



## **Original article**

---

Scand J Work Environ Health [1999;25\(5\):404-409](#)

doi:10.5271/sjweh.452

### **Comparison of four peak spinal loading exposure measurement methods and their association with low-back pain**

by [Neumann WP](#), [Wells RP](#), [Norman RW](#), [Andrews DM](#), [Frank J](#), [Shannon HS](#), [Kerr MS](#)

The following article refers to this text: [2009;35\(4\):245-260](#)

**Key terms:** [biomechanical load](#); [epidemiology](#); [exposure measurement](#); [low-back pain](#)

This article in PubMed: [www.ncbi.nlm.nih.gov/pubmed/10569459](http://www.ncbi.nlm.nih.gov/pubmed/10569459)

---



This work is licensed under a [Creative Commons Attribution 4.0 International License](http://creativecommons.org/licenses/by/4.0/).

## Comparison of four peak spinal loading exposure measurement methods and their association with low-back pain

by W Patrick Neumann, MSc,<sup>1</sup> Richard P Wells, PhD,<sup>1,2</sup> Robert W Norman, PhD,<sup>1</sup> David M Andrews, PhD,<sup>3</sup> John Frank, MD,<sup>2,4</sup> Harry S Shannon, PhD,<sup>2,5</sup> Mickey S Kerr, PhD<sup>2,4</sup>

Neumann WP, Wells RP, Norman RW, Andrews DM, Frank J, Shannon HS, Kerr MS. Comparison of four peak spinal loading exposure measurement methods and their association with low-back pain. *Scand J Work, Environ Health* 1999;25(5):404—409.

**Objectives** This paper examines the performance of 4 different methods of estimating peak spinal loading and their relationship with the reporting of low-back pain.

**Methods** The data used for this comparison was a subset of subjects from a case-referent study of low-back-pain reporting in the automotive industry, in which 130 random referents and 105 cases (or job-matched proxies) were studied. The peak load on the lumbar spine was determined using a biomechanical model with model inputs coming from a detailed self-report questionnaire, a task-based check list, a video digitization method, and a posture and load sampling technique.

**Results** The methods were directly comparable through a common metric of newtons or newton meters of spinal loading in compression, shear, or moment modes. All the methods showed significant and substantial associations with low-back pain in all modes (odds ratios 1.6—2.3). The intraclass correlation coefficients (ICC) showed strong similarities between the checklist and video digitized techniques (ICC 0.84—0.91), moderate similarities between these techniques and the work sampling method (ICC 0.49—0.52), and poor correlations (ICC 0.16—0.40) between the self-report questionnaire and the observer recorded measures.

**Conclusions** While all the methods detected significant odds ratios, they cannot all be used interchangeably for risk assessment at the individual level. Peak spinal compression, moment, and shear are important risk factors for low-back pain reporting, no matter which measurement method is used. Questionnaires can be used for large-scale studies. At the individual level a task-based checklist provides biomechanical model inputs at lower cost and equal performance compared with the criterion video digitization system.

**Key terms** biomechanical load, epidemiology, exposure measurement, low-back pain.

Recently, several studies using high-quality exposure measures have demonstrated strong associations between physical work exposures and a risk of low-back pain (1—3). These and other studies have identified widely differing exposure variables that include heavy work, forceful movements, awkward postures, and moments of force, all of which have different units of measurement (4). Wells et al (5) have identified these differences between the risk factors as a potential problem when the results of different studies are to be compared in, for example, a meta analysis. The use of a “common metric” approach whereby different measurements strategies yield a

consistent and comparable unit of exposure has been proposed (5).

