Original article

Scand J Work Environ Health. 2014;40(3):252–265. doi:10.5271/sjweh.3416

Cost-efficient assessment of biomechanical exposure in occupational groups, exemplified by posture observation and inclinometry

Catherine Trask, PhD,[1, 2] Svend Erik Mathiassen, PhD,[1] Jens Wahlström, PhD,[1, 3] Mikael Forsman, PhD [1, 4]

Trask C, Mathiassen SE, Walhström J, Forsman M. Cost-efficient assessment of biomechanical exposure in occupational groups, exemplified by posture observation and inclinometry. *Scand J Work Environ Health*. 2014;40(3):252–265. doi:10.5271/sjweh.3416

Objectives This study compared the cost efficiency of observation and inclinometer assessment of trunk and upper-arm inclination in a population of flight baggage handlers, as an illustration of a general procedure for addressing the trade-off between resource consumption and statistical performance in occupational epidemiology.

Methods Trunk and upper-arm inclination with respect to the line of gravity were assessed for three days on each of 27 airport baggage handlers using simultaneous inclinometer and video recordings. Labor and equipment costs associated with data collection and processing were tracked throughout. Statistical performance was computed from the variance components within and between workers and bias (with inclinometer assumed to produce "correct" inclination angles). The behavior of the trade-off between cost and efficiency with changed sample size, as well as with changed logistics for data collection and processing, was investigated using simulations.

Results At similar total costs, time spent at trunk and arm inclination angles >60 ° as well as 90th percentile arm inclination were estimated at higher precision using inclinometers, while median inclination and 90th percentile trunk inclination was determined more precisely using observation. This hierarchy remained when the study was reproduced in another population, while inclinometry was more cost-efficient than observation for all three posture variables in a scenario where data were already collected and only needed to be processed.

Conclusions When statistical performance was measured only in terms of precision, inclinometers were more cost-efficient than observation for two out of three posture metrics investigated. Since observations were biased, inclinometers consistently outperformed observation when both bias and precision were included in statistical performance. This general model for assessing cost efficiency may be used for designing exposure assessment strategies with considerations not only of statistical but also cost criteria. The empirical data provide a specific basis for planning assessments of working postures in occupational groups.

Key terms back; baggage handling; ergonomics; musculoskeletal disorder; shoulder.

A suitable exposure assessment is central to occupational health research, not the least as guided by the need to obtain data assuring that statistical power is sufficient to render the study informative. Power in studies comparing exposures between groups or conditions within groups is directly dependent on the precision of the mean exposure estimate, as exemplified by studies of biomechanical exposure (physical workload) (1, 2). This precision, in turn, depends on the sample size, and the variability in exposure caused by differences in behavior between subjects and within subjects across time, as well as by uncertainty associated with the exposure measurement method per se (3, 4). Exposure variability, often expressed in terms of exposure variance components, has been an issue in occupational epidemiology for more than two decades (5–7), with applications reaching far beyond the design of data collection strategies (8). In the field of biomechanical exposures, to which the present paper is specifically devoted, exposure variability was introduced in the mid-1990s (9–11), and

1 Centre for Musculoskeletal Research, Department of Occupational and Public Health Sciences, University of Gävle, SE – 801 76 Gävle, Sweden.
2 Centre for Health and Safety in Agriculture, College of Medicine, University of Saskatchewan, , 103 Hospital Drive, Saskatoon, Saskatchewan, S7N 0W8, Canada.
3 Department of Public Health & Clinical Medicine, Occupational and Environmental Medicine, Umeå University, SE-901 85 Umeå, Sweden.
4 Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

Correspondence to: Catherine Trask, Centre for Health and Safety in Agriculture, College of Medicine, University of Saskatchewan, 103 Hospital Drive, Saskatoon, Saskatchewan, Canada, S7N 0W8. [E-mail: catherine.trask@usask.ca]

a number of papers have discussed exposure assessment strategies, including necessary sample sizes to ensure a certain precision of a mean exposure estimate and/or a sufficient power in comparison studies (1, 3, 12–17). Although these articles have furthered inquiry into efficient exposure assessment from a statistical point of view, they have been criticized for rarely acknowledging the actual costs of exposure assessment. Thus, a 2010 systematic review of literature focusing on cost-efficient collection of exposure data (18) found only nine studies dealing with the trade-off between statistical performance and monetary resources invested in obtaining that performance, even if some studies have appeared after 2010 (19–21). Only some of the publications identified in the 2010 review dealt specifically with occupational or environmental exposures (7, 22–26); only two of these included empirical data to illustrate cost and efficiency (24, 26), and none were devoted to assessment of biomechanical exposure. In the context of study design, the trade-off between cost and statistical performance appears in the form of either one of two questions: (i) for a given research budget, which measurement method and sampling strategy delivers the highest statistical performance with respect to producing unbiased and precise data? and (ii) facing a required statistical performance, for instance in order to obtain sufficient power, which method and sampling strategy is the most cost efficient?

Biomechanical exposure assessment methods are often grouped into three broad categories or "classes" of measurement: direct measurement, observation on-site or from video, and worker self-report. Objective direct measurement methods are generally preferred for accuracy (27–31). Observation has long been believed to represent a middle ground in that they are more objective than self-reports (19, 27, 29), yet less consistent than direct measurements. The suspected inferior precision associated with using observational methods would mainly be a result of variability introduced by within- and between-observer differences in opinions when viewing the same posture (21, 32–34). At the same time, observations have been anecdotally claimed to be cheaper in use than direct measurements (27). This would mean that, for a certain budget, more data could be collected by observation, which might make up for the larger exposure variability, and eventually lead to a more precise estimate of group mean exposure than that obtained by direct measurements at the same total cost. In order to appreciate this trade-off and decide which measurement method to prefer from a cost-efficiency point of view, costs and efficiency must be quantified for each measurement class under realistic data collection scenarios and then compared (18). However, there have been few comprehensive efforts to report the empirical costs of biomechanical exposure assessment, and these preliminary efforts did not include any information on the quality or statistical properties of exposure information delivered (35–37). Recently some studies have combined empirical costs and precision of occupational exposure estimates, although these have addressed cost efficiency in general terms (20), or have focused on a narrow assessment of cost and compared sampling or analysis strategies just within a single method [in casu, observation; (19, 21)].

In this study, empirically collected cost and exposure data were used to assess the cost efficiency of two common classes of methods for assessing working postures: direct measurement by inclinometry using data loggers and video-based posture observation by trained analysts. The objectives were to: (i) quantify and compare the trade-off between costs and statistical performance (ie, cost efficiency) when using inclinometry and observation to retrieve mean exposures to trunk and arm inclination variables in groups; and (ii) identify possible changes to the relative cost efficiency of the two methods when different alternative study design scenarios are considered.

Methods

Study population and sampling

Twenty-seven randomly selected full- or part-time baggage handlers at a large Swedish airport had their trunk and upper-arm postures assessed during three work shifts each using three methods: self-report via questionnaire, observation from video film, and full-shift inclinometer registration using tri-axial accelerometers (only the observation and inclinometer data were included in this paper). Video recordings and inclinometer measurements were made in parallel, and therefore represent the same period of work for each specific subject. All subjects signed an informed consent form and the Regional Ethical Review Board in Uppsala, Sweden, approved the study. Measurements were successfully collected for 3 days from all but one worker who could not complete a third day due to injury, resulting in measurement files from 80 days in total. Data were successfully processed for all these files, except for one inclinometer file which had excessive noise. The data collection methods are described in detail in a separate report (35), as are the data processing methods (37).

