Review article

Disorders characterised by pain: a methodological review of population surveys

Heiner Raspe, Thomas Kohlmann

Abstract

Objective – To review a series of conceptual and methodological problems encountered in surveys primarily devoted to pain disorders.

Criteria for inclusion and exclusion of articles – Published reports were systematically collected by electronic database searches (Medline), citations in existing publications, and through personal contacts. Relevant articles from clinical and epidemiological research on pain were included and special attention was given to epidemiological research on back pain.

Conclusions – Surveys of pain disorders should be based on a multidimensional pain model that includes nociceptive input, pain perception, suffering, and pain behaviour as major components. Because of the limited applicability of diagnostic procedures or genuine “non-specificity” of pain states, or both, epidemiological surveys may result in a considerable proportion of cases without an identifiable pathophysiological basis. Staging and grading procedures for pain disorders (as distinguished from classification) may comprise various aspects of pain perception: regional distribution, pain intensity, temporal characteristics, sensory qualities, and dimensions of cognitive-emotional appraisal. Description of temporal development and chronification (staging) should refer to different components of the multidimensional pain model. Explicit a posteriori procedures for grading are preferable to implicit grading based on question wording. Evidence from several sources suggests that localistic concepts of pain may be misleading. Identification of complex pain syndromes should be one primary target for epidemiological pain surveys. Of the many factors that may impair the reliability and validity of data collected in pain surveys, recall biases seem to deserve special attention.

(J Epidemiol Community Health 1994;48:531–537)

The first survey devoted exclusively to the assessment of multiple sources of pains was published in 1984 by a Canadian group,1 and the number of subsequent studies has been surprisingly small.2–7 Surveys focussing on site or system specific pain, however, have a longer tradition. The first epidemiological survey on musculoskeletal pain was conducted in the late 1950s in the UK.8 Surveys of (chronic) pain face many problems, both conceptual and technical.9 Many of these problems have become evident from surveys of back pain, and indeed this site has been studied for longer and in more detail than any other type of pain (see10 for review). Thus, experience from back pain surveys will be used to clarify the scope and nature of these problems. This review concentrates primarily on concepts and methods and less on a detailed description of results from previous studies: it attempts to reach practical conclusions and recommendations for future epidemiological surveys of painful disorders.

Which dimensions of pain should be measured?

The International Association for the Study of Pain (IASP) defined pain as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage”11. Even superficial recall of a simple pain experience, such as touching a red hot object, reveals its complex nature. Thus, in addition to the description of pain in terms of spatial and temporal characteristics, other facets are of relevance, including: severity, sensory impressions (such as “burning”) and distinct cognitive (“why me again?”), emotional (“horrible”), behavioural (flight, then cooling), and somatic (sweating, increased heart rate) phenomena.

One useful model of different components of pain is that of Loeser12 (fig 1), which includes four dimensions: nociception, pain, suffering, and pain behaviour. In Loeser’s model, “pain” is defined as “the perception of a nociceptive input to the nervous system”. “Suffering” refers to the “negative affective response generated in the higher nervous system by pain”, whereas “pain behaviour” is meant to denote “any type of output from the organism that is commonly understood to suggest the existence of a tissue-damaging stimulus”.

In contrast to the physiological and clinical
concerns about a specific somatic problem involving nociception, behavioural scientists and epidemiologists are constrained by having to start inductively with the assessment of pain reporting behaviour – for example, recording a “yes” or “no” to the question “do you have back pain today?” Thus, epidemiologists concentrate primarily on the most non-specific component of pain biomedically: subjects with totally heterogeneous chronic pain disorders may show very similar pain behaviour patterns.

**Problem of “non-specific” pain**

If the aetopathogenesis of a pain state cannot be identified after a selected set of simple examination techniques such as history taking, psychiatric interview, clinical examination, basic blood tests, and conventional x rays have been done it is usually described as “non-specific” or “idiopathic”. As a consequence, although it seems comparatively easy to collect information on pain related behaviour, affects, and perceptions, it is much more difficult in surveys to address clinically the supposed underlying physical problem(s). The days when participants of epidemiological studies could easily be subjected to x rays and other potentially hazardous diagnostic tests have (fortunately) gone. Examination surveys have basically to rely on a sound clinical examination, which is necessary but insufficient to explain the differential diagnosis of back pain and other painful disorders.

Even more invasive diagnostic measures, however, have their limits. In one German study of 104 carefully and extensively investigated (for example, by discography) back pain patients referred to a university pain clinic, no medical diagnosis could be made for 35. This suggests that the proportion of subjects with “non-specific” pain is a function of both the degree of diagnostic assessment and of the nature of the disorder.

