■ Methodologic Issues in Low Back Pain Research in Primary Care

Lex M. Bouter, PhD,* Maurits W. van Tulder, PhD,* and Bart W. Koes, PhD*

Study Design. Narrative review and discussion of the selected literature.

Objectives. To discuss some important methodologic challenges in low back pain research in primary care.

Summary of Background Data. Many methodologic problems must be confronted when conducting low back pain research. Some of these problems are back pain specific or specific to the primary care setting.

Methods. Methodologic problems related to four research issues will be discussed: study designs, definition of low back pain, determinants of low back pain, and outcome assessment.

Results. Two fundamentally different study designs are frequently used in low back pain research, namely observational studies and experimental studies. The definition of low back pain is typically restricted to a highly variable self-reported symptom, the sensation of pain in the back. There clearly is a need for an evidence-based classification system for low back pain. Because a tenable theoretical framework is lacking, it is difficult to know which determinants of low back pain should be quantified. Low back pain studies focus usually on health-related quality-of-life outcome parameters. The identification of the minimum clinically relevant changes for the most important outcome instruments needs further consideration.

Conclusions. In years to come, low back pain researchers are challenged to overcome some of these (and other) problems to enhance the quality of low back pain research in primary care. [Key words: classification, low back pain, methodology, outcome measures, primary care, study design] Spine 1998;23:2014–2020

Low back pain is clearly an important health problem.[16,44,51] To gain more insight into the problem of low back pain and to reduce the burden of low back pain on patients and society, an impressive and increasing amount of research has been conducted during recent decades. Low back pain research covers a broad range of disciplines: biomechanical, clinical, social, and health sciences. This article focuses on low back pain research in primary care from the clinical epidemiologic perspec-

tive, although several of the topics are relevant in other settings as well.

The main focus of this article is to discuss important methodologic challenges that are currently faced when conducting research projects in this field and to offer tentative solutions for these challenges, when possible. First, an overview is presented of the most prevalent study designs for low back pain research. Second, the operational definition of low back pain is discussed. Third, the problems involved in the measurement of determinants, while adjusting for extraneous factors are outlined. Finally, the topic of outcome assessment is addressed, paying special attention to the quantification of responsiveness.

■ Study Designs for Low Back Pain Research

Most low back pain research in primary care falls within the (clinical) epidemiologic tradition and the main objective is, directly or indirectly, to promote more effective and efficient behavior in patients and health care professionals.[9,51] Epidemiologic studies focus on the causal relation between the occurrence of low back pain (as the dependent variable) and other determinants (as independent variables) such as etiognostic, diagnostic, or prognostic factors, including therapy. The occurrence of low back pain can be expressed categorically (*i.e.*, presence or absence of pain) or as a continuous measurement of the severity of the pain or the related disability (for example, on a visual analog scale [VAS]). The independent variables in low back pain research tend to vary substantially, depending on the study objective. Examples of determinants studied include risk factors (occupational, environmental, or individual factors), diagnostic indicators (medical history or neurologic signs), and predictors of prognosis (previous episodes of low back pain or duration of low back pain). When evaluating the causal correlation between an etiologic, diagnostic or prognostic factor and low back pain, other determinants (confounders and effect modifiers) should always be considered carefully in the design and analysis.

To investigate research questions regarding low back pain, in general, two fundamentally different study designs are used: observational studies and experimental studies. Observational studies focus on the description and analysis of what is happening without interference from the investigator. In a way, the purpose is to learn

from "experiments of nature." Descriptive observational studies include case reports and case series, in which the investigator can only generate hypotheses but cannot evaluate them. Analytic observational studies include cross-sectional, case–control, and cohort studies, which are designed to evaluate causal hypotheses and in which comparison groups are included.[7] Descriptive studies may focus on prevalence or incidence of low back pain,[18] frequency of referrals to paramedical therapists or medical specialists,[47] use of diagnostic testing,[13] and the course of low back pain with time.[27,29,49,50] Analytic observational studies examine, for instance, the impact of work-related risk factors,[8,20,21] the predictive value of a computed tomographic scan,[13] the early markers of chronicity,[14,19] the occurrence of low back pain in relation to radiographic signs of degeneration,[45] the reliability of neurologic tests[28] or radiograph reading,[2] the accuracy of physical examination for diagnosing radiculopathy,[26] and outcome studies linking aspects of care to the course of low back pain.[12,48] Because it is ethically unacceptable to expose people to factors that are assumed to cause low back pain or to affect its course negatively, observational studies are often used to study etiologic or prognostic factors. Some of the disadvantages of observational studies are that they are susceptible to various types of bias, that they may require large sample sizes, that they can be time consuming and expensive, and that the evidence of a causal relation may not be convincing.[7]

