VU Research Portal

A consensus approach toward the standardization of back pain definitions for use in prevalence studies

Dionne, C.E.; Dunn, K.M.; Croft, P.R.; Nachemson, A.L.; Buchbinder, R.; Walker, B.F.; Wyatt, M.; Cassidy, J.D.; Rossignol, M.; Leboeuf-Yde, C.; Hartvigsen, J.; Leino-Arjas, P.; Latza, U.; Reis, S.; del Real, M.T.G.; Kovacs, F.M.; Oberg, B.; Cedraschi, C.; Bouter, L.M.; Koes, B.W.

Link to publication in VU Research Portal

A Consensus Approach Toward the Standardization of Back Pain Definitions for Use in Prevalence Studies

Clermont E. Dionne, PhD,*†‡ Kate M. Dunn, PhD,‡ Peter R. Croft, MD,‡ Alf L. Nachemson,
Rachelle Buchbinder, Bruce F. Walker, Mary Wyatt, J. David Cassidy, Michel Rossignol, Charlotte Leboeuf-Yde,
Jan Hartvigsen, Päivi Leino-Arjas, Ute Latza, Shmuel Reis, Maria Teresa Gil del Real,
Francisco M. Kovacs, Birgitta Öberg, Christine Cedraschi, Lex M. Bouter, Bart W. Koes, H. Susan J. Picavet,
Maurits W. van Tulder, Kim Burton, Nadine E. Foster, Gary J. Macfarlane, Elaine Thomas, Martin Underwood,
Gordon Waddell, Paul Shekelle, Ernest Volinn, Michael Von Korff

Study Design. A modified Delphi study conducted with 28 experts in back pain research from 12 countries.

Objective. To identify standardized definitions of low back pain that could be consistently used by investigators in prevalence studies to provide comparable data.

Summary of Background Data. Differences in the definition of back pain prevalence in population studies lead to heterogeneity in study findings, and limitations or impossibilities in comparing or summarizing prevalence figures from different studies.

Methods. Back pain definitions were identified from 51 articles reporting population-based prevalence studies, and dissected into 77 items documenting 7 elements. These items were submitted to a panel of experts for rating and reduction, in 3 rounds (participation: 76%). Preliminary results were presented and discussed during the Amsterdam Forum VIII for Primary Care Research on Low Back Pain, compared with scientific evidence and confirmed and fine-tuned by the panel in a fourth round and the preparation of the current article.

Results. Two definitions were agreed on a minimal definition (with 1 question covering site of low back pain, symptoms observed, and time frame of the measure, and a second question on severity of low back pain) and an optimal definition that is made from the minimal definition and add-ons (covering frequency and duration of symptoms, an additional measure of severity, sciatica, and exclusions) that can be adapted to different needs.

Conclusion. These definitions provide standards that may improve future comparisons of low back pain prevalence figures by person, place and time characteristics, and offer opportunities for statistical summaries.

Key words: back pain, definitions, consensus, Delphi study. **Spine 2008;33:95–103**

Although back pain research has progressed over the past 20 years, efforts in this field of investigation are still hampered by important methodologic problems, among which is a difficulty in defining back pain prevalence clearly and consistently. The end result of this problem is a large heterogeneity in study findings. This creates limitations or may make impossible in comparing or summarizing results from different studies.

Some of the authors of this article have recently conducted a systematic review on the relationship between age and the prevalence of back pain.[1] The 51 articles assessed for the analyses offered a wide range of definitions of back pain prevalence that made valid comparisons or summary very difficult or impossible. Other authors have reported similar difficulties.[2–4] This problem is even more important considering the other sources of heterogeneity, most of them methodologic (study design, sampling frame, analysis, *etc.*), across back pain prevalence studies. If regional and international comparisons of population surveys are to be used to investigate the causes and consequences of back pain and to determine the influence of different health care systems on the occurrence of back pain, there needs to be standardized information to compare.

To identify standard definitions of low back pain for epidemiological prevalence studies, an international panel of experts in back pain was invited to participate in a Delphi procedure to agree on at least 2 definitions of low back pain:

1. A "minimal" definition, for use in large population-based general surveys, where there are many constraints and space for only 1 or 2 questions, and
2. An "optimal" definition for use in focused studies where the investigators have space or time for multiple questions.