Inclinometer data collection and processing

Trunk and upper-arm posture was assessed as inclination with respect to the line of gravity. An inclinometer with an integrated memory (VitaMove triaxial accelerometer system, 2M Engineering, Veldhoven, the Netherlands) was placed on each upper arm over the medial deltoid, and on

the trunk between the shoulder blades. Inclinometers were set up before the start of the shift, worn for the duration of the shift, and data were downloaded to a laptop computer at the end of the shift. Inclinometer data were calibrated to each subject's recorded "zero" positions. For the upper arms, the zero position was to bend forward and hold a 5 kg weight in the hand. The trunk's reference positions involved bending forward (to define "forward") and then standing upright (to define zero) as described in Teschke et al (38). The raw three-axial data from the inclinometers were calibrated and transformed into degrees by using the same software used by Wahlström et al (39) and Hansson et al (28). In this software, a 5 Hz low-pass filter is used to reject inclination errors due to rapid movements. This software has been validated when used together with a similar type of accelerometer (40, 41). Inclinometer data were also processed to account for extreme trunk and upper-arm angles when workers were resting supine. Data from each measurement day were summarized into daily inclination metrics for both the trunk and each arm: median as a measure of central tendency, 90th percentile as a measure of extreme values, and the percent time spent with flexion (trunk) or elevation (upper arm) >60° as a measure of the occurrence of clearly non-neutral postures. This is consistent with metrics commonly used in other inclinometer studies (39, 42, 43). Since mean exposure values and components of variance were very similar for the left and right arm, only the right arm is presented for the sake of simplicity.

Observation data collection and processing

Workers were video recorded using a single camera following them throughout the airport during their regular work tasks for the first or second half of their work shifts. Video recordings were analyzed by four trained observers using a customized software program similar to that described by Bao et al (44, 45). Software users operated a point-and-click dial to estimate each posture at a 1° resolution with respect to gravity in each still frame. Bao et al (45) has reported the between-rater standard deviation (SD) using a 1° resolution method to be 5.5° for the trunk and 10.8–12.6° for the upper arm. Still frames were selected at 55-second intervals, yielding up to 252 unique frames per the half-shift. Trunk and upper-arm inclination were summarized into the same daily exposure metrics as those selected for the inclinometer (ie, median, 90th percentile, and percent time with the upper arm elevated >60°).

Measurement costs

Comprehensive costs for data collection and processing were tracked throughout these phases and used to developed cost models; these methods have been reported in detail elsewhere (35, 37). In brief, the time for all research staff was recorded and summarized for all major research tasks for both measurement methods, from study planning to worksite measurements to data management. The comprehensive cost, C, combining data collection and processing costs for each method, can be assessed using the following model (fixed costs are denoted by Č and variable costs denoted by Ċ):

Equation 1

$$C = \check{C}_A + \check{C}_R + \check{C}_E + \check{C}_S + \dot{C}_T + \dot{C}_V + \dot{C}_H + \dot{C}_R + \dot{C}_D + \dot{C}_M$$

Where $\check{C}_A$ is the cost of project meetings and administration including documentation, budgeting, and internal correspondence; $\check{C}_R$ is cost related to recruitment, including corresponding with employer and scheduling; $\check{C}_E$ is the capital cost for data collection equipment; $\check{C}_S$ is the cost for developing customized data processing software. In terms of variable costs, $\dot{C}_T$ is the cost of training staff to collect and process data specific to the measurement method; $\dot{C}_V$ is the cost of traveling to the worksite for data collection (depending on the number of trips), $\dot{C}_H$ is the cost of hotel accommodations during overnight trips (depending on the number of nights); $\dot{C}_R$ is the cost of recruiting workers at the worksite (depending on unit cost of recruiting a worker, $\dot{c}_R$, and the number of workers recruited); $\dot{C}_D$ is the cost of onsite data acquisition (depending on the unit cost of a measurement day, $\dot{c}_D$, and the number of data collection days); and $\dot{C}_M$ is the cost of managing the processing of daily data files (depending on the unit cost of processing a data file, $\dot{c}_F$, and the number of data collection days). Equation 1 is an amalgamation of previously-published cost models for data collection (35) and data processing (37). All costs were converted from Swedish kronor to euros using average annual exchange rate for the calendar years data were collected or processed.

Estimating statistical performance

In this study, the statistical performance of a measurement class was measured in two ways: as the precision of the obtained group mean exposure estimate, including only random error sources, and as the root-mean-squared error (RMSE) of the estimate, including both bias and random error of the estimate. Exposure variance components were estimated according to the following model:

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$ *Equation 2*

where:

- x_{ij} is the exposure for worker i on day j;
- μ is the overall group mean value for a given exposure metric (ie, median, 90th percentile, % time >60°);
- α_i is the random effect of worker i (values of i from 1–27); and

- ε_{ij} is the residual error representing the random effect of measurement day j (values of j from 1–3) in worker i.

A similar model has been used to estimate posture variance components in previous studies (39, 46). Based on this model, variance components were estimated using Restricted Maximum Likelihood (REML) via the VARCOMP command in SPSS version 20.0 (IBM, Armonk, NY, USA) with "worker" as a random effect term. The random effect terms α_i and ε_{ij} are then presumed to have a mean of 0 and variances of (between-worker variance) and (within-worker variance), respectively. In the case of our observations, these variances include a non-extractable contribution from within- and between-observer rating variability.

Precision

The variance of the estimated mean, σ_{μ}^2 , for an exposure metric is:

$$\sigma_{\mu}^2 = \frac{\sigma_{bw}^2}{n_W} + \frac{\sigma_{ww}^2}{n_M} \qquad \textit{Equation 3}$$

where:

- σ_{bw}^2 is the between-worker variance;
- σ_{ww}^2 is the within-worker variance;
- n_W is the number of workers; and
- n_M is the total number of measurements.

For descriptive purposes, precision may be expressed in terms of the standard error of the mean (ie, $\sqrt{\sigma_{\mu}^2}$) and statistical performance
– only including random error – conveniently as the inverse of this standard error (ie, $1/\sqrt{\sigma_{\mu}^2}$ (19).

Precision and bias combined

Statistical performance was also quantified using the RMSE of the group mean exposure estimates. RMSE includes both random variance of the estimated mean of the exposure metric, σ_{μ}^2 , and the possible systematic bias inherent in the chosen method, *B*, such that:

$$\mathrm{RMSE}_{\mu} = \sqrt{\sigma_{\mu}^2 + B^2} \qquad \textit{Equation 4}$$

Substituting Equation 3 into Equation 4 expresses the combined effect of bias and (im)precision:

$$\mathrm{RMSE}_{\mu} = \sqrt{\left(\frac{\sigma_{bw}^2}{n_W} + \frac{\sigma_{ww}^2}{n_M}\right) + B^2} \qquad \textit{Equation 5}$$

To be consistent with previous literature describing the hierarchy of exposure assessment methods (27), inclinometry was designated as the method producing "correct" inclination data. Thus, by definition, the results obtained by inclinometry were taken to be unbiased (B=0), and the bias, B, of the results obtained by observation was calculated as the difference between mean exposures obtained by observation and the corresponding results according to inclinometry. The combined precision and bias performance of each posture assessment method was calculated as the inverse of the RMSE_{μ} (19, 21), which is thus, by definition, equal to $1/\sqrt{\sigma_{\mu}^2}$ for inclinometry in the present case.