Experimental studies, controlled clinical trials, or randomized controlled trials are the preferred study designs for the evaluation of clinical efficacy and cost-effectiveness.[38] These designs are well accepted today for examining preventive or therapeutic interventions but may also apply to the evaluation of diagnostic interventions. Diagnostic information is meant to make a difference. Apart from obtaining information regarding diagnosis ("what's wrong ?") and prognosis ("what will happen?"), guidance on further action is the main concern. Whether diagnostic information in fact helps in deciding about further diagnostic testing and in selecting therapeutic interventions, if any, should be studied in randomized controlled trials with relevant end points. Currently, these trials are extremely rare. At issue is the impact on prognosis of diagnostic and therapeutic interventions combined into protocols. Because of the many possible combinations, the design of these management trials will be a major methodologic challenge for the years to come.

■ Definition of Low Back Pain

Perhaps the most important cause of the difficulties in studying low back pain is the lack of a plausible pathophysiologic or pathoanatomic explanation for most cases.[20,21,51] In fact, "caseness" is typically defined by one highly variable self-reported symptom, namely the sensation of pain in the back. Etiologic and prognostic heterogeneity are probably pronounced among low back pain patients. Consequently, there is great difficulty in identifying homogeneous groups of patients who may really benefit from the available options for preventive, diagnostic, or therapeutic interventions.

In general, many disorders lack a clear pathologic basis and are defined as a syndrome consisting of a mix of signs and symptoms. However, low back pain seems to be unusual, in that consensus is still lacking even on the definition of the syndrome, its relevant subcategories, and the operationalization of the various aspects involved. Several parameters may contribute to caseness or its subcategorization. Cases are usually defined by characteristics of the associated pain, such as the absence or presence of low back pain, the intensity of the pain, the duration of the pain episode, and the historic pattern, which results in a label of acute, subacute, or chronic low back pain.[22,41,50] The velocity of onset of the pain or the presence of sciatica are also frequently used to classify cases into subgroups. Besides the characteristics of pain, a certain degree of disability is often considered to be a necessary condition for caseness, as is sickness behavior related to pain or disability. Examples of sickness behavior are consultation with a general practitioner and absence from work because of low back pain.

Although lack of consensus on the definition of the syndrome and its subcategories may explain some of the difficulties low back pain researchers experience, the problems involved in application of the rules for caseness are also substantial. For example, measuring pain and disability in a reliable and valid way is far from easy, and it is therefore not surprising that in this area different methods will lead to different results. Additionally, the retrospective reconstruction of the course of symptoms with time may be subject to a substantial amount of recall bias. In addition, when low back pain patients are recruited while visiting a general practitioner for an incident or prevalent episode, considerable regression to the mean is obviously to be expected.[27,29,53] Most current subclassifications of low back pain are based on the historic pattern of the symptoms. In this context, definitions of the onset and the termination of an episode of low back pain may be an additional source of controversy between study results.[25,50] An example of a simple subclassification of low back pain, focusing on the pattern of pain with time, is presented in Table 1. This classification depends on longitudinal information, preferably prospectively collected during a period of 12 months. Additionally, the investigators define acute low back pain as that which is neither (yet) recurrent nor chronic, and therefore is potentially transient.[50]

Obviously, reaching consensus on the definition of a case of low back pain and identification of subgroups of patients is important. However, the consensus should be sufficiently evidence based. The challenge lies in the identification of homogeneous groups of low back pain patients—homogeneous in the sense that they will benefit