The publication of standardized definitions of low back pain prevalence could have an important impact on the ability to make valid comparisons between low back pain studies in the future, and may thus constitute a major step toward better understanding of this important health problem by increasing the value of individual studies and facilitating the synergy of international research.

■ Materials and Methods

Study Design

This study adopted a modified Delphi approach.[5–7] It was designed, implemented, and coordinated by C.E.D., K.M.D., and P.R.C.

Participants

An international panel of back pain experts was composed from the lists of past International Forums for Primary Care Research on Low Back Pain and from the authors of the 51 articles reviewed for a previous publication on back pain prevalence in older people.[1] In identifying panel members, special attention was given in obtaining a large geographical coverage while keeping the panel small enough to allow efficient exchanges. The final list included 37 investigators from 12 countries. Experts who declined participation were asked to suggest a colleague with similar background to replace them. The list of experts was kept confidential until the workshop.

Data Collection

Round 1. In a first step, the definitions found in 51 articles reporting the results of back pain population-based studies[1] were examined to identify the elements that could be included in a definition of low back pain prevalence. This exercise led to the conclusion that 7 different elements could be distinguished: (1) the time frame of the measure, (2) the site of low back pain, (3) the symptoms observed, (4) the duration of symptoms, (5) the frequency of symptoms, (6) the severity of low back pain, and (7) exclusions. Using the definitions of low back pain found in the 51 articles, 77 different definitions of these elements were identified (time frame: 12, site: 8, symptoms: 26, duration: 13, frequency: 7, severity: 8, exclusions: 3). These items were listed, grouped by element, in a questionnaire that asked Delphi participants to rate each of them on an 11-point rating scale where 0 meant "Not at all suitable for a standard definition of low back pain" and 10 meant "Would definitely use for a standard definition of low back pain." The rating had to be done twice, once for an optimal definition of low back pain and once for a minimal definition. The questionnaire offered the opportunity to write general and specific comments and to add new definitions for each of the elements. The list of all definitions included in the round 1 questionnaire is presented in the Appendix (available online through Article Plus).

In the round 1 questionnaire, experts were also asked if they thought we could use the expression "back pain" to include neck, thoracic, and low back pain.

The round 1 questionnaire was sent by e-mail (Word attachment) on December 19, 2005. E-mail reminders were sent on January 16 and February 3, 2006.

Round 2. Distributions of individual scores of panel members in round 1 were established and items that did not reach an *a priori* determined consensual median score of at least 6 of 10 were excluded. New items suggested by the participants in round 1 were added to the list (these items are identified in the Appendix, available online through Article Plus). The same instructions as for round 1 were used. Median and individual scores of round 1 were provided to each participant. Round 2 questionnaires were e-mailed (Word attachment) on March 28–29, 2006. Two e-mail reminders were sent at 2-week intervals.

Round 3. Distributions of individual scores of the panel's members in round 2 were established and items that did not reach the consensual median score of at least 6 of 10 were excluded from further consideration. In this round, participants were asked to choose only 1 item in each element, for each definition. The round 3 questionnaire was e-mailed (Word attachment) on May 11, 2006, to all those who had answered Round 1. One e-mail reminder was sent 2 weeks later.

Workshop. A workshop was organized at the International Forum VIII for Primary Care Research on Low Back Pain held in Amsterdam (June 8, 2006) to present the results of rounds 1 to 3 and discuss them with the participants. Before the workshop, 2 definitions (minimal and optimal) were built using the items remaining after rounds 1 to 3 (these items are highlighted in the Appendix, available online through Article Plus). Participants of the workshop were provided with these definitions and the list of all items considered in rounds 1 to 3, along with the median scores obtained and specific comments. They were also presented with a summary of the scientific evidence on back pain measurement.

Round 4. Results of the workshop were integrated with those of the first 3 rounds and compared with the scientific evidence. When a definition was not coherent with the scientific evidence after the workshop, a change was suggested to the panel members with an explanation. During this round, which started on October 24, 2006, participants were provided with an online summary document on the study purpose and results and asked to vote for or against 1 minimal definition and 1 optimal definition. They were encouraged to provide specific comments, especially when they voted against a proposal. Two e-mail reminders were sent at 2-week interval.

All questionnaires were pilot tested with 1 research assistant (S.P.) and 2 back pain investigators (N.E.F. and E.T.).

