting and investigators

The study was conducted from November 1996 to December 1997 in the Inala Community Health Centre, a multidisciplinary health service in Brisbane, Australia. It compared the observations of a general practitioner with a diploma and nine years' experience in musculoskeletal medicine with those of one of two physiotherapists of 25 and 20 years' experience respectively.

Recruitment

The selection criteria for this study were determined by a parallel study which compared the prevalence of thoracic spinal signs in pain-free controls with subjects who had had any combination of pain in the back, chest or abdomen in the preceding month. As the parallel study was focusing on non-visceral causes of pain, inclusion criteria included a negative ECG for those with chest pain and an endoscopy for those with abdominal pain. These criteria also aimed to avoid the problem of prevalence bias in the estimation of interobserver reliability by providing a mixed sample of subjects with and without thoracic spinal signs.

Back pain was defined as any combination of thoracic spinal pain and/or lumbar spinal pain as defined by the International Association for the Study of Pain.[6] Subjects could have pain anywhere between the cervicothoracic junction and the iliac crests as it has been shown that somatic referred pain from thoracic spinal segments can be referred to anywhere in this region.[5,7] The age range of subjects was limited to 20-75 years.

Subjects with pain were referred by general practitioners within and near the examining doctor's practice. Subjects were excluded if they had been seen by the examining doctor for back, chest or abdominal pain in the preceding six months. Pain in multiple sites was not an exclusion criterion. Control subjects were chosen and contacted using a stratified random sample from the general practice patient register in the health centre in which the principal investigator worked. The proportion of control subjects in the total sample was approximately one-quarter to reduce the chance of prevalence bias in the calculation of interobserver reliability.

Ethics approval was obtained from the University of Queensland clinical

research ethics committee. Informed consent was obtained from all subjects prior to their participation in the study.

Examination Procedure

The subjects were examined independently by the general practitioner and one physiotherapist in the same morning session, the order of examiners distributed to reduce any fatigue-related bias in subject signs. Two physiotherapists were required as one was not always available for all the subjects throughout the study period.

The examination and very detailed recording of results took 30-45 minutes. Subjects were instructed not to reveal their history until after the examination was completed. This requirement allowed the analysis of the reliability of physical signs unbiased by history and also reduced diagnostic bias in the parallel study on the prevalence of thoracic spinal signs in back, chest and abdominal pain. Subjects were instructed to speak only about matters relating directly to the examination, e.g., "It hurts now when you press on that point".

In recording findings, the following definitions were used:

- Upper thoracic spine – related to the 1st to 4th thoracic vertebrae
- Mid thoracic spine – related to the 5th to 8th thoracic vertebra
- Lower thoracic spine – related to the 9th to 12th thoracic vertebrae.

Tests included in the examination were as follows:

- Inspection of spinal posture in standing with the arms by the side to assess the degree of kyphosis and the presence of any scoliosis in the upper, middle and lower thoracic spine. The degree of kyphosis was recorded as normal, increased or decreased. Scoliosis was described as nil, convex left or convex right.
- Gross active thoracic spinal movement restriction, described as nil (normal), mild (≤ 1/3 restriction), moderate (>1/3 restriction but ≤ 2/3 restriction) and major (> 2/3 to full restriction). Allowance for age-related reduction in normal range of motion was made by examiners during the study. Segmental movements were not tested.
- Pain on gross thoracic spinal movements performed actively and then with overpressure by the examiner at end of range.

Palpation for tenderness over each thoracic costotransverse (CT) joints and zygapophyseal (Z) joint (and associated paraspinal musculature) and over each spinous process in a posteroanterior direction. Each thoracic spinous process was also palpated for tenderness transversely to the left and the right. The landmark for the commencement of palpation was the T1 spinous process, located by finding the lower one of the two prominent spinous processes at the base of the neck.

Palpation for maximally tender points by region. For this purpose the thoracic spine was divided into the following nine regions:

- CT and Z joints on each side of the upper, mid and lower thoracic spine (eight points in each of six regions);
- Upper, mid and lower thoracic spinous processes (four points in each of three regions).

The pressure threshold of the most tender point in each of these regions was recorded using a hand-held pressure threshold meter (Pain Diagnostics and Thermography, New York), which has an interobserver repeatability coefficient of 0.65 to 0.87.[8] Subjects were instructed to report when pain was first felt at the point being tested.

Examiners were instructed to record the pressure threshold of the most tender point with posteroanterior palpation within each of these regions. If no tender point was found within a region, the default point for pressure threshold measurement was used to allow interobserver comparisons for all regions. The default points were the Z joint at T2, T6 and T10 in the CT joint/Z joint regions and the spinous process at T2, T6 and T10 in the spinous process regions.

At the end of the examination the examiner made a diagnosis of intervertebral dysfunction, no intervertebral dysfunction or generalised tenderness. This was based on a synthesis of gross movement restriction, the site(s) of pain with gross movement and the localised site(s) of tenderness on palpation. Where a diagnosis of intervertebral dysfunction was made, the involved segment or segments were recorded. A diagnosis of generalised tenderness was made when tenderness was found in most regions and was poorly localised. A history was taken only after a physical examination diagnosis had been made. The influence, if any, of the history on the diagnosis was then recorded.

Inter-examiner standardisation of examination techniques and of the definition of normal and abnormal findings occurred over a four hour period prior to the commencement of the study. Symptomatic patients were examined jointly in this period. This timeframe was chosen to simulate the amount of peer review that might take place at a postgraduate conference workshop. Further comparison of examination techniques was not permitted during the 12-month study period.

Statistical Analyses

The analysis treated the two physiotherapists as one examiner, as our principal interest was in the degree of agreement between two examinations of the one patient (one by a general practitioner and the other by one of two physiotherapists). The interobserver reliability for each grouping of thoracic spinal signs was calculated using SAS

The Interobserver Reliability of Thoracic Spinal Examination

version 6.12. The difference between observed and expected agreement was expressed using the Kappa coefficient for binary and nominal categorical variables,[9] the weighted Kappa coefficient for ordinal variables[10] and the Spearman's correlation coefficient for continuous variables.[11] The 95% confidence intervals were calculated for all results and are shown in brackets after them.

For ordinal variables, weighted Kappa gives partial credit for responses, which are similar but not in exact agreement. Discordant responses were weighted using the absolute value of the deviation from exact agreement.

Descriptive terms used in this analysis for the various ranges of Kappa and Spearman's correlation coefficients are those proposed by Landis and Koch[12]: <0.00, poor; 0.00-0.20, slight; 0.21-0.40, fair; 0.41- 0.60, moderate; 0.61-0.80, substantial; 0.81-1.00 almost perfect.

Results

Ninety-seven subjects satisfied the inclusion criteria of the study and were examined by both examiners. Their demographic details and categorisation by symptoms appear in Table 1.

The Kappa coefficients for agreement on posture, movement restriction, pain on active movement and pain with overpressure are displayed in Table 2. Kappa coefficients for kyphosis ranged from 0.06 (-0.10, 0.23) to 0.51 (0.35, 0.67) for kyphosis and 0.14 (-0.05, 0.33) to 0.26 (-0.18, 0.70) for scoliosis. Agreement for these assessments was lowest in the lower thoracic spine and highest in the upper thoracic spine. Pain with assessment of movements yielded Kappa coefficients of 0.22 (0.07, 0.37) to 0.39 (0.25, 0.54) for movement restriction, 0.17 (-0.06, 0.39) to 0.47 (0.27, 0.66) for pain with active movements and 0.21 (0.01, 0.41) to 0.63 (0.47, 0.79) for pain with overpressure at end range.

Levels of agreement on regional palpation for tenderness and its precise site within each region are displayed in Table 3. Kappa coefficients for the presence/absence of tenderness in each of the nine regions fell in the range 0.41 (0.22, 0.60) to 0.74 (0.59, 0.89). Where there was agreement that regional tenderness was present, the Kappa coefficients for the precise site of maximal tenderness ranged from of –0.04 (-0.23, 0.16) to 0.31 (0.12, 0.51).

The descriptive levels of agreement for palpation of specific structures are summarised in Table 4. The most common level of agreement for all structures was moderate. Palpation of Z joints had the highest overall level of agreement and CT joint palpation the lowest; however, there was no significant difference in these levels of agreement (Chi-squared = 10.98, p = 0.09).

The Spearman's correlation coefficients for the pressure threshold values of these maximally tender points in each region varied from 0.50 (0.33, 0.67) to 0.82 (0.74, 0.89) (Table 5).

Characteristic	N	%
Sex		
Male	43	44.3
Female	54	55.7
Age		
20 – 29	5	5.2
30 – 39	14	14.4
40 – 49	21	21.6
50 – 59	22	22.7
60 – 69	32	33.0
70 – 75	3	3.1
Pain Category		
Control	24	24.7
Thoracic spine	25	25.8
Chest	18	18.6
Abdomen	6	6.2
Thoracic spine and chest	9	9.3
Thoracic spine and abdomen	8	8.2
Chest and abdomen	5	5.2
Thoracic spine, chest and abdomen	2	2.1
Distribution of order of examinations		
General Practitioner (total)	97	100.0
As 1st examiner	43	44.3
As 2nd examiner	54	55.7
Physiotherapist A (total)	48	49.5
As 1st examiner	25	25.8
As 2nd examiner	23	23.7
Physiotherapist B (total)	49	50.5
As 1st examiner	29	29.9
As 2nd examiner	20	20.6
Physical Examination Diagnoses		
General Practitioner		
Intervertebral dysfunction	57	59.4
No intervertebral dysfunction	28	34.4
Generalised tenderness	11	6.3
Physiotherapists (A & B)*		
Intervertebral dysfunction	57	59.4
No intervertebral dysfunction	33	29.2
Generalised tenderness	6	11.5

*1 missing value

Table 1. Characteristics of the 97 subjects examined in the study

The distribution of diagnoses based on physical examination findings *alone* appears in Table 1. For the basic categorisation into intervertebral dysfunction, no dysfunction or generalised tenderness, the Kappa coefficient was 0.52 (0.36, 0.67). The Kappa coefficient for the side of dysfunction was 0.31 (0.17, 0.44). The Kappa coefficient for the region of dysfunction on the left side was 0.21 (–0.08, 0.34) and on the right side was 0.24 (–0.12, 0.36).

Sign	Number of Observations	Kappa coefficient	95% CI
Kyphosis			
Upper	96	0.51	0.35, 0.67
Middle	96	0.32	0.15, 0.49
Lower	94	0.06	-0.10, 0.23
Scoliosis			
Upper	89	0.26	-0.18, 0.70
Middle	94	0.26	0.03, 0.47
Lower	95	0.14	-0.05, 0.33
Movement restriction			
Flexion	97	0.37	0.17, 0.57
Extension	96	0.22	0.07, 0.37
Sidebending left	93	0.34	0.14, 0.53
Sidebending right	95	0.39	0.25, 0.54
Rotation left	96	0.36	0.19, 0.53
Rotation right	96	0.22	0.04, 0.41
Pain with active movement			
Flexion	97	0.17	-0.06, 0.39
Extension	97	0.36	0.17, 0.55
Sidebending left	97	0.34	0.12, 0.56
Sidebending right	97	0.34	0.14, 0.54
Rotation left	97	0.47	0.27, 0.66
Rotation right	97	0.37	0.18, 0.55
Pain with overpressure			
Flexion	97	0.27	0.07, 0.46
Extension	96	0.30	0.10, 0.50
Sidebending left	97	0.38	0.18, 0.58
Sidebending right	97	0.21	0.01, 0.41
Rotation left	96	0.63	0.47, 0.79
Rotation right	97	0.37	0.19, 0.56

Table 2. Kappa coefficients for the agreement on posture, movement restriction, pain on active movement and pain on overpressure in 97 subjects. Missing observations in some cells due to omissions in recording.

Sign	Number of Observations	Kappa coefficient	95% CI
Presence of Regional Tenderness			
Spinous process			
Upper	97	0.63	0.46, 0.79
Middle	97	0.41	0.22, 0.60
Lower	97	0.53	0.36, 0.70
CT & Z joints			
Upper left	97	0.61	0.43, 0.79
Middle left	97	0.59	0.42, 0.76
Lower left	97	0.54	0.37, 0.71
Upper right	97	0.74	0.59, 0.89
Middle right	97	0.53	0.35, 0.70
Lower right	97	0.60	0.44, 0.76
Site of Maximal Tenderness in Region			
Spinous process			
Upper	57	0.31	0.12, 0.51
Middle	51	0.10	-0.11, 0.30
Lower	34	0.14	-0.10, 0.39
CT & Z joints			
Upper left	67	0.21	0.05, 0.38
Middle left	55	0.29	0.13, 0.45
Lower left	43	-0.04	-0.23, 0.16
Upper right	34	0.14	-0.10, 0.36
Middle right	48	0.12	-0.05, 0.30
Lower right	47	0.09	-0.05, 0.23

Table 3. Kappa coefficients for the agreement on the presence of regional tenderness and, in regions where tenderness was found by both examiners, weighted Kappa coefficients for agreement on the site of maximal tenderness.

Figure 1 illustrates the Kappa coefficients for agreement about intervertebral dysfunction at each thoracic spinal segment. These ranged from –0.04 (-0.08, 0.00) to 0.44 (0.18, 0.70) and were lowest in the mid thoracic spine.

After having taken a history, the general practitioner stated he would have changed his diagnosis in 25% of cases and the physiotherapists would have changed their diagnoses in 19% of cases.

Discussion

There is little literature on the reliability of signs in thoracic spinal pain, and that which does exists offers only an indirect comparison with the results of this study. A small study examining the interexaminer reliability of the assessment of cervicothoracic and shoulder posture[13] reported a Kappa coefficient of 0.611. In the present study, agreement for assessment of kyphosis and scoliosis varied according to the region in the thoracic spine, but on average was only fair. This may have been reduced by the degree of precision expected of examiners, i.e., the detection of any minor variation from normal within each of the three thoracic spinal regions, rather than the thoracic spine as a whole.

Agreement on restriction and pain with thoracic spinal movements was generally fair. Such data as are available relate to the excursion of the trunk as a whole during movements of the lumbar spine, not to the movement range of the thoracic spine. At best they offer a surrogate measure for comparison. They show Kappa scores of 0.10 to 0.58 for agreement between two surgeons, and a surgeon and a physiotherapist assessing pain on lumbar spinal movements.[14]

There was generally a moderate level of agreement for assessment of tenderness by manual palpation of specific structures. There was moderate to substantial agreement about the presence of regional tenderness, but

The Interobserver Reliability of Thoracic Spinal Examination

Structure	No. of structures	Level of Agreement					
		Poor	Slight	Fair	Moderate	Substantial	Almost perfect
CT joint (posteroanterior)	24	0	0	9	14	1	0
Z joint (posteroanterior)	24	0	0	1	20	3	0
Spinous process (posteroanterior)	12	0	0	2	8	2	0
Spinous process (transverse)	24	0	0	3	19	2	0

Table 4. Summary of the descriptive levels of agreement (based on Kappa coefficients) for posteroanterior and transverse palpation of specific structures in the thoracic spine for tenderness.

	Left	Spinous Process	Right
Upper	0.68 (0.57, 0.79)	0.82 (0.74, 0.89)	0.68 (0.56, 0.80)
Mid	0.50 (0.33, 0.67)	0.70 (0.57, 0.82)	0.72 (0.62, 0.82)
Lower	0.60 (0.46, 0.74)	0.74 (0.63, 0.85)	0.60 (0.47, 0.74)

Table 5. Spearman's correlation coefficients (95% CI) for the hand-held pressure threshold meter values of the designated maximally tender (or default) point within each region.

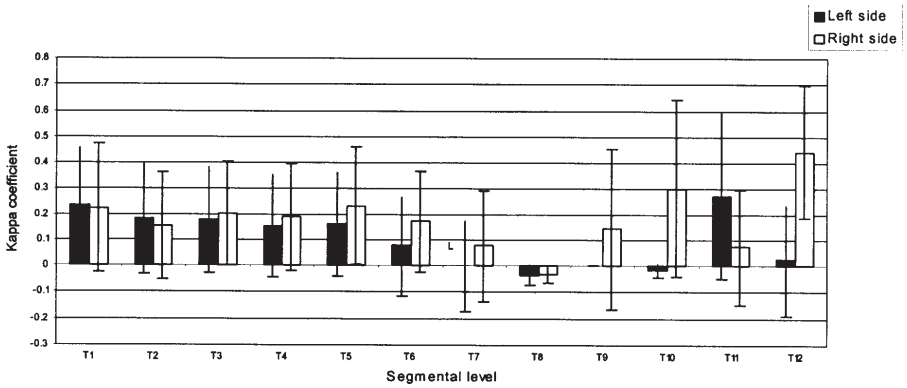


Figure 1. Kappa coefficients (and 95% CI) for agreement on interveterbral dysfunction for each spinal level and side in 97 subjects. Result absent for left T9 level due to unstable table.

only slight agreement about the exact site of maximal tenderness in these regions. Comparative data in the thoracic spine relate only to tenderness in upper trapezius and levator scapulae muscles where a fair to moderate degree of interobserver reliability has been reported.[8] Substantial levels of agreement have been documented for tenderness of lumbar spinous processes (intraclass correlation coefficient of 0.67-0.72)[15] and for cervical spinal joints (Kappa = 0.68).[16] One factor contributing to this is the level of agreement on which spinal level is being palpated. This was not tested here, but in the lumbar spine, a Kappa score of only 0.28 has been reported for agreement on lumbar spinal levels between physiotherapists with similar training to those in this study.[17]

For the localisation of tenderness, as the greater the anatomical precision required, the worse the reliability became. Different examiners could usually agree on whether or not a point or region was tender but showed less agreement about the site of maximal tenderness within a region. More reliable and useful was quantification of the maximally tender point with a pressure threshold meter.

The reliability of diagnosis, based on examination without history, suffered with attempts to increase precision, making diagnosis questionable at anything more than a regional level. Agreement about intervertebral dysfunction was generally slight at a segmental level. In another study of physical signs alone, a fair level of agreement (Kappa = 0.31) has been reported for the diagnosis of thoracic spinal dysfunction based on the presence of deep muscular tension as assessed by four osteopathically trained students examining 15 subjects.[18]

Overall, the results cast doubt on the utility of much of the physical examination of the thoracic spine, but they do not identify the source of poor interexaminer reliability. Kappa coefficients can be paradoxically low despite good agreement if the prevalence of the condition in question is too low or too high.[19-21] A fairer estimate of the representative Kappa coefficient is obtained if the sign is present in about 50% of the study sample. Conversely, Kappa coefficients can be paradoxically elevated when there is a large bias between observers in the number of positive signs found.[19] Analysis of these two paradoxes in this sample, using the method described by Byrt et al[19] gives adjusted Kappa coefficients, which are higher for all but five of the 27 values in Tables 2 and 3. Lower values pertain for assessment of kyphosis in the upper thoracic spine and restriction of active extension. The values for pain on active rotation and tenderness in the lower central and lower left regions remain unchanged. For the diagnosis of intervertebral dysfunction (Figure 1), the same analysis of the effects of prevalence and bias gives higher adjusted values for all 72 Kappa coefficients, principally due to the low prevalence of intervertebral dysfunction at each segmental level.