Table 1. Definition of Phases in the Natural History of Low Back Pain*

Term	Definition
Transient low back pain	An episode in which low back pain is present on no more than 90 consecutive days and does not recur over a 12-month observation period
Recurrent low back pain	Low back pain is present on less than half the days in a 12-month period, occurring in multiple episodes over the year
Chronic low back pain	Low back pain is present on at least half the days in a 12-month period in a single or in multiple episodes
Acute low back pain	Low back pain that is not (yet) recurrent or chronic (as defined above) and whose onset is recent and sudden

* Data from Von Korff M. Spine 1994;19:2041S–2046S.

from specific practical guidance for prevention and management. There is no guarantee that causal homogeneity will automatically also imply prognostic homogeneity. There is urgent need for good ideas about how to identify homogeneous subgroups. These ideas may derive from various sources including clinical practice, biomedical studies, and biologic theories. In observational studies, it can be determined whether the patients of the subgroup at issue who receive the intervention actually profit more than others who did not receive the intervention. Finally, the efficacy of interventions in the subgroups of patients should be studied in randomized controlled trials. This route, of course, takes a long time to travel.

■ **Quantification of Determinants**

Another pitfall is the quantification of determinants, most notably in observational studies on the causes or the prognosis of low back pain. With a view to effective primary or secondary prevention, many researchers attempt to identify risk factors for the occurrence of low back pain or early predictors of chronicity. In a way, this is the Holy Grail of modern back pain research.[20] The general lack of success in this area can be explained in part by methodologic difficulties. Because no well-defined theoretical framework is available, it is difficult to know which determinants to measure.[18] Many potential determinants of occurrence or chronicity can vary substantially with passing time and, consequently, difficult decisions must be made regarding time periods that are relevant to exposure assessment. For example, in a systematic review of observational studies of the correlation between spinal radiographs and low back pain, it was found that most of the included studies were case–control studies in which low back pain status and radiographic findings were assessed at the same time point or in which the radiographs were related to low back pain in the past. If the hypothesis is that degenerative changes cause low back pain, the radiographic examination

should precede the occurrence of low back pain in time.[45] However, how much time there should be between the development of, for example, degenerative changes and the occurrence of low back pain is unclear. The same problem occurs in occupational studies in which the time span between the exposure to a potential risk factor, for example lifting heavy weights, and the occurrence of low back pain is unknown.

The next problem concerns valid and precise measurement of the determinant at issue. There is often an inevitable trade-off between the quality of the information and the feasibility or costs of obtaining it. One example is the quantification of biomechanical risk factors in the workplace, for which questionnaires are often used as cheap proxy instruments for obtaining measurements.[11,21] To make matters worse, retrospective data collection on many determinants is almost impossible, whereas prospective data collection is often unfeasible because of the costs and time involved.

■ **Controlling for Extraneous Factors**

The problem of controlling for extraneous variables is restricted to studies in which causality is at issue and therefore, usually does not apply to diagnostic studies or to etiognostic and prognostic studies that are focused on merely predicting the occurrence or chronicity of low back pain. Regarding these latter studies, a strong predictive power with a small set of easily measurable determinants is all that matters.[19] However, when the focus is on actual or potential preventive or therapeutic interventions, the modeling of causality is, of course, the one and only purpose of data analysis.

Controlling for extraneous factors is problematic in low back pain studies because insight into the multicausal pathway is lacking and the most appropriate determinants are unknown. Apart from the central determinant that is the focus of the study, many other determinants may have an effect on the outcomes at issue. Extraneous variables can function as confounders or as effect modifiers.[39] Confounders influence the outcome through a causal pathway in which the central determinant does not play a role, whereas effect modifiers weaken or strengthen the effect of the central determinant, because of their role in the same causal pathway. Extraneous factors can be controlled for in the study design through restriction, prestratification, or matching. However, because there are no explicit and plausible theories on causal pathways, many confounders and effect modifiers remain unknown and unmeasured. Therefore, control for extraneous factors is often not possible in the study design, and the possibilities for stratified analyses and multivariable regression modelling are limited. Only randomization can be of help in adjusting for unmeasured extraneous factors, making this a potential problem especially in observational studies. However, in randomized controlled trials, extraneous factors can also

Table 2. Some Examples of Outcome Measures and Instruments in Low Back Pain Research