Article. All participants were sent a draft of the current article for review and comments. At this time, they were asked if they supported the final definitions. If they disagreed with the final definitions, they would still be considered among the coauthors, but their disagreement would be mentioned in the article.

■ Results

Participants

Twenty-eight of the 37 experts approached (76%) returned the round 1 questionnaire completed. They represented Australia (n = 3), Canada (n = 2), Denmark (n = 2), Finland (n = 1), Germany (n = 1), Israel (n = 1), Spain (n = 2), Sweden (n = 2), Switzerland (n = 1), the Netherlands (n = 4), UK (n = 6), and US (n = 3). At least 1 expert from each country invited was included among the participants.

Round 1

Overall, 61 of the 77 items were eliminated from further scoring, with a median score <6/10. Sixteen items that got a median score ≥6/10 were left to score in round 2, along with 32 new items suggested by participants (total: 48 items).

The majority of experts (81.5%, n = 22) answered "No" to the question that asked if they thought we could

use the expression back pain to include neck, thoracic and low back pain.

Round 2

Twenty-three of the 28 experts who had answered round 1 replied in round 2 (82% follow-up). At the end of this round, all 16 items from the original list remained, plus 4 of the 32 new items.

Round 3

Twenty-five of the 28 round 1 responders answered and returned the round 3 questionnaire (89% follow-up). Among the comments, several participants (n = 10) mentioned that they would have preferred to check more than one choice, especially for the optimal definition (*e.g.*, time frame: today, 1 month, 1 year). Others raised the question of whether there was any real difference between asking "Today" or "Currently," and some found no good choices for duration. Visual analogue scales (VAS) and numerical rating scales (NRS) were considered equivalent by several participants. An optimal and a minimal definition of low back pain were built from the remaining items and presented at the workshop.

On the basis of comments received during this round, a decision was made to review the existing published evidence for the format of the questions that had been included in the Delphi rounds so far. The results of this were fed back to the group before the workshop and were as follows:

1. Research supports the validity of retrospective reports of pain intensity for at least a 3-month recall period.[8–10]
2. Differences in pain, disability, and psychological status have been described between patients from the following categories of patient-reported symptom duration (time since last pain-free month): 0 to 6 months, 7 to 35 months, 3 years, and more.[11]
3. The traditional division between acute and chronic low back pain has been criticized (*e.g.*, Von Korff *et al* and Waddell)[12,13] and there is a discrepancy between theory and practice regarding the definition of chronic low back pain. The term "chronic" low back pain, as currently used, is equivocal.[14]
4. NRS are more easily understood, more reliable, and responsive than VASs and Verbal Rating Scales.[15–18] NRS have been recommended as the scale of choice to measure pain intensity in patients with low back pain.[19] Jensen *et al*[20] have also shown that an NRS using 11 points is as sensitive as an NRS with more points on the scale. NRS can be administered in written or verbal form, and unlike the VAS, difficulty with the scale does not seem to be associated with age.[21]
5. Brief pain and disability measures that have been well studied and for which there have been extensive assessments of psychometric qualities are the SF-36 Bodily Pain scale,[10,22,23] and the Graded Chronic Pain Scale (GCPS).[10,24–26]

Workshop

The workshop was attended by 24 persons (plus C.E.D., K.M.D., and P.R.C.), of whom 6 (21%) had participated in rounds 1 to 3. Other experts who participated in the rounds did not attend the Forum or were unable to attend due to involvement in concurrent workshops.

Workshop participants mentioned that a minimal definition must be minimal and suggested to leave out the duration and severity criteria. They also proposed that the diagram (body manikin with shaded area for low back pain) should be used when possible. "Past month" was discussed as ambiguous (for instance, on February 15, past month may be interpreted as the period between January 15 and February 14 or the period between January 1 and 31). It was suggested to use "Past 4 weeks" instead.