It could be argued that the study design, which forbade history taking prior to examination, might have re-

duced the levels of agreement, particularly for the diagnosis of dysfunction. The assumption that history taking would have improved the reliability of physical examination lacks substance as history taking styles and subjects' responses to history questions are prone to an unknown level of variation themselves. Our study design allowed an estimate of the reliability of physical signs unbiased by history. This design has been used in many other studies on the reliability of spinal examination.[8,13,17,18,22]

It is possible that the poor agreement for some signs might have resulted from the limited time allocated for comparison of techniques between examiners. This study was designed to simulate the agreement attained by experienced practitioners who have spent the equivalent time of a short postgraduate workshop together comparing techniques for one area of the spine. It reflects the divergence in physical examination techniques which may occur in clinical practice over time, in this case over the ensuing year. A comparison of examiners at the conclusion of the study showed a considerable divergence in techniques and definitions of abnormalities. This was most apparent in mean digital palpation pressures, being 3.2 kg for the doctor, 1.4 kg for one physiotherapist and 6.9 kg for the other physiotherapist. In addition, pressure threshold readings were consistently lower for the doctor than the physiotherapists, averaging 5.2 kg/cm^2 and 6.3 kg/cm^2 respectively.

The available evidence that training examiners to use the same examination techniques can improve reliability comes from lumbar spinal research and is, at best, indirect. It has been shown that the ability of physical therapists to produce specified forces of palpation in the lumbar spine can be improved with daily practice using bathroom scales.[23] Detection of nominated lumbar spinal levels has been shown to be only fair in physiotherapists prior to postgraduate training in manipulative physiotherapy,[17] but almost perfect for a group of physiotherapists with postgraduate diplomas in manipulative physiotherapy.[24] A further reliability study is needed to explore the potential reliability of thoracic spinal examination with a significant period of training and peer review. Signs which show poor reliability after such a process should be abandoned.

Conclusions

The results of this study suggest that there are many thoracic spinal signs which are of questionable reliability and, hence, utility. These include postural assessment, movement restriction and pain with movement. More reliable signs include palpation for regional tenderness and quantification of pressure threshold values of maximally tender points. Reliability is lost when more precise localisation of tender points is required. In the absence of history, diagnosis of thoracic intervertebral dysfunction based on physical examination alone is unreliable.

Acknowledgements

The contribution of the physiotherapists Ms Joie Sumby and Mrs Anne Mitchell to the collection of data, Mrs Anna Nichols to recruitment of control subjects and of Mr Robert Kerr to data entry are gratefully acknowledged.

References

1. Maigne R. *Diagnosis and treatment of pain of vertebral origin.* 1st ed. Paris: Williams and Wilkins, 1996, pp. 99-108.
2. Corrigan B, Maitland G D. Practical and Orthopaedic Medicine. London: Butterworths, 1983, pp. 378-86.
3. Kenna C, Murtagh J. *Back Pain and Spinal Manipulation - A Practical Guide*. Sydney: Butterworths, 1989, pp. 176-85.
4. Dreyfuss P, Tibiletti C, Dreyer SJ. Thoracic zygapophyseal joint pain patterns – a study in normal volunteers. *Spine* 1994; 19: 807-11.
5. Feinstein B, Langton JNK, Jameson RM, Schiller F. Experiments of pain referred from deep somatic tissues. *J Bone Joint Surg Am* 1954; 36-A: 981-97.
6. Merskey H, Bogduk N, eds. *Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms*. Second ed. Seattle: IASP Press, 1994, pp. 112-20.
7. Fukui S, Ohseto K, Shiotani M. Patterns of pain induced by distending the thoracic zygapophyseal joints. *Regional Anaesthesia* 1997; 22: 332-36.
8. Levoska S, Keinänen-Kiukaanniemi S, Bloigu R. Repeatability of measurement of tenderness in the neck-shoulder region by a dolorimeter and manual palpation. *Clin J Pain* 1993; 9: 229-35.
9. Cohen J. A coefficient of agreement for nominal scales. *Psychol Measmt* 1960; 20: 37-46.
10. Cohen J. Weighted Kappa: nominal scale agreement with provision for scaled agreement or partial credit. *Psychol Bull* 1968; 70: 213-20.
11. Armitage P, Berry G. *Statistical Methods in Medical Research.* Third ed. Oxford: Blackwell Scientific Publications, 1994,pp. 443-67.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74.
13. Griegel-Morris P, Larson K, Mueller-Klaus K, Oatis CA. Incidence of common postural abnormalities in the cervical, shoulder, and thoracic regions and their association with pain in two age groups of healthy subjects. *Phys Ther* 1992; 72: 425-30.
14. McCombe PF, Fairbank JC, Cockersole BC, Pynsent PB. Reproducibility of physical signs in low-back pain. *Spine* 1989; 14: 908-18.
15. Maher C, Adams R. Reliability of

The Interobserver Reliability of Thoracic Spinal Examination

pain and stiffness assessments in clinical manual lumbar spine examination. *Phys Ther* 1994; 74: 801-11.
16. Hubka MJ, Phelan SP. Interexaminer reliability of palpation for cervical spine tenderness. *J Manip Physiol Ther* 1994; 17: 591-95.
17. Kenzie AM, Taylor NF. Can physiotherapists reliably locate lumbar spinal levels by palpation? *Physiother* 1997; 83: 235-39.
18. Johnston WL, Allan BR, Hendra JL, et al. Interexaminer study of palpation in detecting location of spinal segmental dysfunction. *JAOA* 1983; 82: 839-45.
19. Byrt T, Bishop J, Carlin JB. Bias, prevalence and Kappa. *J Clin Epidemiol* 1993; 46: 423-29.
20. Cicchetti DV, Feinstein AR. High agreement but low Kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990; 43: 551-58.
21. Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; 43: 543-49.
22. Fjellner A, Bexander C, Faleij R, Strender LE. Interexaminer reliability in physical examination of the cervical spine. *J Manip and Physiolog Therapeut* 1999; 22: 511-16.
23. Keating J, Matyas TA, Bach TM. The effect of training of physical therapist's ability to apply specified forces of palpation. *Phys Ther* 1993; 73: 38-46.
24. Downey J, Taylor NF, Niere KR. Manipulative physiotherapists can reliably palpate nominated lumbar spinal levels. *Man Ther*, 1999; 4(3): 151-56.

Quiz: The Knee

by Peter Watson, Physiotherapist

Q1. A Bakers cyst is located in ..... The typical signs and symptoms are ..... The underlying pathology is .....

Q2. The patello femoral joint may also refer pain to the popliteal region – T/F

Q3. Aspiration of a Bakers cyst is usually curative – T/F

Q4. ..... is commonly confused with Bakers cysts.

Q5. Recurring dislocations and subluxations of the patella are associated with malalignment syndrome. What are some components of this syndrome?

Q6. A history of a jump, turn, twist or pivot on the knee with giving way followed by immediate swelling should be considered ..... unless proven otherwise.

Q7. How are medial collateral ligament tears managed?

Q8. The symptoms of a meniscal tear may occur gradually – T/F

Q9. An 11-year-old soccer player had a three-month history of persistent right thigh pain. He continued to play. Recently cramping pain made participation difficult and he gets flu-like night fevers. What is the possible diagnosis?

Q10. A 16-year-old male basketball player has had knee pain of one year's duration. He has no specific injury. He continues to play but gets effusion of the knee; it clicks and gives way. There are no patellofemoral symptoms. There is a thigh girth difference of 2.5cm. What is the possible diagnosis?

Q11. Women are more affected in knee OA than men – T/F

Q12. A 15-year-old male Rugby player sustained a valgus injury to his left knee. He was in possession of the ball and tackled side-on (on the left side) while his left knee was weightbearing and close to full extension. He was able to limp off the field. He was taken to hospital. X-rays were NAD. He was diagnosed as a medial collateral ligament sprain. He reports pain and effusion. He can manage a normal straight leg raise but limps. The medial joint line is tender but there does not appear to be any anterior or posterior cruciate ligament instability. What should you do?

Answers on p.57.

Autologous Chondrocyte Implantation - Is it the Answer?

Mr E Khoury, Orthopaedic Surgeon, Albury Base Hospital, Albury, New South Wales

Introduction

Full thickness articular cartilage defects of the knee are recognised as a most troublesome problem. William Hunter (1743) stated "from Hippocrates to the present day it is universally allowed that ulcerated cartilage is a troublesome thing and that once destroyed is not repaired".[1]

What is an orthopaedic surgeon to do when faced with a young patient with a painful articular cartilage defect of the knee once conventional treatment has failed? Articular cartilage implantation offers a chance to regenerate and cure traumatic lesions of articular cartilage and osteochondritis dissecans in patients with painful lesions who would otherwise have no prospect of long-term relief of pain.

Clinical Results

It is now well understood that autologous chondrocyte implantation is effective in circumscribed patellar and femoral condylar lesions of traumatic origin and caused by osteochondritis dissecans.[2,3,4] It has to date been difficult to apply to tibial lesions in the knee.

It has been documented that suitable articular cartilage lesions of the knee for articular cartilage implantation make up only 10-11% of all patients who present for arthroscopy.[5,6] The procedure is contraindicated in patients with generalised osteoarthritis, inflammatory disease, crystal disease or chondrocalcinosis. Morbid obesity is also a contraindication as are kissing lesions.[7] Angular deformity must be corrected, as should patello-femoral malalignment.

Numerous papers have now established the clinical efficacy of autologous chondrocyte implantation. The results are good to excellent for 90% of femoral condylar lesions, 74% of femoral condylar lesions with an ACL reconstruction, 84% of osteochondritis dissecans and 69% of patellar lesions with 58% of trochlear and 75% multiple lesions.[4]

Current Pitfalls

Current techniques of treatment of articular cartilage defects include arthroscopy and lavage, drilling, arthroscopy and abrasion techniques, mosaicplasty, osteotomy and microfracture. The success of these procedures is variable. Seventy per cent of patients following arthroscopic debridement alone have been reported to have good to excellent results at three years.[8] Marrow stimulating techniques such as drilling are 70% effective at five years.[9] However, whilst all are effective to a point, none provides a prospect of long-term cure, such techniques promoting development of fibrocartilage which lacks the durability and mechanical properties of hyaline cartilage normally covering articular surfaces.

It is now being established that articular cartilage implantation may be equally, if not more, successful than these techniques at providing pain relief, but in addition also regenerating articular cartilage.

The real test of this procedure in the future will be its reproducibility outside of centres that have developed the technique. Already independent studies are verifying the effectiveness of articular implantation. At the Rizzole Institute of Orthopaedics, Feruzzi et al[10] reported that 82% of their patients had good to excellent results at 1.5 years follow up. J B Richardson et al[11] at the Robert Jones and Agnes Hunt Orthopaedic and District Hospital have also reported successful results with chondrocyte transplantation.

Locally, the success of the procedure rests in part with laboratories being established, such as the Mercy Tissue Engineering Laboratory, which must be able to demonstrate good quality control in the culture of chondrocytes with effective delivery systems to ensure maximum delivery of viable chondrocytes to surgeon and patient. Surgeons must also be trained and educated in the techniques of chondrocyte transplantation and periosteal harvest for the procedure to achieve results comparable to published series. A poor technical procedure is one factor that militates against a successful result.

Achieving Successful Results

In these early stages of developing and assessing autologous chondrocyte implantation techniques, it is vital that the indications and contraindications to the technique are strictly adhered to, and results objectively evaluated and documented via multicentre trials to truly assess the efficacy of this treatment. All surgeons performing this procedure must document results precisely. We need to be able to identify our failures on the basis of histology after second look arthroscopy and biopsy and/or indentrometry.

Another reason for establishing efficacy is the high cost of the procedure, which is currently in the vicinity of $12,500 per patient. If the community is to take on the burden of these costs it is incumbent upon us to demonstrate a benefit.

In the long term, one must establish that the procedure is more successful than previous techniques, that clinical results are maintained, and most importantly that this procedure prevents or alters the outcome of the development of osteoarthritis in joints.

It must also be remembered that this procedure is not without morbidity, requiring one operation for harvest of graft and one operation for re-implantation. Complications such as adhesions and arthrofibrosis, treatment failure, delamination and detachment of the graft and of course joint infection have all been reported in published series.[4]

Patients must understand the benefits and pitfalls of the technique, and

Autologous Chondrocyte Implantation

must be compliant with a prolonged and strict rehabilitation program, otherwise again, failures will occur.

Conclusion

Although the science of articular cartilage implantation is well defined, it is now moving out into the community and being used by the general orthopaedic surgeon who must be vigilant in the initial stages to select only suitable patients and document results carefully.

Refining techniques should be an ongoing objective. Arthroscopic delivery systems and treatment would be preferable to the current open technique, and the application to other joints such as the shoulder and the ankle will probably be expanded as time goes by.

In the future growth factors such as fibroblast growth factor, transforming growth factor and insulin growth factor may well enhance the regeneration of articular cartilage. These combined with synthetic matrices may improve upon current results.

References

1. Hunter W. On the Structure and Diseases of Articulating Cartilage. *Philos Trans Royal Society London* 1743; 42b: 514-21.
2. Brittberg M D, et al. Treatment of Deep Cartilage Defects in the Knee with Autologous Chondrocyte Transplantation. *NEJM* 1994; 331: 889-95.
3. Knutsen G, et al. Autologous Chondrocyte Transplantation in the Knee. Histological and Clinical Follow Up. Presented at ICRS 3rd Symposium, Gothenburg, Sweden, April 2000.
4. Petersen L. Autologous Chondrocyte Transplantation: 2-10 year follow up in 219 patients. Presented AAOS 1998 Annual Meeting, New Orleans, March 21, 1998.
5. Asburn A, et al. Articular Cartilage Lesions of the Knee. A Prospective Arthroscopic Study. Gothenburg, Sweden April 27-29, 2000.
6. Hjelle K. Full Thickness Chondral Defects: A prospective study of 1000 knee arthroscopies, International Cartilage Repair Society, Gothenburg Sweden, 2000.
7. Minas T, Petersen L. Chondrocyte Transplantation., Operative Techniques. *Orthopaed* 1997; 7 (4): 323-33.
8. Minas T, Nehrer S. Current concepts in the Treatment of Articular Cartilage Defects. *Orthopaed* 1997; 20 (6): 525-37.
9. Tippett J W. Articular Cartilage Drilling and Osteotomy in Osteoarthritis of the Knee. In *Operative Arthroscopy*, ed. JB McGinty. New York; Raven Press, 1991, pp. 325-39.
10. Ferruzzi, et al. Autologous Chondrocyte Transplantation: Our Experience at Rizzoli Orthopaedic Institute. Presented at ICRS 3rd Symposium, Gothenburg Sweden, April 2000.
11. Richardson J B, et al. Can the Gothenburg Technique and Results of Autologous Chondrocyte Implantation be Duplicated? Presented at International Cartilage Repair Society, 3rd Symposium, Gothenburg, Sweden, 2000.

Cervical Pain and Regional Relationships

Dr Ron Palmer, Musculoskeletal Physician; Vice -President FIMM; Member, Scientific Committee, FIMM

Cervical pain, or broadly defined, pain in the neck, is not always anatomically associated with neck structures. While many cases of neck pain do relate to structures located in the neck region, a physician must always be on guard for causes that do not relate to the cervical spine and which may not be located within the confines of the neck structure. Put simply, pain in the neck does not necessarily indicate the presence of local cervical pathology. Other regions of the body can refer pain to the neck region.

In evidence based medicine (EBM.) there is a triad of importance: History, Examination, Treatment.

To arrive at a valid diagnosis implies the taking of a precise and correct history and subsequently a thorough and accurate examination. *Without an accurate diagnosis it is not possible to implement the correct treatment.*

One is immediately reminded of the profound statement by Karl Kraus: "Diagnosis is one of the commonest diseases." Think about this. Do not be one of the clinicians who have a list of personal favourite diagnoses.

"As doctors it is important that we recognise that the emotional repercussions of being diagnosed are immense."[1] You can then add another aspect of doctor/patient relationship that also serves a significant role in the patient's diagnosis and subsequent management. "To know what kind of person has a disease is as essential as to know what kind of disease a person has."[2]

It is worth reflecting that while acute neck pain can be caused by serious conditions such as infection, tumours, inflammatory diseases and vascular diseases, EBM indicates that while serious, such lesions are uncommon.

Pain in and from the neck resulting from mechanical factors implies encroachment of space and impairment of movement.[3]

If characteristic pain can be reproduced by a position or a movement and the exact nature of that position or movement is understood, the mechanism of pain production is also understood.[4] As musculoskeletal physicians we are predominantly concerned with the biomechanical function of the tissues and perhaps the Golden Rule is to reproduce the patient's pain. Having now stated the obvious, do not miss other problems by *assuming* that the neck pain is only mechanical.

Cervical pain can arise from either local structures or be referred from distal sources.

Local
Mechanical
Cervical soft tissue
Local neurological (e.g., syringomyelia)
Inflammatory
Infective
Tumours.

Referred from
Vascular (e.g., aorta)
Tumours
Systemic
Postural.

This list of conditions includes Red Flag conditions. It is worth noting that in the recently concluded Australian Initiative, which was the world's largest EBM musculoskeletal trial, not a single Red Flag was missed during the two-year period of the survey. (The full details and results of this trial were reported at the FIMM Chicago International Conference, July 2001). This clearly emphasises the importance of correct history and examination in arriving at the right diagnosis.

Pain referred from the neck to other regions of the body can be summarised as follows:

Radicular pain
Neuralgic pain
Cervical canal stenosis
Brachial plexus pain
Lateral epicondylar pain
Cervicogenic headache
Sympathetic trunk involvement.

Some characteristic history findings can help separate this mixed list of pain generators. Obviously history alone will not enable an exact diagnosis to be established, but features revealed from the history must be matched with the examination findings.

Radicular pain

Irritation to cervical nerve roots leads to radicular pain in one or both upper limbs. It is essential to remember that central pathology can exhibit similar upper limb symptoms. Characteristics of radicular pain are as follows:

- Pain can be proximal
- Frequently distal paraesthesia and pain
- Narrow band of pain
- Pain may be shooting
- Pain often burning in character.

Neuralgic pain

This may be generated by irritation of the dorsal nerve root (sensory portion). Characteristically the pain is extreme and usually has an electric shock component at the periphery of the extremity. It can be associated with paraesthesia and/or dysthesia.

Cervical canal stenosis

Unlike neuralgic pain, *pain in cervical canal stenosis is unusual* and symptoms are often insidious and variable. There is often involvement in the lower limbs as well as the upper ones. While atrophy of the intrinsic hand muscles is frequently seen, beware of a history that indicates numbness in the trunk and extremities, weakness in the legs and an unsteady gait.[5]

Brachial plexus pain

This pain is generated by inflammation at the plexus level. The aetiology is unknown, but the characteristic finding is of excruciating shoulder and arm pain. Frequently there is weakness in the arm, hand or both and this can be

Cervical Pain and Regional Relationships

accompanied by paraesthesia. Symptoms vary as any or all of the nerve roots can be involved.

Lateral epicondylar pain

This can be a great trap in diagnosis for the inexperienced. Pain at the elbow can be referred from C5-6. There may be no complaint of neck pain. Whenever a case of lateral epicondylar pain presents it is mandatory that the cervical spine and related structures be assessed.

Cervicogenic headache

Cervical tension state can lead to prolonged and sustained muscular contraction. In addition to the headache, there are usually reduced cervical movements to be found as a result of the increased muscle tension.

Sympathetic trunk involvement

There are two major components of the sympathetic mechanism in the neck: sympathetic chain and sinuvertebral nerve.

Symptom production has not been defined as yet.

The cervical part of the sympathetic trunk is associated with three ganglia. These are superior cervical, middle cervical, and cervico-thoracic.