Quantifying cost efficiency

The "price for performance" was expressed in euros per $1/\sqrt{\sigma_{\mu}^2}$ and $1/\mathrm{RMSE}_{\mu}$ (ie, $C \times \sqrt{\sigma_{\mu}^2}$ and $C \times \mathrm{RMSE}_{\mu}$) for the precision-only and combined performance measures, respectively. These metrics for quantifying cost efficiency were selected for conceptual convenience, since an increased value will reflect an increased cost relative to the precision delivered. Thus, the most cost-efficient option among compared alternatives can be identified as the one with the smallest value for this metric.

Exploring cost efficiency under alternative study conditions

The cost-efficiency analysis procedure presented here can be used as a study planning tool to compare different research settings and conditions. As an illustration, effects on total study cost and statistical performance were investigated for three different scenarios: (i) the case of the current study, where researchers started from scratch by developing new methods, conducting all data collection, and processing; (ii) immediate duplication, where the data must be collected but equipment, protocols, trained staff and processing methods already exist; as would be encountered if data from an additional population of flight baggage handlers at another airport with similar logistics were collected again immediately; and (iii) inheriting data from another researcher, where the data are already collected and only need to be processed. Scenario 2 differs from the current scenario by removing costs for equipment ($\check{C}_E$), software ($\check{C}_S$) and training ($\dot{C}_T$) in the calculation of total cost (equation 1). Scenario 3 removes all the fixed and variable costs associated with data collection, limiting costs to those associated with data processing (see table 1).

Since the relative impact of fixed and variable costs differs with the sampling strategy, the preferable method in terms of cost efficiency may differ between studies

of different sizes. To investigate this, the current study and the two alternative scenarios were explored further with simulated sample sizes. Thus, for samples ranging from 5–50 participants with three measurement days each (15–150 measurements in total), the comprehensive (fixed and variable) cost was determined [cf equation (1)], together with statistical performance [cf equations (3) and 4)], using the variance components and bias from the current study.

Results

The data collection, data processing, and combined costs for both classes of measurement are reported in table 1; the descriptive statistics, bias, and variance components for each exposure metric are reported in table 2.

Simulated sampling strategies for the trunk and right arm are depicted for all three exposure metrics in the current scenario (figure 1), the case of immediate study duplication (figure 2), and the case of inherited data (figure 3). The ten individual points on each curve illustrate study designs including 3 measurements on each of 5–50 workers, increasing in increments of 5. These figures demonstrate the statistical performance of the methods in terms of inverse standard error (precision) of the mean (ie, $1/\sqrt{\sigma_\mu^2}$) or in terms of inverse RMSE (combined precision and bias, ie, $1/RMSE_\mu$). An increase on the y-axis therefore means an increase in statistical performance. Statistical performance is shown relative to the total study cost in kilo-euros (x-axis); as the number of workers increases from 5 to 50 and the number of measurements from 15 to 150, the total study cost but also performance increases. For example, figure 1a shows precision of median trunk and arm angle as measured by both inclinometer and video observation for a variety of study designs ranging in cost from €50 000–140 000. At the left-most points on the graph with only 15 measurements (5 subjects for 3 days), most of the costs identified for each method are fixed, due to the initial investment in equipment, planning, and training necessary to allow any measurements in the case where all planning, data collection, and data processing must be performed by the researchers (cf. table 1). For the most minimal study depicted in figure

Table 1. Cost components (cf equation 1) in euros for the measurement methods and exposure data sample in the current study.

Cost component	Costs applicable to all methods [a]	Inclinometry (additional costs)	Video observation (additional costs)
Data collection			
Fixed costs			
Administration	4 888		
Recruitment	2 216		
Equipment	2401	6 030	2 095
Training	4 916	1 777	693
Unit costs			
Recruitment (ie, per worker)	22		
Measurement (ie, per work shift)		359	296
Total data collection cost	28 867	36 475	26 494
Data processing			
Fixed costs			
Administration	8 245	5 114	8 639
Software		10 924	3 479
Unit costs			
Per instructor		0	920
Per trainee		0	1,288
Processing (per work shift file processed)		25	50
Total data processing cost	8 245	17 985	22 386
Total study cost [b]		91 584	85 999

[a] Costs that are required to the same extent no matter which method is used.

[b] Represents the total study cost for applying that method, including the total data collection and total data processing costs in the first column.

Table 2. Descriptive statistics for the two body parts, the two measurement methods, and the three posture variables in the current study. [Mean=group mean across all measured shifts; bias =deviation of the mean group exposure from the value obtained using inclinometry; n_W=number of workers; n_M=number of measurements; σ_{BW}=between-worker standard deviation; σ_{WW}= within-worker standard deviation; $\sqrt{\sigma_\mu^2}$ =standard error of the mean; RMSEμ=root mean squared error of the mean].

Descriptive statistic	Right upper arm		Trunk	
	Inclinometry	Video observation	Inclinometry	Video observation
n_W	27	27	27	27
n_M	79	80	79	80
Median inclination				
Mean	25.6°	15.0°	10.2°	4.1°
Bias	.	-10.6° [a]	.	-6.1° [a]
σ_{BW}	4.93°	0.00°	4.77°	0.64°
σ_{WW}	3.94°	5.18°	3.87°	2.97°
$\sqrt{\sigma_\mu^2}$	1.05°	0.58°	1.02°	0.35°
RMSEμ	1.05°	10.61°	1.02°	6.11°
90th percentile				
Mean	47.0°	45.7°	34.3°	20.2°
Bias	.	-1.2° [a]	.	-13.9° [a]
σ_{BW}	4.79°	0.00°	7.91°	4.66°
σ_{WW}	3.96°	16.16°	4.85°	11.66°
$\sqrt{\sigma_\mu^2}$	1.02°	1.81°	1.61°	1.58°
RMSEμ	1.02°	2.19°	1.61°	13.99°
% time >60°				
Mean	4.25	6.16	2.09	1.05
Bias	.	1.90 [a]	.	1.02 [a]
σ_{BW}	1.75	1.93	1.15	0.91
σ_{WW}	1.38	4.73	1.10	2.09
$\sqrt{\sigma_\mu^2}$	0.37	0.65	0.25	0.29
RMSEμ	0.37	2.01	0.25	1.06

[a] Significant difference between inclinometry and observation (P<0.05) on paired t-test.