For the optimal definition, participants suggested that it be built from the minimal definition plus add-ons for each other element. For example, participants wanted to include information on duration, for measuring the prevalence of long-standing back pain, which would involve the minimal definition plus the standard definition of duration. As another example, participants wanted a measure of severity for estimating the prevalence of severe low back pain. It was therefore suggested that a definition of severity be added on to the minimal definition (*e.g.*, 0–10 NRS, with a score >5 indicating severe back pain). People could combine these as they see fit, for example, including duration and severity to get the prevalence of severe long-standing low back pain. Another domain that participants viewed as important was sciatica, so that adding on a question about it could provide an estimate of its prevalence. This resulted in a minimal definition, and a set of add-on characteristics that allowed optimal definitions of back pain prevalence to be produced.

Round 4

The round 4 online questionnaire was filled in by 26 of 27 experts (96% participation—1 expert, A.L.N., was ineligible for this round). Twenty-two (85%) voted in favor of the minimal and 18 (69%) voted in favor of the optimal definitions presented in this round. Several comments were made by the experts and considered in building the final definitions. Changes from the versions submitted to the vote during round 4 were:

For the Minimal Definition. (1) A severity criterion ("bad enough to limit your usual activities or change your daily routine for more than 1 day") was added, following the comment made by several participants that otherwise the minimal definition would result in extremely high prevalence of back pain that would not be meaningful, and (2) the instructions were clarified.

For the Optimal Definition. (1) Alternate time frames were excluded from the formal definition because they yielded too much variability, (2) "Sciatica" was replaced by "pain that goes down the leg," (3) grouping of NRS
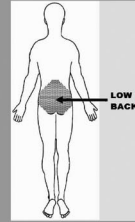
A) For face-to-face interviews and paper or online questionnaires:

<u>In the past 4 weeks</u>, have you had pain in your low back (in the area shown on the diagram)?

Yes ☐   No ☐

If yes, was this pain bad enough to limit your usual activities or change your daily routine for more than one day?

Yes ☐   No ☐

LOW BACK

B) For telephone surveys:

<u>In the past 4 weeks</u>, have you had pain in your low back?

Yes ☐   No ☐

If yes, was this pain bad enough to limit your usual activities or change your daily routine for more than one day?

Yes ☐   No ☐

Figure 1. Final <u>minimal</u> definition of low back pain that results from the Delphi study. The diagram should be used in face-to-face interviews and questionnaires (**A**), and the wording alone used in telephone surveys (**B**). The diagram is used with permission.[29]

scores was changed from ≤5/>5 to <7/≥7 to conform to the most recent scientific evidence,[27,28] (4) categories of duration were further defined (the 0–7 months category was divided in 2 categories—<3 months and 3–7 months—to take into account more acute episodes), (4) the sequence and priority of questions were made clearer, (5) the instructions were clarified (especially for frequency, duration, and severity), and (6) examples of application were added.

Final definitions, worded as questions, are presented in Figures 1, and 2.

■ **Discussion**

This study has reminded us how complicated defining low back pain can be, and how much cultural, linguistic, methodologic, and experiential variability there is in defining back pain prevalence. This further emphasizes the importance of using standard definitions in the back pain research field.

It is very important to stress that the key feature of the approach used in this study is the consensus of experts in the field of back pain prevalence research and the primary care of back pain; hence, the intent was not to find "the best" low back pain definitions or to present the final definitions as "the only" low back pain definitions. The goal was simply to bring leading back pain experts together to agree as much as possible on definitions that could be published and suggested for free use in future studies. As an example, as investigators, when we start a

new prevalence study, this would provide us with definitions that would allow us to compare our study results with those of others. We might decide to add other definitions for different reasons, but by using the standard definitions, we would know that we would most likely be able to compare and summarize our results with those of other prevalence studies, according to person, time, and place.
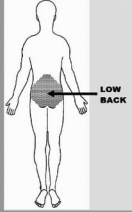
It is also important to remember that the definitions proposed in this article are intended for use in epidemiologic prevalence studies. Consensus studies on the definition of back pain episode[30] and on back pain outcome measures[31–33] have already been published and serve different purposes. Likewise, these definitions are not suitable for detailed clinical studies.

