Scand J Work Environ Health 2001;27(2):125—132

Group-based measurement strategies in exposure assessment explored by bootstrapping

by Marco JM Hoozemans, MSc,[1] Alex Burdorf, PhD,[2] Allard J van der Beek, PhD,[1] Monique HW Frings-Dresen, PhD,[1] Svend Erik Mathiassen, PhD[3,4]

Hoozemans MJM, Burdorf A, van der Beek AJ, Frings-Dresen MHW, Mathiassen SE. Group-based measurement strategies in exposure assessment explored by bootstrapping. *Scand J Work Environ Health* 2001;27(2):125—132.

Objectives The precision of mean exposure to pushing was examined in 2 occupational groups using various combinations of the number of workers and measurements per worker.

Methods The frequency and duration of pushing of the 2 occupational groups was assessed using onsite observation. All data were divided into successive periods of 30 minutes of observation. The precision of the group mean exposure to pushing was expressed by 90% confidence intervals obtained by bootstrapping. The effect on the confidence interval of varying numbers of workers and numbers of periods per worker was examined.

Results For both occupational groups there was little precision to be gained when >10 workers were observed. Within the maximum number of workers used in the bootstrap simulations, it appeared that, beyond 10 workers, the confidence intervals decreased by <5% for every worker that was added, when each worker was observed at least 8 periods of 30 minutes. If workers were observed exactly 4 periods of 30 minutes per worker, an additional 4 workers were required to compensate for the loss of precision. An unbalanced strategy with approximately 8 periods of 30 minutes per worker hardly decreased the precision of the group mean, however.

Conclusions The precision of the group-based mean exposure to pushing is influenced by the number of workers observed and by the number of repeated measurements per worker. In the planning of measurement strategies, it is advisable to account for possible sources of variance in advance and to assess the exposure variability.

Key terms accuracy, epidemiology, manual materials handling, observation, occupation, precision.

The assessment of exposure to risk factors is one of the essential ingredients of the study of the etiology of musculoskeletal complaints. Risk factors for musculoskeletal complaints have often been divided into physical, psychological, and individual factors (1). It is generally thought that physical loading induces stress to the musculoskeletal system that may result in the degeneration of structures and the development of pain (2). Hence considerable attention has been given to the physical loading of the musculoskeletal system and, especially, the manual handling of materials in relation to low-back pain. These associations are not always clear, often because of inadequate exposure measurement (3, 4). Exposure measures are often crudely defined, using job title or a limited number of ordinal levels. This crudeness influences the accuracy and precision of ex-

1 Coronel Institute for Occupational and Environmental Health, A*m*COGG Amsterdam Center for Research into Health and Health Care, Academic Medical Center, University of Amsterdam, The Netherlands.
2 Department of Public Health, Erasmus University Rotterdam, Rotterdam, The Netherlands.
3 Division of Production Ergonomics, Faculty of Technology and Society, Malmö University, Malmö, Sweden.
4 National Institute for Working Life, Stockholm, Sweden.

Correspondence to: Marco JM Hoozemans, Coronel Institute for Occupational and Environmental Health, Academic Medical Center, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands.
[E-mail: m.j.hoozemans@amc.uva.nl]

posure and, as a result, attenuates the association with musculoskeletal complaints (5). To quantify exposure-response relationships, the exposure should be measured at a sufficient level of detail. Several methods are available for assessing mechanical exposure (eg, onsite observation, direct measurements at work, and simulations in the laboratory) (6). Moreover, Winkel & Mathiassen (5) state that mechanical exposure should be assessed by its principle dimensions, namely, duration, repetitiveness (frequency), and amplitude (intensity).

