Scand J Work Environ Health 1993;19(3):208-214
https://doi.org/10.5271/sjweh.1481
Issue date: 01 Jun 1993

Validity of self-reported exposures to work postures and manual materials handling. Stockholm MUSIC I Study Group.

by Wiktorin C, Karlqvist L, Winkel J

Affiliation: Karolinska Hospital and Institute, Department of Occupational Health, Stockholm, Sweden.

The following articles refer to this text: 2017;43(5):415-425; 2018;44(1):58-68; 2018;44(2):163-170; 2023;49(4):283-292; 2024;50(5):317-328

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/8367699

Print ISSN: 0355-3140 Electronic ISSN: 1795-990X

Scand J Work Environ Health 1993;19:208—14

Validity of self-reported exposures to work postures and manual materials handling

by Christina Wiktorin, BS,[1] Lena Karlqvist, PT,[1] Jörgen Winkel PhD,[2] Stockholm MUSIC I study group[3]

WIKTORIN C, KARLQVIST L, WINKEL J, STOCKHOLM MUSIC I STUDY GROUP. Validity of self-reported exposures to work postures and manual materials handling. *Scand J Work Environ Health* 1993;19:208—14. Exposure data from self-administered questionnaires on manual materials handling and work postures were validated in relation to direct measurements and systematic observations on 39 men and 58 women representing 45 different occupations. The agreement was tested at a dichotomous level and, when possible, with quantification of duration or frequency. At the dichotomous level the agreement was "acceptable" for nine variables concerning work postures and the handling of loads weighing >5 kg. No variable showed "acceptable" agreement when the duration or the frequency was quantified in more detail (4- to 6-point scales). Musculoskeletal complaints seemed to introduce a differential bias for some lifting variables. Thus some variables for postures or the handling loads of >5 kg may, under certain conditions, be acceptable for use in epidemiologic studies when the relative risks are high. However, self-reported exposure seems to be too crude if more-detailed information is required.

Key terms: differential bias, epidemiology, ergonomics, evaluation, physical work load, questionnaire.

Classifying exposure correctly is as essential in occupational studies on musculoskeletal disorders as it is in other fields of epidemiology (1). It is also important when improvements made in the work environment are evaluated or when guidelines for hazardous work loads are constructed. For correct classification, an accurate method with satisfactory measurement precision must be used. Many studies indicate that prolonged work in strenuous postures or heavy manual materials handling causes or accelerates musculoskeletal disorders. (For reviews see references 2—5.) Physical exposures have generally been classified indirectly from job titles (2, 5). When work postures or manual materials handling have been measured directly, the data have usually been collected from self-reports (ie, the workers themselves have estimated their exposure) (2). However, only a few and partly contradictory studies have been published concerning the validity of quantitative data on work postures and manual materials handling obtained by self-reports (6—10). The aim of the present study was to determine, by means of a questionnaire, how well subjects quantify their exposure to work postures and manual materials handling.

Our study was part of the work carried out by a more comprehensive research group developing effective methods for physical load assessment in large populations. This "exposure group" then belongs to the investigation being carried out by the Stockholm MUSIC I study group, which is working to validate methods to be recommended for large epidemiologic studies. "MUSIC" is an acronym for "musculoskeletal intervention center" and is comprised of a Stockholm network of 10 departments whose objective is to prevent musculoskeletal disorders.

Subjects and methods

Subjects

Ninety-seven subjects participated in the study. Seventy-two (27 men and 45 women) were randomly selected from the working population of the Stockholm area, 12 were furniture removers (men), and 13 were medical secretaries (women). The two latter groups were included to ensure the occurrence of extreme exposures such as heavy lifting and prolonged sitting. All of the subjects had been employed in their current work for at least three months.