The superior cervical ganglion gives rise to the carotid plexus, a network of postganglionic fibres that follow the ramifications of the carotid arteries. The viscera within the thorax are also supplied from this ganglion. The eye and the salivary glands are similarly supplied.[6]

The sinuvertebral nerve may become irritated and this is considered to be mechanical. Symptoms can be varied and may include:

Headache
Vertigo
Tinnitus
Nasal disturbance
Facial flushing and pain
Pharyngeal disturbance.

Non-mechanical neck pain

Five conditions fit this label, *all* of which constitute Red Flag conditions:[7]

Visceral diseases
Vascular diseases
Tumours
Infections
Fractures.

Visceral diseases

Cervical pain may include pain generated from viscera that is innervated by cervical nerve innervation. This includes some thoracic and abdominal viscera quite remote from the neck. The phrenic nerve (C3,4,5) supplies the mediastinum and the diaphragm.[8]

It is extremely unusual for visceral disease to present with neck pain only. History should reveal the clues to the visceral origin, including pain from the viscera as opposed to posterior neck pain, lumps, tenderness, swelling over the viscera, dysphagia, cough or hoarseness of the voice. Note that lymphadenitis, especially in children, can lack associated features.

When completing a careful history, it is recommended that in patients presenting with neck pain, some consideration should be given to potential visceral pathology that may refer pain into the neck region.

Vascular disease

Vascular pathology is occasionally a hidden tiger that is a major Red Flag and can lead to a disastrous outcome if overlooked. Vascular disease may present with neck pain that is the *sole* initial symptom. Among the common possibilities are:

Arterial dissection
Epidural haematoma
Carotidynia
Aberrant vertebral artery
Subarachnoid haemorrhage.

Coursing through the cervical spine are the vertebral arteries and the common carotid-internal arteries and all are known to be nociceptive.[9]

Neck pain can be the initial feature in patients presenting with arterial dissection. Over 50% of patients with vertebral artery dissection have only cervical pain as the initial symptom and this may mean no occipital headache as a feature in the very early stage. As the lesion advances, occipital headache becomes a paramount symptom.[10]

The likely anatomical positions for such dissection are:

Internal carotid artery
Vertebral artery
Aortic artery.

The following symptoms and signs must cause immediate concern:

Sudden onset of severe headache
TIA or strokes
Horner's syndrome, tinnitus, palsies
Photophobia, nausea, vomiting
Dysarthria, dysphagia, diplopia
Paraesthesia.

Arterial dissection should be considered in a patient with neck pain who does not exhibit a restricted range of head movement and has no cervical tenderness.[11]

Epidural haematoma may result following a very minor aetiological event; for example, nothing more serious than a simple sneeze. On history the patient may complain of neck pain only. Pathologically there will be an increase in spinal epidural pressure, but diagnosis is usually made from a CT scan or MRI. Needless to say, this diagnosis is a neurosurgical emergency.[12]

Carotidynia is a difficult diagnosis. The patient presents with tenderness over the region of the carotid bifurcation. There is usually associated migraine or cluster headaches. This entity suggests disease of the carotid artery or visceral disease of the throat.[13] Although considered an entity, carotidynia should not be accepted as a diagnosis.[11] It is a symptom that invites further consideration for the cause of the pain.

An aberrant vertebral artery, while always mentioned in textbooks, is very rare.

There is usually compression of the C1 dorsal root and the spinal accessory nerve by the aberrant artery. This impingement produces neuralgic neck pain and usually pain referring into a shoulder. Diagnosis is by MRI.[14]

Subarachnoid haemorrhage presents with the classical explosive headache and is accompanied by meningismus and finally loss of consciousness. Before the intervention of the loss of consciousness, the sole presenting symptom may be neck pain only.[15] A rupture of an aneurysm is the underlying aetiology.

Tumours

Cervical radiculopathy and radicular pain should be considered separately from any neck pain. Their clinical assessment and investigations are distinctly different.[16] Investigations may include nerve conduction studies, CT scan, or MRI.

Although tumours are uncommon, the Mayo Clinic reported 179 cases between 1914-50. Weakness may be in the lower limbs and not the upper ones. Presentation is normally with neurological signs as well as neck pain. Important conditions are these:

Neurofibroma
Meningioma
Glioma
Chordoma
Vascular tumour.

Infection

Infection is usually accompanied by other symptoms, such as temperature and malaise. History will indicate a non-traumatic onset. It is usual to have neck stiffness in the following conditions:

Meningitis
Encephalitis
Osteomyelitis.

Fractures

Fractures generally have a history of trauma. Stress fracture in the neck is exceedingly rare, but where headgear is worn as a regular feature and the head is placed in a position of prolonged muscular stress, such a diagnosis should be remembered as a possibility. Whenever fracture is considered a possibility as the cause for cervical pain, radiology is indicated. The Canadian C-spine rules are a useful adjunct.[17]

In summary, the most likely causes of neck pain are reasonably straightforward and uncomplicated conditions. Unfortunately there lurk a number of Red Flag conditions that should always be on the list of possible diagnoses. Like all Red Flag conditions, the outcome for the patient can be very serious if they are missed. The practice of EBM evokes the principle that the doctor takes both a full history as well as a correct and thorough examination before arriving at a final diagnosis. An incorrect diagnosis made in haste may result in the wrong treatment being commenced. Yet if the above simple approach is followed, then very few serious problems will be missed.

References

1. Aust Doc 2001 (19 January): 53.
2. J Med Ed 1962; 37: 495.
3. Cailliet R. *Neck and Arm Pain,* 3rd ed. 1998.
4. Ibid.
5. Hoff, et al. *Spine* 1997; 2: 2.
6. Clinical Neurology and Neurophysiology, 8th ed. 1992.
7. Australian MSM Initiative, 1997, Appendix NA 1.
8. Kelly, Harris, Ruddy, Sedge. *Textbook of Rheumatology* 1993: 397-416.
9. Ray, Wolff. *Arch Surg* 1940; 41: 813-56.
10. Neurol 1995, 45: 1517-1522.
11. Australian MSM Initiative, 1997, Acute Neck Pain.
12. Ann Emerg Med 1994; 23: 1368-70.
13. Roseman, Carotidynia. *Arch Otolaryngol* 1967; 85: 81-84.
14. Yano et al. *Surg Neurol* 1993; 40: 43-46.
15. Schattner. Pain in the Neck. *Lancet* 1996; 348: 411-12.
16. Australian MSM Initiative,1997, Appendix NA 3.
17. Stiell I, et al. *JAMA* 2001; 286: 1841-48.

Knowledge of Musculoskeletal Medicine at Undergraduate and Postgraduate Levels

Dr Kerry Vlahos, MB BS, FRACGP ; Professor Norman A Broadhurst, MB BS, PhD, FAFRM (Orthopaedics); Dr Malcolm J Bond, PhD (Psychiatry); all from School of Medicine, Flinders University, Adelaide

Abstract

Background. Deficiencies in musculoskeletal competence among general medical practitioners is commonly acknowledged. The degree of this deficiency is thought to be widespread. A recent study by Freedman and Bernstein sought to quantify this deficiency using a 25-item questionnaire. The goal of the current study was to determine whether their results would be replicated in the Australian setting.

Methods. Two of their items were changed to reflect the local curriculum. Face validity and importance ratings for the questionnaire were provided by consultant orthopaedic surgeons following the procedure described by Freedman and Bernstein. The questionnaire was administered to both interns during their first roster and to a randomly chosen group of general practitioners.

Results. Australian importance ratings ranged between 2.6 and 4.7. There was only a modest correlation between American and Australian ratings (r = .50). Mean competence scores were 77 (± 10.9) for GPs and 69.4 (± 12.0) for interns, these being statistically different (p = .001). Sixty-eight per cent of GPs and 39% of interns reached the desired standard, with females performing worse than males. This sex difference was significant only among interns (p = .023).

Conclusions. Musculoskeletal knowledge among recent medical graduates has again been found wanting. The need for further musculoskeletal education has been established. Implementing strategies to correct the deficiency has yet to be addressed.

Introduction

In the absence of any formal assessment of musculoskeletal medicine as an adjunct to the core disciplines of medical curricula, an appropriate level of competency in musculoskeletal knowledge is arguably difficult to acquire. It is likely to be ignored in favour of more easily acquired skills demanded by disciplines that deal with more life-threatening conditions.[1-3] Yet it has been established that musculoskeletal problems in family practice are second only to respiratory presentations,[4] hence the need for all medical graduates to demonstrate an appropriate skill level.

A recent study by Freedman and Bernstein[5] revealed a significant deficiency in the musculoskeletal knowledge base of recent medical graduates in Pennsylvania. The authors also implied that such a situation was likely to be more widespread than the population they sampled. The 25-item questionnaire developed by Freedman and Bernstein is the first attempt to set a standard of what might reasonably be used by all medical schools to assess the musculoskeletal knowledge base of their students.

The first aim of the current study was therefore to administer this questionnaire to recent graduates from the two medical schools in South Australia. The purpose was to compare their knowledge base to that of their counterparts in the United States. Second, because the Flinders Medical School is the tertiary referral centre for the southern region of the major city, Adelaide, the study also surveyed local GPs, who frequently refer patients for treatment. A comparison of the knowledge base of recent graduates with those already in family practice was thought informative to ongoing curriculum development.

Methods

Construction of the Australian Competency Examination

The 25-item Pennsylvania examination was modified slightly to acknowledge differences in Australian musculoskeletal practice and hence curriculum. Two questions were substituted: "What is the difference between osteoporosis and osteomalacia?" and "Why, in elderly patients, are displaced fractures of the femoral neck typically treated with joint replacement, whereas fractures near the trochanter are treated with plates and screws?" were omitted. They were replaced with "Why, in a radial nerve palsy, is the upper limb internally rotated and the hand flexed?" (Answer: Pronator teres and flexors are unopposed), and "Rupture of the biceps at the elbow results in weakness of what two functions?" (Answer: Supination and elbow flexion). The full 25 items, correct answers, importance ratings, and the mean percentage of correct responses for both GPs and interns are presented in Table 1 (see next page).

Validation of the Australian Examination

The 31 orthopaedic surgeons responsible for teaching undergraduate musculoskeletal medicine in South Australia were sent the Australian Examination and invited to score the importance of each of the 25 items using a response scale ranging from 0 (not important) to 5 (extremely important).

Administration of the Australian Examination

The general practitioners (GPs) formed a random sample drawn from the register of the Southern Division of General Practice in South Australia, comprising 600 GPs. The nature and purpose of the examination were explained to the selected GPs by telephone, at which time verbal informed consent was obtained. They then received a copy of the examination by mail. In all correspondence they were instructed to complete the examination without reference to either literature or colleagues. Non-responders were followed up after two weeks. Completed examinations were returned by 47 of the 100 GPs sampled. Seventy-two per cent of the final sample was male, and 68% had been in general practice at

Knowledge of musculoskeletal medicine at undergraduate and postgraduate levels

Table 1. Australian Musculoskeletal Medicine Competency Examination

	Question	Answer	Rating	GPs Mean %	Interns Mean %
1.	What orthopaedic problem must always be assessed in the newborn?	Congenital dislocation of the hip	9.4	100	100
2.	A patient is very tender over the snuffbox 6 weeks after a fall on outstretched arm. What condition must be considered?	Scaphoid fracture	9.2	97	97
3.	A patient dislocates a knee in an MVA. What non-bony structure must be evaluated?	Popliteal artery	8.8	49	39
4.	A 25-year-old motor cyclist involved in an MVA is lying with hip in flexion, internal rotation and adduction. What is the diagnosis?	Hip dislocation	8.2	62	67
5.	A 12-year-old boy presents with a badly swollen and tender lateral malleolus. What 2 pathologies are likely?	Sprain, fracture	6.6	91	82
6.	A 30-year-old has a displaced fracture near the fibular neck. What structure is at risk of injury?	Common peroneal nerve	8.4	66	89
7.	A 20-year-old injures a knee at basketball. Significant swelling is noted very soon after. What 3 structures may be responsible for the swelling?	Ligament tear, fracture, peripheral meniscal tear	7.0	97	77
8.	What are the 5 most common organs that produce metastatic bone disease?	Breast, prostate, lung, kidney, thyroid	7.0	88	87
9.	What bony malignancy is not readily detected by a bone scan?	Multiple myeloma	5.2	36	26
10.	What nerve is compressed in carpal tunnel syndrome?	Median nerve	9.4	96	100
11.	How is the motor function of the median nerve tested?	Metacarpophalangeal finger flexion, thumb opposition, flexion or abduction	8.0	60	79
12.	What is a compartment syndrome?	Increased pressure in a closed fascial space	9.4	96	92
13.	How is acute compartment syndrome treated?	Fasciotomy	8.8	94	92
14.	A non-traumatic acutely swollen knee may be due to which 3 conditions?	Gout, infection, reactive synovitis	7.2	96	84
15.	A patient injures the ulnar side of his hand in a fight. The area is tender and he can't make a fist. List 3 common injuries related to such an action.	Fracture, dislocation, ligament tear	5.4	86	61

Knowledge of musculoskeletal medicine at undergraduate and postgraduate levels

16.	A 23-year-old male has a long history of low back pain and stiffness. What is the probable diagnosis and definitive investigations?	Ankylosing spondylitis. CT of sacroiliac joints	6.4	52	39
17.	A patient has a 5th lumbar radiculopathy. How is the motor function tested?	Dorsiflexion or plantar flexion of the great toe	8.0	47	17
18.	Name 2 differences between rheumatoid arthritis and osteoarthrosis.	Inflammatory vs degenerative, proximal inter-phalangeal joint vs distal interphalangeal joint	8.2	79	67
19.	A patient presents for the first time with low back pain. List 4 criteria that would lead to consider a plain lumbar x-ray.	Previous tumour, night sweats, night pain, failure to respond to conservative treatment	5.8	83	27
20.	An elderly patient gets buttock pain and lower limb pain on walking less than 50 metres. What 2 conditions need investigating?	Vascular claudication, central canal stenosis	8.4	94	66
21.	What is the function of the anterior cruciate ligament of the knee?	To prevent anterior displacement of the tibia on the femur	6.8	94	57
22.	What muscle group is involved in lateral epicondylitis?	Wrist extensors	6.8	76	52
23.	Why, in a radial nerve palsy, is the upper limb is internally rotated and the hand flexed?	Pronator teres and flexors are unopposed	5.2	45	41
24.	Rupture of the biceps at the elbow results in weakness of what 2 functions?	Supination and elbow flexion	6.6	61	75
25.	Name the rotator cuff muscles.	Subscapularis, infraspinatus, supraspinatus, teres minor	6.0	69	90

least 16 years.

The intern sample was obtained by personally visiting all teaching hospitals in South Australia during the first intern rotation of the year. In this way, all interns were included. Verbal informed consent was obtained prior to the examination being completed under supervision. There were 66 interns available to complete the examination, of which 62% were male.

Scoring of the Australian Examination

Examination results were determined anonymously according to a predetermined answer key. A maximum of one point was given for each correct answer, with partial credit being given where applicable. These raw scores were then weighted according to the average importance rating given to each item by the orthopaedic surgeon educators; that is, relatively more credit was given for correct answers to questions deemed to be of more importance, and vice versa. The weighted scores were then converted to percentages. In accord with the Pennsylvania study, a pass mark of 73.1% was applied.

Statistical Analyses

A Pearson correlation was used to compare the importance ratings applied in the current study to those used in the Pennsylvania study. Examination scores were compared using two-tailed student t tests. Pass/fail propor-

Knowledge of musculoskeletal medicine at undergraduate and postgraduate levels

tions were compared using chi square analyses.

Results

Validity of the Australian Examination

The importance ratings applied to the 25 items ranged from 2.6 to 4.7, with a mean of 3.7 (SD = 0.6, median = 3.5). Twenty-two items (88%) received a rating of at least 3 (that is, important). The correlation between the ratings assigned to the 23 items included in both the USA and Australian studies yielded a coefficient of .50 ($p < .001$). That is, there were only 25% of shared variance (r^2) in the two sets of ratings.

Australian Examination Scores

The mean score among GPs was 77.0 ±10.9, compared with 69.4 ±12.0 for the interns. This difference was statistically significant ($p = .001$). A similar difference was noted when pass/fail results were analysed ($p = .003$). A higher proportion of GPs passed the examination (68% vs 39%). There were no sex differences in the overall mean scores of either the GPs ($p = .703$) or interns ($p = .129$). However, a greater proportion of male interns (51%) than female interns (22%) achieved the pass mark ($p = .023$). This was not true of the GPs (73% vs 54%, $p = .195$).

Discussion

Although the scores for the South Australian medical graduates were marginally higher than those in Pennsylvania, the baseline competence was not attained to the degree desired. This suggests that a similar pattern of deficiency exists in both states. This raises a cause for concern to improve the musculoskeletal base of our graduates. The question will always be "what can be exchanged for more musculoskeletal teaching". The answer from medical schools is likely to be "very little". There are few options available which are likely to resolve this dilemma. One option is that graduates headed for family practice could be targeted for skills training. Other options are limited by time and pressures from other disciplines as well as the lack of capable and dedicated clinicians to spend the time required bringing all graduates to a suitable level. Indeed, it is realised that there is a finite amount of time available in each day and to overload the student reflects poorly on the curriculum. An extensive elective period within the medical program would help overcome these deficiencies but would be at the expense of another area of the medical curriculum.

However it is achieved, an improved knowledge base in musculoskeletal medicine would bring with it significant cost savings in the delivery of healthcare, demonstrable by at least five likely changes to the practice of musculoskeletal medicine. First, there would be less use of imaging services which are often ordered "just in case" or "to see if". It has been established that more than 90% of plain lumbar films have no influence on the management of low back pain. Other more expensive scans are often ordered to try to find a source of pain instead of pursuing the problem with appropriate and proven clinical skills. It takes time and practical experience to be able to correlate the presenting complaint with the imaging services needed for appropriate management of the problem. Exposure to a range of musculoskeletal clinics would be of considerable benefit to those destined for family practice.

Second, there would be less prescribing of medications, particularly the nonsteroidal anti-inflammatory drugs. Most musculoskeletal problems can be managed with simple to moderate analgesics rather than anti-inflammatory drugs; the cause of the pain is more likely to be mechanical than inflammatory. Historically, indiscriminate use of anti-inflammatory medication has produced a variety of gastrointestinal problems, with deaths occurring out of proportion to the severity of the complaint. Third, given a comprehensive knowledge of the presenting complaint, there would be fewer referrals to specialist services where there is often pressure for even more investigations in place of the explanation and reassurance which are the real needs.

Fourth, a full appreciation of the natural course of the injury or disease process allows the implementation of self-management exercises without recourse to unnecessary medications, investigations and therapies. This approach minimises the unwanted side effects of drugs and investigations while providing the patient with mechanisms for ongoing management, especially should the condition recur. Finally, rural family practitioners need more skills than those in an urban practice, as access to specialist services on a regular basis is often very limited. Thus knowledge of the natural course of the injury or disease process provides confidence instead of the practitioner being stressed or frustrated at not being able to access already over-used city services.

It must be remembered that a knowledge base in musculoskeletal medicine is not preparing students to become orthopaedic surgeons but simply to manage musculoskeletal problems at the coal face. Unless the family physician can manage musculoskeletal problems better, there will be the continuing slow drain away from conventional medical practice toward quasi-alternative therapies. Practitioners of such therapies are not medically trained and hence likely to miss the Red Flag symptoms of a life threatening condition. It must therefore be considered as an obligation to find ways to ensure medical graduates are competent in managing musculoskeletal problems.