Besides the choice of measurement technique and parameters, the selection of workers and the variation of the exposure measure between and within workers are important aspects of measurement strategies (7). According to Burdorf & Van Riel (7), an efficient measurement strategy aims at reducing the number of measurements required. At the same time, the number of measurements should be sufficient to achieve the required level of precision of the exposure measure. Hence, the number of workers to be measured should be considered, as well as the number of repeated measurements per worker and the duration of each measurement. Another fundamental aspect of the measurement strategy is whether the measurement strategy should be individual-based or group-based. The data presented in this paper were gathered in a comprehensive epidemiologic study aimed at pushing and pulling in relation to musculoskeletal complaints at the group level. The mean level of exposure has been assessed for subgroups, and all the individuals within a subgroup were assumed to have the same exposure level. Generally, group-based strategies generate less precise but essentially unbiased estimates of the risk estimate when compared with individual-based strategies (8). Therefore, within a group-based epidemiologic study an accurate (zero biased) and precise (small random error) estimate of exposure is essential for the biological and statistical significance of the exposure-response relationship. The question arises as to how much measurement effort is needed to arrive at an estimate of the group mean of an exposure measure that is relatively unbiased with respect to the true group mean. Therefore, the objective of the present study was to examine the precision of the group-based mean exposure to pushing using various combinations of the number of workers within the group and the number of repeated measurements per worker.

Subjects and methods

Participants

Observational data of 2 occupational groups were used in the present study. Their characteristics are described in table 1. The first group consisted of 15 train stewards out of a total of 97 train stewards of a rail company. The train stewards daily pushed a 135-kg cart to provide train passengers with food and drinks. Beforehand, after a walkthrough and interviews, shift (early or late) and city of departure were identified as possible sources of variance with respect to the exposure to pushing and pulling. The participants and their shifts were randomly chosen from the group of train stewards that worked at the city of departure. The number of participants observed at each city of departure depended on the relative number of train stewards on point-duty at the city in comparison with the total number of train stewards. All the train stewards received information concerning the study, and none of the approached participants refused to take part.

The second group consisted of 18 nurses out of a total of 136 from a nursing home. Their work consisted of, for instance, patient-handling activities, bed making, providing medical assistance, and feeding patients. Two levels of need for care in the ward (somatic or psychogeriatric care), shift, and work during the weekend were expected to be possible sources of variance. To obtain representative samples of both types of care, a predetermined number of observation days for each type of care was set depending on shift and workday. The participants were randomly chosen from the nurses that were scheduled to work at that time. The study was intensively encouraged among all the employees of the nursing home, which did not object to participation. All the participants signed an informed consent before the observations.

Exposure assessment

All the participants were continuously observed at their workplace for a full workday on a real-time basis using TRAC (task recording and analysis on computer) (9). The observations were continued during breaks and unexpected events. Preceding the observations, the 2 participating observers were trained to improve inter- and

Table 1. Age, height, and weight of the 33 workers categorized into two occupational groups.

Occupational group	Number of workers	Age (years)		Height (m)		Weight (kg)	
		Mean	SD	Mean	SD	Mean	SD
Train stewards	15	33.7	8.6	1.75	0.07	71.6	6.9
Nurses	18	36.7	13.5	1.64	0.09	70.4	20.6

intraobserver reliability. During a week of intense training, the percentage of agreement and the Cohen's kappa between and within the observers were assessed for all the variables. At the end of the training period, it was ensured that all the variables had a percentage of agreement of at least 80% and at the same time a Cohen's kappa of at least 0.50, which is acknowledged to be an acceptable standard of observer reliability (14). Each observer was assigned to one of the occupational groups. Tasks, activities, and materials handled were observed. The activities were divided into lifting, carrying, pushing, pulling, standing, sitting, walking, and kneeling. Only pushing was used for the analyses in this study. Since observations were recorded on a real-time basis, the absolute frequency and duration of pushing could be assessed.

Exposure characteristics of the occupational groups

The average total observation time for the train stewards was 8 hours 11 minutes (SD 103 minutes), and for the nurses it was 8 hours 3 minutes (SD 40 minutes). For the train stewards an average of 17% of the time was spent pushing [duration of pushing 4882 (SD 1321) seconds], while the nurses spent only 2% of their worktime pushing objects [duration of pushing 538 (SD 338) seconds]. The average daily frequency of pushing was 219 (SD 66) times for the train stewards and 57 (SD 32) times for the nurses. Generally, for both occupational groups, pushing occurred more frequently than other manual materials handling activities, including pulling.