The data presented in this paper were obtained as part of a large epidemiologic study of risk factors for reported low-back pain in the automotive industry. It identified psychosocial, biomechanical, and psychophysical variables all as independent, significant, and substantial risk factors (6). Further analysis of the biomechanical exposure data bases of this study identified the following 4 separate risk factors that are independently associated with the risk of low-back pain: peak spinal load, integrated spinal load, trunk kinematics, and exter-

1 Department of Kinesiology, University of Waterloo, Waterloo, Ontario, Canada.

2 Institute for Work and Health, Toronto, Ontario, Canada.

3 Department of Kinesiology, McMaster University, Hamilton, Ontario, Canada.

4 Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada.

5 Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada.

Reprint requests to: Mr WP Neumann, Faculty of Applied Health Sciences, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. [E-mail: pneumann@healthy.uwaterloo.ca]

nal (or hand) forces (1). Exposure to spinal loads during work can be assessed using biomechanical models that calculate loading in terms of newtons of compression, newton meters ( $N \cdot m$ ) of moment, or newtons of shear; these are metrics that are independent of the method used to collect the model input data. The data required to drive the biomechanical assessment can be obtained by several different methods, including questionnaires, pencil and paper observation approaches, and several video digitization or more technological methods.

We have applied 4 different measurement strategies within this epidemiologic study of reported low-back pain. All 4 methods (digital video analysis, a paper-based posture and load checklist, work posture and load sampling, and a self-report posture and load questionnaire) were designed to generate inputs for the same biomechanical model, which yields spinal loading estimates for each measure's input. The peak spinal load estimates resulting from input data collected using each method could then be compared directly in their common metric forms. The purpose of this paper was to determine if all 4 methods of estimating peak spinal loading on the job are equally capable of identifying risk and to determine a preferred approach for future research.

## Methods

The Ontario Universities Back Pain Study (OUBPS) was conducted between 1993 and 1995 at a large automotive assembly facility with a study base of over 10 000 hourly paid workers, including production and nonproduction (eg, maintenance) workers. The study used a case-referent design. Cases were defined as persons who, in the absence of a similar event in the previous 90 days, reported to the plant nursing station with low-back pain. If a case was not available for a work-site analysis then proxies, who performed the same work as the case, were substituted (3). Referents were randomly selected from the employee rosters and were screened to exclude those with nursing-station low-back pain reports within the previous 90 days. All the subjects received a detailed test battery including interview and self-administered questionnaires and a comprehensive multi-method analysis of physical loading on the job.

Each of the methods used for this comparative study was designed to generate data for input into a biomechanical model. The model is a quasidynamic, 15-member, link segment model that uses joint center coordinate, hand force amplitude and direction, and basic body anthropometric data to calculate estimates of compression, moment, and shear loading forces on the L4-L5 spine. While we recognize the assumptions and limitations of such models, we have attempted to incorporate as much content validity into the model as possible. The validity

of biomechanical models remains an issue in biomechanical research. (See reference 1 for details.) For our study the same model was used for all the methods and all the participants; thus any model bias would be the same for all the methods and participants, regardless of whether they were cases or referents. For each method, once model inputs were generated and subsequent model outputs for each loading mode were compiled, the highest instance of loading for each participant was taken to be the peak spinal load of that method. These methods are described elsewhere (5), but the methods used to obtain biomechanical model inputs are reviewed here in brief.

### *Video digitization*

During the course of the assessment each worker was video-taped in the sagittal plane while performing regular work duties. Research team members would then identify individual frames which they estimated would be the instant of peak loading for each of the tasks analyzed by the observer in the field. Their subjective, preliminary estimates were based upon their biomechanical knowledge of load handling, trunk postures, and hand forces, and estimated low-back moment of force. Any additional situations the team member identified as potentially relevant were also included. These specific frames would then be captured by the computer, and the workers' joint centers would be manually digitized using a computer mouse interface. If the video image was not suitable for digitization then a manipulable computer mannequin would be used to produce a simulation of the worker's posture in the instant of high loading. Load amplitude was assessed using a force transducer to measure loads or applied forces. The highest resulting spine load estimated from all the tasks and frames analyzed was recorded as the peak spine load for that job. Video digitization personnel were trained on the computer system until they were able to perform 5 consecutive analyses within a lumbar compression force of 170 N and a lumbar shear force of 50 N from the criterion values as determined by expert consensus. The video digitization method provides the most precisely quantified input for the biomechanical modeling procedure and is, for comparison purposes, deemed to be the criterion, or "gold standard" measure.