Questionnaire

A questionnaire with nine illustrated questions concerning different work postures and eight questions

1 Karolinska Hospital and Institute, Department of Occupational Health, Stockholm, Sweden.
2 National Institute of Occupational Health, Division of Applied Work Physiology, Solna, Sweden.
3 Karolinska Hospital and Institute, Departments of Epidemiology/IMM, Occupational Health, Orthopaedics, Rheumatology, and Rehabilitation and Physical Medicine, Stockholm, Sweden; National Institute of Occupational Health, Divisions of Applied Work Physiology, Occupational Medicine, Psychophysiology, and Work and Environmental Physiology, Solna, Sweden; National Institute of Psychosocial Factors and Health, Stockholm, Sweden.

Reprint requests to: Ms C Wiktorin, Department of Occupational Health, Karolinska Hospital, Box 60 500, S-104 01 Stockholm, Sweden.

about manual materials handling was evaluated (appendix 1, part A). The questions described risk factors suggested in previous studies. Three different quantification scales were used (appendix 1, part B). The time scale had previously been developed by Statistics Sweden (11), and the frequency scale was derived from questionnaires used in other studies.

Reference measurements

Four questions (numbers 1, 2, 3, and 17 in appendix 1, part A) were validated by direct simultaneous measurements made during an entire workday. Question 17, concerning walking distance, was validated with a pedometer (Fitty 3 Electronic, Germany), which records the number of steps taken (12). The walked distance was obtained by multiplying the number of steps by the step length. The step length was obtained from the number of steps taken when the subject walked a distance of 5 m. Question number 1, on sitting, was validated with a posimeter (Posimeter 100, Sweden) (Selin et al, unpublished results). A posimeter registers the time spent in one of two positions relative to the line of gravity. The sensor of the posimeter was attached to the thigh of a standing person at an angle of 45 degrees to the horizontal plane, and the box for the data collection was placed on a waist band. Two questions on trunk flexion (appendix 1, part A, numbers 2 and 3) were validated with an inclinometer, a trunk flexion analyzer (13). It registers the time spent in each of seven 18-degree intervals relative to gravity. Time spent in 18—54 degrees of forward flexion was compared with self-reports in question 2, and time spent in more than 54 degrees was compared with the responses to question 3 (appendix 1, part A). The remaining 13 questions (appendix 1, part A, numbers 4—16) were validated with a computerized observation method (14). The duration and frequency of these exposures were registered on portable computers (Olivetti or Psion) by experienced ergonomists.

Procedure

Each subject was studied during an ordinary workday. The pedometer, the posimeter, and the inclinometer were attached to the subject, and the registrations lasted the entire day — except for a lunch break — or as long as the subject accepted wearing the equipment. Observations were made for one event at a time and lasted as long as the ergonomist could maintain her concentration. The average recording time for each of the reference methods is shown in table 1. At the end of each recording period the subjects answered the question corresponding to the previously registered event. The subjects were informed when each measurement started but not what kind of exposure they should quantify at the end of the recording period.

Data treatment and statistical analysis

The relationship between the data obtained from the self-reports and reference methods was analyzed for the pooled study population (N = 97). The reference measurements revealed that some of the exposure variables had a short duration or a low frequency. These exposures were treated separately and called group I variables. A group I variable was defined as an exposure for which at least 80% of the subjects obtained reference measurements corresponding to scale item 1 or 2 — see appendix 1, parts B and C. For group I variables the analyses were reduced to test the ability of the subjects to discriminate between being "exposed" (scale items 2 or greater) and "unexposed" (scale item 1) by calculating a kappa coefficient. Kappa was calculated only if at least 10 subjects belonged to each of the two categories "exposed" and "unexposed" with regard to the reference measurements.

The remaining exposure variables (group II) were evaluated from calculations of the rank-order correlation coefficient (Spearman's coefficient) and a kappa coefficient. Before the kappa coefficients were calculated, the reference measurements were classified according to the scale items used in the self-reports (appendix 1, part C). The kappa coefficient was used to evaluate the accuracy of the self-reports by describing the proportion of the subjects that reported time or frequency in the same class as the reference measurements. To evaluate the accuracy in quantifying the time or frequency, we used the following kappa values: <0.40 = poor accuracy, 0.40—0.75 = fair to good accuracy, >0.75 = excellent accuracy (15). If the accuracy was poor, the quantification classes were pooled and the kappa was recalculated for fewer and wider classes. When the scales were dichotomized, the cutoff was made between scale item 1 and 2 of the original scales (appendix 1, parts B and C), as well as between the two items offering the highest kappa value. The latter procedure was done to find a possible nonlinear relation between the self-reports and the reference methods. The correlation coefficient (r_s) added information about the relationship, as it took the degree of misclassifica-

Table 1. Number of subjects studied, number of questions validated, and average recording time for each of the reference methods applied in the study.