What is the appropriate level of investigation in surveys? Since no epidemiological study will solve the enigma of non-specific pain in general, it seems more useful to concentrate on a few selected conditions than to use a wide range of unselective examinations with no clear target. The ultimate answer depends critically on the research question: an aetologically oriented study has to specify its target condition(s) as strictly and narrowly as possible, whereas a study that analyses the health care usage of back pain sufferers or the reactions of primary care physicians can concentrate solely on the symptom.

**Classification of pain states**

Given the heterogeneity of most painful disorders it is generally useful to classify cases ascertained. In medicine, “classification” refers to the assignment of cases or patients to predefined groups (“species”) within a taxonomy of entities. Ideally, each entity should show an identical clinical appearance, pathophysiology, course, prognosis, and therapeutic response, and the taxonomy should have an aetio-pathogenetical basis. In practice, such a system does not and probably will never exist. This is particularly true for heterogeneous disorders such as back pain, and a number of different schemes has been proposed.

In one approach, Talo et al. subjected a set of psychological variables to cluster analysis and, using a concept proposed by Turk and Rudy, identified three classes of back pain patients – “interpersonally distressed”, “adaptive copers”, and “dysfunctional patients”. The distress and risk assessment method offers a related classification of low back pain patients based on four variables, which can, in principle, be applied within an epidemiological study. These are two psychological variables – “modified somatic perception” and a depression score, and the presence of “non-organic signs” and “inappropriate symptoms”. Subjects are assigned to four groups – those showing “no psychological distress”, those “at risk of developing major psychological overlary”, “distressed-depressive”, and “distressed-organic”.

An alternative approach to classification systems based on manifestations of back pain is to distinguish between specific and non-specific back pain as defined earlier. The *International Classification of Diseases* (ICD) in its 9th and 10th revisions, may be used as a catalogue of disorders and diseases predisposing to back pain, for instance, deforming dorsopathies, inflammatory spondylopathies, intervertebral disc disorders. Different sets of classification criteria are now available to identify some of them as, for example, a group of spondylarthropathies or the fibromyalgia syndrome. By contrast there are no accepted classification schemes based either on pathological processes or the anatomical structures thought to be involved.

Non-specific back pain could be further subclassified in terms of its manifestation, for instance as purely organic, almost purely organic, mixed, or purely psychiatric, as suggested on the basis of a standardised medical and psy-
chiatric examination of clinical back pain patients.\textsuperscript{29} The heterogeneity of the different approaches is striking, and surprisingly, so far only few have been tested in population studies.

Whereas "classification" relies on a system of fixed nominal classes or types of pain, and "staging" implies a temporally dynamic concept of a developmental process, "grading" refers to an ordinal scale with different levels of severity independent of class or stage. The basis for staging and grading are variables derived from an analysis of the perception of (back) pain (see below).

**Dimensions of pain perception**

Four important structural dimensions will be considered here: regional distribution, intensity, temporal characteristics, and sensory qualities including cognitive and emotional appraisal. All are, as exclusively subjective phenomena, accessible in questionnaire and interview surveys.

**REGIONAL DISTRIBUTION**

Population based surveys usually rely on verbal descriptions of painful sites (such as back pain, chest pain, facial pain) or organs (as joint pains, muscle pains), leaving the decision on where to locate the back, for example, entirely to the respondents. Recent studies offer region of interest (ROI) drawings or complete pain mannequins in which the affected regions can be marked (fig 2). ROI drawings are obviously useful but still lack standardisation. Thus, the recently completed East German Health Survey used a drawing that covered the total area below C7 and the gluteal folds (fig 2, right). An English version (fig 2, left) followed a medical definition of the low back — that is, the area "between the lower costal margins and the gluteal folds".\textsuperscript{15,30} The Standardised Nordic questionnaire for the analysis of musculoskeletal symptoms\textsuperscript{31} (fig 2, middle) presents an even smaller region. How the different definitions match with the subjective body concepts of lay persons is unknown.