Many of the family physicians in the Southern region of Adelaide achieved

Knowledge of musculoskeletal medicine at undergraduate and postgraduate levels

an acceptable standard on our test. If the undergraduate level was inadequate, how did the GPs improve their knowledge base? Assuming respondents did not consult other sources, adequate learning had to be in-house or by osmosis of some kind. A knowledge base in any area of medicine increases with exposure to relevant conditions, as well as communication with other practitioners. Perhaps the best way forward might be for undergraduates to align themselves with mentors who are prepared to engage in educational pursuits and then to be linked to a craft group which has a meaningful continuing education program.

The more searching question is how such a level of competency can be validated. Is the Pennsylvania questionnaire a gold standard? It is certainly a reasonable first step toward a reliable standard being established. Is there a definable body of musculoskeletal knowledge that reflects best practice in the discipline? If it is agreed that the current standard is poor, whose responsibility is it to improve the knowledge and skill base? The answer must involve clinicians practising in the musculoskeletal disciplines (that is, orthopaedics, rheumatology, rehabilitation, and occupational medicine). Perhaps it is time to produce orthopaedic physicians!

Key learning points

- Adelaide participants were just as deficient as their USA counterparts.
- There is a need to improve the knowledge base and associated skills in musculoskeletal medicine.
- There is difficulty allocating time for additional activities within existing medical curricula.
- The use of mentors may be a worthwhile pursuit in the attempt to improve proficiency.

Funding

No external funding was received for this study.

Contributors

The principal investigator was NAB, who initiated and supervised the study. KV was responsible for implementation of the study, including data collection. MJB provided assistance with data management, statistical analysis, and preparation of the manuscript.

References

1. Morrison MC. Teaching of musculoskeletal medicine: a survey of general practitioners and deans. *Med Educ* 1993; 27: 245-49.
2. Sneiderman C. Orthopaedic practice and training of family physicians: a survey of 302 North Carolina practitioners. *J Fam Pract* 1977; 4: 267-70.
3. Crotty M, Ahern MJ, McFarlane AC, Brooks PM. Clinical rheumatology of Australian medical students. A national survey of 1991 graduates. *Med J Aust* 1993; 158: 119-20.
4. Royal College of General Practitioners. *Office of Population Censuses and Statistics. Morbidity Statistics from General Practice: Third National Study*. London: HMSO; 1982.
5. Freedman KB, Bernstein MD. The adequacy of medical school education in musculoskeletal medicine. *J Bone Joint Surg* 1998; 80-A: 1421-27.

Reliability and Validity Studies of Diagnostic Procedures in Manual/ Musculoskeletal Medicine

Protocol Formats, 2nd edition

Scientific Committee, FIMM

Preface to the Reproducibility and Validity Protocol 2nd Edition

Based on an internal discussion within the Scientific Committee (SC) of the International Federation for Manual/Musculoskeletal Medicine (FIMM), a second edition became necessary. It became clear that the first protocol showed shortcomings with respect to the logistic performance of reproducibility studies and the prevalence problem. This second edition has been changed in two main aspects. The first protocol has been rewritten as a more practical manual for performing reproducibility studies. Attention is paid to the logistic aspect of a reproducibility study.

In contrast to the first protocol: in the 2nd edition an additional subject "the overall agreement phase" has been incorporated. To clarify and/or explain different aspects of the kappa value, different items of the first protocol have been elaborated in more detail.

The 2nd edition has been developed for reproducibility studies not only of the lumbar region but also for the cervical region.

The Scientific Committee of the FIMM is aware that developing this kind of protocol is a continuous process.

By publishing the 2nd edition on the website of the FIMM, the Scientific Committee hopes that scientists who use this protocol will send their comments to the chairman of the Scientific Committee. In this way, we hope to improve the present protocol.

The SC asks scientists who receive this protocol to disperse it to their fellow scientists. The protocol becomes thus accessible for all practitioners in the field of M/M Medicine.

This protocol is the end product of the energy of all the members of the SC.

Dr Jacob Patijn, MD, PhD, Neurologist, Physician for Manual/Musculoskeletal Medicine, Chairman of the Scientific Committee of the FIMM, Responsible member for the Reliability Group of this Committee

Scientific Committee FIMM
Chairman:
Dr Jacob Patijn, Eindhoven, The Netherlands (email *jpat@sane.azm.nl*)

Members:
Dr Jan van Beek, The Hague, The Netherlands
Dr Stefan Blomberg, Stockholm, Sweden
Professor Boyd Buser, Biddeford, United States
Dr Paul Cohen, Geneva, Switzerland
Dr Richard Ellis, Salisbury, United Kingdom
Dr Jean Yves Maigne, Paris, France
Dr Ron Palmer, Herston Australia
Dr Lars Remvig, Holte, Denmark
Dr Jan Vacek, Prague, Czech Republic
Professor Robert Ward, Michigan, United States
Professor Lothar Beyer, Jena, Germany
Advisor:
Professor Dr Bart Koes, Epidemiologist, Erasmus University Rotterdam

I. Introduction from chairman Scientific Committee

This is the second edition of one of the first two scientific protocols (Reproducibility and Validity Studies: a protocol format, and Efficacy Trial: a protocol format, 2001) of the Scientific Committee of FIMM (SC). They concern a standardised format for reproducibility, validity, sensitivity and specificity studies and efficacy trials in manual/musculoskeletal medicine (M/M Medicine) for diagnostic procedures in M/M Medicine.

In future, improved scientific protocols will be developed. When necessary, single protocols of particular regions of the locomotion system such as the thoracic shoulder regions and extremities will be published by the SC.

The SC's reason for developing these protocols has been extensively discussed in previous reports of the SC for the General Assembly and has been published in *FIMM News*.

To provide a short background of these protocols a brief overview of the past SC activities is given.

The SC formulated the problem with respect to diagnostic procedures in M/M Medicine, which is summarised in the following statement.

There are too many different schools in manual/musculoskeletal medicine in many different countries of the world, with too many different diagnostic procedures and too many different therapeutic approaches.

The consequences of this state-

Reliability and Validity Studies of Diagnostic Procedures in M/M Medicine

ment are five-fold:

1. Most schools within M/M Medicine have not yet validated their own characteristic diagnostic procedures in the different regions of the locomotion system. Therefore, reproducibility, validity, sensitivity and specificity of these diagnostic procedures is still lacking.

2. All the different schools within M/M Medicine still coexist. Because of lack of good reproducibility, validity, sensitivity and specificity studies, mutual comparison of diagnostic procedures is impossible. Scientific information exchange and fundamental discussions between these different schools, based on solid scientific methods, is almost possible in the present situation.

3. Absence of validated diagnostic procedures in M/M Medicine leads to heterogeneously defined populations in efficacy trials. Therefore, comparison of efficacy trials, with the same therapeutic approach (for instance, manipulation), is impossible.

4. If the present situation is allowed to continue, it will lead to a slowing down of the badly needed process of professionalisation of M/M Medicine.

5. Non-validated diagnostic procedures of different schools, ill-defined therapeutic approaches and low quality study designs are the main causes for the weak evidence of a proven therapeutic effect of M/M Medicine.

It is the opinion of the SC that the committees should create conditions for exchange of scientific information between the various schools in M/M Medicine. This information exchange must be based on results of solid scientific work. By comparing the results of good reproducibility, validity, sensitivity and specificity studies, performed by different schools, a fundamental discussion will arise. The main aim of this discussion is not to conclude which school has the best diagnostic procedure in a particular area of the locomotion system, but to define a set of validated diagnostic procedures which can be adopted by the different schools and become transferable to regular medicine.

The SC wants to provide the national societies of FIMM with standardised scientific protocols for future studies.

The SC thought that the best forum for creating a discussion platform would be to organise every other year a SC conference in cooperation with a particular national society. Details will be published later.

As chairman of the SC, I want to emphasise that good reproducibility, validity, sensitivity and specificity studies have the first priority. These kinds of studies are easy and cheap to perform and form the best base for mutual discussion between schools in M/M Medicine. They are also essential for defining a homogeneous population in efficacy studies.

Cooperation and active involvement of the national societies of FIMM are indispensable and crucial for the future work of the SC.

In providing the first protocols to the national societies of FIMM, the SC hopes to attribute a substantial contribution to the professionalisation of M/M Medicine.

Dr Jacob Patijn, MD, PhD, Neurologist

II Reproducibility and validity

Nomenclature

One of the major problems in medicine and in research is the fact that different names are used for the same definition. Therefore, we thought it important first to provide the reader of this protocol with an overview of the definitions used here. In clarifying the definitions in advance we hope to make reading easier.

1.0 Reliability can be divided into Precision and Accuracy.

1.1 *Precision* also mentioned *Reproducibility.*

In the case of reproducibility of an observation made by one observer on two separate occasions, we call it the intra-observer variability or the intra-observer agreement.

In the case of reproducibility of an observation by two observers on one occasion, we call it the inter-observer variability or the inter-observer agreement.

In this protocol, we will use the terms *reproducibility*, *intra-observer agreement* and *inter-observer agreement*.

Reproducibility studies of diagnostic procedures in M/M Medicine evaluate whether two observers find the same result of a diagnostic procedure in the same patient population, or whether a single observer finds the same result of a diagnostic procedure in the same patient population on two separate moments in time.

1.2 *Accuracy*, also mentioned *Validity*.

In this protocol, we will use the term *validity*.

Validity studies measures the extent to which the diagnostic test actually does what it is supposed to do. More precisely, validity is determined by measuring how well a test performs

against the gold or criterion standard.

When a diagnostic test has to be evaluated with respect to what it is supposed to do, one needs a gold standard as reference. This is a major problem in medicine. Sometimes, radiological findings, post-mortem findings or findings during operation can act as gold standard. In the case of subjective quantification of range of motion, the gold standard can be the results a quantitative method performed in a normal population. Gold standards are needed for estimation of the sensitivity and specificity of a test (see V.1).

2.0 Index condition

2.1 The *index condition* is synonymous with the diagnosis of a patient. This diagnosis must be based on reproducible diagnostic procedures with a proven validity.

2.2 The *prevalence of the index condition* is the frequency of the index condition in a particular population at a particular moment.

It is essential to realise that the prevalence of an index condition can vary in different institutes, countries, and from time to time.

In this protocol, we will use the terms *index condition* and *prevalence of the index condition and/or positive test procedure.*

In reproducibility studies, the prevalence is assessed with regard to the number of tests judged positive by the observers.

In the 2x2 contingency table hereunder, a theoretical example of the results of a reproducibility study of two observers A and B is shown.

		Observer B: Yes	Observer B: No	total
Observer A	Yes	a	b	a + b
	No	c	d	c + d
total		a + c	b + d	n

Figure 1. 2x2 contingency table.

The squares with a and b represent the number of tests judged positive by observer A. The squares with a and c represent the number of tests judged positive by observer B. The squares with a,b and c are the tests judged positive by both observers in total n patients.

The prevalence is calculated by the formula for the prevalence (P):

$$P = \frac{[a + (b + c)/2]}{n}$$

3.0 Overall agreement

The overall agreement reflects the percentage of the patients in which both agree about the judgement of the test. Based on Figure 1, both observers agree in a and d (respectively positive and negative). In the squares with b and c, the observers disagree.

Overall agreement P_o is calculated by the formula:

$$P_o = \frac{[a + d]}{n}$$

4.0 Sensitivity and Specificity

4.1 The *sensitivity* of a test is defined as: the proportion of the cases that have the index condition that the test correctly detects.

4.2 The *specificity* of a test is defined as: the proportion of the cases that do not have the index condition that the test correctly detects.

In this protocol the "Nosographic Sensitivity and Specificity" is identical with the terms "*Sensitivity and Specificity*" used.

4.3 To translate the statistics of sensitivity and specificity figures into daily practice, the physician has to know whether a positive test in the individual patient is truly positive as opposed to false-positive. This is expressed respectively as the "positive predictive value of a test" and "negative predictive value of a test".

In contrast to the "Nosographic Sensitivity and Specificity", the positive predictive value of a test and negative predictive value of a test are also called the "Diagnostic Sensitivity and Specificity".

In this protocol the "Diagnostic Sensitivity and Specificity" is identical with the terms "*positive and negative predictive value* of a test".

5.0 Kappa value: interpretation

Kappa value is a statistical measurement for the intra-observer and inter-observer agreement corrected for chance.

The kappa value can be either negative or positive and ranges between –1 and +1. Some authors (Landis, Koch, Biometrics 1977; 33: 159-74) use 0.6 as cut off point, others (Bogduk) use 0.4. The cut off level is arbitrary. We use the cut off point 0.6.

III. Starting points in reproducibility protocol of diagnostics in M/M

To perform reproducibility studies for diagnostics in M/M Medicine several points are important to consider to start with.

1.0 Character of the diagnostic procedure

Before starting a reproducibility study in M/M Medicine, it important to be clear about what kind of diagnostic procedure we are dealing with.

In general we have two kinds of diagnostic procedures: (a) Qualitative Procedures, (b) Quantitative Procedures

1.1. *Qualitative Procedures*

Qualitative diagnostic procedures in M/M Medicine are characterised by subjective outcomes of observer and/

Reliability and Validity Studies of Diagnostic Procedures in M/M Medicine

or patient. Typical examples of this kind of procedure in M/M Medicine are end feeling and pain provocation under different conditions (provoked by observer, provoked by movements of the patient).

1.2 *Quantitative Procedures*

In subjective quantitative diagnostic procedures in M/M Medicine methods are usually involved which subjectively quantify the results of the diagnostic procedure performed (restricted yes or no). Typical examples of these kinds of procedure in M/M Medicine are subjective range of motion or motion patterns.

When a real quantitative method with certain developed devices is being used, test/retest procedures and normative values are needed.

2.0 Aim of the diagnostic procedure

In studying the reproducibility of a diagnostic procedure in M/M Medicine one has to be clear about the aim of the test.

2.1. In M/M Medicine, in evaluating a single diagnostic test, information about the reproducibility of the test procedure only is obtained.

In the vast majority of single diagnostic tests, no information is obtained about a specific diagnosis based on that single diagnostic test. Therefore, a single diagnostic test seldom differentiates between normal subjects and patients. In general, in the absence of a gold standard, sensitivity and specificity studies are useless if they are based on a single reproducible diagnostic test.

2.2. In M/M Medicine, sometimes, a combination of several positive tests is related to a particular phenomenon or clinical finding. In evaluating a combination of test procedures, only information about the reproducibility of the combination tests is obtained. The positive findings are non-specific and can be seen not only in specific and non-specific LBP and CPS but also in normal subjects. In the absence of a gold standard, sensitivity and specificity studies are useless when based on a combination of reproducible diagnostic tests alone.

2.3. Reproducing a test in time, by repeating the test after a time interval, can help differentiate between patients and normal subjects and can be used to estimate the sensitivity and specificity of a test. These tests, when combined with other clinical data can also differentiate between patients and normal subjects. In the vast majority, no information is obtained about a specific diagnosis based on this combination. In general, only in the presence of a gold standard are sensitivity and specificity studies useful, based on a combination of valid test procedures.

2.4. In M/M Medicine, sometimes, a combination of several positive tests is related to a specific diagnosis or syndrome. In evaluating a combination of test procedures, information about the reproducibility of the combination of tests is obtained. The positive findings are related to a specific diagnosis or syndrome. In this case, a gold standard is needed (for instance, findings of imaging procedures). In general, in the presence of a gold standard, sensitivity and specificity studies based on a combination of valid test procedures, are useful.

3.0 Number of tests to be evaluated

Reproducibility studies in non-specific LBP sometimes show evaluations of a large number of tests. Many of the tests show low kappa values and therefore are judged of no clinical importance by the authors. Since prevalence and overall agreement figures are frequently lacking, such a definite conclusion about the reproducibility of the tests cannot be drawn. Since a heterogeneous study population consists of different subgroups with an unknown frequency, there is a risk that some positive tests show a low prevalence, because of the small size of a particular subgroup.

The test to be evaluated must have a relationship to the characteristics of the study population. For example, evaluating the reproducibility of radicular provocation tests in LPB patients without any signs of sciatica, have no sense.

In the case of a population with sciatica, evaluating the reproducibility of radicular provocation tests, one can decide on a minimal number of positive tests needed to make the diagnosis of a lumbar radicular syndrome.

Another aspect of too many tests is the mutual dependency of tests that are supposed to assess the same clinical feature. For example, many SI-tests are supposed to test a SI-dysfunction or hypomobility of the SI-joint. This dependency was shown in a re-

Test	Obsv.	I	II	III	IV	V	VI
I	A						
	B						
	C					Kappa values	
II	A	-0.09					
	B	+0.02					
	C	+0.36					
III	A	+0.25	-0.01				
	B	+0.34	+0.17				
	C	+0.36	+0.22				
IV	A	+0.34	-0.29	+0.25			
	B	+0.06	-0.05	+0.15			
	C	+0.22	-0.01	+0.36			
V	A	+0.61	-0.12	+0.28	+0.43		
	B	+0.33	+0.39	+0.34	+0.01		
	C	+0.10	+0.21	+0.21	+0.32		
VI	A	+0.61	-0.22	+0.18	+0.43	+0.89	
	B	+0.23	+0.19	+0.21	-0.15	+0.52	
	C	+0.21	+0.32	+0.24	+0.27	+0.84	

Figure 2. Mutual dependency of six SI-tests (I-VI) in three observers A,B and C. The kappa values > 0.60 reflect a mutual dependency.

Reliability and Validity Studies of Diagnostic Procedures in M/M Medicine

producibility study of six SI-tests (Deursen van, Patijn). By calculating the kappa values between different SI-tests for one observer, the mutual dependency is illustrated. Figure 2 shows these mutual kappa values of six SI-tests (I-VI) in three observers (A,B,C).

For example, SI-test V versus SI-test VI in the last right column showed in all three observers A,B,C high kappa values of, respectively, +0.89, +0.52 and +0.84. This means that all three observers unconsciously judged SI-test VI positive after they had judged SI-test V as positive. In this study only SI-tests II, III and IV were independent (2nd 3rd and 4th column).

This aspect is very important for reproducibility studies when selecting tests for the same clinical feature.

In kappa studies, besides evaluating the reproducibility of the test, sometimes the interobserver agreement of the diagnosis based on these tests is evaluated.

From the same study, it became clear that with too many tests observers use only a few for their final diagnosis. By calculating the mutual kappa value of the single tests (I-VI) and the final diagnosis in all three observers A, B, and C this phenomenon is illustrated (see Figure 3).

Note that in the far right column "SI-Diagnosis" all three observers use only SI-test V and VI for their judgement of the SI-diagnosis. In all three observers A, B and C SI-tests I-IV contributed not at all to the final SI-diagnosis.

In general it is advisable to evaluate a maximum of three tests for the same clinical feature.

4.0 Number of observers

There is no real statistical reason for performing a reproducibility study with more than two observers. In some studies, more observers are involved to evaluate the effect of the observers' experience on the interobserver agreement. The problem with experienced observers is that they probably have developed a personal performance and interpretation of the test. Most of these studies lack a proper training period for standardisation of the performance of the test procedure and its interpretation. The results of these kinds of studies inform us more about the skills and/or the quality of the educational systems of the observers, rather than about the reproducibility of the evaluated tests. The same is true for reproducibility studies which estimate kappa values of tests done in the in-vivo condition, in which no standardisation of the test procedures was carried out (to mimic the daily practice of a test).

In principle, reproducibility studies, using the format discussed below, provide us with the potential reproducibility of a test procedure. If the reproducibility of a test procedure is established, a second study can be performed to evaluate the effect of observers' characteristics on the reproducibility.

A second flaw of using too many observers in a reproducibility study is the possibility of a therapeutic effect of the test procedure. If in a single patient, a passively performed procedure (passive cervical rotation) is performed too many times by different observers in a row, a therapeutic effect of the procedure may influence the range of motion and therefore the results of the last observer.