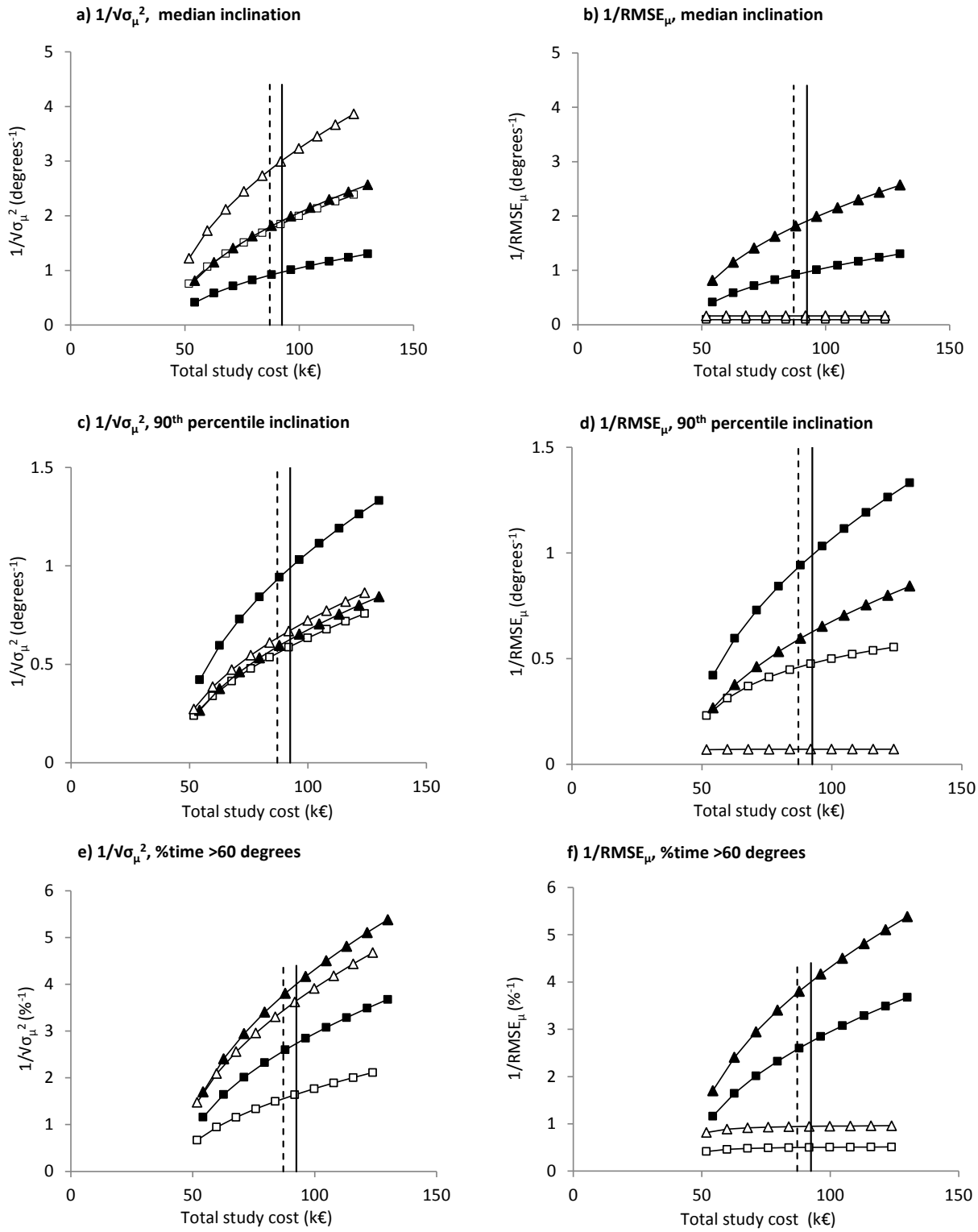

Figure 1. Statistical performance in terms of precision only (a, c, e), and combined precision and bias (b, d, f) by study cost for trunk and right arm inclination across the three investigated posture variables for the current study design. Closed triangles (▲) represent trunk inclinometry; open triangles (△) represent trunk observation; Closed squares (■) represent arm inclinometry; open squares (□) represent arm observation. Solid and dashed vertical lines indicate the cost of the current study for inclinometry and observation, respectively.

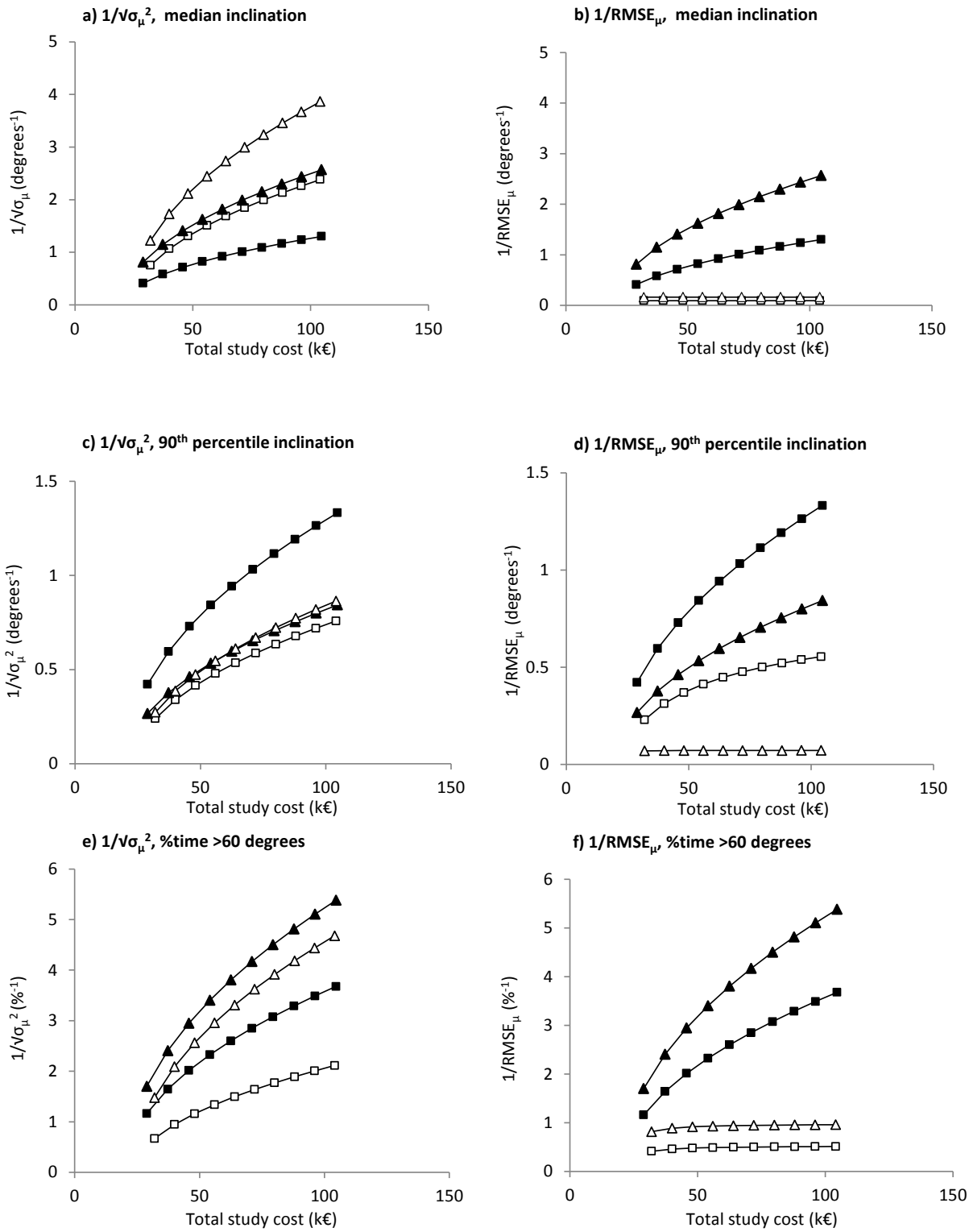

Figure 2. Statistical performance in terms of precision only (a, c, e), and combined precision and bias (b, d, f) by study cost for trunk and right arm inclination across the three investigated posture variables for the case of immediate study duplication. Closed triangles (▲) represent trunk inclinometry; open triangles (△) represent trunk observation; Closed squares (■) represent arm inclinometry; open squares (□) represent arm observation.