Means of validation	Subjects: Men (N)	Subjects: Women (N)	Questions (N)	Average recording time
Pedometer	38	56	1	5 h 48 min
Posimeter	36	53	1	5 h 48 min
Trunk flexion analyzer	33	50	2	3 h 36 min
Systematic observation	39	50	13	0 h 26 min

tion into consideration, as well as systematic over- or underestimation.

Subjects with musculoskeletal disorders may overestimate or underestimate their exposure more than subjects without such symptoms. This possibility could introduce a differential bias when self-reports are used in epidemiologic studies (16). Thus subjects reporting complaints during the last year from the lumbar spine, the shoulder or neck region, or the knees, according to the Nordic questionnaire (17), were compared with subjects without complaints from the corresponding parts of the body. To reduce the amount of data without loosing important information, we concentrated the analysis on biases that would change the size of the incidence rate ratios in an epidemiologic study. For each exposure variable showing at least acceptable agreement, the prevalence odds ratio was calculated both from the self-reports (POR_{quest}) and from the reference measurements (POR_{ref}). The prevalence odds ratio was calculated because it could be assumed to be a valid estimate of the incidence rate ratio if one assumed the same duration of disease in both the exposed and unexposed groups (16). If the size of POR_{quest} was similar to POR_{ref}, we assumed that musculoskeletal complaints did not cause any differential bias in the self-reports. If the size of POR_{quest} was higher than POR_{ref}, a differential bias was indicated. A lower POR_{quest} than POR_{ref} indicated either a differential or a nondifferential bias in the self-reported exposures.

Results

Agreement for variables with a short duration and low frequency (group I)

Twelve variables belonged to group I. The ability to discriminate between unexposed and exposed subjects was excellent for the kneeling-squatting question (kappa 0.76) and acceptable for the question concerning the trunk being bent forward more than 60 degrees (kappa 0.43) and for all questions concerning the handling of weights of more than 5 kg (kappa 0.50—0.66) (table 2). The ability to discriminate between unexposed subjects and those with a short length of exposure was poor for the questions concerning head rotation, hands above shoulder level, and carrying, pushing or pulling loads of 1—5 kg (kappa 0.17—0.26). Trunk rotation of more than 45 degrees, carrying, pushing or pulling loads of more than 45 kg, and lifting more than 45 kg were such uncommon exposures that no calculations were made.

Table 2. Relation between the self-reports and reference measurements for the 12 group I variables according to the kappa coefficients and the percentage of full agreement (full %) for all of the subjects participating in the study (N = 97).

Variables	Kappa	Full %
Trunk bent forward >60° (time)	0.43	88
Trunk rotation >45° (time)	·	61
Hands above shoulder level (time)	0.17	57
Head rotation (time)	0.17	58
Kneeling or squatting (time)	0.76	88
Carrying, pushing or pulling		
1—5 kg (time)	0.26	63
6—15 kg (time)	0.50	79
16—45 kg (time)	0.64	90
>45 kg (time)	·	90
Lifting		
6—15 kg (frequency)	0.66	86
16—45 kg (frequency)	0.65	93
>45 kg (frequency)	·	90

Agreement for variables with varied duration and frequency (group II)