In general, using the proposed format in this protocol, two observers are sufficient to estimate the potential reproducibility of a test.

5.0 Hypothesis of a test

It is very important for a reproducibility study of a test to discuss and analyse what the test is supposed to test. For range of motion there is no problem. For mobility, for instance hypomobility of the SI-joint, there is a problem. In many reproducibility studies of the SI-joint, the hypothesis for the various tests was that they were supposed to test the mobility of the SI-joint. Although SI-mobility is proven, based on cadaver studies, it is impossible, even for the most experienced observer, to test manually the mobility of the SI-joint. This incorrect belief is probably the reason for the low kappa values of SI-tests in the literature. Looking critically at the substantially different procedures of the large number of SI- tests, we have to question whether all these procedures can test the hypomobility of the SI-joint. In reproducibility studies, the observer has to forget the hypothesis of the tests and has to concentrate on all the different aspects of the test procedure. For instance, according to the literature, the Patrick test for the SI-joint is supposed to test the mobility of a SI-joint.

SI-Tests		I	II	III	IV	V	VI	SI-Diagnosis
	Obsv.							Kappa
I	A							**-0.61**
	B							**+0.23**
	C							**+0.21**
II	A	-0.09						**-0.22**
	B	+0.02						**+0.19**
	C	+0.36						**+0.32**
III	A	+0.25	-0.01					**+0.18**
	B	+0.34	+0.17					**+0.21**
	C	+0.36	+0.22					**+0.24**
IV	A	+0.34	-0.29	+0.25				**+0.43**
	B	+0.06	-0.05	+0.15				**-0.15**
	C	+0.22	-0.01	+0.36				**+0.22**
V	A	+0.61	-0.12	+0.28	+0.43			+0.89
	B	+0.33	+0.39	+0.34	+0.01			+0.52
	C	+0.10	+0.21	+0.21	+0.32			+0.84
VI	A	+0.61	-0.22	+0.18	+0.43	+0.89		+1.00
	B	+0.23	+0.19	+0.21	-0.15	+0.52		+1.00
	C	+0.21	+0.32	+0.24	+0.27	+0.84		+1.00

Figure 3. Mutual dependency of six SI-tests (I-VI) with the final SI-diagnosis in three observers A,B and C. The kappa values > 0.60 reflect a mutual dependency.

Reliability and Validity Studies of Diagnostic Procedures in M/M Medicine

Looking critically at the test procedure, the observers can decide that the Patrick test, measuring end feeling and motion restriction, evaluates only increased muscle tension of a certain group of muscles related to the hip joint. The effect of the hypothesis on the reproducibility on SI-test was illustrated in two studies. The first study which assessed six SI-tests supposed to evaluate SI-mobility, resulted in very low kappa values. In the second study, three tests supposed to test muscle hypertonia in different muscle groups around the lumbosacral-hip region resulted in a kappa value of 0.7.

Whatever tests are selected for a reproducibility study, the observers have to investigate step by step the whole test procedure and agree about what the test really tests.

Based on this agreement, the observers can define a more plausible hypothesis for the test, which can completely contradict the hypothesis stated in the literature.

Full agreement of the observers about a more plausible hypothesis of a test can lead to better results in reproducibility studies.

6.0 Blinding procedures

In every reproducibility study, blinding procedures are essential not only for the patient/observer condition but also for both observers and must be well defined.

7.0 Test procedure and test judgement

As already argued under item 6.0, the observers have to standardise the whole test performance and the way they judge the result of a test. In the protocol format discussed below, the training period is essential for standardisation in a reproducibility study. The consensus about the definition of the test procedure and its assessment must be discussed in the final publication. To prevent observers "personal interpretation" during the study, we also advise that the standardised procedures and test assessments are printed on the forms used in the study.

8.0 Selection and number of subjects

In reproducibility studies, the primary source population, out of which the subjects are selected, must be defined, and selection procedures must be very clear.

In general, for simple reproducibility studies 40 subjects are sufficient. This number makes such a reproducibility study easy and cheap to perform and not restricted to large institutes.

IV Statistics Reliability and Presentation

1.0 Statistics in reproducibility studies: the kappa value

In reproducibility studies, with two observers evaluating dichotomous tests (Yes/No), estimation of the kappa values is the method of choice (see below).

1.1 *Kappa dependency on prevalence.* In the literature many reproducibility studies judge diagnostic tests with kappa values below 0.6 as clinically irrelevant. However, in the vast majority of reproducibility studies no information is presented about the corresponding prevalence and overall agreement of the kappa value. This is essential, because the kappa value is dependent on the prevalence and the overall agreement.

Published reproducibility studies which present evaluations of tests with low kappa values as clinically worthless or of minor importance, without mentioning any figures about prevalence and overall agreement, are misleading.

Low kappa values can reflect high as well as low prevalences!

Figure 4 shows the dependency of the kappa value on the prevalence.

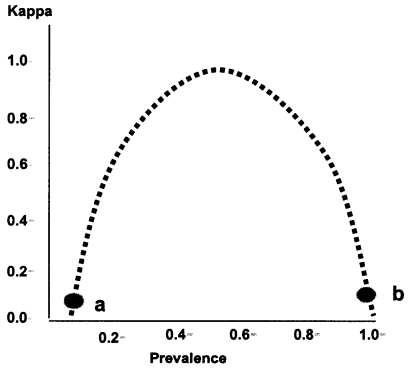


Figure 4. Relationship between kappa values and prevalences.

Note that in case of very low (a) and very high prevalences (b) the kappa value becomes very low.

1.2 *Kappa dependency on overall agreement (P_O).* In Figure 5 the maximum kappa value is 1.0 and the minimum kappa value is nearly 0. This range is dependent of the overall agreement P_O of the two observers. The lower the overall agreement in a reproducibility study, the lower the maximum and minimum kappa values become. In Figure 3 this relationship is shown. Note that in the prevalence/

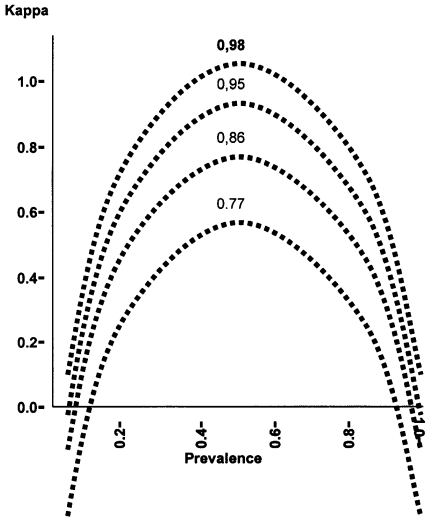


Figure 5. Relationship between the different kappa/prevalence curves and the different overall agreements ranging from 0.77 to 0.98.

kappa curves with a low overall agreement P_O (0,86 and 0.77), the minimum kappa values become negative.

The dependence of the kappa both on the prevalence P and also on the overall agreement P_O illustrates the fact that a kappa value can be interpreted in a proper fashion only when both prevalence and overall agreements are mentioned in a reproducibility study report.

1.3 Influencing the overall agreement and prevalence in advance

When performing a reproducibility study the end result may be a low kappa value because of two factors.

First, an overall agreement less than 0.85 has the risk of resulting in a low kappa value.

Therefore, it is essential that in the training phase of the study (see below) observers try to achieve a substantial overall agreement P_O preferably above 0.90. In this way the effect of the P_O on the final kappa value is under control.

Secondly, as shown above, very high and very low prevalences of the index condition result in low kappa values. Therefore, we developed a theoretical method to influence the prevalence of the index condition in advance.

In Figure 6 the prevalence/ kappa curves are presented for the overall agreements P_O ranging from 0.83 till 0.98. Note the two lowest curves (P_O 0.83 and 0.86) are located beneath the line of the kappa value of 0.6. The curves with a $P_O > 0.90$ have a substantial area above the 0.6 kappa line.

To prevent our getting a low kappa value because of too high or too low prevalences, we prefer to have a prevalence of the index condition of 0.50. The kappa values of prevalence of 0.50 are always located at the top of the curves.

Suppose we have achieved in the overall agreement period (see below) an overall agreement P_O of 0.85. We have 40 subjects in whom we can study the reproducibility of a index condition. Both observer A as well as observer B have each selected 20 subjects, and each sends his/her 20 cases to the other observer.

Each observer sends 10 subjects whom he judged to have a positive result of the test and 10 subjects whom he judged to have a negative result of the test to the other observer. Based on an overall agreement of 0.85, both observers will agree in 85% of the positive and negative judged tests and disagree in 15 %. In Figure 7 the scheme is presented.

Based on the number of subjects of agreement and disagreement in Figure 7 a kappa value can be calculated. In Figure 8 a 2x2 contingency table shows the results. The prevalence is 0.50 with a overall agreement of 0.85, resulting in a kappa value of 0.70.

By performing a overall agreement period in a reproducibility study, one can influence the prevalence in advance resulting in a substantial kappa value of a test procedure.

The easiest way of calculating the kappa value is to use a spreadsheet in which the formulae are integrated. In this way only the basic data have to be filled in and the kappa value is automatically calculated (see appendix 1). On the FIMM website a spreadsheet file can be downloaded.

P_O	RANGE PREVALENCE
0.83	0.32 -0.69
0.86	0.22 -0.78
0.89	0.17 -0.84
0.92	0.11 -0.89
0.95	0.07 -0.94
0.98	0.02 -0.96

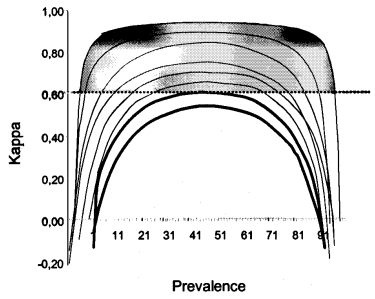


Figure 6. Kappa/prevalence curves of different overall agreements (0.83 – 0.98). The line through a kappa value of 0.60 demarcates the acceptable kappa over this line.

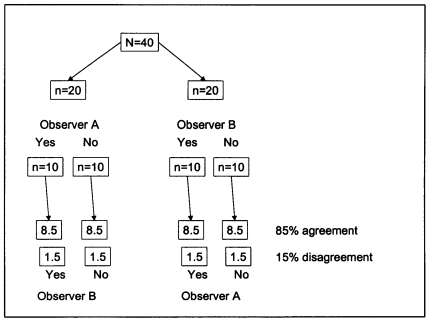


Figure 7. Scheme presenting the number of 40 subjects with an overall agreement of 0.85, trying to get a prevalence of the index condition of 0.50

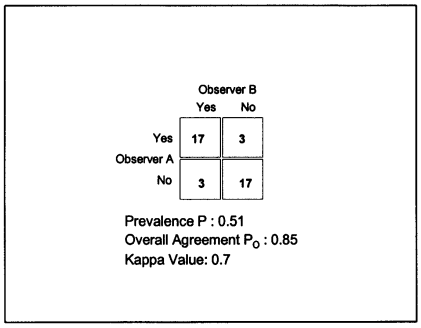


Figure 8. 2x2 contingency table based on the results of Figure 6.

2.0 Presentation kappa studies

In publishing the results of a reproducibility study, all aspects discussed under items 1-8 have to be presented. Furthermore, 2x2 contingency tables (see Figure 9), the overall agreements and the prevalences are essential in a

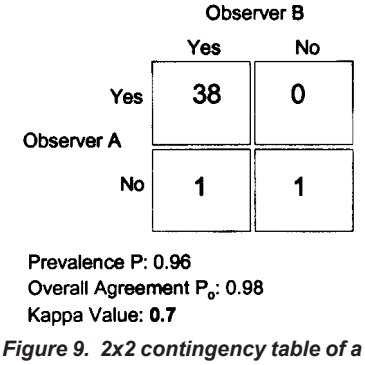


Figure 9. 2x2 contingency table of a reproducibility study of 40 subjects.

publication. In this way the reader of a paper can easily judge on what data the conclusion is based.

Figure 9 shows an example of a 2x2 contingency table, with the calculation of the kappa value shown.

V. Golden rules for a reproducibility study

In Figure 10 a scheme is presented of the different aspects and stages of a reproducibility study on which the golden rules are based.

Reproducibility studies are easy to perform and not restricted to large institutes like universities. Private practices or other institutes with two or more practitioners in M/M Medicine are very suitable for this kind of study.

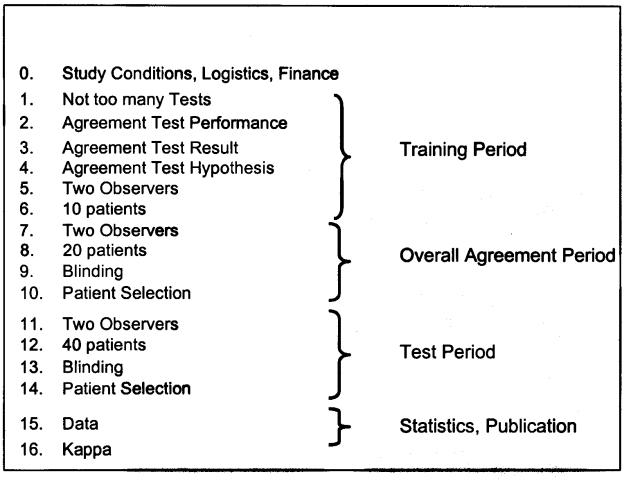


Figure 10. Plan of a reproducibility study.

Rule 1. Create a clear logistic and responsibility structure for the reproducibility study.

In a study one person must be responsible for the whole process of the study.

This person is responsible for the study logbook in which all agreements and disagreements are recorded and can be used as a reference cadre in group discussions. This person is responsible for the final format of the protocol. All participants have to sign this final protocol.

Rule 2. Always create a training period before performing a reproducibility study.

In this training period, participants have to decide what tests and how many tests they are going to evaluate.

In the training period, it is essential for the future observers of a reproducibility study to discuss and define what tests and how many tests they are going to select for the reproducibility study. The decision on how many tests one wants to evaluate is dependent of the aim of the reproducibility study.

In the training period, participants have to agree about the detailed performance of the test(s) that they are going to use for the reproducibility study.

Ten patients can be used to discuss the precise sequence of procedure of the test(s). Finally, they have to agree about the precise performance of the test(s) and make sure that each observer in a written protocol has a standardised definition of the test procedure.

Participants have to agree how to define the outcome of the test(s) they are going to use for the reproducibility study. Participants have to perform the test(s) on the same 10 patients and discuss the precise conclusions of the test(s). Finally, they have to agree about the precise judgement of the test(s) and make it sure that each observer in a written protocol has a standardised definition of the test result.

Where a combination of tests is being studied, define the minimum number of positive tests for a final positive result of the test procedure.

Participants have to agree about the hypothesis of the test(s) they are going to use for the reproducibility study. Whatever test(s) are selected for a reproducibility study, the observers have to investigate step by step the whole test procedure and agree about what the test(s) really test.

Rule 3. Always create an overall agreement period before performing a reproducibility study.

This period is essential to achieve a substantial overall agreement > 0.85. If the overall agreement is less than 0.85, participants have to discuss their agreements of the training period again.

Rule 4. Always use a blinding procedure in a reproducibility study.

In the protocol it must be clear how the blinding is achieved not only with respect to the observers but also with respect to the patients. In most protocols, except with items such as pain, blinding is guaranteed when no information is exchanged either between observer and patient or between both observers.

Rule 5. The source population from which the subjects are selected must be defined.

This is essential to prevent bias in selection of patients.

Rule 6. In publishing a paper on a reproducibility study always mention the definition of the source population, selection method, blinding procedure, definition of test procedure and test results.

Rule 7. In publishing a paper on a reproducibility study always show a 2x2 contingency table with the prevalence and overall agreement figures.

VI. Validity

1. Gold or criterion standard

After achieving good reproducibility of a test procedure (the extent to which two observers agree about a test in the same population), the validity of a test has to be assessed. Validity measures the extent to which the test actually does what it supposed to do. More precisely, the validity is determined by measuring how well a test performs against the gold or criterion standard. This is a major problem as much for diagnostics in general medicine as in M/M Medicine.

In M/M Medicine many characteristic diagnostic procedures, using for instance the end feeling in a passively performed test, are supposed to evaluate the mobility of the anatomical structure being examined. In the vast majority, only a hypothesis is available. For many tests in M/M Medicine, the gold or criterion standard has yet to be developed.

The criterion standard for a clinical test can be a radiological or surgical finding, or defined abnormal quantitative criterion based on data out of a normal population.

In M/M Medicine different kinds of diagnostic procedures are available.

Qualitative clinical tests which evaluate the observers' subjective estimate of range of lumbar motion has to be compared with the result of a quantitative method of range of lumbar motion in the same population, in order to estimate the validity of the qualitative clinical test.

Prior to this validity study, the quantitative method has to be evaluated in normal subjects to estimate the normal lumbar ranges. The evaluation of the quantitative method has also to include a test/retest procedure, to see whether the procedure shows the same data in the same normal subject on two different occasions.

The same arguments are true for tests such as the trunk list, lumbar motion patterns and mutual positions of bony structures such as pelvic distorsion.

In M/M Medicine many tests are used to estimate the mobility of a joint by means of the end feeling. In this case two different policies can be followed. First, one can develop a quantitative method to evaluate the end feeling. In this case the end feeling is validated clinically. Secondly, one can develop a quantitative method to estimate mobility of a joint. In this case, the mobility aspect of a clinical test is evaluated. In subjective testing of lumbar muscle hypertonicity, electromyographic findings can act as a gold standard.

So far, imaging techniques such as x-ray, CT and MRI are inconclusive in M/M Medicine because a large number of normal subjects show abnormalities with these techniques.

In special cases, such as the Slump Test, which evaluates dural sac irritation, for example, from postoperative lumbar adhesions, MRI with gadolinium contrast can act as the gold standard.

For some pain-provoking tests in M/M Medicine for LBP AND CPS, the criterion standard is the effect of local anaesthesia in that particular area. The problem with this kind of criterion standard is that one is never sure about the systemic effect of local anaesthetics, whether we are dealing with a referred pain area, and whether we are sure that the pain is related to the anatomical structure we want to investigate, etc.

The list of examples is far from complete, but illustrates the way a gold standard can be developed.

In the absence of a well-defined criterion standard, sometimes a consensus view of experts using other tests is used as a criterion standard. The problem with the consensus view is, that the experts are agreeing only about a test procedure based on hypothesis, while the real validity of a test remains uncertain.

In M/M Medicine, much energy has to be spent on defining criterion standards for many commonly used diagnostic procedures.

2. Sensitivity and specificity

In validity studies, 100 subjects are sufficient.

The same group of 100 patients is assessed with the test in question and with the criterion standard (see 2x2 contingency table below). Cases **a** and **d** are correct, cases **c** and **b** are respectively false positive and false negative. A good test has to have few false-positive and false-negative results.

	Criterion Standard: positive	Criterion Standard: negative	
Result of test: positive	a	b	a+b
Result of test: negative	c	d	c+d
	a+c	b+d	n=a+b+c+d

The prevalence of the index condition is illustrated by the formula:

(a+c)/ n.

It is essential to realise that the prevalence of an index condition can vary in different institutes, countries and from time to time.

The sensitivity of a test is defined as: the proportion of the cases that have the index condition **(a+c)** that the test correctly detects. In formula: **a / (a+c).**

The specificity of a test is defined as: the proportion of the cases that do not have the index condition (**b+d)** that the test correctly detects. In formula this is: **d / (b+d).**

Both sensitivity and specificity are needed to determine the validity of a test and always have to be presented together in a paper.