Although we had reached a consensus on a minimal definition that included only 1 question ("In the past 4 weeks, have you had pain in your low back?"), it was finally decided, on the suggestion of many experts, to add a minimum severity criterion. Without such a criterion, many thought that the prevalence measured would have been extremely high, but that it would have included many instances of nonsignificant pain. Because the criterion "bad enough to limit your usual activities or change your daily routine for more than 1 day" was in competition with VAS and NRS in the "Severity" category from the start, it was discarded in favor of VAS and NRS in the third round (where participants could only choose 1 item per category). However, in the first 3 rounds, this item was rated just below VAS and NRS. It must be men-

Figure 2. Final <u>optimal</u> definition of low back pain that results from the Delphi study. The diagram is used with permission.[29] Elements can be combined as investigators see fit to provide different specific definitions (see examples in Figure 3). [1]The diagram should be used in face-to-face interviews and paper or online questionnaires and omitted in telephone surveys, as detailed in the minimal definition (Figure 1). [2]Questions on frequency, duration and severity can be used for sciatica by replacing "low back pain" by "pain that goes down the leg." [3]For reporting, categories are: Mild = <7/10 and Severe = ≥7/10. ***The SF-36 Bodily Pain Scale[10,22,23] and Graded Chronic Pain Scale (GCPS)[10,24−26] are also suggested as alternative optimal definitions because they have been well studied and there has been extensive assessment of their psychometric qualities. The GCPS has been often used in back pain prevalence studies.

Time frame: <u>**In the past 4 weeks,**</u>

Site and symptoms:[1] **have you had pain in your low back (in the area shown on the diagram)?**

Yes ☐   No ☐

If yes, was this pain bad enough to limit your usual activities or change your daily routine for more than one day?

Yes ☐   No ☐

Sciatica: **have you had pain that goes down the leg?**

Yes ☐   No ☐

If yes, has this pain spread below the knee?

Yes ☐   No ☐

Exclusions: **Please do not report pain from feverish illness or menstruation.**

Frequency:[2] **If you had pain in your low back** <u>**in the past 4 weeks,**</u> **how often did you have the pain?**

On some days ☐   On most days ☐   Every day ☐

Duration:[2] **If you had low back pain** <u>**in the past 4 weeks,**</u> **how long was it since you had a whole month without any low back pain? (Please tick only one box).**

Less than 3 months   ☐
3 months or more but less than 7 months   ☐
7 months or more but less than 3 years   ☐
3 years and more   ☐

Severity:[2,3] **If you had low back pain** <u>**in the past 4 weeks,**</u> **please indicate what was the usual intensity of your pain on a scale of 0 to 10, where 0 means "no pain" and 10 means "the worst pain imaginable"? (Please circle your answer).**

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
No pain   Worst pain

---

tioned that using this second question in the minimal definition makes it more specific, but does not preclude comparisons on the responses to the first question if wanted.

The use of the time frames "Today" and "In the past year" was claimed essential in some investigations by several experts. Maintaining multiple time frames within the standard definitions has implications for the comparability and summary of studies using the different time frames, and in the wording of the supplementary elements of the definitions. For example, when using "To-

day," the wording of the questions would need to be changed (to use the present tense of the verbs), and the minimal severity criterion ("bad enough to limit your usual activities or change your daily routine for more than 1 day") and the question on frequency would have to be omitted. Alternatively, questions on duration and severity were not considered to provide valid answers when used with the time frame "In the past year." We thus recommend to use "In the past 4 weeks" for the standard optimal definition and to add other time frames if necessitated by the study purposes, settings, and methods.

Example 1. If the investigator is interested in the prevalence of severe long-standing low back pain and conducts a survey using postal questionnaires, the following questions should be asked:

Q1-In the past 4 weeks, have you had pain in your low back (in the area shown on the diagram)? Please do not report pain from feverish illness or menstruation.

Yes ☐   No ☐

Q2-If yes, was this pain bad enough to limit your usual activities or change your daily routine for more than one day?
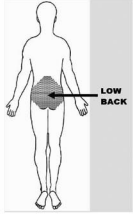
Yes ☐   No ☐

Q3-If you had low back pain in the past 4 weeks, how long was it since you had a whole month without any low back pain? (Please tick only one box).

Less than 3 months ☐
3 months or more but less than 7 months ☐
7 months or more but less than 3 years ☐
3 years and more ☐

Q4-If you had low back pain in the past 4 weeks, please indicate what was the usual intensity of your pain on a scale of 0 to 10, where 0 means "no pain" and 10 means "the worst pain imaginable"? (Please circle your answer).

0   1   2   3   4   5   6   7   8   9   10
No pain   Worst pain

Example 2. If the investigator is interested in the prevalence of sciatica and conducts a telephone interview, the following questions should be asked:

Q1-In the past 4 weeks, have you had pain that goes down the leg?