Data analysis

The observational data of each participant was divided into successive periods of 30 minutes. This way a workday of 8 hours was divided into 16 periods, which were considered repeated measurements within workers. For the 15 train stewards a total of 235 observation periods of 30 minutes were gathered, while for the 18 nurses 272 periods were available.

The variation of exposure to pushing within the 2 occupational groups was described by the 5th and 95th percentile of the 30-minute averages of the 15 and 18 workers with respect to the frequency and duration of pushing. Thus all 30-minute periods of observation of a worker were averaged, and all 15 or 18 means were averaged again, which can be described as a mean of means approach. These results can be compared with the average and the 5th and 95th percentiles for all of the 235 and 272 observations of 30 minutes, independent of worker. This procedure gives some indication of the between- and within-worker variance, and also of the differences introduced by the method of averaging the exposure of the occupational group.

The influence of the number of workers and the number of periods on the precision of the group mean of the exposure measures was studied using a bootstrap method (10, 11). According to Briggs et al (10), the bootstrap method estimates the sampling distribution of the exposure measure through a large number of simulations, based on sampling with replacement from the original observational data. For instance, the distribution of the exposure measure among the unknown real population can be estimated by drawing 5 times with replacement of a value out of 15 values (representing the average exposure values of 15 persons, which is a random sample taken from the real population). The 5 values are averaged, and this procedure is repeated, for instance, 1000 times. The 1000 average values represent an estimate of the distribution in the population with a mean of the 1000 average values that is more or less equal to the average of the 15 values. This distribution can be compared with a distribution when values are drawn 6 times with replacement. The latter, of course, will have a somewhat smaller distribution. The bootstrap method does not rely on parametric assumptions concerning the underlying distribution. To determine the precision of the group-based mean exposure to pushing when the assessment of the exposure is varied among the number of workers within the group and the number of repeated measurements per worker, a nested bootstrapping procedure was performed. First, a predetermined number of workers was drawn with replacement, and, second, for each of the selected workers a predetermined number of periods was drawn with replacement. Each simulation consisted of 1000 replications of this whole procedure. The predetermined number of workers was increased from 1 to the maximum number of workers of the occupational group (ie, 15 or 18 for the train stewards and the nurses, respectively). The predetermined number of periods was 2, 4, 8, and 16 periods per worker. Of the 1000 replications the mean and a measure of precision could be deduced. The precision of the empirical estimate of the sampling distribution of the exposure measures was defined using 90% confidence intervals. The confidence intervals were calculated using the bias-corrected percentile method as described by Efron & Tibshirani (11) with adjustment for possible bias in the bootstrap estimate of the mean compared with the mean of the observations. For each bootstrap distribution the confidence interval was calculated using the range between the bias-corrected 5th and 95th percentiles.

Within the practical setting of observing workers at their workplace it is often difficult to arrive exactly at a predetermined number of measurements per worker. To explore the effect of obtaining different numbers of 30-minute periods per worker, bootstraps were performed allowing the number of 30-minute periods to vary.

A hypothetical distribution of the obtained periods was made in which 4 was the most common and 2, 3, 5, and 6 were obtained in fewer amounts (figure 1). Again, a predetermined number of workers was first drawn with replacement. Then, for each of these workers, the number of periods per worker was drawn with replacement from the hypothetical distribution. Compared with the approach using an exact number of periods per worker, the procedure using the hypothetical distribution of periods will be referred to as an unbalanced procedure because, for a predetermined number of workers, the number of periods per worker is unequal. Because they served as an example, these bootstraps were only performed on the frequency of pushing for the train stewards and for unbalanced procedures with approximately 4 and 8 periods per worker (figure 1). The bootstrap procedures were performed using MATLAB (The MathWorks, Inc).