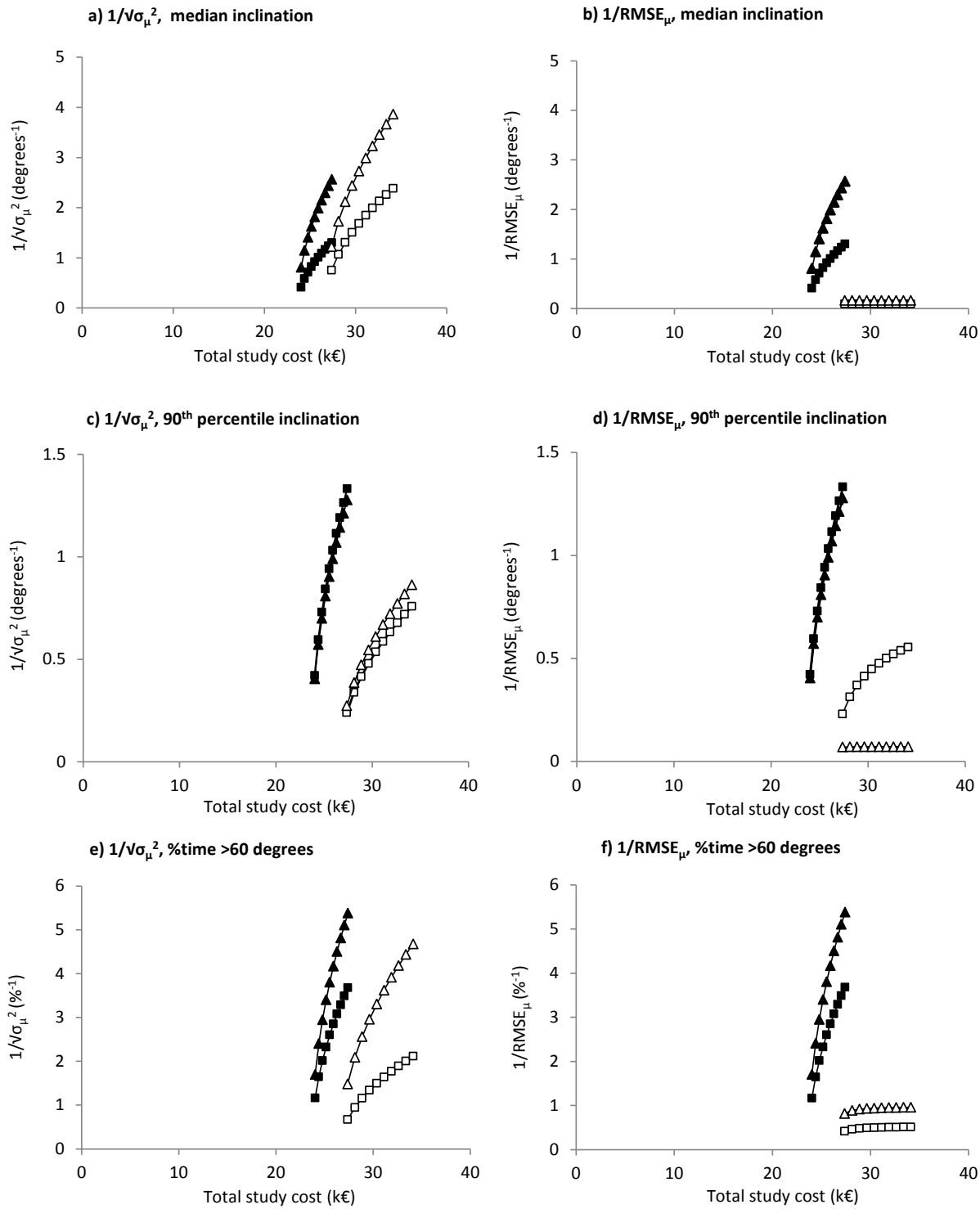

Figure 3. Statistical performance in terms of precision only (a, c, e), and combined precision and bias (b, d, f) by study cost for trunk and right arm inclination across the three investigated posture variables for the case of inherited data. Closed triangles (▲) represent trunk inclinometry; open triangles (△) represent trunk observation; Closed squares (■) represent arm inclinometry; open squares (□) represent arm observation. Note that the scale on the x-axis differ from that in figures 1 and 2.

1a costing roughly €50 000, the performance ($1/\sqrt{\sigma_\mu^2}$) for trunk observation data (black squares) of ~0.41 on the y axis corresponds to a standard error of the mean of approximately 2.43° (1/0.41). Statistical performance increases steeply with an increase in investment towards the right-hand side of the graph, resulting in a value of roughly 1.30 (ie, a standard error of the mean of approximately 0.77°) at the study including 50 subjects, which costs €140 000.

Scenario 2 (figure 2) has intermediate fixed costs, since the costs of equipment and training for both data collection and processing are assumed to be zero. For example, figure 2a also shows trunk observation to have the best statistical performance in terms of precision, but the same levels of performance ($1/\sqrt{\sigma_\mu^2}$ of 0.41 and 1.30) which entailed prices of €50 000 and €140 000 in scenario 1 can now be obtained for a lower study cost: €32 000 and €104 000, respectively.

Scenario 3 (figure 3) eliminates all the costs of data collection, but the fixed costs for software development, training, and administration costs are still over €20 000; this fixed cost comprises the bulk of the study cost on the left-most data point with only 15 measurements. Moving towards the right-hand side of the graph, total study costs increase as determined by the unit cost of additional measurements to a maximum of €33 000 with 150 measurements.

Discussion

Comparing inclinometer and observation: the effect of bias and precision

For both median arm and trunk inclination, observation demonstrated better precision when estimating the group mean exposure (ie, smaller $\sqrt{\sigma_\mu^2}$, table 2), while inclinometry was more precise for the other two arm inclination variables. For trunk 90th inclination percentile and % time inclined >60°, the two methods had almost the same precision (table 2). When costs were also considered and precision alone was used as the indicator of statistical performance, observation still delivered the best performance (ie, the largest $1/\sqrt{\sigma_\mu^2}$) for median inclinations at any particular cost (figure 1a), while inclinometry was more cost-efficient for the % time >60° (figure 1e). For 90th percentiles, inclinometry was more cost-efficient for the arm but less cost-efficient for the trunk (figure 1c). The cost-efficiency relationship between inclinometry and observation was maintained at all simulated sample sizes (figures 1a, 1c, 1e), indicating that the optimal method was consistent over the investigated range of sample sizes.

For several observed metrics, between-worker variability was estimated to be zero (table 2), which is one reason that observations performed well in terms of precision (equation 3). "Zero" between-worker components of variance have been reported previously in observation studies of lumbar posture (47), and they may be an artifact associated with the algorithms used for distributing the overall variance in data between different sources. Thus, if within-worker variability is very large, little or no variability will be "left" for the between-worker component, especially if this component is, truly, small (48). In our study, observer variability may have contributed to inflating within-worker variability, as indicated by most of these variabilities being larger in observed data than in the corresponding inclinometer measurements (table 2).

However, when bias was also considered, the performance of inclinometry relative to observations improved substantially, since the inclinometers were considered to give correct data. The bias associated with observations was substantially higher than the random exposure variability for the body parts and exposure metrics studied, resulting in greater cost efficiency for inclinometry (figures 1b, 1d, 1f). The advantage of inclinometry was maintained across all investigated sample sizes. This result clearly illustrates the importance of recognizing possible bias when documenting the statistical performance of different measurement methods. This is rarely done in occupational epidemiology, probably because the extent of bias is unknown, especially in studies addressing only one exposure assessment method. The status of inclinometry as "best performer" in the present case is due entirely to the a priori decision to designate inclinometers as giving "correct" data, which is standard in occupational studies (38, 49, 50). If observations were designated as the method producing correct results from which to calculate bias, the observation method would have had better relative performance than inclinometers when bias and precision were combined. If both observations and inclinometer results were, in fact, biased compared to some unambiguous but (as of today) unknown gold standard for posture measurement, their rank in terms of cost efficiency would remain as found here if both were either negatively or positively biased, while the relationship might change if their biases had different signs (ie, if the gold standard result were somewhere in-between the results obtained by inclinometry and observation). In the absence of an unambiguous gold standard, the decision to identify inclinometers as producing correct results is consistent with established thinking regarding the hierarchy of method classes in biomechanical exposure assessment (27, 29). The result that observation appears to estimate the group mean of median inclination with greater precision than inclinometers is, however, novel.