In clinical populations, it has been found that pain drawings using an outline of a human body on to which patients mark the painful areas allow the identification of the spatial distribution of pain.\textsuperscript{12-34} In addition they may support a gross estimate of its organic versus its non-organic nature\textsuperscript{15-37} and predict poor treatment outcome.\textsuperscript{38} However, their usefulness in identifying patients whose pain has a "psychological overlay"\textsuperscript{29,41} has not been unequivocally supported.\textsuperscript{42-44} The advantage of pain drawings (compared with ROI drawings) is that they allow the researcher to apply different definitions of "the back", to compare these with verbal responses, and to identify pain patterns affecting several and possibly distant regions. The pain can then be rated as monotypical, oligotopical, polytopical, or widespread;\textsuperscript{45} symmetrical or asymmetrical; periarticular or axial; etc. This promising technique has not been widely used in pain surveys and its validity and reliability have to be studied further.\textsuperscript{45}

**PAIN INTENSITY**

Different techniques have been adopted for measuring pain intensity, with visual analogue scales and verbal or numerical rating scales being most often used in clinical pain research and in pain surveys. Comparative studies have shown that different methods of measurement yield similar results. In practice, visual analogue scales seem to be more difficult to understand by respondents.\textsuperscript{46-49} Verbal rating scales (for example, "almost imperceptible", "like wearing tight clothes", "like being stiff after exercise", etc.) might work well in one country but might be difficult to translate into other languages. We therefore use and recommend numerical rating scales (for example an 11 point scale anchored at "no pain" = 0 and "intolerable pain" = 10) which have been successfully applied (with slight modifications) to pain surveys (for example\textsuperscript{39}).

**TEMPORAL CHARACTERISTICS**

Assessing temporal characteristics of pain perception is a highly complex matter: the duration and frequency of episodes as well as acuteness of and time since onset are relevant parameters of the temporal structure of pain. No general framework for its measurement has been developed so far. The most simple question seems to be that for present pain: "do you have (back, head, neck etc) pain today". Under certain assumptions, it leads to an estimate of point prevalence. Less simple questions refer to the duration of the actual episode: "for how many days without interruption have you had these back pains, that is without one or several days free of pain?\textsuperscript{51} or to the number of days in pain: "on roughly how many days (if any) in the past 12 months have you had that type of pain?\textsuperscript{2} Retrospective questions may differ in the time frame they use ("pain during past ... days/weeks/months" or "ever"). Other questions may clarify the longitudinal pattern (for example, episodic, intermittent, recurrent, persistent), the general trend (increasing/decreasing within the last year), or age at and mode of first manifestation (acute or slow).
SENSORY QUALITIES AND COGNITIVE-AFFECTIVE CHARACTERISTICS

Sensory qualities and characteristics of the cognitive and affective appraisal have been studied in numerous clinical samples of patients suffering from chronic pain disorders. Evidence of the usefulness of particular measures in epidemiological studies is lacking. The McGill pain questionnaire (MPQ) is probably the most widely used instrument for the assessment of sensory, affective, and evaluative dimensions of pain. Yet its discriminant validity (even in a clinical context) has been questioned in a recent multicentre evaluation study. Specific scales for the measurement of pain related coping strategies, pain related cognitions and self statements, or pain locus of control have been developed for use in clinical samples, but have not yet been included in population based (back) pain surveys.

Staging of pain and defining chronicity

Staging is widely used in clinical medicine and is based on a temporally dynamic model of developmental processes with identifiable stages to which cases can be assigned. In pain research, a basic distinction is made between acute, subacute, and chronic pain. Pain is arbitrarily and variably defined as acute if it is present for less than six weeks, though others adopt limits of three or even six months. Subacute pain is said to last between six weeks and six months, whereas chronic pain has to persist for more than six months.

This confusion is only increased by defining chronic pain in terms of its persistence beyond what is deemed a normal healing time. This definition focusses on the supposed physical basis of the problem and seems unable to differentiate adequately between chronic pains with ongoing nociceptive input and those in which persistent noxious stimuli and chronic nociception cannot be identified. To clarify this issue, different aspects of chronicity, each with its own determinants and temporal requirements, can be distinguished, based on the separate features of the physical problem/nociception, pain perception, suffering, and pain behaviour. A fifth aspect may be added: the social consequences of pain. An appropriate temporal definition of acute versus chronic can then be used for each of the five levels. As an example, the Quebec Task Force has defined activity related back pain as chronic if absence from work exceeds more than seven weeks. This model may also be helpful in developing a more detailed understanding of chronicity or "augmentation" of pain problems. The idea of augmentation has given rise to the concept of chronic "dysfunctional" pain, which is not only persistent or recurrent but is regularly accompanied by an augmentation process with a (possibly distinct) sequence of cognitive, affective, and behavioural changes.

At present a simple temporal definition, which retrospectively assesses the duration of the actual (however multifaced) pain episode, seems preferable. The identification of the total disease duration since first onset is less useful. Thus, in Germany, the lifetime prevalence of back pain reaches nearly 90% in subjects aged 25 to 29 and even more than 40% of 12-year-old schoolchildren report back pain.