Yes ☐   No ☐

Q2-If yes, was this pain bad enough to limit your usual activities or change your daily routine for more than one day?

Yes ☐   No ☐

Q3-If you had pain that goes down the leg in the past 4 weeks, has this pain spread below the knee?

Yes ☐   No ☐

Q4-If you had pain that goes down the leg in the past 4 weeks, how often did you have the pain?

On some days ☐   On most days ☐   Every day ☐

Figure 3. Examples of optimal definitions of low back pain built from the items presented in Figure 2.

Of all elements, it is the definition of "Duration" that has been most difficult to reach consensus on. Although the most highly rated items were "Total duration of this event" and "Acute (<3 months)/Chronic (≥3 months)," neither reached consensus for the optimal definition and both were the source of strong comments from the participants. The suggestion made by some panel members to use "how long since you had a whole month without any low back pain" was thus followed and agreed on by the majority, but it was mentioned several times that the first category of the initial question (0–7 months) was too broad and hence susceptible to missing out more acute low back pain episodes. To account for this point, the first category was split into 2 (<3 months and 3–7

months). Although the original classification had been validated and found useful to distinguish patients according to their prognosis,[11,34] the new definition will need to be tested further.

The SF-36 Bodily Pain Scale[10,22,23] and GCPS[10,24–26] are suggested as alternative optimal definitions because they have been well studied and there has been extensive assessment of their psychometric qualities. The GCPS has been often used in back pain prevalence studies, which would allow for comparisons. Despite that, these definitions did not reach consensus. Indeed, several experts consider them too complex for use as simple brief tools for population surveys, which may explain why they are not used more often.

It is clear that no single definition, whether minimal or optimal, can meet the needs of all studies. We suggest that where researchers find that the standard definition does not include their preferred items, they should consider using them in addition to the standard definition, rather than instead of, so that comparisons remain facilitated. This would also allow further research comparing different versions of the questions and elements, and permit review of these standard definitions in the future.

As a whole, this work should be considered as a step toward better standardization of definitions of low back pain. The use of these definitions will allow the comparison of low back pain prevalence figures for different countries, age groups, settings, and occupational groups, among others, and will facilitate meta-analysis of results, which is currently difficult or impossible. Validation work and use will allow researchers to test these definitions and improve them as evidence for the validity of specific elements emerges. At present, maintaining the wording and presentation of the current definitions will allow maximum benefit to be gained from the research conducted.

One important strength of the study is that it included several international back pain investigators, many of whom are multilingual, therefore maximizing the possibilities for cross-cultural translation. It is, however, currently a limitation that the definitions are only published in English. It will also be important to develop and publish standard translations. The importance of translation should not be underestimated. For instance, in the German language, there is no generally accepted word for "low back pain"; pain in any part of the spine can be back pain, which contradicts the consensus on the meaning of back pain. Readers who are interested to contribute to specific translations are invited to contact the first author. Validated translations of the definitions will be posted on the SPINE website as they become available.

Last, but not least, it is essential to report in scientific publications precisely which definition(s) of back pain prevalence has(ve) been used, so that apples can be compared with apples, and oranges with oranges, and that each one can be distinguished from another.

■ Conclusion

Using a modified Delphi approach, 2 standard definitions of low back pain for use in epidemiological studies have been developed and agreed on. The aim for their use is to improve the synergy and potential for comparison between back pain studies internationally. Their use is not intended to detract from, or discourage, the additional use of other validated questions and instruments in epidemiological studies. Widespread dissemination of these definitions will optimize their usefulness.

■ Key Points

- There is large heterogeneity in reports of back pain prevalence in the general population that limits or renders impossible the valid comparison or summary of the results from multiple investigations.
- A modified Delphi study was conducted with 28 experts from 12 countries to identify standardized definitions of low back pain for use in epidemiological studies.
- For the majority (82%) of experts, the expression "back pain" only means "low back pain" and could not be used to include neck, thoracic, and low back pain.
- Two definitions were agreed on (minimal and optimal), providing standards that may improve the validity of future comparisons of low back pain prevalence figures and facilitate statistical summaries.