Video observation consistently underestimated pos-

tures compared to inclinometer measures for all metrics and body parts (table 2). This is opposite to what has been reported in a study of hairdressers' upper-arm postures, where video observers overestimated inclination angles (21). A lab-based study identified an observation over-estimation in higher exposure ranges for trunk flexion when compared to opto-electric tracking (51); validation studies of neck flexion also showed observations to over-estimate when compared to an inclinometer (52). Lab tests rarely imitate the exposure distributions seen in occupational studies, since experiments generally favor a balanced design with equal proportions of measurement throughout the exposure range. Real work exposure is more likely to be skewed with fewer high exposure instances. In a comparison of field-based observation and inclinometry of trunk postures during real work, Village et al (33) showed that observation underestimated the proportion of time spent in trunk extension, in lateral bending, and in 20-45° of flexion (the most frequent non-neutral category), while overestimating the time spent in neutral postures. It may be that the sampling rate of both the Village et al study and the current observation design are not adequate to capture the relatively less frequent extreme postures. In this vein, a study of the effect of sampling size on bias and precision of estimates of upper-arm elevation obtained using inclinometers showed that 90th percentile inclinations were underestimated when samples were short, and thus less likely to capture rare, extreme posture events (53). If the same effect were also found for work sampling at low frequencies, one would expect the current study's results to show more bias for the 90th percentile than the median, but this was, in fact, only the case for trunk inclination.

It should also be noted that the observations were based on a single-camera image. This may introduce error since a three dimensional phenomenon is being represented by a two-dimensional image. It is ideal to have a perpendicular view of the trunk to assess flexion/extension of the trunk flexion and upper arm, since cameras placed in line with the direction of movement (as opposed to perpendicular) demonstrate poor inter-observer agreement (45), and presumably lower accuracy as well. Although video recording with synchronized cameras at orthogonal angles has been performed in manufacturing contexts (45), it was not practical for full-day recordings of baggage handling work given its dynamic nature, nor would it be appropriate with similar constraints in construction, warehousing, agriculture, or resource industries. Also, it would probably entail an increased cost for collecting and processing data. Furthermore, the effect of changing camera angles from 0 to 90° on the exactness of posture observation has been shown to be small (54), and single cameras are acceptable for determining trunk posture in many cases (45).

The observation method used here employs smaller exposure bins (ie, of 1°) than many other observation methods where bins of 15 or 30° are typically used. This was done to allow for treatment of inclination as a continuous variable, both when obtained from observation and from inclinometers, and thus facilitate direct comparison of observation and inclinometer measurements at the same level of resolution. Data on a continuous scale are also a prerequisite for calculating two of our three exposure variables (ie, the median and the 90th percentile). The cost efficiency of categorical observation, which may be performed faster than the high-resolution observation applied by us, is an interesting issue for additional research.

Comparing different body parts: trunk versus upper arms

In the current study, median trunk and arm inclination are more precisely estimated using observation, while inclinometry is more precise for arm 90th percentile and % time >60° (table 2). The two methods have almost the same precision for trunk 90th percentile and % time >60°. Inclinometry performs better for all arm and trunk posture variables if bias is also taken into consideration. The performance by total study cost is very similar for observation of both body parts when considering combined bias and precision (figure 1b, 1d, and 1f). Given the scale of the bias, it seems likely that (especially at larger sample sizes) the measurement performance will be more similar for different body parts using the same method than for the same body part using different methods. That is, once a method is selected it would appear that additional body parts could be assessed with similar performance. Although this appears mostly true for the trunk and the upper arm, these are also two of the largest and easiest-to-assess body segments. Observation error can increase for smaller body parts (45), when the viewing angle is altered (54), and when higher resolution is demanded. Thus, our cost-efficiency results are more likely to be representative for large body parts moving through large ranges of motion, for example hip flexion or knee flexion rather than movement of smaller segments like ulnar deviation or finger flexion.

Considering alternative research scenarios

Scenario 2 acknowledges the savings gained when a research team has the protocols, trained staff, and equipment to immediately embark on a similar study without the labor of methods development and study planning and the capital cost of equipment. Since total study cost is lower, the statistical performance ($1/\sqrt{\sigma_{\mu}^2}$ or $1/\mathrm{RMSE}_{\mu}$) obtained at a certain budget is larger in scenario 2 (figure 2) than scenario 1 (figure 1). Performance at a particular

sample size does not change in scenario 2 compared to scenario 1, so cost efficiency for both inclinometry and observation is better in scenario 2 than in scenario 1, while the relationship between the two methods is maintained.

It is not unusual for researchers to repurpose previously collected data to answer an additional research question. This might be in the form of resampling existing direct measurements (1, 12, 13), or repeated observation of video film to investigate inter-rater differences or compare different processing schemes (21, 32). Scenario 3 describes a situation where no new data collection is required, resulting in a substantial reduction in total study cost. The reduction in cost is more pronounced for the inclinometer method, since the data processing costs for the chosen observation protocol make up a large proportion of the total cost of that method. Due to this decreased total study cost, inclinometer demonstrates greater performance at a certain budget, when considering either only precision or bias and precision combined. One exception is median trunk angle, where observation performs better in terms of precision than inclinometer for study costs over €30 000 (figure 3a). As with the other two scenarios, inclinometer shows greater advantages in cost efficiency when combined bias and precision are considered.

Choosing a measurement method

In this study, we rank exposure assessment methods on the basis of cost efficiency. However, there are additional qualitative considerations when selecting a measurement method. Although the current study has compared several exposure metrics for postural angles, a comprehensive biomechanical assessment could involve measures of force, muscle activity, manual materials handling, and exposure to contact forces or vibration. Observation has been identified as a more versatile method than direct measurements; if assessment of multiple types of exposure is desired, a single observation protocol might more easily assess several of these aspects (27). Because of this, observation is uniquely suited for screening; checklists and screening instruments can be used to identify potentially harmful situations or particularly exposed body parts worthy of more in-depth analysis using a direct measurement method (47). In contrast, a single inclinometer can deliver only data on postural angle of the body part on which it is mounted and derivatives thereof, such as movement velocities.

The results presented here suggest that, in most scenarios, inclinometry was the more cost-efficient alternative for data collection, which confirms the presumed immutable "exposure assessment hierarchy". It is not clear, however, whether our specific results can be extrapolated to other techniques for direct measurement and observation and to what extent they apply even to other posture variables. While, for instance, movement frequency, forces applied, and loads in manual handling tasks are difficult to assess at all by observation, several exposure metrics may be more easily observed than monitored by technical equipment; for instance frequency of lifting or nature of manual handling (i.e. push vs. pull vs. lift vs. carry). Thus, precision, bias, and cost efficiency is probably highly specific to the exact method used for exposure assessment.