3. Positive and negative predictive value

To translate the statistics of sensitivity and specificity figures to daily prac-

Reliability and Validity Studies of Diagnostic Procedures in M/M Medicine

tice, the physician has to know in the individual patient the chances of a positive test being truly positive as opposed to false-positive. This is expressed in the so-called "positive predictive value of a test". In the 2x2 contingency table above, the formula of positive predictive value of a test is: **a / (a+b).** One has to realise that the positive predictive value of a test is dependent of the prevalence of the index condition **(a+c)/ n**.

Suppose we have 1000 subjects with a sensitivity and specificity of respectively 0.8 and 0.7 and a prevalence of the index condition is 10% (see 2x2 contingency table above).

This means when **n**=1000, that **a+c =** 0.10 X 1000 = **100**.

In the case of a given sensitivity of 0.8:
a/(a+c) = 0.8
a/100 = 0.8
a = 80
Since a+c = 100:
80+c = 100
c= 20
Now:
b+d = n-(a+c)
b+d = 1000 - 100
b+d = 900

In the case of a given specificity of 0.7:
d/(b+d) = 0.7
d/900 = 0.7
d = 630
So b+ 630 = 900
b = 900 - 630
b = 270

The positive predictive value of a test in this case is:
a / (a+b) = 80 / (80 + 270) = 0.22

Where there is a larger prevalence of the index condition **(a+c)/ n**, the positive predictive value of a test **a / (a+b)** also rises with the same sensitivity and specificity figures. Therefore, the positive predictive value of a test reflects the prevalence of the index condition only and not the property of the test itself.

4. Likelihood ratio

For estimation of the predictive power of a test, independently of the prevalence of the index condition, the likelihood ratio has to be calculated. By definition the likelihood ratio in formula is:

$$\text{Likelihood ratio} = \frac{\text{Sensitivity}}{1 - \text{Specificity}}$$

Tests with likelihood ratios close to 1 or < 1 are completely useless for daily practice.

First, some remarks about this likelihood ratio and its use in calculating the diagnostic confidence odds.

Normally, we are accustomed to think of percentages like prevalence or true positive figures. The likelihood ratio does not operate on percentages, but on odds based on prevalence and diagnostic certainty.

Odds are the ratio of changes in favour of a condition versus the chances against that condition being present.

For example, if a condition has a prevalence of 60%, the prevalence odds of the test being correct is 60 : 40 = 3 : 2. These odds can be changed again in decimal terms. If the prevalence odds are 3:2, the chances in favour are 3/ (3+2) = 0.6.

By mathematical calculation, the diagnostic confidence odds are calculated by multiplying the likelihood ratio and the prevalence odds.

[Prevalence Odds] X [Likelihood ratio] = [Diagnostic Confidence Odds]

To illustrate the importance of a large likelihood ratio in relation to the prevalence of a condition an example is shown.

Suppose a condition has a prevalence of 60% in your practice. Based on reproducibility and validity studies you know that the sensitivity 0.8 and the specificity are 0.98.

Based on the formula:

$$\text{likelihood ratio} = \frac{\text{Sensitivity}}{1 - \text{Specificity}}$$

the likelihood ratio is 40.

If a patient with a particular condition enters your practice, with a known prevalence figure of 60%, the chance of having this condition is 40%.

The prevalence odds in favour of having the condition are 6 : 4.

The odds for diagnostic confidence is 6/4 x 40 = 60.

Diagnostic confidence odds = 60 : 1.

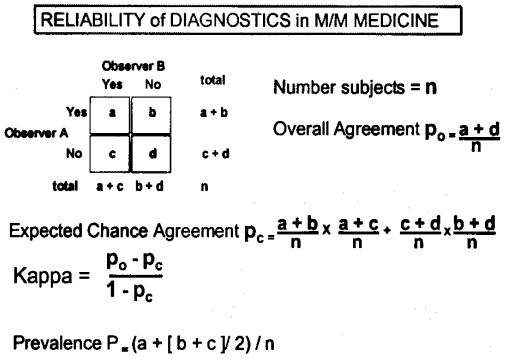
Diagnostic confidence is 60/60+1= 0.98=98%. This means that you have improved your confidence from 60% to 98%. This is a good test.

When calculating for the same prevalence of 60%, but with a likelihood ratio of 0.6, the diagnostic confidence will be only 0.47 or 47%. This is less than the chance of 60% of having the condition for a patient when entering your practice. This is a bad test.

Validity studies are important to show us the usefulness of clinical tests. Ir the likelihood ratio is high, the test will be good regardless of the prevalence of the index condition. If the likelihood ratio is close to one, the test is unhelpful.

(See Appendix 1 next page)

Appendix 1



In a spreadsheet the following columns can be defined (see figure):

Only data **a,b,c,d** has to be filled in:

Column A: data **a (see 2x2 contingency table)**	
Column B: data **b (see 2x2 contingency table)**	
Column C: data **c (see 2x2 contingency table)**	
Column D: data **d (see 2x2 contingency table)**	
Column E: data **n**	Formula =A1+B1+C1+D1
Column F: data **a+b**	Formula =A1+B1
Column G: data **a+c**	Formula =A1+C1
Column H: data **c+d**	Formula =C1+D1
Column I: data **b+d**	Formula =B1+D1
Column E: data **a+d**	Formula: =A1+D1
Column K: Prevalence	Formula: =A1/E1+B1/2*E1+C1/2*E1
Column L: Overall agreement P_0	Formula: =J1/E1
Column M: (a+b)/n	Formula: =F1/E1
Column N: (a+c)/n	Formula: =G1/E1
Column O: (c+d)/n	Formula: =H1/E1
Column P: (b+d)/n	Formula: =I1/E1
Column Q: data column M X N	Formula: =M1*N1
Column R: data column O X P	Formula: =O1*P1
Column S: Expected change agreement P_c	Formula: =Q1+R1
Column T: $P_0 - P_c$	Formula: =L1-S1
Column U: $1 - P_c$	Formula: =1-S1
Column V: Kappa value	Formula: =T1/U1

The Effects of a Pragmatic Approach to Low Back Pain Including Manual Therapy and Steroid Injections in Three Randomised Controlled Trials

Stefan Blomberg, MD, PhD, Uppsala University, Department of Public Health and Caring Sciences, Family Medicine Section and Stockholm Clinic of Manual Medicine, Stockholm, Sweden; and J Bogefeldt, MD; M Grunnesjö, DN; K Svärdsudd, MD, PhD

Dr Stefan Blomberg is a world recognised leader in musculoskeletal medicine research and will be one of the keynote speakers at our annual scientific meeting in Melbourne October 17-20. He will also conduct a two-day workshop on his specific style of managing low back pain.

Original Results Successfully Replicated in a New Study

1. The original study: A Pragmatic Approach to Low Back Pain Including Manual Therapy and Steroid Injections: A Multicentre Study in Primary Health Care (Doctoral dissertation at Uppsala University 1993 from the Department of Family Medicine, University Hospital, S-751 85 Uppsala, ISBN 91-554-3030-9, ISSN 0282-7476)

The thesis was based on the following papers:

1. Blomberg S, Svärdsudd K, Mildenberger F. A controlled, multicentre trial of manual therapy in low back pain; initial status, sick leave and pain score during follow-up. *Scand J Primary Health Care* 1992;10:170-78.

2. Blomberg S, Svärdsudd K, Tibblin G. A randomized study of manual therapy with steroid injections in low back pain; telephone interview follow up of pain, disability, recovery and drug consumption. *Eur Spine J* 1994;3: 246-54.

3. Blomberg S, Hallin G, Grann K, Berg E, Sennerby U. Manual therapy with steroid injections — a new approach to treatment of low back pain; a controlled multicenter trial with an evaluation by orthopedic surgeons. *Spine* 1994; 19: 569-77.

4. Blomberg S, Tibblin G. A controlled, multicentre trial of manual therapy with steroid injections in low back pain; functional variables, side effects and complications during four months follow-up. *Clin Rehab* 1993;7: 49-62.

5. Blomberg S, Svärdsudd K, Tibblin G: Manual therapy with steroid injections in low-back pain; improvement of quality of life in a controlled trial with four months' follow-up. *Scand J Primary Health Care* 1993; 11: 83-90.

The main conclusions were:

- That a pragmatic approach to low-back pain including manual therapy, muscle stretching and steroid injections, is superior to standardised, optimised conventional activating treatment in Swedish primary health care in:
  — facilitating recovery
  — reducing sick-leave
  — reducing pain
  — improving everyday function
  — reducing clinical findings on physical examination
  — reducing drug consumption
  — reducing common symptoms
  — improving quality of life.
- That:
  — long-term efficacy of the experimental treatment was shown. Together with an *increase* in outcome differences between the two groups for many major efficacy measures between the two and four months' follow ups, there was still a difference in sick leave in favour of the experimental group after eight months.
  — manual therapy seems to be particularly important in the long term for activities which make great demands of the lower back function such as athletics, heavy lifting, running, heavy physical work and making a bed.
  — the results concerning common symptoms and quality of life constitute indirect evidence against the belief that low back pain virtually always is of primary psychogenic origin.
  — the experimental patients were generally more content with their treatment.
  — the difference in treatment outcome in favour of the experimental group cannot be due to steroid effects only. It was also dependent on other items from the therapeutic arsenal.
  — manual therapy can achieve large public cost savings, due to reduced sick leave and drug consumption, and a considerably smaller treatment volume than in the conventionally treated group.

However, the results should be reproduced in future studies before definite conclusions are drawn.

Abstract

Blomberg S. A pragmatic approach to low back pain including manual therapy and steroid injections: A multicentre study in primary health care. Acta Universitatis Upsaliensis, Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1993; 394, 148 pp.

101 outpatients with acute or subacute low-back pain were randomly allocated to one of two treatment groups. One group was given standardised, conventional optimal activating treatment by primary health care teams. The other group received, according to a pragmatic approach, another treat-

ment program including manipulation, specific mobilisation, muscle stretching, autotraction and cortisone injections.

The two groups were similar in most of the pretrial variables, including age, sex, occupation, education, previous low back pain problems, previous treatment, sick leave, findings at the physical examination, quality of life score, presence of common symptoms, disability rating and pain score.

After one month in the study, the proportion of patients on sick leave was six times larger in the conventionally treated group than in the group receiving the specific manual treatment. The difference diminished over time but was significant concerning the average number of days on sick leave per patient during the eight months of follow-up. Significant differences in favour of manual therapy were also shown for pain score, disability rating, recovery score, drug consumption, quality of life and prevalence of common symptoms. There was also a difference in objective findings, assessed by blinded, independent and unbiased orthopaedic surgeons, in favor of the group receiving manual treatment. The blinding procedure in the objective evaluation proved to be successful. The outcome difference increased during the four months follow up for many of the efficacy measures.

As expected, the trial treatment was, due to injections and muscle stretching, more painful than the conventional treatment but only a handful of patients rated the manipulation and specific mobilisation as painful. Rare, mild, transient side effects but no major complications were reported in the two groups.

The experimental patients had a more positive opinion of the treatment than the patients did in the conventionally treated group. Parallel therapy other than the intended treatment program in the two groups was closely supervised and there was no such treatment in the experimental group and negligible parallel treatment in the conventionally treated group.

In conclusion, all of the applied efficacy measures indicated that manual therapy is superior to standardised, optimised conventional activating management of patients suffering from low back pain. Assuming that the results can be reproduced in future studies, it may also be concluded that manual therapy can reduce public costs for low back problems, since the treatment volume, drug consumption and sick leave were considerably less in the experimental group than in the conventionally treated group.

2. The reproducibility study: A Randomized Clinical Trial Comparing Four Different Treatment Regimens: Manual Therapy Including Steroid Injections, Manual Therapy, Muscle Stretching and Orthopaedic Care

Abstract

Study Design. A pragmatic, prospective, randomised controlled trial with a 2-year follow-up of return to work rates, sick leave volumes, x-rays and neuroradiological investigations and incidence of back surgery. A 10-week follow up of pain, functional variables, pain drawings, quality of life and psychosomatic symptoms.

Objectives. To replicate the results of a previous study, to investigate whether the evaluated manual treatment program is communicable to other therapists, and to evaluate the efficiency of four different treatment regimens. The present study population is compared to the earlier study population.

Summary of Background Data. Manual therapy is a subject of controversy and further studies are needed to evaluate its efficiency. A previously published study in which beneficial effects in favor of a new pragmatic manual treatment approach including steroid injections had to be replicated before more firm conclusions could be drawn.

Methods. One hundred and sixty outpatients with acute or subacute LBP were randomly allocated to one of two experimental groups (manual therapy with or without steroid injections) or two control groups (standardised orthopedic care with or without muscle stretching). Pain, disability and quality-of-life were recorded using 100 mm visual analog scales (VAS). Information on the prevalence of 27 mainly psychosomatic symptoms was obtained by questions answered by "yes" or "no". Sick leave information was obtained from the social insurance offices and other sources. The incidence of low back surgery and the rates of radiological examinations were recorded. Planned subgroup analyses in patients with radiating pain were performed.

Results. With regards to regression of over time, the manual therapy, in combination with steroid injections, was consistently superior to the three other treatments. Manual therapy without steroid injections scored better than the control treatments, while the control treatment including muscle stretching scored worse than the other control group at five weeks. No differences were found between the control groups after 10 weeks.

Conclusions

The complete pragmatic manual therapy concept combined with steroid injections is superior to the other three evaluated treatments, and are in agreement with the results from a similar and previously published study.

The method was even more effective in patients with pain radiating to the leg(s) than in patients with no radiation.

The favourable results of the experimental treatments were even more evident if the pain radiated below the knee(s).

The method is communicable to other

The Effects of a Pragmatic Approach to Low Back Pain

physicians and physical therapists, and a modest amount of education in the evaluated treatment model is enough to achieve beneficial effects.

Many of the differences in favour of the manual treatment increased during the follow up period, implying long-term treatment effects.

Steroid injections are particularly important in helping patients whose everyday life and/or work place great demands on lower back function.

The manual treatment without steroid injections is more successful than both control treatments.

Therefore, the differences in favor of the complete pragmatic treatment concept in comparison to the control treatments cannot solely be due to steroid effects but that it was also dependent on other items of the therapeutic arsenal.

The addition of muscle stretching to the basic control treatment, as a manual single therapy, seems to postpone recovery during the first weeks.

Email:
stefan.blomberg@allmmed.uu.se

Correspondence to:
Dr Stefan Blomberg
Department of Public Health and Caring Sciences
Family Medicine Section
Uppsala Science Park
SE-751 85 Uppsala, Sweden

Difficulties in Management: Chronic Patients

Dr Chris Hogan, Dr Scott Masters

Much has been written about "Heartsink Patients", that is, patients who cause their doctor's heart to sink when their name is seen in the day's appointment book. They are usually patients with ill-defined chronic problems, often involving pain management. We submit the following as a pragmatic approach to these patients

Awareness

Foremost, we need to be aware why we allow ourselves to become frustrated and overwhelmed by our patient's problems. These reasons commonly include:

- Feeling we cannot help the patient, yet the patient returns.
- The patient does not follow suggested treatment but returns anyway.
- The many-headed hydra syndrome - one problem is solved only to have two more occur in its place.

Reflection

There are several useful thoughts and questions that are useful to reflect on in these situations.

- Feeling we cannot help the patient: Might we be helping anyway? Constructive listening may assist patients in finding their own answers.
- The patient does not follow suggested treatment but returns anyway: Is it that the patient does not believe we are treating the real problem, but nevertheless believes we will eventually help?
- The many-headed hydra syndrome: What is the underlying problem- the major problem? Are their other issues the patient hasn't raised or that we haven't heard?
- Ensure the current diagnosis is correct: Have we missed one of the numerous masquerades in primary practice?

The passage of time is a useful diagnostic tool. Even in chronic undiagnosed pain, the GP's best friend, the passage of time, reveals that the diagnosis can become obvious.

Mental illness and undiagnosed pain

While working with patients with chronic undiagnosed pain it became obvious to us that around 1-2% a year would "declare" themselves as either a malignancy or a neurodegenerative disorder. A recent study in the UK[1] confirmed this suspicion and also revealed an increased risk of dying from accidents, suicide and violence.

Sadly, many of these patients have been told that the problem was all in their heads or that they were mentally ill. One of the things we can do for them is to make a definite diagnosis of sanity. There is no place for presuming a psychiatric disorder without formally checking for it. Depression can often be a reaction to pain rather than a cause. Continuously being disbelieved or thought to be crazy can also tend to make people react very defensively to the point of paranoia. If they have also had to endure the run-around of the insurance/legal game, depression and paranoia are doubly common.

Mental illness is a rare cause of chronic undiagnosed pain but it does happen. Psychiatric referral is helpful either to exclude mental illness or to help deal with it.

Drug-seeking patients

A thorough history and examination backed up by access to the notes of previous doctors is essential to determine the existing diagnoses. Note we use the plural here as dual or even multiple diagnoses are common in drug-seeking patients. As well as panic disorder and depression, many patients are victims of childhood sexual assault. This is a diagnosis they will often neither accept, remember nor acknowledge for years. It needs to be approached with patience and sensitivity.

Merely feeding a dependence helps nobody. A polite refusal to prescribe the medication sought, the offer of ongoing support and an appropriate referral for shared care is appropriate.

Emotionally dependent patients

Encouraging and supporting independence with a strict policy on contact with the surgery and visiting the surgery is helpful. Once again, previous childhood sexual assault is common. They often cause despair in the doctor due to the emotionally negative responses they give. This is typified by the adage "Misery Loves Company". Being alert to our responses to such people often helps us maintain a clinical equilibrium and prevent such an occurrence getting worse.

Dealing with these patients is difficult and wearing as the doctor can often become depressed and will need peer support through a divisional doctor-for-doctors group or such like.

Some folk are merely lonely and isolated without being clinically depressed.

Management Tips

What can be done when nothing else can be done?

There are four things to do when nothing else can be done:

1. *Education.* Ensure the patient has as much information as they need about their illness.
2. *Diet.* Ensure they receive information on an appropriate diet.
3. *Exercise.* Ensure they receive information on an appropriate exercise program. Even the dying will receive some benefit from maintaining activity levels.

4. Stress management. There are several components to stress management:

   External

   Pharmacotherapy
   Massage
   Aromatherapy
   Music therapy

Difficulties in Management: Chronic Patients

Acupuncture

Internal

Relaxation exercises and procedures: positive imaging, the anti-torture techniques used by SAS and SEALS, which include breathing, muscle contraction and relaxation exercises

Spiritual: reconciliation and forgiveness for those who have wronged them

Relationship counselling

Meditation

Prayer

Sexual counselling: even the ill are entitled to a good sex life.

Summary

The above outline attempts to enable practitioners to formulate an approach for even the most difficult undifferentiated presentations to primary practice. At all times we must remember, it is the patient who has the problem. We need to actively avoid being dragged into the patient's problem. If we stand by their side, empathise and keep realistic and positive actions foremost in their cognitions, we will have achieved much.

Correspondence to:
Dr Chris Hogan, GP
Sunbury, Victoria 3429
and
Dr Scott Masters
Caloundra Sports Medicine Centre, Suite 4 "Trinity House", 39 Minchinton St., Caloundra Qld 4551

Practice Tip: "Pseudo" Gout not Pseudo-Gout

Dr Peter Jackson, Musculoskeletal Physician, Brisbane

Occasionally, in clinical practice, a patient presents with pain in the region of the first metatarsophalangeal joint. The good doctor, silently patting himself on the back for making the easy diagnosis of gout, prescribes appropriate medication and refers the patient for a serum uric acid assay.