### *Task-based posture and load checklist*

The observers used a pencil-and-paper-based checklist for a detailed assessment of the physical loading exposures on each worker's job. They received intense training in the application of the checklist in the field until their results qualitatively matched those of a senior team leader. In performing the checklist assessments, the observer identified key tasks and performed a detailed assessment of each work component. For each identified task's instant of peak loading, determined similarly to the video-digitized method, the observer used posture

**Table 1.** Descriptive statistics and t-test results for 4 different measures of peak compression, peak moment, and peak shear forces at the L4/L5 spinal level, obtained using a self-report questionnaire, a detailed checklist, video digitization, and work sampling methods.

	Cases			Random			t-test probability
	N	Mean	SD	N	Mean	SD	
Compression (N)							
1. Video	105	3402	1429	129	2744	1070	0.0001
2. Checklist	105	3752	1768	130	2969	1332	0.0002
3. Questionnaire	61	3467	1827	81	2852	1489	0.0289
4. Work sampling	104	3293	1275	130	2675	1011	0.0001
Moment (N · m)							
1. Video	105	180	85	129	141	63	0.0001
2. Checklist	105	201	101	130	154	77	0.0001
3. Questionnaire	61	201	111	81	144	108	0.0028
4. Work sampling	104	178	77	130	140	61	0.0001
Shear (N)							
1. Video	105	462	178	129	354	159	0.0000
2. Checklist	105	465	189	130	363	161	0.0000
3. Questionnaire	61	439	161	81	327	230	0.0009

**Table 2.** Univariable odds ratios (OR) and 95% confidence intervals (95% CI) and estimated variance accounted for [R<sup>2</sup>(adj)] calculated for exposure differences equal to the interquartile spread from the random referents(Q3-Q1) for 4 different measures of peak compression force, peak moment of force, and peak shear force at the L4/L5 spinal level, obtained using a self-report questionnaire, a checklist, video digitization, and work sampling methods.

	Q3-Q1	OR	95% CI	R <sup>2</sup> (adj)
Compression (N)				
1. Video	1325	1.8	1.3—2.4	0.088
2. Checklist	2032	2.0	1.4—3.0	0.083
3. Questionnaire	2074	1.6	1.1—2.5	0.045
4. Work sampling	1129	1.7	1.3—2.3	0.091
Moment (N · m)				
1. Video	76	1.8	1.3—2.4	0.091
2. Checklist	109	2.0	1.4—2.9	0.088
3. Questionnaire	138	1.9	1.3—3.1	0.084
4. Work sampling	71	1.8	1.3—2.4	0.096
Shear (N)				
3. Video	190	2.1	1.5—3.0	0.126
2. Checklist	225	2.2	1.5—3.2	0.107
1. Questionnaire	294	2.3	1.4—4.1	0.096

**Table 3.** Intraclass correlation coefficients for measures of peak compression, peak moment, and peak shear at the L4/L5 spinal level, obtained using a self-report questionnaire, a detailed checklist, video digitization, and work sampling methods.

	1	2	3	4
Compression (N)				
1. Video	1.00			
2. Checklist	0.84	1.00		
3. Questionnaire	0.36	0.40	1.00	
4. Work sampling	0.50	0.49	0.16	1.00
Moment (N · m)				
1. Video	1.00			
2. Checklist	0.85	1.00		
3. Questionnaire	0.31	0.36	1.00	
4. Work sampling	0.52	0.51	0.16	1.00
Shear (N)				
1. Video	1.00			
2. Checklist	0.91	1.00		
3. Questionnaire	0.27	0.30	1.00	

sketches that represented categorical scales to record back posture (6 angles), upper-arm posture (8 angles), elbow angle (4 angles), and load direction (4-point scale). Load amplitudes were assessed using the same force transducer and measurement process as the video digitization technique.