Only five variables belonged to group II (table 3). The accuracy in quantifying the duration or frequency was poor for all five variables for which this ability could be tested (kappa 0.06—0.35). However, there was a statistically significant ($P<0.01$) correlation between the self-reports and reference measurements for four of these variables, namely, sitting (r_s 0.85), head bent forward (r_s 0.41), walking distance (r_s 0.59), and lifting 1—5 kg (r_s 0.63). The exposure variable sitting, which had the highest correlation, achieved acceptable accuracy when the six-point time scale was pooled into three points (kappa 0.52). Head bent forward achieved acceptable accuracy when the time was dichotomized (kappa 0.41). Walking distance showed poor agreement when the scale was dichotomized between <1 km and ≥1 km (kappa -0.02) but was acceptable when it was dichotomized between ≤2 km and >2 km (kappa 0.41). This finding indicates a nonlinear relationship. There was a correlation (r_s 0.63) between the self-reports and the reference measurements for the exposure variable lifting 1—5 kg, but the frequencies were systematically underestimated. The accuracy was poor wherever the scale was dichotomized (kappa 0.22—0.32). The exposure variable trunk bent forward 20—60 degrees showed no significant correlation between the self-reports and the reference measurements (r_s 0.10), and accuracy was poor wherever the scale was dichotomized (kappa -0.10—0.16).

Influence of musculoskeletal complaints

The prevalences of musculoskeletal complaints were 49% for high and low back regions, 55% for the shoulders and neck, and 30% for the knees. The prevalences of exposure, referring to the reference measurements, were also high (mean 40%, range 10—70%) and rather equally distributed among the subjects with and without complaints. For the nine exposure variables with a kappa greater than 0.40, the prevalence odds ratios derived from the questionnaire and the reference measurements are compared in table 4. The ratio between POR_{quest} and POR_{ref} differed from unity in five of the eighteen comparisons.

Table 3. Relation between the self-reports and reference measurements for the five group II variables according to the rank-order correlation coefficient, Spearman (r_s), the kappa coefficients (kappa), and the percentage of full agreement (full %) for the full graded and pooled number of scale points. The results include all of the subjects participating in the study (N = 97).

Variables	Scale points	r_s	Kappa	Full %
Sitting (time)	6	0.85**	0.35	48
	3 (0—37%; 38—62%; 63—100%)		0.52	68
	2 (0—37%; 38—100%)		0.77	89
Trunk bent forward 20—60° (time)	6	0.10	0.12	31
	2 (0—17%; 18—100%)		0.16	66
Head bent forward (time)	6	0.41**	0.06	26
	3 (0—37%; 38—62%; 63—100%)		0.31	57
	2 (0—37; 38—100%)		0.41	76
Lifting 1—5 kg (frequency)	4	0.63**	0.12	35
	2a) (<1 h^{-1}; ≥1 h^{-1})		0.32	
	2b) (≤30 h^{-1}; >30 h^{-1})		0.38	88
Distance walked (km)	5	0.59**	0.17	41
	2a) (<1 km; ≥1 km)		−0.02	37
	2b) (≤2 km; >2)		0.41	76

** P<0.01.

Table 4. Prevalence odds ratio (POR) for musculoskeletal complaints from the low back and from the shoulder and neck region when exposure was estimated by self-reports (POR_{quest}) and by reference measurements (POR_{ref}). The ratio between the two prevalence odds ratios (POR_{quest} : POR_{ref}) and the 95% confidence interval (95% CI) are also presented.

Variable	Complaints last 12 months: Low-back region				Shoulder and neck region			
	POR_{quest}	POR_{ref}	POR_{quest} : POR_{ref}	95% CI	POR_{quest}	POR_{ref}	POR_{quest} : POR_{ref}	95% CI
Sitting	0.5	0.5	1.0	0.6—1.5	1.3	1.0	1.3	0.9—1.9
Trunk bent forward <60°	1.8	2.6	0.7	0.5—1.0	1.6	1.6	1.0	0.7—1.5
Head bent forward	0.9	0.7	1.2	0.8—1.7	1.3	1.0	1.3	0.8—2.0
Kneeling or squatting	1.2	0.8	1.5	0.9—2.4	0.8	0.8	1.0	0.6—1.5
Walking distance	1.0	1.4	0.7	0.4—1.3	0.9	0.6	1.5	0.8—2.7
Carrying, pushing or pulling 6—15 kg	1.6	1.3	1.2	0.7—2.1	1.5	0.8	1.9	1.2—3.0
Carrying, pushing or pulling 16—45 kg	1.0	0.7	1.4	0.7—2.7	0.6	0.4	1.5	0.7—3.0
Lifting 6—15 kg	2.6	1.4	1.8	1.0—3.1	1.1	0.7	1.6	0.9—2.9
Lifting 16—45 kg	1.4	0.5	2.7	1.2—6.1	1.0	0.4	2.5	1.1—5.6

