International Journal of Industrial Ergonomics 39 (2009) 333–340

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: www.elsevier.com/locate/ergon

Force measurement in field ergonomics research and application

Stephen Bao*, Peregrin Spielholz, Ninica Howard, Barbara Silverstein

Safety and Health Assessment and Research for Prevention (SHARP) Program, Washington State Department of Labor and Industries, PO Box 44330, Olympia, WA 98504, USA

ARTICLE INFO

Article history:
Received 14 January 2008
Received in revised form
13 March 2008
Accepted 14 March 2008
Available online 19 May 2008

Keywords:
Force matching
Self-reporting
Observational methods
Exposure assessment

ABSTRACT

This paper addresses the issue of quantifying forceful exertions of lifting, pushing/pulling, pinch and power griping with several commonly used methods (direct measurement, force-matching, ergonomist estimation based on observation and worker's self-report). The aims were to study differences of ergonomists in making decisions of collecting forceful exertion data, ability of the studied force quantification methods in detecting exposure differences between jobs, and relationships between measurements obtained by different methods. Seven hundred and thirty-three (733) subjects participated in the study, and 2482 forceful exertions were quantified with the selected force quantification methods. Results showed that the determination of whether a forceful exertion was considered important enough to be measured was very subjective. More objective criteria need to be developed. Although the different force quantification methods could detect exertion-level differences between different types of jobs, the sensitivity of detecting the difference varied. Direct measurement of lifting and pushing/pulling forces seems to be slightly more sensitive than ergonomists' estimates, and the force matching of grip forces may be more sensitive than workers' self-reports. Ergonomists estimations through observation seem to be a good alternative for measuring forceful exertions. Pearson correlation coefficients between force estimations obtained by the different methods were between 0.28 and 0.71. The degrees of the correlations between the different measurement methods varied. This indicates that the different methods might quantify different aspects of the forceful exertions. These methods may not necessarily be used interchangeably in job evaluation tools. Different weights and/or cut-points may need to be developed when different force quantification methods are used.

Relevance to industry: Objective criteria of collecting forceful exertion data in job evaluations need to be developed. Different forceful exertion quantification methods may be used in job evaluations. However, different cut-points classifying jobs risk levels should be used when different exertion quantification methods are used.

1. Introduction

Forceful exertion has been considered one of the most important exposure factors contributing to upper extremity work-related musculoskeletal disorders (National Institute for Occupational Safety and Health (NIOSH), 1997). Several widely used exposure assessment methods use estimates of force as a primary measure for determination of injury risk. The Strain Index (Moore and Garg, 1995) uses six exposure parameters (intensity of exertion, duration of exertion, frequency of exertion, hand/wrist posture, speed of work and duration of work) to calculate a single index (the Strain Index, or SI) which quantifies upper extremity exposures. The intensity or level of exertion has the highest weight in the calculation among the six parameters. Another example is the ACGIH hand activity level threshold limit values (TLV) method of the American Conference of Governmental Industrial Hygienists (ACGIH, 2001), which uses peak hand force in combination with repetitiveness to determine the exposure level. Forceful exertions are also quantified in many different ways in musculoskeletal disorder epidemiological studies (Burdorf and van der Beek, 1999; Hoozemans et al., 1998; van der Beek et al., 1999). Analyses of baseline data from a recently completed prospective epidemiological study of work-related upper extremity musculoskeletal disorders showed that forceful exertions, including high handgrip forces, lifting, and pushing/pulling forces, were significantly associated with the prevalence of carpal tunnel syndrome, lateral epicondylitis, and rotator cuff syndrome (Fan et al., 2007; Silverstein et al., 2007).

Different methods have been used to quantify forceful exertion in ergonomics application and research. Hand forces applied by a worker have been collected by direct measurement of forces, for example, by mounting a force gauge or gauges in tool handles, or by

* Corresponding author. Tel.: +360 902 5676; fax: +360 902 5672.
E-mail address: baos235@lni.wa.gov (S. Bao).