*Self-report questionnaire*

The participating workers completed the self-report questionnaire on their own time without assistance. As part of this questionnaire, the workers were asked to identify the instant of their job that was heaviest on their back. They were then asked to describe, for this peak loading instant, their trunk and arm postures using stick figure diagrams (7). External load direction and amplitude were determined using 4-point and 6-point categorical scales, respectively, which were compatible with those used in the task-based checklist. Whenever possible, the questions used graphic images to assist the participant in selecting appropriate postures, which were then used to generate the required input to the biomechanical model.

*Work sampling*

The work sampling method used in this study involved over 250 observations randomly spaced at 10- to 20-second intervals over a 2-to-4-hour period. Observations were made using a pencil and paper notation scheme in which the worker's back posture, distance of arms from the torso, amplitude, and direction of force applied were recorded on categorical scales. The back posture categories included those from the checklist plus lateral, twisting, and sitting postures. The hand position was categorized as *close*, *medium* or *far*, where far represented arms horizontally outstretched. Force amplitude and direction were categorized on the same scales as the questionnaire. A biomechanical simulation, using median male and female anthropometrics (8), was run for all possible combinations of posture and load in order to provide

“look-up” tables of lumbar compression, moment, and shear forces. Interobserver reliability for peak compression force for this tool has previously been assessed to be high with an intraclass correlation coefficient of 0.83 (9).

All of the techniques described in this paper were applied to each worker simultaneously (with the exception of the self-report questionnaire) during a single physical loading exposure assessment as they performed their regular work duties. The assessments ranged from 2 to 8 hours with most assessments lasting about 4 hours (half a shift).

### Analysis

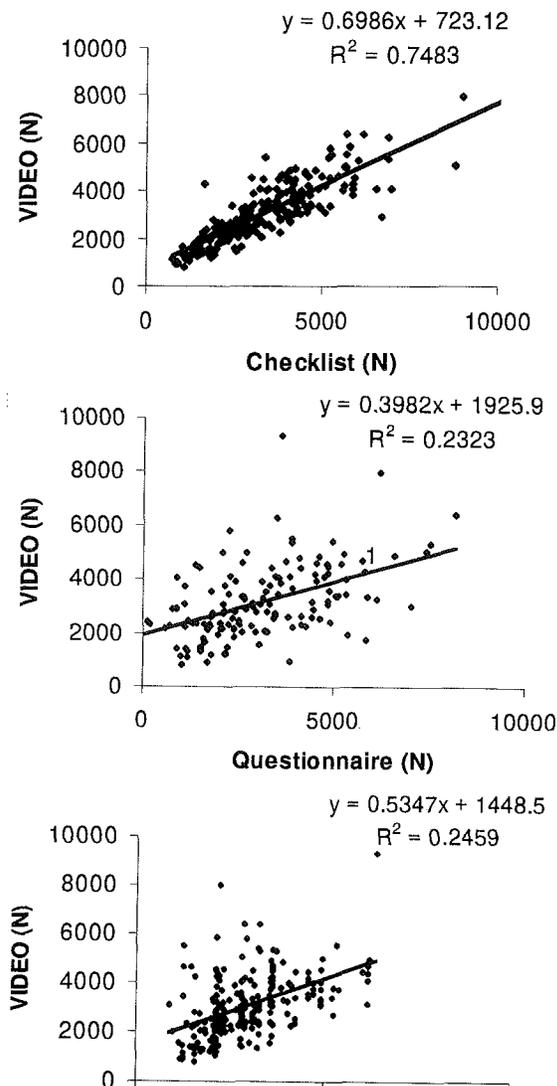
Differences between the cases and referents were first examined using Student's *t*-test for each method. No attempts were made to remove the variability within each of the methods as they would normally be applied by ergonomics practitioners. The shear loading force from the work sampling technique was excluded from the comparisons as the calculation was conceptually different from that of the other methods. Univariable odds ratios (OR) were calculated for all the variables using a logistic regression procedure, and they were calculated conservatively using an exposure difference equal to the interquartile spread of the random reference group. The similarity of each method at the individual level was assessed for each mode of loading using an intraclass correlation coefficient (10). An analysis of the peak compression data was performed using a linear regression analysis to explore the differences between video digitization and the other methods.