Results

The variation in exposure to pushing during 30 minutes as a result of averaging over the workers' means of the 30-minute periods (mean of means approach) was compared with the variation as a result of averaging over the total number of 30-minute periods (table 2). For both occupational groups, as well as for both the frequency and duration of pushing, standard deviations and the range between the 5th and 95th percentile were smaller for averaging over workers. This finding indicates that a reasonable part of the variance can be explained by the within-worker variance (ie, the worker means were not that different from each other, while the 30-minute periods within each worker differed because exposure to pushing was not equally divided over the workday). Furthermore, for the nurses, the distribution of the total number of 30-minute periods of the frequency, as well as the duration, of pushing appeared to be skewed to the right. This finding indicates that the high values of these exposure measures were incidental over the workday for the nurses.

Figure 2 presents the results of the bootstraps that were performed by randomly taking 2, 4, 8, or 16 periods of 30 minutes per worker. For each of these numbers of periods the number of workers was increased. Generally, the gain in precision from increasing the number of workers was considerable at the low initial numbers. Thereafter the precision dropped off rapidly as the number of workers increased (figure 2). Within the maximum number of workers that was observed in the present study, it appeared that, beyond 10 workers, the inclusion of an additional worker improved the precision by <5%. On the basis of trivial statistical considerations (12), it was expected that, beyond 10 workers, an additional 30 workers would be needed to increase the precision of the estimate of the group mean by 50%. Furthermore, observing 2 random periods of 30 min-

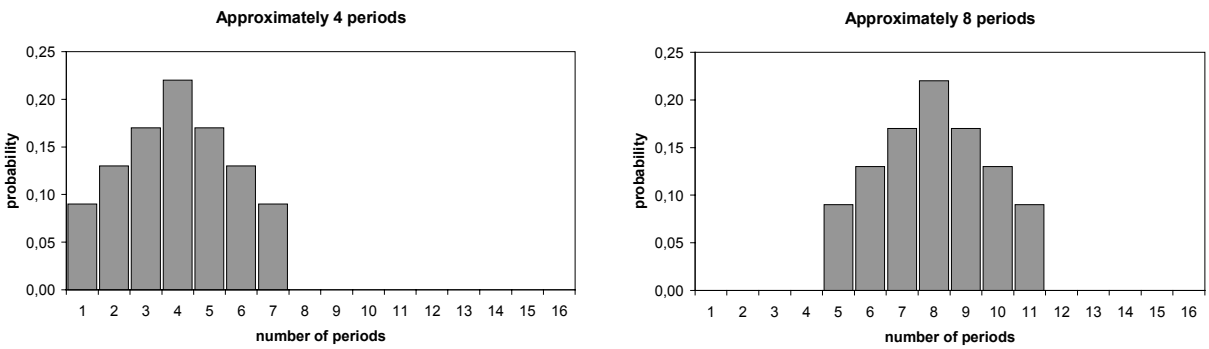

Figure 1. Hypothetical distributions to explore the effect of obtaining different numbers of 30-minute periods per worker. The distributions were used to obtain approximately 4 and 8 periods of 30 minutes per worker.

Table 2. Frequency and duration of pushing for the train stewards and nurses.

	Pushing frequency				Pushing duration			
Occupational group	Mean	SD	5th %ile	95th %ile	Mean	SD	5th %ile	95th %ile
Train stewards								
Exposure of individual workers (N = 15)	13.4	3.2	8.8	17.6	297.4	63.5	197.1	379.2
Total sample of 30-minute periods (N = 235)	13.4	8.8	0.0	29.0	295.6	176.7	0.0	578.0
Nurses								
Exposure of individual workers (N=18)	3.6	2.0	0.4	6.3	34.2	20.6	3.3	65.6
Total sample of 30-minute periods (N=272)	3.6	4.2	0.0	11.0	34.2	40.3	0.0	108.1