Outcome Measures	Domain	Instruments
Symptoms	Individualized	Pain intensity (VAS or NRS)
	LBP specific	Presence or absence of radiation
	Generic	Overall improvement (VAS or NRS)
Disability	Individualized	Functional limitation at baseline (VAS or NRS)
	LBP specific	Roland Disability Questionnaire, Oswestry Scale
	Generic	Sickness Impact Profile (SIP), Medical Outcome Study Short Form 36 (SF 36)
Role functioning	Generic	Work absenteeism (no. of days), medical consumption, health care utilization

VAS = visual analog scale; NRS = numerical rating scale; LBP = low back pain.

play an important role after randomization. Some examples of this are differences among the groups concerning cointerventions, compliance, and placebo effects and the success of blinding.

Problems in the precise and valid quantification of determinants in back pain research can also be a major drawback in managing the extraneous variables in causal analyses. Some examples of potentially relevant extraneous factors that are difficult to measure are: work-related biomechanical risk factors,[11] individual susceptibility, specific psychosocial stressors,[8] and the influence of the social security system. The solution to problems regarding controlling for extraneous factors, although clear, is difficult to realize: Identify the factors to be measured, and measure them with high-quality instruments.

■ Outcome Measurement

Substantial attention has been given to the design and the clinimetric properties of outcome parameters in low back pain research.[6,17,24,35,40] Whereas in etiognostic and diagnostic studies caseness or one of its subclassifications usually constitutes the dependent variable, the objective of most prognostic studies is to quantify outcome in a more subtle way. Intervention research, geared to the needs of primary care, typically focuses on health-related quality-of-life outcome measures.[17] These measures include generic, low back pain–specific, and individualized varieties. This last category focuses, for instance, on the severity of the main complaint at baseline, which may be different for each patient.[5,43]

Outcome measures may consist of a single global rating or of multiple items to be combined in an aggregate score. The instruments at issue may cover symptoms, disability, and role functioning, of which some examples are provided in Table 2.

In choosing outcome measures for a study, several methodologic issues should be subjected to careful consideration, such as reliability, validity, responsiveness, applicability, practicality, and comprehensiveness.[40] Many authors prefer to use their own *ad hoc* outcome measures and instruments. However, information on the clinimetric properties of the candidate instrument should be available and should indicate favorable characteristics for the instrument used. Although a certain amount of information can often be obtained, some popular instruments have escaped scrutiny, and direct comparisons between instruments are still relatively rare.[3] For example, the Roland and the Oswestry disability questionnaires, which were constructed without using a conceptual approach or empirical methods of item development, analysis, and selection, are the most widely used scales for measuring disability in back pain patients. The Quebec Back Pain Disability Scale was recently developed according to clinimetric standards and showed measurement properties similar to those of the other two scales.[36,37] The current investigators have compared the Dutch translations of the Roland and the Quebec disability scales and have concluded that both scales seemed to be reliable, valid, and responsive.

Usually, more than one outcome parameter is used in a study. If this is the case, their hierarchy must be considered carefully, or a way must be found to combine the multiple end points before analyzing the data. The same applies to the timing of the outcome assessments, possibilities for blinding, and optimal methods for data analysis. The authors' extensive experience in reviewing randomized controlled trials on low back pain indicates that in these areas there is certainly still much room for improvement in the design and the reporting of studies.[1,31–34,46]

■ Responsiveness

The methodologic literature can also be confusing. An example is the topic of responsiveness. The basic idea is simple. When comparing follow-up severity scores with baseline, three conclusions are possible: no change, deterioration, or improvement. A suitable outcome parameter should detect changes larger than a clinically defined minimum. However, for this minimum clinically relevant change there is usually no criterion or gold standard available, and the absolute responsiveness of an outcome parameter therefore cannot be established in many instances. Using a surrogate criterion—typically, extending the "no change" zone—leads to inflated estimates of relative responsiveness, which is perfectly acceptable when the sole purpose is to compare the responsiveness of several outcome parameters within a study.[3,4,10,24] The quantification of responsiveness consists typically of the calculation of one or more of the many effect sizes proposed in the methodologic literature.[42] The idea is to calculate a "signal-to-noise ratio," in which the signal is the change to be detected and the noise is the variability in the absence of change. In addition, a receiver operating characteristics (ROC) curve is often plotted to show the accuracy of detecting change according to the crite-