A higher POR_{quest} than POR_{ref} was found for low-back complaints in relation to lifting 6—15 kg and lifting 16—45 kg and for shoulder and neck complaints in relation to carrying, pushing or pulling 6—15 kg and lifting 16—45 kg. This finding indicates a differential bias for the self-reports. A lower POR_{quest} was found for low-back complaints in relation to trunk bent forward greater than 60 degrees. Trunk bent forward greater than 60 degrees was equally good or poor for subjects with low-back complaints (specificity and sensitivity 0.4 and 0.9, respectively) as for subjects without such symptoms (specificity and sensitivity 0.6 and 0.9, respectively). This finding indicated a nondifferential bias.

The prevalence odds ratio for knee complaints in relation to kneeling or squatting was similar whether it was calculated from the self-reports (POR_{quest} 0.8) or from the reference measurements (POR_{ref} 0.9).

Discussion

At a dichotomous level the agreement between self-reports and reference measurements was acceptable for 9 of 17 questions. Five questions showed poor agreement, and the remaining three were not analyzed because they concerned exposures that occurred for less than 10% of the study group. The ability to quantify the duration (6-point scale) or the frequency (4-point scale) in more detail was, however, poor for all five questions for which this ability could be tested. These results were achieved when the time interval under study varied from 26 min to almost 6 h. Musculoskeletal complaints introduced a differential bias to some questions about manual materials handling.

Comparison with other studies

In general, our results are in accordance with those of previous studies showing a weak correlation between self-reports and reference measurements. However, some important discrepancies can be summarized. Thus Baty et al (6) did not find any correlation between self-reported and observed time sitting although the difference between the self-reports and reference measurements was less than 5% in their study. The lack of correlation may partly be due to lack of contrast; only one occupational group, nurses, with a narrow range of values was studied. Kuorinka & Kilbom (9) found no relation between self-reported and observed occurrence of a kneeling position. They had separated the positions kneeling and squatting into two questions, and it may be difficult to differentiate between these two postures. Our study showed no correlation between the self-reported and observed occurrence of hands above shoulder level, while Kuorinka & Kilbom's study (9) did. In our study misclassification could have been due to an unclear phrasing or lay-out of the question, the reference measurements may have been invalid, or the underreporting could have been due to too short a recording period.

Classification of exposure

To quantify physical exposure, the following three dimensions are needed: the amplitude of forces (eg, newton, newtons per centimeter squared, newton meter), their duration (eg, seconds, hours, days, years), and their repetitiveness (eg, second^{-1}, hour^{-1}, day^{-1}, year^{-1}) (5, 18). In our study all of the questions except one, that concerning walking distance, were two dimensional, concerning amplitude and duration or amplitude and frequency (repetitiveness). The amplitude of the exposure was defined in the phrasing of each question, for example, lifting 5—15 kg (appendix 1, part A). The duration was quantified in proportions of the whole observation time and the frequency in times per hour (appendix 1, part B). Thus, before agreement could be achieved between the self-reports and the reference measurements, both the amplitude and the duration or frequency had to correspond with the results obtained by the reference measurements.

Sources of error in the self-reports

Failure to assess exposure by means of the questionnaire could have been due to the question, the report scale, the length of the recording period, or to the subjects themselves (eg, musculoskeletal disorders). The strength of agreement varied between the questions. The agreement was higher for questions concerning positions or manual materials handling representing a higher amplitude of load than a lower amplitude. It may be easier for subjects to remember if they have lifted burdens of more than 5 kg than it is for them to remember if they have lifted loads that weighed less; it may also be easier for them to perceive bending their trunk markedly forward (kappa 0.43) than bending it only moderately forward (kappa 0.16).