Sometimes, when the result returns, it is found to be within the normal limits. The doctor, with a sickly smile, after remembering a clinical elite telling him during his residency that "you can still have gout with a normal uric acid level" tells the usually still suffering patient that he has gout despite a normal test.

What he has probably forgotten from his second or third year anatomy classes is that muscles that move joints share a common innervation. There are other causes for pain referred to this area, namely, muscles with local insertions. The most common is the tibialis anterior which inserts into the plantar surface of the medial cuneiform and the first metatarsal. The offending trigger points (TrPs) are located a long way proximally in the belly of the muscle adjacent to the lateral surface of the upper one-third of the tibia where usually five or so exquisite trigger points can be palpated. Surprisingly, patient's are unaware of this tenderness until the palpatory experience occurs.

If the examination of this muscle proved futile the other common offending muscles are the flexor hallucis brevis and the abductor hallucis. If you haven't looked lately, they both insert into the medial sesamoid bone and the base of the proximal phalanx. Whilst there are insertion enthesis trigger points, the major ones are again found in the middle of the muscle belly.

Treatment involves mechanically deactivating the TrPs with a wet or dry needle, depending on your preference, and teaching your patient how to post-isometrically stretch the dysfunctional muscles.

Journal Abstracts

This section aims to update the reader with some of the more significant musculoskeletal research published in the last year which is listed on the Medline and CINAHL databases.

SPINAL

Hogeboom CJ, Sherman KJ, Cherkin DC.Variation in diagnosis and treatment of chronic low back pain by traditional Chinese medicine acupuncturists. *Comp Ther in Med* 2001:9: 154-66.

Objectives. To assess interrater reliability of Traditional Chinese Medicine (TCM) diagnosis and treatment of chronic low back pain.

Design. Under a Latin square design, six TCM acupuncturists evaluated the same six patients on the same day.

Setting. Northwest Institute of Acupuncture and Oriental Medicine, Seattle, Washington.

Interventions. Assessment only.

Outcome measures. TCM diagnosis, acupoint prescriptions, auxiliary treatment recommendations.

Results. Twenty diagnoses and 65 acupoints were used at least once. The diagnosis of Qi/Blood Stagnation with Kidney Deficiency and the acupoint UB23 were used for every patient by most acupuncturists. However, consistency across acupuncturists regarding diagnostic details and other acupoints was poor. No diagnoses, and only one acupoint, were used preferentially for a subgroup of patients. Some diagnoses and treatment recommendations were dependent more on the practitioner than on the patient. Fine-grained diagnoses and most acupoints were unrelated to either patient or practitioner.

Conclusions. TCM diagnoses and treatment recommendations for specific patients with chronic low back pain vary widely across practitioners. Acupuncture clinical trials using an individualised treatment arm may be difficult to replicate or evaluate because of low concordance among acupuncturists. Comparison of individualised treatment with a thoughtfully developed standardised approach is warranted to determine which, if either, is superior.

Acupuncture meridian abbreviations: UB, Urinary Bladder.

Comment. Problems with interrater reliability thwart attempts to assess validity of acupuncture diagnosis and treatment. This seems to be a problem common to many forms of segmental pain diagnosis and therapy. – *Dr David Roselt*

Hunt JL, Winkelstein BA, Rutkowski MD, et al. Repeated Injury to the Lumbar Nerve Roots Produces Enhanced Mechanical Allodynia and Persistent Spinal Neuroinflammation. *Spine* 2001: 26: 2073-79.

Study Design. A lumbar radiculopathy model investigated pain behavioural responses after nerve root reinjury.

Objectives. To gain a further understanding of central sensitisation and neuroinflammation associated with chronic lumbar radiculopathy after repeated nerve root injury.

Summary of Background Data. The pathophysiologic mechanisms associated with chronic radicular pain remain obscure. It has been hypothesised that lumbar root injury produces neuroimmunologic and neurochemical changes, sensitising the spinal cord and causing pain responses to manifest with greater intensity and longer duration after reinjury. However, this remains untested experimentally.

Methods. Male Holtzman rats were divided into two groups: a sham group having only nerve root exposure, and a chromic group in which the nerve root was ligated loosely with chromic gut suture. Animals underwent a second procedure at 42 days. The chromic group was further divided into a reinjury group and a chromic-sham group, in which the lumbar roots were only re-exposed. Bilateral mechanical allodynia was continuously assessed throughout the study. Qualitative assessment of spinal cord glial activation and interleukin-1beta expression was performed.

Results. Mechanical allodynia was significantly greater on both the ipsilateral and contralateral sides after reinjury ($P < 0.001$), and the response did not return to baseline after reinjury, as it did with the initial injury. There were also persistent spinal astrocytic and microglial activation and interleukin-1beta expression.

Conclusions. The bilateral responses support central modulation of radicular pain after nerve root injury. An exaggerated and more prolonged response bilaterally after reinjury suggests central sensitisation after initial injury. Neuroinflammatory activation in the spinal cord furthers supports the hypothesis that central neuroinflammation plays an important role in chronic radicular pain.

Comment. As stated, "central sensitisation refers to an enhanced responsiveness of the central nervous system to afferent input and is defined as a decreased threshold, an increased response to suprathreshold stimuli (i.e. hyperalgesia) and ongoing spontaneous activity in the dorsal horn." This study of *reinjury* in the same rat lumbar nerve root is the first described, according to the authors. Activation of contralateral dorsal horn neurones is thought to evoke *mirror pain*, which has been described after unilateral cordotomy for intractible pain and indicates central sensitisation. This is implicated in persistent pain behaviour. – *Dr David Roselt*

Sherman KJ, Hogeboom CJ, Cherkin DC. How traditional Chinese medicine acupuncturists would diagnose and treat chronic low back pain: results of a survey of

Journal Abstracts

licensed acupuncturists in Washington State. *Comp Ther in Med* 2001: 9: 146-53.

Objectives. This survey was undertaken to learn how Traditional Chinese Medicine acupuncturists diagnose and treat patients with chronic low back pain in order to develop a standardised treatment for a clinical trial of that condition

Design. We surveyed a randomly selected group of 56 acupuncturists in Washington State, USA, about styles of acupuncture they used for treating chronic low back pain, diagnoses made, and key features of treatment for this condition.

Results. While substantial variability existed among practitioners, there was agreement on several broad features of treatment including the use of local and distal acupuncture points (86% of practitioners), the use of acupuncture points on the meridians traversing the back (especially the UB meridian, 90%) the use of acupoints determined by palpation (82%), the importance of eliciting de qi (60%), and of providing up to eight treatments for achieving therapeutic results (79%).

Conclusion. The use of practitioner surveys can enhance the systematic development of acupuncture treatment protocols and should be part of this process in future clinical trials of common conditions.

Comment. Western medical acupuncture assessment and treatment has traditionally been eclectic in nature and individualised to the patient. This paper shows how a representative acupuncture treatment protocol may be arrived at for use in randomised controlled trials. Most treatments seem to involve the use of locally tender points that when needled could have segmental spinal gating effects involving met-enkephalin. Distal points are purported to have other neuromodulatory effects by increasing descending inhibition via serotonergic and noradrenergic pathways that also involve met-enkephalin in the dorsal horn, and releasing beta-endorphin centrally. – *Dr David Roselt*

Wittenberg RH, Oppel S, Rubenthaler FA, Steffen R. Five-Year Results From Chemonucleolysis With Chymopapain or Collagenase. *Spine* 2001; 26: 1835-41.

Study Design. A 5-year clinical follow-up assessment of a prospective randomised study of chemonucleolysis using chymopapain (4000 IU) or collagenase (400 ABC units) was performed.

Summary of Background Data. Intradiscal therapy can be performed for patients with contained discs by chemonucleolysis, percutaneous discectomy, or laser ablation. The oldest intradiscal therapy is chemonucleolysis with chymopapain.

Objective. The purpose of this study was to compare prospectively the efficacy of chymopapain and collagenase for intradiscal injection.

Methods. In this study, 100 patients with indication for intradiscal therapy were prospectively randomised to treatment with either chymopapain or collagenase. All the injections were performed by the double-needle technique with the patient under general anaesthesia. The mean age of the patients was 35.5 years in the chymopapain group and 38 years in the collagenase group. An equal number of injections were performed at L4-5 and L5-S1.

Results. After 5 years, good and excellent results were observed in 72% of the chymopapain group and 52% of the collagenase group when the surgically treated and lost patients were graded as poor. Using a scale of 0 (no pain) to 10 (intractable pain), the pain level dropped from 8.5 to 0.7 in the chymopapain group and from 8.6 to 0.9 in the collagenase group. Microdiscectomy at the injected level was required for 23 patients (14 in the collagenase group and 9 in the chymopapain group).

Conclusions. After 5 years, no deterioration had occurred, as compared with the 1-year follow-up assessment. Chymopapain has proved to be safe, with one minor anaphylactic reaction, and effective even over the long term. Collagenase may need further study and can not be recommended at this time.

Comment. Chymopapain is an attractive option for contained discs causing radicular pain or radiculopathy. The study indications were for "sciatic leg pain stronger than back pain" and "clear clinical signs for nerve root irritation" with CT and MRI correlation implying radicular pain. Collagenase (which purportedly has less allergenic potential than chymopapain) results were not as good and 18% experienced a neurological deficit following injection that usually responded to remedy with microdiscectomy. This was more of a problem than the relatively minor allergic problems seen with chymopapain – 12% overall (mostly flushing and itch) but only 2% were clinically obvious and significant in the form of an anaphylactic reaction resulting in slight laryngospam and flushing. Chemonucleolysis with chymopapain is much safer than surgery for radicular pain with contained discs when surgical and peri-operative complications are examined, and microdiscectomy is usually successful (65-75%) after failed chemonucleolysis and could be reserved for this situation. – *Dr David Roselt*

Vad VB, Bhat AL, Lutz GE, Cammisa F. Transforaminal Epidural Steroid Injections in Lumbosacral Radiculo-

pathy. A Prospective Randomised Study. *Spine* 2002; 27: 11-16.

Study Design. A prospective study randomised by patient choice from the private practice of a single physician affiliated with a major teaching hospital was conducted.

Objectives. To compare transforaminal epidural steroid injections with saline trigger-point injections used in the treatment of lumbosacral radiculopathy secondary to a herniated nucleus pulposus.

Summary of Background Data. Epidural steroid injections have been used for more than half a century in the management of lumbosacral radicular pain. At this writing, however, there have been no controlled prospective trials of transforaminal epidural steroid injections in the treatment of lumbar radiculopathy secondary to a herniated nucleus pulposus.

Methods. Randomised by patient choice, patients received either a transforaminal epidural steroid injection or a saline trigger-point injection. Treatment outcome was measured using a patient satisfaction scale with choice options of 0 (poor), 1 (fair), 2 (good), 3 (very good), and 4 (excellent); a Roland-Morris low back pain questionnaire that showed improvement by an increase in score; a measurement of finger-to-floor distance with the patient in fully tolerated hip flexion; and a visual numeric pain scale ranging from 0 to 10. A successful outcome required a patient satisfaction score of 2 (good) or 3 (very good), improvement on the Roland-Morris score of 5 or more, and pain reduction greater than 50% at least 1 year after treatment. The final analysis included 48 patients with an average follow-up period of 16 months (range, 12-21 months).

Results. After an average follow-up period of 1.4 years, the group receiving transforaminal epidural steroid injections had a success rate of 84%, as compared with 48% for the group receiving trigger-point injections ($P < 0.005$).

Conclusion. Fluoroscopically guided transforaminal injections serve as an important tool in the nonsurgical management of lumbosacral radiculopathy secondary to a herniated nucleus pulposus.

Comment. Well-designed study giving us long-term data on epidurals for radicular leg pain. This gives doctors more reassurance about the efficacy of this injection technique. The lack of a blinded control is the major weakness of this trial. - *Dr Scott Masters*

PERIPHERAL JOINTS

Morgan-Jones R, Watson AS, Cross MJ, Saldanha JD. The Meniscal "Pseudocyst" A Clinical Sign of a Torn Meniscus. *Am J Sports Med* 2001; 29: 543-54.

We report a study of 636 patients requiring knee surgery, all of whom underwent detailed preoperative assessment. Fifty-eight patients had a clinical sign of a lump on the joint line when the knee was examined at 45° of flexion, which has been thought to indicate a meniscal cyst. Of these 58 patients, however, only 30 patients had a meniscal cyst demonstrated at surgery. The remaining 28 patients had a meniscal tear without a cyst. In these 28 cases, the clinical sign of a lump protruding from the joint line was termed a "pseudocyst." This new clinical sign is important because of its frequency of occurrence and the complete correlation with meniscal tears requiring surgical intervention.

Comment. Add meniscal tear to the differential diagnosis for lumps in or near the joint line in the knee. Other non-meniscal possibilities include loose bodies, bursae and exostoses. All "pseudo-cysts" in this series involved the lateral joint line, and were most prominent at 45° flexion. – *Dr David Roselt*

Struijs PAA, Smidt N, Arola H, et al. Orthotic devices for tennis elbow: a systematic review. *Br J Gen Pract* 2001; 51: 924-29.

Lateral epicondylitis (tennis elbow) is a frequently reported condition. A wide variety of treatment strategies have been described. As yet, no optimal strategy has been identified. The aim of this review was to assess the effectiveness of orthotic devices for treatment of tennis elbow.

An electronic database search was conducted using MEDLINE, EMBASE, CINAHL, the Cochrane Controlled Trial Register, Current Contents, and reference lists from all retrieved articles.

Experts on the subjects were approached for additional trials. All randomised controlled trials (RCTs) describing individuals with diagnosed lateral epicondylitis and assessing the use of an orthotic device as a treatment strategy were evaluated for inclusion. Two reviewers independently assessed the validity of the included trials and extracted data on relevant outcome measures. Dichotomous outcomes were expressed as relative risks and continuous outcomes as standardised mean differences, both with corresponding 95% confidence intervals. Statistical pooling and subgroup analyses were intended. Five small-size RCTs (n = 7-49 per group) were included. The validity score ranged from three to nine positive items out of 11. Subgroup analyses were not performed owing to the small number of trials. The limited number of included trials present few outcome measures and limited long-term results. Pooling was not possible owing to the high level of heterogeneity of the trials. No definitive conclusions can be drawn concerning effectiveness of orthotic devices for

Journal Abstracts

lateral epicondylitis. More well-designed and well-conducted RCTs of sufficient power are warranted.

Comment. Included in this study were treatment with braces, splints, casts, bands and straps. It is impossible to blind the caregivers and patients so these items score negatively in assessing methodological quality with validity scores for the trials. *– Dr David Roselt*

IMAGING

Oatridge A, Herlihy AH, Thomas RW, et al. Magnetic resonance: magic angle imaging of the Achilles tendon. *Lancet* 2001; 358: 1610-11.

Tendons do not normally produce detectable signals with conventional magnetic resonance techniques and are recognised as dark signal voids. However, if tendons are examined at 55° to the static magnetic field (the "magic angle"), signals become detectable and the tendons can become the brightest structure on the image. We have used this approach to establish tendon relaxation times and magnetisation transfer ratios and to show contrast enhancement. We have also shown more detail of acute and chronic tendon rupture by this method compared with images made with the tendon parallel to the static magnetic field.

Comment. This is an additional means of imaging tendons and ligaments (using what was a recognised artefact formerly avoided)in those instances where clinical examination with or without ultrasound are insufficient to guide management. It involves unconventional patient positioning but is otherwise simple and could be used on any clinical MR system. *– Dr David Roselt*

Pfirrmann CWA, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic Resonance Classification of Lumbar Intervertebral Disc Degeneration. *Spine* 2001; 26: 1873-78.

Study Design. A reliability study was conducted.

Objectives. To develop a classification system for lumbar disc degeneration based on routine magnetic resonance imaging, to investigate the applicability of a simple algorithm, and to assess the reliability of this classification system.

Summary of Background Data. A standardised nomenclature in the assessment of disc abnormalities is a prerequisite for a comparison of data from different investigations. The reliability of the assessment has a crucial influence on the validity of the data. Grading systems of disc degeneration based on state of the art magnetic resonance imaging and corresponding reproducibility studies currently are sparse.

Methods. A grading system for lumbar disc degeneration was developed on the basis of the literature. An algorithm to assess the grading was developed and optimised by reviewing lumbar magnetic resonance examinations. The reliability of the algorithm in depicting intervertebral disc alterations were tested on the magnetic resonance images of 300 lumbar intervertebral discs in 60 patients (33 men and 27 women) with a mean age of 40 years (range 10-83 years). All scans were analysed independently by three observers. Intra- and interobserver reliabilities were assessed by calculating kappa statistics.

Results. There were 14 Grade I, 82 Grade II, 72 Grade III, 68 Grade IV, and 64 Grade V discs. The kappa coefficients for intra- and interobserver agreement were substantial to excellent: intraobserver (kappa range, 0.84-0.90) and interobserver (kappa range, 0.69-0.81). Complete agreement was obtained, on the average, in 83.8% of all the discs. A difference of one grade occurred in 15.9% and a difference of two or more grades in 1.3% of all the cases.

Conclusion. Disc degeneration can be graded reliably on routine T2-weighted magnetic resonance images using the grading system and algorithm presented in this investigation.

Comments. As stated by the authors, "the signal characteristics of the disc in T2-weighted MRIs reflect changes caused by ageing or degeneration," and reflects the proteoglycan content of the nucleus. Previous reliability studies looking at degree of disc extension beyond the interspace gave lower kappa values of 0.58 to 0.71. The reliability in this non-invasive study is in keeping with studies grading gross morphology on pathology specimens. There is a good association between the finding of a high intensity zone and IDD (internal disc disruption) but even then disc stimulation is required to confirm a particular disc as being a source of pain. We look forward to the validity studies. *– Dr David Roselt*

PAIN

Mantyselka P, Kumpusalo E, Ahonen R, Takala J. Patients' versus general practitioners' assessments of pain intensity in primary care patients with non-cancer pain. *Br J Gen Pract* 2001; 51: 995-97.

Summary. Pain is a major reason for visiting a primary care physician. There are, however, few studies on the assessment of pain patients at the primary care level. The aim of this cross-sectional study was to investigate the concordance between general practitioners (GPs') and patients' assessments of pain intensity and whether this assessment is influenced by the duration or intensity of pain. Seven hundred and thirty-eight pa-

tients aged 16 to 75 years who were visiting a GP because of pain participated. Both the patients and the GPs rated pain intensity using the horizontal 100 mm Visual Analogue Scale (VAS). Means and correlations were calculated using non-parametric tests. The VAS scales were arbitrarily divided into five grades (one unit = 20 mm) to investigate the concordance between GPs' and patients' assessments of pain intensity. Spearman's correlation coefficient between GPs' and patients' assessments was 0.31 for non-chronic pain (of duration less than six months) and 0.20 for chronic pain. GPs evaluated graded pain intensity at least one unit lower than the patients in 37% of the visits. In one-fifth of the visits (2 0.5%), the GP's rating was at least two units lower than the patient's rating. The more severe the pain as assessed by patients the greater the non-concordance between patients' and GPs' assessments. There was considerable non-concordance between GPs' and patients' assessments of pain intensity. GPs tended to estimate their patients' pain intensity as clinically significantly lower than the patients themselves, particularly in chronic and severe pain.

Comment. Pain is a major reason for presentation in primary care, and chronic benign musculoskeletal pain is a modern epidemic. Patients with these problems are generally less satisfied with their care than those with other chronic conditions and this may be due to underestimation of pain leading to inadequate management. This may lead to development of chronic pain. The visual analogue scale (VAS) is a simple validated technique suitable for use in primary care as part of the clinical assessment for measuring a patient's subjective experience of pain. Pain needs to be measured so that it can be managed. – *Dr David Roselt*

INJURY

Norton K, Schwerdt S, Lange K. Evidence for the aetiology of injuries in Australian football. *Br J Sports Med* 2001; 35: 418-23.