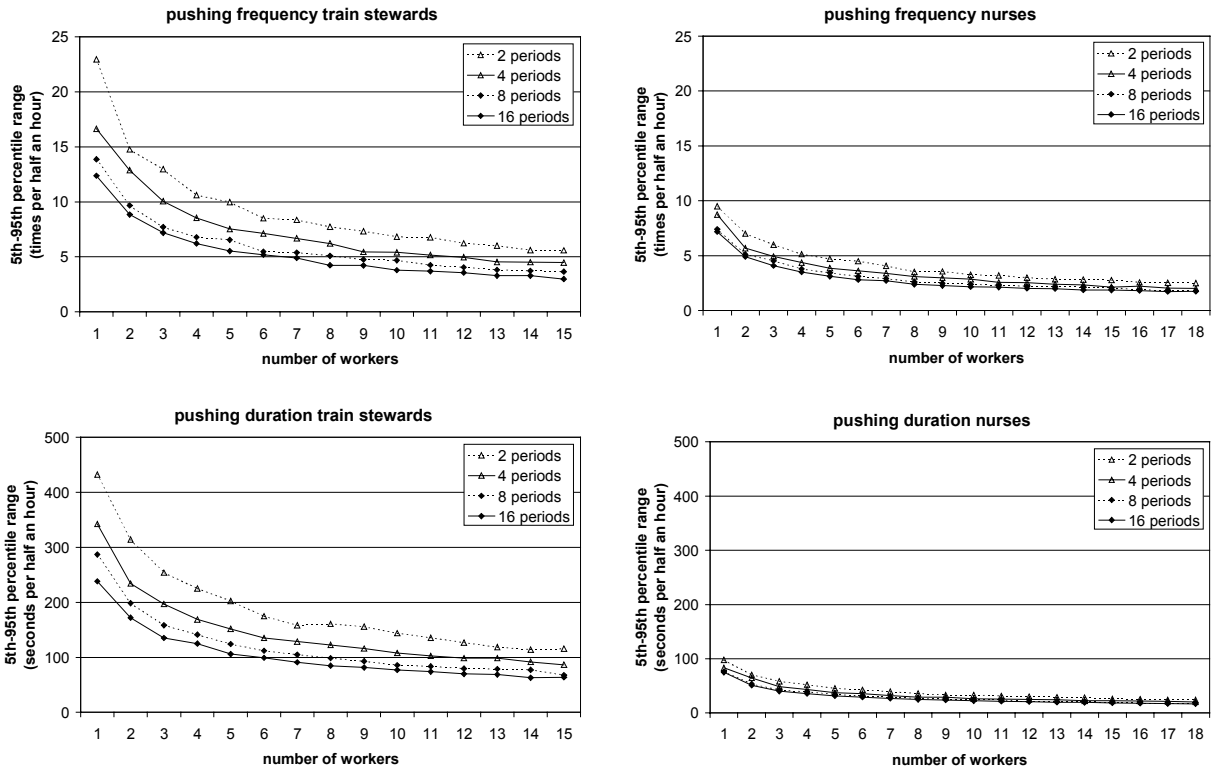

Figure 2. Results of the bootstraps performed by randomly taking 2, 4, 8 or 16 periods of 30 minutes per worker with replacement for 1000 replications. For both occupational groups and for both the frequency and duration of pushing the 5th—95th percentile range is presented.

Table 3. Number of workers needed to arrive at a specified value (10, 7.5, and 5) of the 5th—95th percentile range if different numbers of periods (2, 4, 8, and 16) are sampled per worker. The results are presented for the frequency of pushing for the train stewards.

	Necessary number of workers		
Number of 30-minute periods	10.0 of 95th—5th %ile	7.5 of 95th—5th %ile	5.0 of 95th—5th %ile
2 periods	5	9	[a]
4 periods	3	5	12
8 periods	2	3	8
16 periods	2	3	7

[a] This level of precision was not reached within the number of workers used in the simulations (N=15)

utes yielded the largest 5th—95th percentile interval. Using 4 periods resulted in more precise estimates of the group mean, and also the observation of 8 periods per worker resulted in improvement in the precision. In comparison with the random observation of 8 periods per worker, there was little to be gained by observing 16 periods of 30 minutes (full workday).

From the point of view of reaching a certain level of precision, table 3 presents the number of workers that had to be observed with respect to the frequency of pushing by train stewards. The table is based on the corresponding figure in figure 2. For instance, to reach a 5th—95th percentile range of 10, 5 workers had to be observed 2 random periods of 30 minutes per worker. To reach the same level of precision, 2 workers had to be observed if these workers were randomly observed for 16 periods of 30 minutes. A more precise estimate with a 5th—95th percentile range of 5 could be reached if 8 workers were observed with randomly 8 periods per worker. Observing 2 random periods per worker would not reach this level of precision within the maximum number of workers available in the present bootstrap simulations (N=15).