Table 3. Baseline Score, Mean Change at 5 Weeks, Effect Size Statistic and Area Under the ROC Curve for Patients Classified as Improved (n = 38)*

Instrument	Baseline Score	Mean Change at 5 wks	Effect Size†	AUC
Oswestry (0–100)	26.2	11.9	0.8	0.76
Roland (0–100)	50.4	32.6	2.0	0.93
Pain last week (0–100)	55.7	36.6	1.6	0.91
Main complaint (0–100)	71.4	41.3	1.6	0.82

* Data from Beurskens AJHM et al. Pain 1996;65:71–76.
† Within patient mean change score divided by the standard deviation of the mean change score.

rion for different cutoff points in the change score of the outcome parameter.[15]

Table 3 shows an example of a study involving a comparison of the responsiveness of the Oswestry low back pain disability questionnaire, the Roland disability questionnaire, pain during last week measured on a 0 to 100 VAS, and severity of the main complaint measured on a 0 to 100 VAS.[4] For the sake of convenience, all instruments are transformed to express scores on a scale from 0 to 100. Included were 81 patients with nonspecific low back pain of at least 6 weeks' duration. Outcomes were measured at baseline and after 5 weeks of treatment. Patients were classified as improved because they indicated much improvement or complete recovery (n = 38), according to their self-reported scores on a 7-point scale. Thirty-eight additional patients were classified as unimproved also based on their self-reported ratings. Five patients' status deteriorated, and their data were excluded from the analysis. Using effect size statistics and area under the ROC curve analysis, it was concluded that in this study the Roland disability questionnaire and the pain rating (0–100 VAS) were more responsive than the Oswestry scale and the severity of the main symptom measured on a 0 to 100 VAS.[4] Unfortunately, the various indexes of responsiveness do not always point in the same direction.[55] That the reason is not yet fully understood makes this another methodologic challenge. Similarly, there is clearly a need for further development of methods to identify minimum clinically relevant change,[23,30,52,54] including sample size calculations and statistical tests to compare the responsiveness of different outcome parameters.[3]

■ Conclusion

This list of methodologic topics relevant for low back pain research in primary care is far from complete. Summarizing, the main methodologic challenges for the near future concern the design of diagnostic efficacy studies, the identification of homogeneous groups, a more accurate prediction of the occurrence and chronicity of low back pain, a more sophisticated modeling of causality,

and the identification of indexes of responsiveness and identifying the minimum clinically relevant changes of the most important outcome measures.

Most of the methodologic problems discussed will continue to be prevalent in the near future. Identification of these problems is only a first step in overcoming the challenges. However, there are promising advances in many aspects of low back pain research to date. Step by step, the research effort will contribute to the reduction of the burden of low back pain on patients and society.

References

1. Assendelft WJJ, Koes BW, Heijden GJMG van, Bouter LM. The effectiveness of chiropractic for treatment of low back pain: An update and attempt at statistical pooling. J Manipulative Physiol Ther 1996;19:499–507.

2. Assendelft WJJ, Bouter LM, Knipschild PG, Wilmink JT. Reliability of lumbar spine radiograph reading by chiropractors. Spine 1997;22:1235–41.

3. Beaton DE, Hogg–Johnson S, Bombardier C. Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol 1997;50:79–93.

4. Beurskens AJHM, de Vet HCW, Köke AJA. Responsiveness of functional status in low back pain: A comparison of different instruments. Pain 1996;65:71–6.

5. Beurskens AJHM, de Vet HCW, Köke AJA, et al. A patient-specific approach for measuring functional status in low back pain. In: Beurskens AJHM, ed. Low Back Pain and Traction. Maastricht, The Netherlands: Rijksuniversiteit Limburg, 1996:83–96. Thesis.

6. Beurskens AJHM, de Vet HCW, van der Heijden GJMG, Knipschild PG. Measuring the functional status of patients with low back pain: Assessment of the quality of four disease-specific questionnaires. Spine 1995;20:1017–28.