📄 *document*

Appendix available online through Article Plus.

this study as a panel member. Professor Nachemson died on December 4, 2006.

Panel of 28 international experts on back pain: Rachelle Buchbinder, Monash Department of Clinical Epidemiology at Cabrini Hospital, and Department of Epidemiology and Preventive Medicine, Monash University, VC, Australia; Bruce F. Walker, School of Chiropractic, Division of Health Sciences, Murdoch University, Murdoch, WA, Australia; Mary Wyatt, Department of Epidemiology and Preventive Medicine, Monash University, VC, Australia; J. David Cassidy, Division of Health Care Outcomes and Research, Toronto Western Research Institute, University Health Network, and Faculty of Medicine, Department of Public Health Sciences, University of Toronto, Toronto, ON, Canada; Michel Rossignol, Department of Epidemiology and Biostatistics, McGill University, Montreal, QC, Canada; Charlotte Leboeuf-Yde, The Back Research Centre and University of Southern Denmark, Denmark; Jan Hartvigsen, Clinical Locomotion Science, Institute for Sports Science and Clinical Biomechanics, University of Southern Denmark, Denmark; Päivi Leino-Arjas, Finnish Institute of Occupational Health, Musculoskeletal Research, Helsinki, Finland; Ute Latza, Institute of Occupational Medicine, University of Hamburg; Hamburg State Department for Social Affairs, Family, Health, and Consumer Protection, Hamburg, Germany; Shmuel Reis, Departments of Family Medicine and Medical Education, Ruth & Bruce Rappaport Faculty of Medicine, Technion—Israel Institute of Technology, Haifa, Israel; Maria Teresa Gil del Real, Spanish Back Pain Research Network, Kovacs Foundation, Madrid, Spain; Francisco M. Kovacs, Spanish Back Pain Research Network, Kovacs Foundation, Palma de Mallorca, Spain; Birgitta Öberg, Department of Health and Society, Linköpings Universitet, Linköping, Sweden; Christine Cedraschi, Division of General Medical Rehabilitation & Multidisciplinary Pain Centre, and Division of Clinical Pharmacology and Toxicology, University Hospitals, Geneva, Switzerland; Lex M. Bouter, EMGO Institute, VU University Medical Center, Amsterdam, The Netherlands; Bart W. Koes, Department of General Practice, Erasmus Medical Center, Rotterdam, The Netherlands; H. Susan J. Picavet, Centre for Prevention and Health Services Research, National Institute of Public Health and the Environment, Bilthoven, The Netherlands; Maurits W. van Tulder, Institute for Research in Extramural Medicine and the Department of Clinical Epidemiology & Biostatistics, Vrije Universiteit Medical Centre, The Netherlands; Kim Burton, Spinal Research Unit, University of Huddersfield, West Yorkshire, UK; Nadine E. Foster, Primary Care Musculoskeletal Research Centre, Primary Care Sciences, Keele University, Staffordshire, UK; Gary J. Macfarlane, Aberdeen Pain Research Collaboration (Epidemiology Group), University of Aberdeen, Scotland, UK; Elaine Thomas, Primary Care Musculoskeletal Research Centre, Primary Care Sciences, Keele University, Staffordshire, UK; Martin Underwood, Centre for Health Sciences, Barts and the London, Whitechapel, London, UK; Gordon Waddell, UnumProvident Centre for Psychosocial and Disability Research, Cardiff University, Cardiff, UK; Paul Shekelle, Greater Los Angeles Veterans Affairs Healthcare System, Los Angeles, CA, USA; Ernest Volinn, Pain Research Center, Department of Anesthesiology, University of Utah, Salt Lake City, UT, USA; and Michael Von Korff, Center for Health Studies, Group Health Cooperative, Seattle, WA, USA.