The results of the pushing frequency of the train stewards were further explored considering that, in practice, not every worker in the population could be observed exactly 4 or 8 periods. The results obtained by using an unbalanced procedure with approximately 8

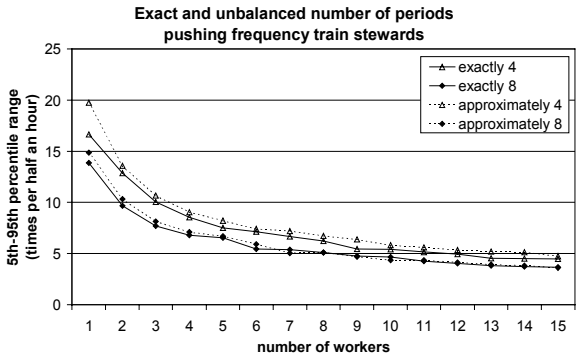

Figure 3. Results of the bootstraps performed by randomly taking an exact or a variable number of periods of 30 minutes per worker with replacement for 1000 replications. The 5th—95th percentile range is presented for bootstraps on the frequency of pushing among train stewards. A comparison is made between drawing exactly 4 or 8 periods per worker and drawing following the unbalanced procedure as described in figure 1.

periods per worker hardly differed from those using a strategy with exactly 8 periods per worker (figure 3). Only when the number of workers was small (fewer than 6), was the 5th—95th percentile range slightly wider for the unbalanced strategy. When the same comparison was made for 4 periods per worker, the unbalanced strategy showed wider percentile ranges for all numbers of workers.

Discussion

The influence of the number of observed workers and the number of repeated observations per worker on the precision of the average group exposure to pushing was studied. Two occupational groups were examined, which were, on the average, exposed differently to pushing. For both occupations, when more than a random sample of 10 workers was observed, the precision of the estimate of the group mean did not dramatically increase for every worker added to the occupational group. The same held true when each of the workers was observed for >8 random periods of 30 minutes. Although the occupations showed different exposures, there were no remarkable differences in the results of the bootstrap procedures. For other manual materials handling activities, such as pulling, lifting, and carrying, which were not presented in the present paper, the same patterns as those for pushing were found in both occupational groups.

The results of our study were somewhat different from those reported by Burdorf & Van Riel (7). They concluded that, for the duration of worktime with trunk flexion over 20 degrees, between 15 and 25 workers must be observed to estimate the group mean exposure. At least 3 factors may explain the differences. First, each worker in the study of Burdorf & Van Riel (7) was observed for 4 periods of 30 minutes on 2 separate days compared with our observations of 1 full workday. Figure 2 and table 3 show that fewer workers had to be observed for the same level of precision when the number of 30-minute periods was increased from 4 to 8 per worker. Second, Burdorf & Van Riel observed the relative time spent with the trunk flexed using 1 observation every 20 seconds. In our study, real-time (continuous) observation was used to calculate the absolute time and frequency of pushing. The within- and between-worker variance of the relative measure trunk flexion, in comparison with the absolute values of activities such as pushing, may be different and may, therefore, have a direct influence on the relation between the number of workers and repeated measurements and the precision of the estimate of the group mean exposure (13). Third, with a low number of observed workers the precision increased relatively fast for every worker that was added. Beyond a certain point little precision is gained for every worker added. The interpretation of this point may be prone to subjectivity depending on the shape of the figure. Furthermore, for instance, statistical or epidemiologic considerations influence the decision of whether the precision of the estimate of the group mean can be assumed to be sufficient.