7. Bombardier C, Kerr MS, Shannon HS, Frank JW. A guide to interpreting epidemiologic studies on the etiology of back pain. Spine 1994;19(suppl):S2047–56.

8. Bongers PM, de Winter CR, Kompier MAJ, Hildebrandt VH. Psychosocial factors at work and musculoskeletal disease. Scand J Work Environ Health 1993;19:297–312.

9. Borkan JM. An agenda for primary care research on low back pain. Spine 1996;21:2880–4.

10. Bronfort G, Bouter LM. Responsiveness and concurrent validity of the COOP Charts and the SF-36: A comparison of two short functional health status instruments used as outcomes in a randomized clinical trial on chronic low-back pain. In: Bronfort G, ed. Efficacy of manual therapies of the spine. Amsterdam: EMGO Institute, 1997:69–79. Thesis.

11. Burdorf A. Exposure assessment of risk factors for disorders of the back in occupational epidemiology. Scand J Work Environ Health 1992;18:1–9.

12. Carey TS, Garret J, Jackman, McLaughlin C, Fryer J, Smucker DR. The outcomes and costs of care for acute low back pain patients seen by primary care practitioners, chiropractors, and orthopedic surgeons: The North Carolina back pain project. N Engl J Med 1995;333:913–7.

13. Cherkin DC, Deyo RA, Wheeler K, et al. Physician varia-

tion in diagnostic testing for low back pain: Who you see is what you get. Arthritis Rheum 1994;37:15–22.

14. Coste J, Delecoeillerie G, Cohen de Lara A, Le Parc JM, Paolaggi JB. Clinical course and prognostic factors in acute low back pain: An inception cohort study in primary care. BMJ 1994;308:577–80.

15. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Control Clin Trial 1991;12:1425–585.

16. Deyo RA. Practice variations, treatment fads, rising disability: Do we need a new clinical research paradigm? Spine 1993;18:2153–62.

17. Deyo RA, Andersson G, Bombardier C, et al. Outcome measures for studying patients with low back pain. Spine 1994;19(suppl):S2032–6.

18. Deyo RA, Phillips WR. Low back pain: A primary care challenge. Spine 1996;21:2826–32.

19. Dionne CE, Koepsell TD, Von Korff M, Deyo RA, Barlow WE, Checkoway H. Predicting long-term functional limitations among back pain patients in primary care settings. J Clin Epidemiol 1997;50:31–43.

20. Frank JW, Brooker AS, DeMaio SE, et al. Disability resulting from occupational low back pain, II: What do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. Spine 1996;21:2918–29.

21. Frank JW, Kerr MS, Brooker AS, et al. Disability resulting from occupational low back pain, I: What do we know about primary prevention? A review of the scientific evidence on prevention before disability begins. Spine 1996;21:2908–17.

22. Frymoyer JF. Back pain and sciatica. N Engl J Med 1988;318:291–301.

23. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: Development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. OMERACT Committee. J Rheumatol 1993;20:561–5.

24. Gommans I, Koes BW, van Tulder MW. Validity and responsiveness of the Dutch adaptation of the Roland Disability Questionnaire. In: van Tulder MW, Koes BW, Bouter LM, eds. Low Back Pain in Primary Care: Effectiveness of diagnostic and therapeutic interventions. Amsterdam: EMGO Institute, 1996:57–70.

25. van den Hoogen JMM, Koes BW, Devillé W, van Eijk JTM, Bouter LM. The prognosis of low back pain in general practice. Spine 1997;22:1515–21.

26. van den Hoogen JMM, Koes BW, van Eijk JTM, Bouter LM. On the accuracy of history, physical examination and erythrocyte sedimentation rate in diagnosing low-back pain in general practice: A criteria-based review of the literature. Spine 1995;20:318–27.

27. van den Hoogen JMM, Koes BW, van Eijk JTM, Bouter LM, Devillé W. On the course of low back pain in general practice: One-year follow-up study. In: van den Hoogen JMM, ed. Low back pain in general practice [thesis]. Amsterdam: EMGO Institute, 1996:47–65.

28. van den Hoogen HJM, Koes BW, Devillé W, van Eijk JTM, Bouter LM. The inter-observer reproducibility of Lasègue's sign in patients with low back pain in general practice. Br J Gen Pract 1996;46:727–30.