Objectives. To determine in Australian football (a) the influence of ground hardness and playing grade (level) on game speed and structure, and (b) player movement patterns throughout the game and across levels.

Methods. The design consisted of several studies. Seventeen games played on grounds of different hardness in 2000 were used to determine game speed and structure. Four first grade and four second grade grand final games (1994, 1996, 1997, 1999) were used to determine the game speed and structure on the same ground but at different levels. Fifty-one players (44 first grade and seven second grade) were used to measure movement patterns within games and across levels during the 2000 season.

Results. There was a significant relation between ground hardness and game speed, which could lead to higher injury rates when the ground is harder. There was a 6.7% difference in game speed between the first and second grade levels reflecting differences in injury incidence. The first grade games were also characterised by a greater number of shorter, high intensity play periods and longer stop periods than the second grade games. Midfield players in the first grade games covered about 24% greater distance than their second grade counterparts, and there was a significant difference in their playing speeds.

Conclusions. Over the past 40 years, the game speed in the top level of Australian football has approximately doubled. Over the same time, the number of collisions and the estimated injury incidence have also doubled. This study provides additional support to the suggestion that these variables are strongly linked. Factors such as ground hardness, playing level, and time during the game influence game speed and are therefore important in injury development in Australian football.

Comments. Australian football is a collision sport and injury rates are amongst the highest for any sport played in Australia. It may impact on long-term health with chronic injuries and incapacitation in retired players. Ground hardness can be measured with a "penetrometer" apparently that measures the depth of penetration in centimetres. Harder surfaces were associated with faster game speeds and other things being equal, higher collision impact forces. Possible solutions mooted include moulding the game into a slower one such as in the hybrid international rules game against Ireland or use of protective equipment. – *Dr David Roselt*

The idea of ground hardness and injury rates has been studied by Dr John Orchard in the context of AFL. He found that ACL injury is related to ground hardness. He postulates that the likely mechanism is shoe-surface traction. It's possible that ground watering and softening, playing games during winter months, using natural grasses such as perennial ryegrass and using shorter studs on boots all may help reduce non-contact lower limb injuries. – *Dr Scott Masters*

OSTEOARTHRITIS

Sowers M. Epidemiology of risk factors for osteoarthritis: systemic factors. *Curr Opinion in Rheumatol* 2001, 13: 447-51.

Osteoarthritis (OA) appears to be a mechanically driven but chemically mediated disease process in which there is attempted (or aberrant) repair. Well-established risk factors for OA include ageing, obesity, gender, and,

Journal Abstracts

in selected subgroups, congenital anomalies. This review addresses less well-established risk factors for OA that can impact joints through their effect on systemic metabolism rather than their contribution to local joint geometry and structure. These systemic risk factors include obesity; bone and bone density; nutrients, particularly those that function as antioxidants; and genetic factors. There is great opportunity for new prevention and intervention strategies as we expand our understanding of the role of these systemic risk factors.

Comment. Prospective data on women suggests the risk for knee OA is increased by 15% for each additional kg/m^2 of body mass index above 27. OA is a genetically heterogeneous disorder. Understanding of systemic risk factors is still in its infancy but hopefully will lead to new intervention strategies. – *Dr David Roselt*

FIMM Report

Dr Ron Palmer, Vice-President FIMM

As you are all probably aware, FIMM is the international organisation representing manual/musculoskeletal medicine. As the membership transverses so many nations, 26 in all, there is often some difficulty in arriving at a general consensus in some fields. In an endeavour to overcome many of the root problems, the Education and Scientific Committees have put in place a program to attempt to standardise the acceptable knowledge base and the teaching protocol. Australia is well represented on these committees and through myself as vice-president, has some input into the direction of the general thrust. As with our current state in Australia, there is a battle with other medical groups and a reluctance of governments to accept our legitimate claim as a specialist group.

There is a vast degree of difference in the standard of medicine practised around the world. It is with awareness of this fact that the committees are working to arrive at a standard that is acceptable to all members and one that is sufficiently advanced to justify our being a specialty group in our own right. Unfortunately, as is our own experience here in Australia, there are no "short cuts" to gaining recognition and the changes in government can greatly vary the outcomes. New Zealand has specialist recognition and Australia has not. The same situation applies overseas where Russia has full recognition, Switzerland part recognition, and the UK none. Unfortunately this does not mean that the MSM standard in the UK is inferior to that which exists in Russia, exactly the same as our own standards are not inferior to those of the Kiwis. The fight for our independent specialist group must continue and just as rightly so, the world body through FIMM must continue to fight for specialist ranking.

The second edition of 'Reliability and Validity Studies' has just been released by the Scientific Committee. A copy has been forwarded to the editor and may appear in print in this journal. When reading such a review it can be so easily assumed that not a lot of effort or time is consumed in this type of publication. That would be so wrong. To compile so lengthy a scientific document does not happen overnight. Further, while science itself (or evidence based medicine) does not have national boundaries, it nevertheless must be acceptable in theory for all member countries or there is no hope of upgrading the standard of MSM across the board. Likewise, the Education Committee has to produce teaching material that reflects the Scientific Committee's work and to have this successfully taught around the world implies that it must be accepted by all nation members. You can now perhaps see some of the difficulties we face in unifying our field of medicine on the grand scale. Norm Broadhurst is Australia's representative on the Education Committee, having taken over when Phil Watson resigned.

FIMM, like the AAMM, has a printed list of definitions of terminology in use within MSM. This was completed three years ago in the UK and is currently being upgraded.

There is always little in the way of bouquets for this type of tedious work and probably few current AAMM members will remember the effort contributed by Wade King in our local version. Unfortunately this behind-the-scene work is essential if our group is to progress.

The Policy Committee met last December in Prague and while Australia has no direct representative on this committee, I had some input as financing FIMM costs was a major issue. Victor Dvorak (Switzerland) had similar thoughts to my own and we are now investigating the possibilities. Running an organisation like FIMM is not cheap. Airfares and accommodation eat into the budget. Yet if you wish to be democratic, there is no easy way around the problem. Of course most work is done via e-mail, but there still has to be a meeting once a year where individual differences can be thrashed out. No, a telephone hook up as is done with the Otago Diploma Course would not work. While everyone speaks English there is sufficient accent and word interpretations to make a "round table" meeting an essential item. Therefore there is not a simple answer to grossly reducing expenditure. We now hold meetings at universities and seek the cheapest accommodation available. All committee meetings are kept to the bare minimum.

The next Scientific Committee meeting is scheduled for Prague in May. Prague is virtually central to most European nations and costs there are greatly inferior to say UK or France. It makes logical sense to use such a base. The only long airflights to get there are the USA and Australia. We will be holding this meeting at the Charles University (home base of Einstein). We are not under the illusion of being in quite the same street. The topic for this meeting is the shoulder girdle, as we have already completed the cervical spine, the lumbar region and reviewed validity and reliability as well as efficacy. The data from these reviews are presented at MSM conferences around the world and are to be used as a basis for the Education Committee work. Also much of it is available on our FIMM website.

The Scientific Committee has been invited to present a full half-day of our work to the Swiss Society's conference at Interlaken in November/December this year. The exact material has not been finalised, but will be done in Prague in May.

My own topic for the May meeting is "Shoulder Pathology and its Relation to Distal Regions of the Locomotion System". Hopefully this PowerPoint presentation will be converted into written form for possible publication later in the year.

FIMM Report

In closing I would like to point out that MSM as taught in Australia is as good as anywhere else on the planet and probably a lot better than in most countries. It is still in our interest to push hard to increase the standard of medical practice across the world. We are now closing distance via electronic communication and we increasingly need to upgrade our own standards as well as those of other countries. The Australian Initiative run by Nik Bogduk was presented at the Chicago Triennial Conference and was well received. The standards set by this material need to be introduced everywhere. FIMM, via its committees, is attempting to do this. The next Triennial will be held in Bratislava, Slovakia, in 2004. There will be update on material about this in the next journal.

Letters to the Editor

Australasian Faculty of Musculoskeletal Practice Standards and Protocols, November 2001

Dear editor

The authors of these guidelines are to be commended on their call for rigour and painstaking conscientiousness in diagnosis and management of "mechanical" lumbar and cervical pain. However, one or two important omissions regarding selection of patients might mislead the more naive physician to believe that these guidelines could or should be applied in the real world.

It surely does not go without saying, under "Selection of patients" that "... all such patients should have undergone a comprehensive trial of manual/manipulative treatment with appropriate exercise prescription, and under "Relative Contraindications" that "... patients who exhibit signs of abnormal illness behaviour, somatisation and/or central sensitisation phenomena of a chronic pain disorder are not suitable candidates for this protocol.

Furthermore, in the UK, the "real world", MRI is used only as a presurgical investigation and not a routine first-line investigation as shown in the algorithm.

One might also question the usefulness of pursuing an exact tissue diagnosis, that is, discogenic pain, C0-1, C1-2, C2-3 joint pain when there is no evidence based treatment available unless one is selecting patients for a clinical trial of an experimental treatment. A clinical judgement surely must be made on every individual, taking into account the patient's expectations, level of pain, degree of disability and likelihood of successful therapeutic outcome before even embarking on this costly and labour-intensive exercise.

Yours sincerely

John Tanner
Musculoskeletal and Sports Physician
West Sussex, UK

Reply from Professor Bogduk

Any confusion stems from stems the double negtives involved.

Revel DID NOT find POSITIVE features that correlated with positive responses to diagnostic blocks. He found that patients whose pain was aggravated by coughing, flexion, etc, DID NOT get relief from blocks of the Z Joints. Accordingly he found that extension-rotation was not a predictor of Z joint pain, contrary to what others have asserted.

One way of making sense of this is to imagine that coughing, flexion, and extension catch are features of IDD. Therefore, when present, these features are indicative of NOT Z joint pin. When present they mean IDD and NOT Z joint pain, but consequently and reciprocally, when ABSENT these features are not diagnostic of Z joint pain, but their ABSENCE increases the likelihood of Z joint.

Professor Nikolai Bogduk
Newcastle Bone and Joint Institute

Answers to Quiz

(Questions on page 22)

Q1. The medial popliteal fossa. Pain in the popliteal fossa and mechanical symptoms. It is sometimes asymptomatic. Either osteoarthritis, inflammatory arthritis or meniscal tear.

Q2. True, though this is not common.

Q3. False. The cyst settles usually when the underlying pathology is treated. Aspiration is difficult and the cyst often recurs. Excision of the cyst is rarely required.

Q4. Pes anserinus bursitis. This bursa is located between the conjoined distal tendons of the sartorius, gracilis and semitendinosus muscles and the tibial insertion of the medial collateral ligament. Symptomatic treatment and cortisone usually help. Synovial sarcoma is rare but could potentially mimic cysts and bursitis around the knee. Nocturnal pain and other related symptoms warrant investigation of unusual cysts.

Q5. Femoral anteversion, genu valgus, hyperpronation of forefeet, small high flattened patella, loose medial structures and tight lateral structures, shallow femoral groove, poorly developed vastus medialis, high Q angle (the angle formed by the intersection of lines drawn from the anterior illiac spine and tibial tubercle through the mid-point of the patella. The normal Q angle is less than 20°). It is more often abnormal in females and those with endomorphic body type especially if there is recurvatum and valgus knees. These factors may also contribute to ongoing retropatellar pain without episodes of subluxation.

Q6. Anterior cruciate ligament tear. Sixty to eighty percent of acute haemarthrosis is due to anterior cruciate trauma. Approximately 80% of patients experience a popping or tearing sensation. The two most important functions of the anterior cruciate ligament is to cause the screw home external rotation of the tibia on the femur as the knee extends. In the sidestep, excessive internal rotation of the tibia occurs and the ACL is tightened. The second most important function of the ACL is to resist anterior displacement of the tibia. Posterior cruciate ligament injuries are less common and harder to detect.

Q7. Grade 1 injuries are assisted to normal weightbearing within a week and with the assistance of physiotherapy are usually back at sport in 3-6 weeks. Grade 2 injuries with a significant laxity may require a period of partial weightbearing with or without a knee range immobilising brace for four weeks. With the assistance of therapy they can be back at sport between six and eight weeks. Grade 3 injuries usually don't require surgery if other structures are intact and are treated similarly to Grade 2. Initially the patient may non-weightbear for 5-10 days. By four weeks normal weightbearing is resumed and, depending on progress, return to sport within 6-10 weeks (Crichton et al, 1992). Studies (Meislin, 1996) report that immobilisation significantly diminishes the ultimate load capacity of the MCL and increases osteoclastic activity at the tibial MCL insertion sites. X-rays should be performed to rule out intra-articular fractures.

Q8. True. Older patients who have a degenerative meniscal tear may not recall a specific incident.

Q9. Femoral sarcoma, osteomyelitis, septic arthritis and discitis. Clinical recognition in young patients with osteomyelitis may be challenging because classic symptoms of chills, fever, pain and swelling may be absent. Osteomyelitis may also mimic soft tissue infection. Osteomyelitis in children most commonly occurs in the distal metaphysis of long bones such as tibia and femur. X-ray findings may not be detectable up to two weeks post-infection, 30-50% of bone mineral must be lost before plain x-rays reveal osteomyelitic changes. Delayed therapy leads to poor outcome but early treatment often leads to return to competitive activity without disability.

Q10. Meniscal lesion and/or osteochondritis dissecans (OCD). With OCD clinical findings are inconclusive but one consistent finding is thigh atrophy. Meniscal stress tests are often positive. The cause is multifactorial such as repetitive stress, skeletal abnormalities due to endocrine dysfunction and abnormal ossification of the epiphyseal cartilage. The end result is avascular necrosis of subchondral bone and changes in the overlying articular cartilage. The less the separation of the lesion the less the severity of the mechanical symptoms, hence those in this category still possess sufficient proprioception and agility for sport. OCD may go undiagnosed for a long time and should therefore be considered in differential diagnosis of all diffuse adolescent knee pain. Conservative management is frequently successful in skeletally immature patients.

Q11. True.

Q12. Rule out distal femoral physeal fracture. Any valgus strain on a skeletally immature patient can cause a distal femoral physeal fracture. Displaced physeal fractures are obvious radiologically but undisplaced fractures look normal on initial x-rays. In such cases stress x-rays have been suggested but referral to an orthopaedic surgeon is suggested. MRI can help evaluate if any valgus laxity is from physeal injury or the MCL. The attachment of the MCL to the distal femoral epiphysis often means that the physis is fully exposed to valgus stress on an extended knee. In a skeletally immature patient the cartilage is much weaker than the MCL. What appears to be valgus instability from MCL damage has been demonstrated on MRI with some patients to have intact MCL. The post-injury weight-bearing regime in managing MCL and physeal damage is radically different. This physis accounts for 70% of the femurs longitudinal growth and 40% of the lower extremities, so accurate diagnosis of these injuries is essential.

Musculoskeletal Medicine Educational Activities

MASTERS, DIPLOMA AND CERTIFICATE COURSES IN MUSCULOSKELETAL MEDICINE

FLINDERS UNIVERSITY DIPLOMA/CERTIFICATE IN MUSCULOSKELETAL MEDICINE

DATE	TITLE/KEY RESOURCE PERSON	VENUE	PROVIDER	CONTACT	CME POINTS
15-22/6/02	Module Three - Vertebral Column	Flinders Medical Centre, Adelaide	Flinders University, South Australia	A/Prof Norm Broadhurst Ph 08 8295 1890	50 per module
21-28/9/02	Module four - Appendicular Skeleton	Flinders Medical Centre, Adelaide	Flinders University, South Australia	A/Prof Norm Broadhurst Ph 08 8295 1890	50 per module

UNIVERSITY OF NEWCASTLE MASTERS IN PAIN MEDICINE

DATE	TITLE/KEY RESOURCE PERSON	VENUE	PROVIDER	CONTACT	CME POINTS
2002	Masters in Pain Medicine	Internet	University of Newcastle	Prof Nikolai Bogduk Ph + 61 2 4923 6172 Fax + 61 2 4923 6103 *mgillam@mail.newcastle.edu.au*	N/A

UNIVERSITY OF OTAGO DIPLOMA/CERTIFICATE IN MUSCULOSKELETAL MEDICINE

DATE	TITLE/KEY RESOURCE PERSON	VENUE	PROVIDER	CONTACT	CME POINTS
29/7/02- 2/8/02 29/7/02- *2/8/02*	Part 2 - Clinical Diagnosis (New Zealand) Part 2 - Clinical Diagnosis (New Zealand)	Christchurch Christchurch	University of Otago As above	V McGroggan Ph + 64 3 364 1086 Fax + 64 3 364 0909 *veronica.mcgroggan@chmeds.ac.nz* Or, Geoff Harding Ph + 61 7 3269 5522 Fax + 61 7 3269 6407 *geoffharding@uq.net.au* *www.chmeds.ac.nz/go/dept-orthop*	50 points* (*total Pt 1 & 2)
7/02 - 10/02	MSME 707 - Musculoskeletal Rehabilitation	Fortnightly teleconferences on Tuesdays	As above	Same contacts as above	NZ - on application Aust - 50 points
7/02 - 10/02	MSME 710 - Recreational & Sports Injuries	Fortnightly teleconferences on Tuesdays	As above	Same contacts as above	NZ - on application Aust - 50 points

Musculoskeletal Medicine Educational Activities

OTHER MUSCULOSKEELTAL MEDICINE EDUCATIONAL ACTIVITIES

DATE	TITLE/KEY RESOURCE PERSON	VENUE	PROVIDER	CONTACT	CME POINTS
20-21/04/02	Members Forum - Case Discussions and Peer Review	Sheraton Hotel, Brisbane	Aust. Medical Acupuncture College	Dr U Jayaswal Ph 07 3286 9269 *ujamac@yahoo.com*	N/A
7-10/5/02	Rehabilitation: A Global Perspective - Annual Scientific Meeting of Australasian Faculty of Rehabilitation Medicine (Preconf. course: State of the Art Review of Back Pain on 7/5/02	Sheraton Hotel, Brisbane	Australasian Faculty of Rehabilitation Medicine	DC Conferences Ph 02 9439 6744 Fax 02 9439 2504 *mail@dcconferences.com.au*	To be announced
20-21/7/02	Australian Medical Acupuncture College AGM and Educational Meeting	Novotel, Brisbane	Australian Medical Acupuncture College	Dr U Jayaswal Ph 07 3286 9269 *ujamac@yahoo.com*	N/A
17-20/10/02	A Pain in the Butt: Groin, Hip and Pelvic Pain. Annual Scientific Meeting of the AAMM & AFMM	Melbourne venue tba	AAMM	Vic Wilk Ph + 61 3 9596 7211 Fax + 61 3 9596 7871 *vicwilk@smart.net.au*	To be announced
9-13/3/03	A Fresh Approach to Pain Management Annual Scientific Meeting of Aust. Pain Society and NZ Pain Society	Sydney Convention Centre	Australian Pain Society	DC Conferences Ph 02 9439 6744 Fax 02 9439 2504 *mail@dcconferences.com.au* *www.apsoc.org.au*	To be announced
15-19/6/03	Pain in Childhood: The BigQuestions International Symposium on Paediatric Pain	Sydney Convention Centre	Paediatric Pain Medicine Unit, Sydney Children's Hospital	DC Conferences Ph 02 9439 6744 Fax 02 9439 2504 *mail@dcconferences.com.au* *www.apsoc.org.au*	To be announced