Both the measurement technique (onsite observation) and the bootstrap method could have had a direct effect on the results of our study. It is expected that, when onsite observation is used, frequency and duration are assessed with a reasonable level of accuracy (14, 15). The bootstrap method is an empirical method for estimating the population's mean exposure. Under the assumption of normality, confidence intervals of the group means may also be obtained by analytical methods based on between- and within-worker variance components (12). In large samples, the mean values are normally distributed irrespective of the underlying distribution. Because of the relatively small sample size of our observational data the normality assumption can be questioned. The large standard deviations relative to the group mean, especially with respect to the nurses, are the result of large differences in pushing frequency and duration between workers and within workers (days). Hence, the use of the bootstrap method is justified. Moreover, the bootstrap procedure is particularly appropriate for exploring the unbalanced measurement strategy, the precision of which cannot be estimated by any available analytical method. Another consideration to be taken into account when the bootstrap method is applied is the number of replications. The results of our study may have been biased when the number of replications was too low, but the 1000 replications applied in our study can be considered sufficient (11).

The bootstrap procedure considers the observations as a true representation of the population. It is very important that the sample of observational data used for bootstrapping be representative of the larger population if the results are to be used to establish an efficient measurement strategy. In this study the participants were selected using a stratified sampling procedure in order to take into account a priori the most important sources of variation in exposure. Within each of the strata, for instance, late shift, workers were selected randomly. In addition, the number of workers within each of the sources of variance was in proportion to the number of workers that worked within these sources in relation to the total population. Although the variance caused by these sources of variance was not studied further, accounting for possible sources of variance should be an important part of the measurement strategy, especially when the number of observations is kept to the minimum. Our considerations as to measurement efforts assumed that the observational data were correct and invariable. Acceptance of a stochastic variability of the data would increase the estimates of the measurement efforts. Such acceptance is acknowledged in standard procedures for power assessment.

An interesting question is whether many workers should be observed for a short period of time or a few workers for a longer period of time to arrive at a relatively precise estimate of the group mean. The first hypothetical answer is shown in table 3. To reach a 5th—95th percentile range of 5 for the pushing frequency of train stewards, 12 workers must be observed when each worker is observed for 4 random periods, which account for a total of 48 periods. The same level of precision is reached when 8 workers are observed for 8 periods per worker. This situation accounts for a total of 64 periods. As an example, table 3 can be rewritten into table 4. The general message is that it is favorable to observe more workers for a short period of time (a small number of repeated measurements). This procedure would take a smaller total number of periods and could reduce costs and measurement effort. However, there are some practical considerations as to observing more workers for shorter periods of time. For practical reasons it is convenient to switch between workers on the day of observation. The occupational setting determines whether or not such a switch can be made. For our study, for instance, it was impossible to switch between train stewards because only 1 train steward worked at a time on a specific train. Observing more than 1 nurse per day would have been feasible, because they all worked in the same building or ward.

Another practical consideration was already shown in figure 3. Due to various reasons it is not always possible to get exactly 4 or 8 periods per worker. Since the precision of the group mean can be reduced by unbalanced sampling, the total number of periods needed to arrive at a certain precision would be increased in this case. Table 4 demonstrates this effect in quantitative terms. While an unbalanced procedure with approximately 8 periods per worker results in an equal total number of periods in comparison with exactly 8 periods per worker, an unbalanced procedure with approximately 4 periods increases the total number of periods compared with exactly 4 periods. In addition, an increase in the desired precision results in an increase in the difference between exactly 4 periods and the unbalanced strategy. The practical unbalanced concept was only applied to 4 and 8 periods per worker. Applying the same methods to 2 and 16 periods per worker would be predictable. The precision of an unbalanced procedure with approximately 16 periods will decrease in comparison with the precision for exactly 16 periods because 16 periods would be an upper limit for the number of periods obtained per worker. Thus, with this strategy, the average number of periods per worker is smaller than 16. An unbalanced strategy at 2 periods will increase the precision due to the opposite effect, namely, the lower limit is 2 periods per worker, and hence the average will be larger than 2.

In conclusion, the results of our study show that the statistical precision of the average group exposure to pushing is influenced by the number of workers observed, and at the same time by the number of repeated measurements per worker. For the 2 occupational groups under study, an efficient measurement strategy would be to observe 10 workers randomly for at least 8

Table 4. Number of 30–minute periods that must be observed to arrive at a specified value (10, 7.5, and 5) of the 5th—95th percentile range for the frequency of pushing for the train stewards. Results are presented for an exact number of periods per worker, as well as for an unbalanced number of periods per worker.