The report scale could also contribute to the strength of agreement. It is much easier to give a correct yes or no answer than to quantify the duration or frequency in more detail. A six-point time scale, a five-point distance scale, and a four-point frequency scale were used in this study. They all seemed to be too detailed for the present exposure variables. Only one variable, sitting, could acceptably be quantified at a three-point level (kappa 0.52).

The items in the time scale were described as proportions of the total observation time. Subjects may consider it more difficult to estimate proportions of time than to estimate minutes or hours. Moreover, the time scale used in this study may have been less appropriate for several of the questions. Presumably, exposure to extreme positions (eg, trunk bent forward >60 degrees, hands above shoulder level, or handling loads of >5 kg) does not occur for more than short periods during a workday in any population in Sweden. Before a report scale is designed, the distribution of the true physical load should be known. The length of the recording period could also affect the strength of agreement.

For most of the exposure variables the recording period was only about 30 min. A short recording period minimizes memory problems. On the other hand the recording period could have been too short. It may be difficult to discriminate between not being exposed at all and only a short length of exposure. A given posture might initially be overlooked and not recognized until after it starts to become strenuous. Therefore the agreement between an objectively measured and subjectively perceived duration of a posture or of manual materials handling may be nonlinear. Perception of exertion due to physical exposure is usually exponential (19). For the five variables for which the duration or frequency varied over all of the scales, a nonlinear relationship may partly explain why the demonstrated relationship between the self-reports and reference measurements differed according to where the cutoff was made on the scale. Thus all 13 questions concerning a short observation time should be evaluated during whole days in groups where the duration of the exposure is longer and frequencies higher.

We conclude that it is important to know the distribution of the true physical work loads and the psychophysiological relationships before questions are designed to quantify these loads. The validation studies should also be performed during whole workdays for an appropriate quantification of physical load in terms of duration and frequency.

Sources of error in the reference measurements

A trunk flexion analyzer, a posimeter, and a pedometer were used to validate the questions concerning the trunk being bent forward 20—60 degrees, trunk being bent forward more than 60 degrees, sitting, and walking distance. The trunk flexion analyzer has been used elsewhere and was considered valid. The accuracy of the posimeters and pedometers was tested and considered acceptable for the present purpose. The average difference between the posimeter recordings and the observed percentage of time sitting was 3 (range: 0—15)% (12). The median difference between the pedometer recording and the observed number of steps was 6 (range: 1—27)% (12). However, the pedometer counted the steps and not the distance walked. In some jobs (eg, floor cleaning work) the subject may be moving on the same spot. The subjective rating may then be more valid. Among the observations made, trunk rotation, head rotation, and head bent forward were difficult to estimate. This result could have been due to poorly defined reference lines forming the actual angles. Trunk rotation, for example, was measured from observations of the shoulders in reference to the hips. The reference line through the hips was difficult to observe, and an angle of 45 degrees between these reference lines could be obtained without motion in the trunk. The self-reports may then be more valid than the reference measurements. Moreover, the viewing angles were sometimes poor and therefore may have impaired the estimation of the angles (20). In addition, it was difficult to estimate force by observation. Whenever possible, these forces were also measured by a dynamometer. Estimated body angles and forces obtained by observation should be further validated.

Influence of musculoskeletal complaints

To determine whether subjects with musculoskeletal complaints overreport, underreport, or report more thoroughly than subjects without complaints, the specificity and sensitivity of self-reports relative to reference measurements should be analyzed for both groups. Such an analysis creates much data. We thought it would be more interesting to concentrate on biases that would change the size of the incidence rate ratios in an epidemiologic study. Therefore we compared the prevalence odds ratio calculated from the self-reports with the "true" prevalence odds ratio, calculated from the reference measurements.