### Results

One hundred and five cases (including 20 proxy subjects) and 130 referents received direct physical loading assessments while performing their regular work duties. The descriptive statistics and *t*-test results are presented in table 1. Significant differences between the cases and referents were found at the group level for all variables on all methods with small intermethod differences in the mean scores for each group. The cases had higher exposures than the referents for all the comparisons. The questionnaire method had a relatively large number of missing data points due to incomplete answers to questions needed for the peak spinal loading estimates. The univariable odds ratios, derived from the logistic regression analysis, ranged from 1.6 to 2.3 at the interquartile spread (table 2). When the odds ratios were calculated using the full exposure range observed for the random referents, they ranged from 5.0 to 13.2, from 10.9 to 23.3, and from 13.2 to 21.8 for peak compression, peak moment, and peak shear forces, respectively. All of the variables

studied showed significant and substantial relationships with risk at the case and referent group level regardless of the method used to assess exposure.

The intraclass correlation coefficients (ICC) (table 3) indicated that the similarity of the methods at the individual level ranged from almost perfect (ICC >0.8) to low (ICC <0.2). The checklist and digitized video were the most similar methods with consistently high ICC values; the questionnaire data tended to be least similar to the other methods, having only fair to slight similarity. The work sampling method was moderately similar to the checklist and video methods but was not similar to the questionnaire. Scatter plots of the regression analysis, in which the compression force from each method was compared with the video digitization, confirmed the ICC results (figure 1).



**Figure 1.** Scatter plots and linear regression equations comparing a self-report questionnaire, checklist, and work sampling methods to the criterion video digitization method for peak compressive load on the L4/L5 spine (units are in Newtons).

**Discussion**

All the methods used to assess peak spinal loading showed significant case-referent differences and substantial odds ratios for the reporting of low-back pain. Furthermore, all 3 modes of lumbar spine load were significant risk factors regardless of whether compression, shear, or lumbar moments of force were considered. The amount of variance accounted for by each variable in this study ranged from 5% to 13% of the total variance (table 2). While this is a small fraction, peak spinal load represents only 1 variable in the multifactorial etiology of low-back pain (6). The amount of variance accounted for by each of the single variables was similar to that reported for the final predictive model by Bigos et al in the Boeing study (11). In a recent comprehensive review of the epidemiologic literature, conducted by the US National Institute for Occupational Safety and Health (NIOSH), 13 of the 18 studies examined found a positive relationship between back pain and work-related lifting and forceful movements (4). While the units of measurement for these risk factors varied between studies, and hence cannot be directly compared, the studies using more sophisticated measures of biomechanical exposure of individuals, rather than group classification, generally found higher risk estimates.

The intraclass correlation matrix indicated that, with the exception of the checklist and video digitization methods, these methods cannot be used interchangeably. While all the methods detected differences at the case-referent group level, the questionnaire and work sampling data did not correlate well with each other or with the video digitization method. Several factors could be contributing to the differences observed between the methods, including (i) the selection of different tasks to be analyzed, (ii) differences in the selection of the instant to be analyzed for each task, (iii) differences in the accuracy of the posture or load inputs to the biomechanical model, (iv) the precision of the estimate, and (v) effects due to interobserver variabilities.