	Necessary number of 30–minute periods					
	10.0 of 95th—5th %ile		7.5 of 95th—5th %ile		5.0 of 95th—5th %ile	
Number of 30–minute periods	Exact	Unbalanced	Exact	Unbalanced	Exact	Unbalanced
4 periods	12	12	20	24	48	60
8 periods	16	16	24	24	64	64

periods of 30 minutes per worker, because there was little to be gained in the precision of the estimate of the average group exposure when more workers or more observations per worker were added. If workers are observed for 4 random periods per worker, at least 12 workers must be observed, and for 2 random periods per worker about 20 workers have to be observed, to arrive at the same level of precision. This measurement strategy could not be considered as a general rule. Between- and within-worker variance will affect the number of workers and repeated measurements needed to arrive at a certain relative precision of the estimated group mean exposure. Although it is often recommended to conduct a pilot study in order to get the exposure variability data needed to assess necessary measurement efforts in the main study (16), this may not explicitly reduce costs and measurement effort. Therefore, it is advised to account for possible sources of variance in advance and to assess the exposure variability during the actual measurements.

Acknowledgments

The authors would like to thank Tineke Piersma, Nienke van der Meulen, and Jitske Groothausen for their support during the data acquisition.

References

1. Armstrong TJ, Buckle PW, Fine LJ, Hagberg M, Jonsson B, Kilbom Å, et al. A conceptual model for work-related neck and upper limb musculoskeletal disorders. Scand J Work Environ Health 1993;19:73—84.
2. Van Dieën JH, Toussaint HM. Evaluation of the probability of spinal damage caused by sustained cyclic compression loading. Hum Factors 1997;39:469—80.
3. Burdorf A, Sorock GS. Positive and negative evidence of risk factors for back disorders. Scand J Work Environ Health 1997;23:243—56.
4. Kuiper JI, Burdorf A, Verbeek JHAM, Frings-Dresen MHW, Van der Beek AJ, Viikari-Juntura E. Epidemiological evidence on manual materials handling as a risk factor for back disorders: a systematic review. Int J Ind Ergon 1999;24:389—404.
5. Winkel J, Mathiassen SE. Assessment of physical work load in epidemiologic studies: concepts, issues and operational considerations. Ergonomics 1994;37:979—88.
6. Van der Beek AJ, Frings—Dresen MHW. Assessment of mechanical exposure in ergonomic epidemiology. Occup Environ Med 1998;55:291—9.
7. Burdorf A, Van Riel M. Design of strategies to assess lumbar posture during work. Int J Ind Ergon 1996;18:239—49.
8. Seixas NS, Sheppard L. Maximizing accuracy and precision using individual and grouped exposure assessments. Scand J Work Environ Health 1996;22:94—101.
9. Frings-Dresen MHW, Kuijer PPFM. The TRAC-system: an observation method for analyzing work demands at the workplace. Saf Sci 1995;21:163—5.
10. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ 1997;6:327—40.
11. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1986;1:54—77.
12. Samuels SJ, Lemasters GK, Carson A. Statistical methods for describing occupational exposure measurements. Am Ind Hyg Assoc J 1985;46:427—33.
13. Tielemans E, Kupper LL, Kromhout H, Heederik D, Houba R. Individual-based and group-based occupational exposure assessment: some equations to evaluate different strategies. Ann Occup Hyg 1998;42:115—9.
14. Van der Beek AJ, Van Gaalen LC, Frings-Dresen MHW. Working postures and activities of lorry drivers: a reliability study of on-site observation and recording on a pocket computer. Appl Ergon 1992;23:331—6.
15. Kilbom Å. Assessment of physical exposure in relation to work-related musculoskeletal disorders — what information can be obtained from systematic observation. Scand J Work Environ Health 1994;20 special issue:30—45.
16. Burdorf A. Reducing random measurement error in assessing postural load on the back in epidemiologic surveys. Scand J Work Environ Health 1995;21:15—23.

Received for publication: 14 July 2000