References

1. Dionne CE, Dunn KM, Croft PR. Does back pain prevalence really decrease with increasing age? A systematic review. *Age Ageing* 2006;35:229–34.
2. Loney PL, Stratford PW. The prevalence of low back pain in adults: a methodological review of the literature. *Phys Ther* 1999;79:384–96.
3. Leboeuf-Yde C, Lauritsen JM. The prevalence of low back pain in the literature. A structured review of 26 Nordic studies from 1954 to 1993. *Spine* 1995;20:2112–18.
4. Bressler HB, Keyes WJ, Rochon PA, et al. The prevalence of low back pain in the elderly. A systematic review of the literature. *Spine* 1999;24:1813–19.
5. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32:1008–15.
6. Powell C. The Delphi technique: myths and realities. *J Adv Nurs* 2003;41:376–82.
7. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ* 1995;311:376–80.
8. Salovey P, Seiber W, Smith A, et al. *Reporting chronic pain episodes on health surveys. Vital Health Statistics* 6. National Center for Health Statistics; 1992.
9. Carey TS, Garrett J, Jackman A, et al. Reporting of acute low back pain in a telephone interview. Identification of potential biases. *Spine* 1995;20:787–90.
10. Von Korff M. Epidemiologic and survey methods: chronic pain assessment. In: Turk DC, Melzack R, eds. *Handbook of Pain Assessment*. 2nd ed. New York: Guilford Press; 2001:603–18.
11. Dunn KM, Croft PR. The importance of symptom duration in determining prognosis. *Pain* 2006;121:126–32.
12. Von Korff M, Deyo RA, Cherkin D, et al. Back pain in primary care. Outcomes at 1 year. *Spine* 1993;18:855–62.
13. Waddell G. *The Back Pain Revolution*. Edinburgh: Churchill Livingstone; 1998.
14. Cedraschi C, Robert J, Goerg D, et al. Is chronic non-specific low back pain chronic? Definitions of a problem and problems of a definition. *Br J Gen Pract* 1999;49:358–62.
15. Bolton JE, Wilkinson RC. Responsiveness of pain scales: a comparison of three pain intensity measures in chiropractic patients. *J Manipulative Physiol Ther* 1998;21:1–7.
16. Croft P, Raspe H. Back pain. *Baillieres Clin Rheumatol* 1995;9:565–83.
17. Ferraz MB, Quaresma MR, Aquino LR, et al. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *J Rheumatol* 1990;17:1022–4.
18. Grotle M, Brox JI, Vollestad NK. Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. *J Rehabil Med* 2003;35:241–7.
19. Strong J, Ashton R, Chant D. Pain intensity measurement in chronic low back pain. *Clin J Pain* 1991;7:209–18.
20. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain* 1994;58:387–92.
21. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain* 1986;27:117–26.
22. Ware J. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: The Health Institute, New England Medical Center; 1993.
23. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160–4.
24. Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain. *Pain* 1992;50:133–49.
25. Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. *Spine* 1999;24:1104–12.
26. Penny KI, Purves AM, Smith BH, et al. Relationship between the chronic pain grade and measures of physical, social and psychological well-being. *Pain* 1999;79:275–9.

27. Jensen MP, Chen C, Brugger AM. Interpretation of visual analog scale ratings and change scores: a reanalysis of two clinical trials of postoperative pain. *J Pain* 2003;4:407–14.

28. Bodian CA, Freedman G, Hossain S, et al. The visual analog scale for pain: clinical significance in postoperative patients. *Anesthesiology* 2001;95:1356–61.

29. Kuorinka I, Jonsson B, Kilbom A, et al. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Appl Ergon* 1987;18:233–7.

30. de Vet HC, Heymans MW, Dunn KM, et al. Episodes of low back pain: a proposal for uniform definitions to be used in research. *Spine* 2002;27:2409–16.

31. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine* 1998;23:2003–13.

32. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 2000;25:3100–3.

33. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders. Introduction. *Spine* 2000;25:3097–9.

34. Dunn KM, de Vet HCW, Hooper H, et al. Measurement of back pain episode inception in questionnaires: a study combining quantitative and qualitative methods. *J Musculoskelet Pain* 2006;14:29–37.

35. Jensen MP, Turner LR, Turner JA, et al. The use of multiple-item scales for pain intensity measurement in chronic pain patients. *Pain* 1996;67:35–40.

36. Anderson JA. Epidemiological aspects of back pain. *J Soc Occup Med* 1986;36:90–4.

37. Dickinson CE, Campion K, Foster AF, et al. Questionnaire development: an examination of the Nordic Musculoskeletal questionnaire. *Appl Ergon* 1992;23:197–201.

38. Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.

39. Borkan J, Reis S, Hermoni D, et al. Talking about the pain: a patient-centered study of low back pain in primary care. *Soc Sci Med* 1995;40:977–88.