Musculoskeletal complaints did not seem to cause any differential biases in the self-reported exposures to postures. However, shoulder or neck complaints seemed to cause a differential bias in two of four variables concerning manual materials handling. Subjects with shoulder or neck complaints had a tendency to overreport, and subjects without symptoms had a tendency to underreport, exposure in the self-reports. Low-back complaints seemed to cause a differential bias in two exposure variables concerning lifting. Subjects without low-back complaints had a tendency to underreport lifting activities. A lift has a very short duration and may therefore easily be overlooked, especially if the activity does not cause any harm to the subject. It should be emphasized that only modest complaints occurred among the subjects, all of whom were able to carry out their ordinary tasks. Moreover, the complaints concerned the last 12 months and not the day when the recordings were performed. Further investigations are necessary to clarify whether musculoskeletal complaints influence self-reported exposure.

Concluding remarks

The ability to discriminate between being exposed and unexposed to certain postures or manual materials handling occurring during an ordinary workday or part of a day seemed to be acceptable. The ability to quantify these exposures in more detail seemed to be poor. Musculoskeletal complaints seemed to cause a differential bias for some self-reported exposures but not for others. However, further investigations are necessary to clarify whether musculoskeletal complaints influence self-reported exposure. Thus self-reported exposure to certain postures and also to the handling of loads of more than 5 kg may, under certain conditions, offer sufficient accuracy to be used in epidemiologic studies in which the relative risks are high. Self-reported exposure may be too crude to be used in studies in which more detailed information is required, such as in epidemiologic case-referent studies of limited size with low relative risks, in ergonomic intervention studies, or in studies aimed at the construction of guidelines. A better designed questionnaire may give better and more-detailed exposure data. It should also be stressed that, when a questionnaire is used for both the exposure and the outcome (eg, complaints), a correlation of errors may arise that might cause severe bias (21). Further investigation in this field is encouraged.

Acknowledgments

This study was supported by grants from the Stockholm County Council, the Swedish Work Environment Fund, and the Folksam Insurance Company.

We are grateful to Professor L Alfredsson and Professor A Ahlbom for their valuable comments on the statistical analyses and to T Bellander, PhD, and M Köster, BS, who helped us with the analyses. We are also grateful to ergonomists Ms E Melin, Ms K Selin, and Ms A Willars for their highly qualified work in the data collection and to the whole exposure group for their generous cooperation.

References

1. Ahlborg GA Jr. Validity of exposure data obtained by questionnaire: two examples from occupational reproductive studies. Scand J Work Environ Health 1990;16:284—8.
2. Burdorf A. Exposure assessment of risk factors for disorders of the back in occupational epidemiology. Scand J Work Environ Health 1992;18:1—9.
3. Hagberg M. Exposure variables in ergonomic epidemiology. Am J Ind Med 1992;21:91—100.
4. Riihimäki H. Low-back pain, its origin and risk indicators. Scand J Work Environ Health 1991;17:81—90.
5. Winkel J, Westgaard R. Occupational and individual risk factors for shoulder-neck complaints: part II — the scientific basis (literature review) for the guide. Int J Ind Ergon 1992;10:85—104.
6. Baty D, Buckle PW, Stubbs DA. Posture recording by direct observation, questionnaire assessment and instrumentation: a comparison based on a recent field study. In: Corlett N, Wilson J, Manenica I, ed. The ergonomics of working postures: proceedings of the First