Implicit and explicit grading of back pain

Some grading system for back pain is necessary given its frequency in the general population. In West Germany for instance its point prevalence presently reaches 40%, its one year prevalence 70%, and its lifetime prevalence exceeds 80%. Its severity, however, varies within a wide range between hardly perceptible and intolerable, disabling pain.

Each back pain survey has used an idiosyncratic key question, mostly with a temporal or an intensity restriction, or both - that is with an implicit grading (and/or staging). Some examples may illustrate this: The US National Health and Nutrition Examination Survey (NHANES-1) identified subjects with back and other types of pain only if they had suffered from it "most days for at least 1 month in the 12 months preceding the interview". The NHANES II survey enquired about "pain in your back on most days for at least 2 weeks" ever. To be included in further analyses, respondents to the Puget Sound Study in western Washington State had to report pain "that had lasted a whole day or more, or that had occurred several times in a year". The Canadian Household Survey aimed at identifying subjects who were "often troubled with pain". In a recent British study only low back pain which, "lasted for more than 24 hours and which was not associated exclusively with febrile illness, menstrual problems, or pregnancy" was recorded.

There are, however, screening questions which are less specific. In a Danish study respondents were asked: "Have you ever/within the last 12 months experienced pain or other symptoms in the lower part of your back?". The corresponding question in The British Health and Lifestyle Survey was: "Within the last month have you suffered from any problems with a bad back?". In three recent studies the authors have used as a first postal screening question: "Do you have back pain today?".

This diversity makes it probable that different surveys cover different parts of the total severity spectrum of back pain and hence explain, at least partly, some of the variation in reported prevalence rates. This unsatisfactory situation could be overcome if different groups either used an identical implicit grading system (which is unlikely to be achieved) or employed a comparably sensitive baseline question which would permit a post hoc explicit grading. Various post hoc grading procedures have been proposed, each of which is based on different concepts. Some were derived using strict psychometric criteria and were thoroughly tested in epidemiological studies. The authors applied two basic dimensions of such a model and produced a simple scheme for grading current pain status of subjects with back pain by using current intensity (measured by a 10 point numerical rating scale) and current functional limitations in everyday life.
assessed by a short, back pain-specific questionnaire (ADL list). Back pain cases were assigned to one of three ordered categories.

Other approaches to grading could be based on the temporal characteristics of the index episode,69 the level of concomitant symptoms (see below), and variables that reflect the pain-related cognitive and emotional appraisal. The development of a more complex grading system, however, is a time consuming project, the additional benefit of which may be questioned.

Localised pain as part of a wider complaint syndrome
Most site specific pain surveys have neglected the importance of concomitant symptoms. These single-site focused, questionnaire/interview techniques are blind to the fact that perhaps most types of pain are likely to be part of complex complaint syndromes.

This can best be illustrated with results from pain surveys: The Nuprin Pain Report enquired about the occurrence of seven distinct types of pain within the past year. Only 10% of all subjects with pain in the back had experienced back pain only, whereas 90% reported (unspecified) combinations of back and other pains.7 In the Puget Sound Study, 47% of the pain subjects recorded simultaneous pains in two or more (of a total of five listed) sites.72 Similar results have been found in other surveys.873 Most recently, these results have been extended by showing a close association between chronic, widespread pain and a number of somatic and affective symptoms.7 In one postal survey, 31% of 4037 German residents of Hannover, Germany, aged 25 to 74 had back pain "today" of whom 81% suffered from neck pain, joint pain, morning stiffness and/or subjective swelling around joints simultaneously. In another German postal survey, these pain/complaint patterns were investigated more systematically.75 The pain questionnaire was accompanied by a validated German symptom check list with 24 items76 and a score range from 0 to 72. Among people who reported current back pain, the mean symptom score (adjusted for age and sex) was 22.5 (95% CI, 21.0-24.4) compared with the score of 13.4 for those with no back pain "today" and 5.6 for the small subgroup who reported never having had back pain. Subjects with back pain had significantly higher scores for 23 of the 24 items, notably for tired legs, dyspnoea, restless legs, vertigo, and fatigue.

In summary, these results suggest that the concept of "regional back pain"77 may be questioned: back pain seems to imply more than pain in the back!