A qualitative comparison of the tools used in this

study is presented in table 4. Self-report questionnaires are commonly used in epidemiologic studies since they can be applied to very large numbers of subjects with relatively little cost. Andrews et al (12) have shown the questionnaire used in this study to have good repeatability but decreasing accuracy for complex postures. This study shows that it is possible to get peak spine load estimates using a very detailed self-administered questionnaire. Such detail may come at the expense of a reduced response rate and increased measurement error however. An interview-assisted approach may provide better response rates but would also be substantially more expensive. The work-sampling technique requires roughly 2 hours to complete, depending on the diversity of the work tasks observed, and has a data entry time that is equivalent to that of a detailed questionnaire, although computerized implementation could speed up the work sampling procedure. Work sampling is particularly well suited to highly diversified work situations where task-breakdowns, as performed for the checklist approach, are difficult or inappropriate. Modification of this tool is recommended to include the ability to trap infrequently occurring instances of substantial loading, which might be missed in the sampling process.

There is a number of limitations to our case-referent epidemiologic study, the first of which concern limitations, due to the design itself, to the ability to infer causality between the risk factors identified and the outcome of interest (ie, the reporting of low-back pain). While steps were taken to limit the awareness of the field study teams to the worker's case-referent status, formal blinding was not feasible. However there was no reason to believe that systematic bias of the data analysts was present. Other potential limitations of this study include the context-sensitive criteria for the classification of cases and referents, based on reports of pain to the plant nursing staff, the use of proxy subjects for 20% of the cases, and measurement errors during the 2-to-8-hour field data collection sessions. While they remain potential sources of error, these factors would likely be random and affect both groups equally and thereby reduce, rather than over-

**Table 4.** Qualitative summary comparison of key aspects of the 4 data collection methods as implemented in this study for assessing peak spinal load in the workplace. This summary is based on the assessment of a standard production job, and the actual requirements may vary considerably depending on the application.

Comparison parameter	Method			
	Video	Checklist	Work sampling	Questionnaire
Time demands - participant	✓✓	✓✓	✓	✓✓✓✓
Time demands - operator	✓✓✓✓	✓✓✓	✓✓✓	✓
Data collection, equipment cost	✓✓✓✓	✓✓✓	✓✓	✓
Technological complexity	✓✓✓✓	✓✓	✓✓	✓
Operator training level	✓✓✓✓	✓✓✓	✓✓	✓
Observation, success rate	✓✓✓	✓✓✓✓	✓✓✓	✓✓
Precision of model inputs	✓✓✓✓	✓✓✓	✓	✓✓
Other considerations	Permanent video record; resistance to video in some areas	On-site collection; other risk factors assessable	Posture and load distributions available; good for nonrepetitive work	Worker perceptions available; other risk factors assessable

estimate, the likelihood of observing differences between the cases and referents (13). In spite of these limitations, consistent differences and significant odds ratios emerged no matter how spinal loading was estimated. The between-method consistency of the biomechanical findings, revealed through direct comparison of common metrics, serves to strengthen our confidence in these findings.

### Concluding remarks

Considered univariably, peak spinal load showed significant associations with low-back pain reporting for all 4 of the exposure measurement approaches. These methods could be compared directly through the use of the common metrics of lumbar force for peak compression (newtons), peak moment (newton meters), and peak shear (newtons), and they all provided very similar mean exposure estimates. However, with the exception of the task-based observer checklist and the video digitized analysis methods, these measurement approaches did not produce sufficiently similar results at the individual level to suggest that all 4 methods could be used interchangeably in data collection protocols.

The choice of a preferred tool is application dependent. Any of the 4 methods used can be used to identify risk factors at the group level. Self-report questionnaire methods can be used successfully to generate biomechanical model inputs, and consequently spinal load exposure outputs, on large numbers of workers with minimal cost. Work sampling approaches are particularly useful in complex work where the identification of individual tasks is problematic. For applications where individual estimates are needed, the task-based posture and load checklist approach provides an effective method of assessing risk in an industrial workplace.