An additional factor of concern when deciding for measurement instrument(s) in a particular exposure assessment is the possible need to involve highly educated staff when dealing with collection and, especially, processing of directly measured exposures. In the present study, it was not possible to estimate the cost of training the inclinometer processers, as these were the same researchers who developed and tested the inclinometer processing software so the two costs were intertwined. It could be argued that this expertise is accounted for in the hourly wage of skilled researchers, but there may be other situations where researchers do not have specific expertise in the instrumentation at hand and will need to spend specific time training. Where this is the case, the fixed cost of inclinometer would increase, thereby skewing the cost-efficiency results to favor observation. The results of the present study do quantify separate costs for observation software development and training, so the comparison between measurement methods and the extrapolation to other situations should be considered carefully within the local context. It should be noted that as technology and methods develop, skills must be maintained and developed, otherwise one can imagine a "depreciation" of skills and training over time, just as seen with equipment and other capital.

In the future, advancements in technology towards cheaper instruments and trends of more highly automated data processing can have an impact both on the performance of direct measurements and their cost. Technological advancements may also have an impact on video analysis through the use of image pattern recognition, allowing for partial or complete automation and decreased processing costs. The numerical results presented here can be considered a detailed quantification of cost efficiency only as far as technology, equipment, and labor costs remain stable; large changes in these factors would require a recalculation using our methodology with updated cost and performance data as inputs.

Concluding remarks

By the comparison of two different methods for assessing trunk and upper-arm inclination (inclinometry and video-based observation) in terms of their cost

efficiency, this study joins an overall limited body of research that quantifies and compares the cost efficiency of exposure assessment methods. The results demonstrate that at a certain total study cost, higher precision estimates of median exposure were delivered using observation, while time spent at angles >60° as well as the 90th percentile arm angle were determined with better precision using inclinometer. When adopting the standard assumption of inclinometers providing correct angular data, observations were shown to be biased. Adding this bias into the metric for statistical performance led to inclinometers consistently outperforming observations for all posture variables, irrespective of scenario.

While this study has been devoted mainly to assessment of working postures, we propose that assessment of cost efficiency is an important step in informed epidemiologic study design in general. To that end, we recommend our approach of calculating costs using a comprehensive cost model, assessing statistical performance on the basis of exposure variance components and suspected bias, and then merging the two in a quantitative relationship between cost and performance.

Acknowledgments

The authors would like to thank AFA Insurance (grant # 100071) and the Swedish Research Council for Health, Working Life and Welfare (Forte; Dnr 2009-1761) for their financial support, as well as study partners SAS Ground Services and the Vocational Training & Working Environment Council (Transport Trades, TYA). We gratefully acknowledge the data collection efforts of Jennie Jackson and Carl Lind, the data processing efforts of Margaretha Marklund, Kerstin Nilsson, Lars Holm, Karin Holmkvist, Johan Hultman, and Niklas Lindfors, as well as the software development efforts of Per Gandal. We would particularly like to thank all of the flight loaders who participated in this study.

The authors declare no conflicts of interest.

References

1. Mathiassen SE, Burdorf A, van der Beek AJ. Statistical power and measurement allocation in ergonomic intervention studies assessing upper trapezius EMG amplitude. A case study of assembly work. J Electromyogr Kinesiol. 2002;12(1):45-57. http://dx.doi.org/10.1016/S1050-6411(01)00028-1

2. Mathiassen SE, Moller T, Forsman M. Variability in mechanical exposure within and between individuals performing a highly constrained industrial work task. Ergonomics. 2003;46(8):800-24. http://dx.doi.org/10.1080/0014013031000090125

3. Liv P, Mathiassen SE, Svendsen SW. Theoretical and empirical efficiency of sampling strategies for estimating upper arm elevation. Ann Occup Hyg. 2011;55(4):436-49. Epub 2011/04/14.

4. Chen CC, Chuang CL, Wu KY, Chan CC. Sampling strategies for occupational exposure assessment under generalized linear model. Ann Occup Hyg. 2009;53(5):509-21. Epub 2009/05/23.

5. Kromhout H, Symanski E, Rappaport SM. A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. Ann Occup Hyg. 1993;37(3):253-70. http://dx.doi.org/10.1093/annhyg/37.3.253

6. Rappaport SM. Assessment of long-term exposures to toxic substances in air. Ann Occup Hyg. 1991;35(1):61-121. http://dx.doi.org/10.1093/annhyg/35.1.61

7. Armstrong B. Study design for exposure assessment in epidemiological studies. Sci Total Environ. 1995;168(2):187-94. Epub 1995/06/16.

8. Loomis D, Kromhout H. Exposure variability: concepts and applications in occupational epidemiology. Am J Ind Med. 2004;45(1):113-22. http://dx.doi.org/10.1002/ajim.10324

9. Burdorf A. Sources of variance in exposure to postural load on the back in occupational groups. Scand J Work Environ Health. 1992;18(6):361-7. http://dx.doi.org/10.5271/sjweh.1561

10. Burdorf A. Reducing random measurement error in assessing postural load on the back in epidemiologic surveys. Scand J Work Environ Health. 1995;21(1):15-23. http://dx.doi.org/10.5271/sjweh.3

11. van der Beek AJ, Kuiper JI, Dawson M, Burdorf A, Bongers PM, Frings-Dresen MH. Sources of variance in exposure to nonneutral trunk postures in varying work situations. Scand J Work Environ Health. 1995;21(3):215-22. http://dx.doi.org/10.5271/sjweh.30

12. Trask C, Koehoorn M, Village J, Johnson P, Teschke K. How long is long enough? Evaluating sampling durations for low-back EMG assessment. J Occup Environ Hyg. 2008;5(10):664-70. http://dx.doi.org/10.1080/15459620802313840

13. Mathiassen SE, Burdorf A, van der Beek AJ, Hansson GA. Efficient one-day sampling of mechanical job exposure data--a study based on upper trapezius activity in cleaners and office workers. AIHA J (Fairfax, Va). 2003;64(2):196-211. http://dx.doi.org/10.1080/15428110308984809

14. Fethke NB, Anton D, Cavanaugh JE, Gerr F, T.M. C. Bootstrap exploration of the duration of surface electromyography sampling in relation to the precision of exposure estimation. Scand J Work Environ Health. 2007;33(5):358-67. http://dx.doi.org/10.5271/sjweh.1155

15. Hoozemans MJ, Burdorf A, van der Beek AJ, Frings-Dresen MH, Mathiassen SE. Group-based measurement strategies in exposure assessment explored by bootstrapping. Scand J Work Environ Health. 2001;27(2):125-32. http://dx.doi.org/10.5271/sjweh.599

16. Andrews DM, Callaghan JP. Determining the minimum sampling rate needed to accurately quantify cumulative spine loading from digitized video. Appl Ergon. 2003;34(6):589-95. http://dx.doi.org/10.1016/S0003-6870(03)00077-2

17. Lampa EG, Nilsson L, Liljelind IE, Bergdahl IA. Optimizing occupational exposure measurement strategies when estimating the log-scale arithmetic mean value--an example from the reinforced plastics industry. Ann Occup Hyg. 2006;50(4):371-7. Epub 2006/02/10.

18. Rezagholi M, Mathiassen SE. Cost-efficient design of occupational exposure assessment strategies--a review. Ann Occup Hyg. 2010;54(8):858-68. Epub 2010/10/12.

19. Mathiassen SE, Liv P, Wahlström J. Cost-efficient measurement strategies for posture observations based on video recordings. Appl Ergon. 2013;44:609-17. Epub Jan 2013. http://dx.doi.org/10.1016/j.apergo.2012.12.003

20. Mathiassen SE, Bolin K. Optimizing cost-efficiency in mean exposure assessment - cost functions reconsidered. BMC Med Res Methodol. 2011;11:76. Epub 2011/05/24.

21. Rezagholi M, Mathiassen SE, Liv P. Cost efficiency comparison of four video-based techniques for assessing upper arm postures. Ergonomics. 2012;55(3):350-60. Epub 2012/03/14.