Sampling and non-sampling errors
Finally, it is important to consider the impact of a great variety of potential biases on pain surveys. These may be attributable to characteristics of the respondents, the interviewer, or the method used for data collection.78-81

The necessity for systematic investigation is obvious from the many problems discussed earlier as well as from a range of counter intuitive results in virtually any back pain survey (see for more details). Three prominent examples were:

(1) In several studies the prevalence of current and past back pain (of all grades combined) shows a distinct "paradoxical" age pattern, with a peak around the age of 60 and a decrease among the elderly.66-68 We and others7 have observed, even more surprisingly, a peak in one year and/or lifetime prevalence among the youngest. In contrast, severe back pain usually follows a pattern with an increasing prevalence and a peak in the highest age group.51,67 (2) The curves for the age specific, one-year and lifetime prevalence rates run, on a different level but strictly parallel; the difference between the two is often around or less than 10%.66-68 (3) Incidence estimates derived from cross sectional and cohort studies lead to incredible results: Valkenburg and Haanen,73 for instance, reported incidence rates of "some 11% annually" in their youngest age group (20-25 years) with a decrease to 2% in the highest; Biering-Sorensen recorded an annual incidence rate of even 17%. However, any rate above 4% seems implausible.

It is still not clear which of these results has a plausible explanation for which specific sampling and non-sampling errors can be identified. Firstly, all population based back pain studies suffer from incomplete response, mostly in the range of 15% to 30%. Thus, selection biases might be in effect. Secondly, most relevant surveys of pain, and specifically back pain, have used retrospective reports of past pain experiences. The accuracy of retrospective reports of episodes of pain depends on the respondent's ability to recall the episode, to date it with sufficient precision, and to recall its intensity, duration, and other qualities. Memory decay, "telescoping", and additional response errors can introduce considerable recall biases.8,92 Factors which have been found to contribute to recall bias are characteristics of the actual pain experience, time interval and degree of detail, significance of events, personal (for example, sociodemographic) characteristics of the respondent, and interviewing technique.73 There have recently been very promising approaches to studying the general problem of "memory for health related events" from a social cognition perspective.93 In the epidemiology of back pain, it is generally assumed that more recent episodes are recalled better than earlier events and that the same holds true for more severe (versus mild) episodes.10 Recall bias in interview data of the incidence and prevalence of back pain can be illustrated by the results of Westrin,87 who compared work incapacity data held by the General Sickness Insurance in Gothenburg with self reported sick listing because of low back pain. For a 12 year period, the self report data showed comparatively good reliability. Under-reporting occurred in only 3% of the cases with multiple spells of incapacity to work, yet it was present in 28% of those with only one spell of sick leave listing because of low back pain. The duration of the episode as well
as certain personality traits were associated with accuracy of recall. A population study on the lifetime incidence of low back pain in a sample of 40–47 year old Swedish men, identified 27% among those without an interview history of low back pain who had actually claimed sick benefit because of low back pain during the preceding 22 years. Walsh et al. used a self administered questionnaire to obtain data on a history of back pain and associated disability in a sample of 225 men and women. Over a 12 month period they observed good agreement of data on whether subjects had ever suffered back pain or had consulted a general practitioner because of back pain. Other variables, for example, on the speed of onset of symptoms, presence or absence of sciatica, and disability, however, were less reliably measured.

Summary
Back pain is a very common and extensively studied type of pain. It can be used as a paradigmatic condition in order to identify conceptual and methodological problems of pain surveys in general. Because of the complexity of any pain experience, surveys for pain disorders should be based on a multidimensional pain model. The limited applicability of diagnostic procedures in epidemiological studies or genuine non-specificity of pain states, or both, in surveys may result in a considerable proportion of cases without an identifiable pathophysiological basis.

Classification, staging, and grading are related but different operations which have to be distinguished clearly. While classification could employ "manifestational" or pathogenetic concepts, staging and grading are usually based on various aspects of pain perception: regional distribution, pain intensity, temporal characteristics, sensory qualities and characteristics of cognitive and affective appraisal. Description of temporal development (staging) should consider different levels of chronicity. For grading, explicit a posteriori methods are preferable to variable screening questions with an implicit grading.

The typical coincidence of site specific pains (for example, back pain) with pain at other sites or other symptoms, or both, indicates that strictly "local" concepts may be misleading. More emphasis should be placed on the identification of complex complaint syndromes.

Finally, a multitude of factors may negatively affect the reliability and validity of data collected in pain surveys. Recall biases seem to deserve special attention.

The authors are greatly indebted to Alan J Silman for his encouragement and support in writing and editing this paper.

Disorders characterized by pain


69 Croft PR, Rigby AS. Socioeconomic influences on back problems in the community. J Epidemiol Community Health 1994;48:166-70.


72 Dworkin SF, Von Korff M, LeResche L. Multiple pains and psychiatric disturbance. Arch Gen Psychiatry 1990;47:239-44.