International Occupational Ergonomics Symposium; 1985 April 15—17. London and Philadelphia: Taylor & Francis 1986:283—91.
7. Burdorf A, Laan J. Comparison of methods for the assessment of postural load on the back. Scand J Work Environ Health 1991;17:425—9.
8. Hildebrandt VH, Bongers PM. Validity of selfreported musculoskeletal workload. Rev Epidémiol Santé Publique 1992:40 suppl 1:124.
9. Kuorinka I, Kilbom Å. Självrapportering av arbetsställningar och arbetsmetoder samt deras relation till muskuloskeletala besvär [Self-reported working postures and working methods and their relation to musculo-skeletal complaints]. Stockholm: National Institute for Occupational Health, 1990. (Undersökningsrapport 10.) (summary in English.)
10. Woodcock Webb K. Self-reports in ergonomics: agreement between workers and ergonomist. In: Aghazadeh F, ed. Trends in ergonomics/human factors V. Amsterdam: Elsevier Science Publishers 1988:387—94.
11. Wikman A. Att utveckla sociala in dikatorer, en surveyansats belyst med exemplet arbetsmiljö [Developing social indicators — an effort with a survey method illustrated with the example of working environment] [dissertation]. Stockholm: Stockholm University, Stockholm Statistics, 1991. (Urval 21.)
12. Selin K, Winkel J, Stockholm MUSIC I Study Group. Evaluation of two instruments recording sitting/standing and number of foot steps. Appl Ergon. (In press.)
13. Nordin M, Örtengren R, Andersson GBJ. Measurement of trunk movements during work. Spine 1984;9:465—9.
14. Fransson C, Gloria R, Kilbom Å, Karlqvist L, Nygård C-H, Wiktorin C, et al. Presentation and evaluation of a portable ergonomic observation method. In: Quéinnec Y Daniellou F, ed. Designing for everyone: proceedings of the XI Congress of the International Ergonomics Association; vol 1. London: Taylor & Francis, 1991:242—44.
15. Fleiss JL. The measurement of interrater agreement: statistical methods for rates and proportions. 2nd ed. New York, NY: John Wiley & Sons, Inc, 1981:212—36.
16. Rothman KJ. Modern epidemiology. Boston/Toronto: Little, Brown and Company, 1986.
17. Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sörensen F, Andersson G, et al. Standardised Nordic questionnaires for analysis of musculo-skeletal symptoms. Appl Ergon 1987;18:233—7.
18. Mathiassen SE, Winkel J. Quantifying variation in physical load using exposure-vs-time data. Ergonomics 1991;34:1455—68.
19. Borg G. Psychophysical scaling with applications in physical work and the perception of exertion. Scand J Work Environ Health 1990;16 suppl 1:55—8.
20. Douwes M, Dul J. Validity and reliability of estimating body angles by direct and indirect observations. In: Quéinnec Y Daniellou F, ed. Designing for everyone: proceedings of the XI Congress of the International Ergonomics Association. London: Taylor & Francis, 1991;2:885—7.
21. Kristensen P. Bias from nondifferential but dependent misclassification of exposure and outcome. Epidemiology 1992;3:210—5.

Appendix 1

A. Condensed version of the seventeen questions in the questionnaire

	Variable	*Scale*
1	Sitting?	6 points
2	Trunk bent forward 20—60°?	6 points
3	Trunk bent forward >60°?	6 points
4	Trunk rotation 45°?	6 points
5	Hands above shoulder level?	6 points
6	Head bent forward?	6 points
7	Head rotation?	6 points
8	Kneeling or squatting?	6 points
9	Carrying, pushing or pulling using force corresponding to 1—5 kg?	6 points
10	Carrying, pushing or pulling using force corresponding to 6—15 kg?	6 points
11	Carrying, pushing or pulling using force corresponding to 16—45 kg?	6 points
12	Carrying, pushing or pulling using force corresponding to >45 kg?	6 points
13	Lifting weight 1—5 kg?	4 points
14	Lifting weight 6—15 kg?	4 points
15	Lifting weight 16—45 kg?	4 points
16	Lifting weight >45 kg?	4 points
17	Walking distance?	5 points

B. The three report scales used for the questions in A

Scale	*Scale items*					
6 points	Not at all [ ]	About 1/10 of the time [ ]	About 1/4 of the time [ ]	Half the time [ ]	About 3/4 of the time [ ]	Almost all the time [ ]
4 points	<1 h^{-1} [ ]	1—10 h^{-1} [ ]	11—30 h^{-1} [ ]	<30 h^{-1} [ ]		
5 points	<1 km [ ]	1—2 km [ ]	3—5 km [ ]	6—10 km [ ]	>10 km [ ]	

C. Classification of duration obtained by the reference methods in accordance to the six-point time scale used in the questionnaire

Item number	1	2	3	4	5	6
Questionnaire	"Not at all"	"A little. About 1/10 of the time"	"About 1/4 of the time"	"Half the time"	"About 3/4 of the time"	"Almost all the time"
Reference methods	0%	1—17%	18—37%	38—62%	63—87%	88—100%

Received for publication: 2 September 1992