22. Armstrong BG. Optimizing power in allocating resources to exposure assessment in an epidemiologic study. Am J Epidemiol. 1996;144(2):192-7. Epub 1996/07/15.

23. Duan N, Mage DT. Combination of direct and indirect approaches for exposure assessment. J Expo Anal Environ Epidemiol. 1997;7(4):439-70.

24. Whitmore RW, Pellizzari ED, Zelon HS, Michael LC, Quackenboss JJ. Cost/variance optimization for human exposure assessment studies. J Expo Anal Environ Epidemiol. 2005;15(6):464-72. Epub 2005/05/12.

25. Lemasters GK, Shukla R, Li YD, Lockey JE. Balancing cost and precision in exposure assessment studies. J Occup Environ Med. 1996;38(1):39-45. Epub 1996/01/01.

26. Shukla R, Luo J, Lemasters GK, Grinshpun SA, Martuzevicius D. Sampling over time: developing a cost effective and precise exposure assessment program. J Environ Monit. 2005;7(6):603-7. Epub 2005/06/03.

27. Winkel J, Mathiassen SE. Assessment of physical work load in epidemiologic studies: concepts, issues and operational considerations. Ergonomics. 1994;37(6):979-88. http://dx.doi.org/10.1080/00140139408963711

28. Burdorf A, van der Beek AJ. In musculoskeletal epidemiology are we asking the unanswerable in questionnaires on physical load? Scand J Work Environ Health. 1999;25(2):81-3. http://dx.doi.org/10.5271/sjweh.409

29. van der Beek AJ, Frings-Dresen MH. Assessment of mechanical exposure in ergonomic epidemiology. Occup Environ Med. 1998;55(5):291-9. http://dx.doi.org/10.1136/oem.55.5.291

30. Burdorf A, Derksen J, Naaktgeboren B, van Riel M. Measurement of trunk bending during work by direct observation and continuous measurement. Appl Ergon. 1992;23(4):263-7. http://dx.doi.org/10.1016/0003-6870(92)90154-N

31. David GC. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. Occup Med (Lond). 2005;55(3):190-9. http://dx.doi.org/10.1093/occmed/kqi082

32. Weir PL, Andrews DM, van Wyk PM, Callaghan JP. The influence of training on decision times and errors associated with classifying trunk postures using video-based posture assessment methods. Ergonomics. 2011;54(2):197-205. Epub 2011/02/05.

33. Village J, Trask C, Luong N, Chow Y, Johnson P, Koehoorn M, et al. Development and evaluation of an observational Back-Exposure Sampling Tool (Back-EST) for work-related back injury risk factors. Appl Ergon. 2009;40:538-44. http://dx.doi.org/10.1016/j.apergo.2008.09.001

34. Paquet V, Punnett L, Buchholz B. Validity of fixed-interval observations for postural assessment in construction work. Appl Ergon. 2001;32(3):215-24. http://dx.doi.org/10.1016/S0003-6870(01)00002-3

35. Trask C, Mathiassen SE, Wahlstrom J, Heiden M, Rezagholi M. Data collection costs in industrial environments for three occupational posture exposure assessment methods. BMC Med Res Methodol. 2012;12:89. Epub 2012/06/29.

36. Trask C, Teschke K, Village J, Chow Y, Johnson P, Luong N, et al. Measuring low back injury risk factors in challenging work environments: an evaluation of cost and feasibility. Am J Ind Med. 2007;50(9):687-96. http://dx.doi.org/10.1002/ajim.20497

37. Trask C, Mathiassen SE, Wahlstrom J, Jackson J. Data processing costs for three posture assessment methods. BMC Med Res Methodol. 2013;13(124):1-14.

38. Teschke K, Trask C, Village J, Johnson P, Koehoorn M. Measuring posture for epidemiology: Comparing inclinometry, observations, and self-reports. Ergonomics. 2009;52(9):1067-78. http://dx.doi.org/10.1080/00140130902912811

39. Wahlstrom J, Mathiassen SE, Liv P, Hedlund P, Ahlgren C, Forsman M. Upper arm postures and movements in female hairdressers across four full working days. Ann Occup Hyg. 2010;54(5):584-94. Epub 2010/04/14.

40. Bernmark E, Forsman M, Wiktorin C. Head movements during two computer work tasks assessed by accelerometry. Appl Ergon. 2011;42(2):309-13. Epub 2010/08/27.

41. Hansson GA, Asterland P, Holmer NG, Skerfving S. Validity and reliability of triaxial accelerometers for inclinometry in posture analysis. Medical & biological engineering & computing. 2001;39(4):405-13. Epub 2001/08/29.

42. Kazmierczak K, Mathiassen SE, Forsman M, Winkel J. An integrated analysis of ergonomics and time consumption in Swedish 'craft-type' car disassembly. Appl Ergon. 2005;36(3):263-73. Epub 2005/04/28.

43. Palmerud G, Forsman M, Neumann WP, Winkel J. Mechanical exposure implications of rationalization: a comparison of two flow strategies in a Swedish manufacturing plant. Appl Ergon. 2012;43(6):1110-21. Epub 2012/05/12.

44. Bao S, Howard N, Spielholz P, Silverstein B. Two posture analysis approaches and their application in a modified rapid upper limb assessment evaluation. Ergonomics. 2007;50(12):2118-36. Epub 2007/09/14.

45. Bao S, Howard N, Spielholz P, Silverstein B, Polissar N. Interrater reliability of posture observations. Human factors.

2009;51(3):292-309. Epub 2009/09/16.

46. Hansson GA, Arvidsson I, Ohlsson K, Nordander C, Mathiassen SE, Skerfving S, et al. Precision of measurements of physical workload during standardised manual handling. Part II: Inclinometry of head, upper back, neck and upper arms. J Electromyogr Kinesiol. 2006;16(2):125-36. Epub 2005/08/17.

47. Burdorf A, Van Riel M. Design of strategies to assess lumbar posture during work. Int J Ind Ergon. 1996;18(4):239-49. http://dx.doi.org/10.1016/0169-8141(95)00035-6

48. Searle S, Casella G, McCulloch C. Variance components Hoboken: John Wiley & Sons, Inc.; 2006. 501 p.

49. Lowe BD. Accuracy and validity of observational estimates of shoulder and elbow posture. Appl Ergon. 2004;35(2):159-71. Epub 2004/04/24.

50. Spielholz P, Silverstein B, Morgan M, Checkoway H, Kaufman J. Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors. Ergonomics. 2001;44(6):588-613.

51. Leskinen T, Hall C, Rauas S, Ulin S, Tonnes M, Viikari-Juntura E, et al. Validation of Portable Ergonomic Observation (PEO) method using optoelectronic and video recordings. Appl Ergon. 1997;28(2):75-83. Epub 1997/04/01.

52. Juul-Kristensen B, Hansson GA, Fallentin N, Andersen JH, Ekdahl C. Assessment of work postures and movements using a video-based observation method and direct technical measurements. Appl Ergon. 2001;32(5):517-24. Epub 2001/09/06.

53. Mathiassen SE, Wahlstrom J, Forsman M. Bias and imprecision in posture percentile variables estimated from short exposure samples. BMC Med Res Methodol. 2012;12:36. Epub 2012/03/27.

54. Sutherland CA, Albert WJ, Wrigley AT, Callaghan JP. The effect of camera viewing angle on posture assessment repeatability and cumulative spinal loading. Ergonomics. 2007;50(6):877-89. Epub 2007/04/26.

Received for publication: 25 June 2013