### Acknowledgments

This work was funded by the Institute for Work & Health, whose core funding is provided by the Workplace Safety & Insurance Board of Ontario, Canada.

The authors would like to acknowledge all of the members of the Ontario Universities Back Pain Study (OUBPS) working group: Beaton DE, Bombardier C, Ferrier S, Hogg-Johnson S, Mondloch M, Peloso P, Smith J, Stansfeld SA, Tarasuk V, Dobbyn M, Edmondstone MA, Ingelman JP, Jeans B, McRobbie H, Moore A, Mylett J, Outerbridge G, Woo H. The OUBPS working group would like to thank General Motors of Canada Ltd, Mr Elmer Beddome, the Canadian Auto Workers' Union, Mr John Graham, the occupational health nurses, and all of the GM medical staff, as well as the study participants themselves.

### References

1. Norman R, Wells R, Neumann P, Frank J, Shannon H, Kerr M. A comparison of peak vs cumulative physical work exposure risk factors for the reporting of low back pain in the automotive industry. *J Clin Biomech* 1998;13:561—73.
2. Marras WS, Lavender SA, Leurgans SE, Rajulu SL, Allread WG, Fathallah MS, et al. The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders — the effects of workplace factors, trunk position, and trunk motion characteristics on risk of injury. *Spine* 1993;18(5):617—28.
3. Punnett L, Fine LJ, Keyserling WM, Herrin GD, Chaffin DB. Back disorders and nonneutral trunk postures of automobile assembly workers. *Scand J Work Environ Health* 1991;17:337—46.
4. National Institute for Occupational Safety and Health. Musculoskeletal disorders and workplace factors — a critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back. Baltimore (MD): US Department of Health and Human Services, 1997. Document 97—141.
5. Wells R, Norman R, Neumann P, Andrews D, Frank J, Shannon H, Kerr M. Assessment of physical work load in epidemiologic studies: common measurement metrics for exposure assessment. *Ergonomics* 1997;40(1):51—61.
6. Kerr S, Shannon HS, Frank JW, Norman RWK, Wells RP, Neumann P. The relative importance of biomechanical and psychosocial risk factors in a case-control study of low-back pain. In: Seppälä P, Luopajarvi T, Nygård C-H, Mattila M, editors. Proceedings of the International Ergonomics Association's 13th Triennial Congress 1997, Tampere, Finland. Helsinki: Finnish Institute of Occupational Health, 1997:64—5.
7. Andrews DM, Norman RW, Wells RP, Neumann P. The Accuracy of self-report and trained observer methods for obtaining estimates of peak load information during industrial work. *Int J Ind Ergon* 1997;19:445—55.
8. Government of Canada. Canadian Standardized Test of Fitness (CSTF) operations manual. 3rd edition. Ottawa (Ontario): Government of Canada — Fitness and Amateur Sport, 1986.
9. Edmondstone MA, Neumann P, Wells R, Norman R. Interobserver reliability in work sampling: a measurement system for back posture and spinal loading. In: Proceedings of the 28th annual conference of the Human Factors Association of Canada. Mississauga (Ontario): Human Factors Association of Canada, 1996:79—84.
10. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86(2):420—8.
11. Bigos SJ, Battie MC, Spengler DM, Fisher LD, Fordyce WE, Hansson T, et al. A longitudinal, prospective study of industrial back injury reporting. *Clin Orthop* 1992;279:21—33.
12. Andrews DM, Norman RW, Wells RP. Accuracy and repeatability of low back spine compression force estimates from self-reports of body posture during load handling. *Int J Ind Ergon* 1996;18:251—60.
13. Kerr MS. A case-control study of biomechanical and psychosocial risk factors for low-back pain reported in an occupational setting [dissertation]. Toronto (ON): University of Toronto, 1997.

Received for publication: 24 September 1998