29. van Hoogen HJM, Koes BW, van Eijk JTM, Bouter LM, Devillé W. Pain and health status of primary care patients with low back pain. J Fam Pract 1997;44:187–92.

30. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. J Clin Epidemiol 1994;47:81–7.

31. Koes BW, Bouter LM, Heijden GJMG van der. Methodological quality of randomized clinical trials on treatment efficacy in low back pain. Spine 1995;20:228–35.

32. Koes BW, Scholten RJPM, Mens JMA, Bouter LM. Efficacy of epidural steroid injections for low-back pain and sciatica: A systematic review of randomized clinical trials. Pain 1995;63:279–88.

33. Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM. Spinal manipulation for low-back pain: An updated systematic review of randomized clinical trials. Spine 1996;21:2860–73.

34. Koes BW, Scholten RJPM, Mens JMA, Bouter LM. Efficacy of non-steroidal anti-inflammatory drugs for low back pain: A systematic review of randomised clinical trials. Ann Rheum Dis 1997;56:214–23.

35. Kopec JA, Esdaile JM. Spine update: Functional disability scales for back pain. Spine 1995;20:1943–9.

36. Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood–Dauphinee S, Lamping DL, Williams DI. The Quebec Back Pain Disability Scale: Measurement properties. Spine 1995;20:341–52.

37. Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec Back Pain Disability Scale: Conceptualization and development. J Clin Epidemiol 1996;49:151–61.

38. Miettinen OS. The need for randomization in the study of the intended effect. Stat Med 1983;2:267–71.

39. Rothman KJ. Modern Epidemiology. Boston: Little, Brown, 1986.

40. Schoppink LEM, van Tulder MW, Koes BW, Beurskens AJHM, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. Phys Ther 1996;76:268–75.

41. Spitzer WO, LeBlanc FE, Dupuis M, et al. Scientific approach to the assessment and management of activity-related spinal disorders: A monograph for clinicians. Report of the Quebec Task Force on spinal disorders. Spine 1987;12(suppl 7):S1–S59.

42. Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. Phys Ther 1994;74:528–33.

43. Tugwell P, Bombardier C, Buchanan W, et al. Impact on quality of life assessed by traditional standard-item and individualized patient preference health status questionnaires. Arch Intern Med 1990;150:59–62.

44. Van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in the Netherlands. Pain 1995;62:233–40.

45. Van Tulder MW, Assendelft WJJ, Koes BW, Bouter LM. Spinal radiographic findings and non-specific low back pain: A systematic review of observational studies. Spine 1997;22:427–34.

46. Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic non-specific low back pain: A systematic review of randomized controlled trials of the most common interventions. Spine 1997;22:2128–56.

47. Van Tulder MW, Koes BW, Bouter LM, Metsemakers JFM. Management of chronic nonspecific low back pain in primary care: A descriptive study. Spine 1997;22:76–82.

48. Von Korff M, Barlow W, Cherkin D, Deyo RA. Effects of

practice style in managing back pain. Ann Intern Med 1994;
121:187–95.

49. Von Korff M, Saunders K. The course of back pain in primary care. Spine 1996;21:2833–9.

50. Von Korff M. Studying the natural history of back pain. Spine 1994;19(suppl):S2041–6.

51. Waddell G. Low back pain: A twentieth century health care enigma. Spine 1996;21:2820–5.

52. Wells GA, Tugwell P, Kraag GR, Baker PRA, Groh J, Redelmeier DA. Minimum important differences between patients with rheumatoid arthritis: The patients' perspective. J Rheumatol 1993;20:557–60.

53. Whitney CW, Von Korff M. Regression to the mean in treated versus untreated chronic pain. Pain 1992;50:281–285.

54. Wright JG. The minimally important difference: Who's to say what is important? J Clin Epidemiol 1996;48:1221–2.

55. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239–46.

Address reprint requests to

Lex M. Bouter, PhD
Institute for Research in Extramural Medicine
Faculty of Medicine
Vrije Universiteit Amsterdam
Van der Boechorststraat 7
1081 BT Amsterdam
The Netherlands
E-mail: lm.bouter.emgo@med.vu.nl