0169-8141/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ergon.2008.03.005

simulating work using instrumented handles (McGorry et al., 2004). Estimates of muscle effort, the internal response to forceful exertions, have been used in some studies by measuring electromyography from specific muscles of the hand and forearms (Cook et al., 1998). Indirect measurement of forceful exertions using a force-matching method (Bao and Silverstein, 2005; Casey et al., 2002) has been used often by ergonomics practitioners. This method asks a worker who performs a task involving gripping actions to mimic the amount of force exerted using a dynamometer, immediately following the actual task performance. The estimate on the dynamometer is considered the hand force applied during the task. Forceful exertions have also often been estimated by trained evaluators based on observations of task performances of workers using a rating scale, such as the one developed by Latko (1997), or by self-reports from workers after exerting a force using a scale such as the Borg scale (1990). The Latko scale is one of the recommended methods used in the ACGIH hand activity level TLV method (ACGIH, 2001).

There is a misconception among some ergonomists that one force measurement method can be used interchangeably with another without a significant difference in results. The SI (Moore and Garg, 1995) and the ACGIH hand activity level TLV method (ACGIH, 2001) have listed a number of method options for forceful exertion quantification without clarification on the differences between methods and impact on the resulting assessment. This gives users a false impression that one might be able to choose among the suggested methods for a job evaluation and obtain equally robust results.

Forceful exertion quantification in many situations is a surrogate measure of certain aspects of the exertion. For example, using forearm muscle EMG to quantify forceful hand exertion assumes that there is a relationship between the handgrip force and the forearm muscle activities. In reality, work techniques, hand postures and worker's individual factors all can influence this relationship. Therefore, in many situations, no perfect relationships could be found. A recent study (Bao et al., 2006b, 2006c) using a partial dataset of the present study examined the repetitiveness aspects of forceful hand exertions according to two different repetitiveness definitions and estimated forceful exertion efforts of task performance with three different methods. A correlation analysis was also performed between the three different forceful exertion quantification methods. Initial results showed that the methods used seem to be able to quantify forceful exertion level and repetitiveness of hand activities (Bao et al., 2006c), and the Spearman rank-order correlation coefficients between the different forceful exertion estimation methods were around 0.30–0.75 (Bao et al., 2006a), suggesting different measurement methods may quantify different aspects of forceful exertions. However, that paper did not focus in details on issues related to decision on forceful exertion data collection and comparisons of different forceful exertion quantification methods. In addition, a larger dataset is available for the present analysis.

The goal of the present study was to further examine different forceful exertion estimation techniques used in a large epidemiological studies of upper extremity musculoskeletal disorders for lifting, pushing/pulling, pinch grip and power grip forces. This paper focused on the aspect of forceful exertion level estimations. Specific aims of the study were: (1) to study empirically differences between ergonomists in making decisions on whether an exertion was important to be measured, (2) to compare the ability of differentiating exertion levels between jobs when different forceful exertion estimation methods were used; and (3) reexamine the relationships between the different forceful exertion estimation methods using a larger dataset.

2. Materials and methods

2.1. Subjects

Seven hundred and thirty-three subjects (female/male: 350/383, median age: 42.0/36.6 years, range of age: 18.1–64.1/18.1–64.9) from 12 different worksites in manufacturing ($N=634$ or 86%) and health care ($N=99$ or 13.5%) participated in the study. About 14.6% of the subjects (107) had hand/wrist problems. The data used in the analyses included all forceful exertion measurements conducted during a 3-year period among these subjects.

This study was approved by the Washington State Institutional Review Board (IRB). Written consent from each subject was obtained prior to the data collection. The consent forms were available in different languages and verbal translations provided by professional interpreters, were given if needed.

2.2. Field data collection

Subjects participated in this study were randomly assigned to one of the three ergonomists involved in the project without the knowledge of their forceful exertions and health conditions. The ergonomists in turn observed them performing tasks at their worksites. All ergonomists were certified professional ergonomists with formal ergonomics education and many years experience in assessing physical exposures in workplaces. Work elements involving forceful exertions were noted during the observation and then quantified using several different methods on-site. Four types of forceful exertion were examined for this paper: (1) pinch griping, (2) power griping, (3) lifting, and (4) pushing/pulling. The forceful exertion data collection was a part of extensive data collection and analysis procedures used in a large prospective epidemiological study on work-related upper extremity musculoskeletal disorders. The ergonomists interviewed the workers and supervisors about the tasks, and selected a period of at least 15 min for each worker when typical task activities were performed. These task activities were then observed on-site by the ergonomists. During the observation the ergonomists identified all exertions related to the task performances, and made a determination whether an exertion was considered "forceful exertion" and important enough to be estimated. The definition of "forceful exertion" was similar to that used by Stetson et al. (1991), where it was defined as "a conspicuous application of force by the hand and included using the hand to hold, manipulate, trigger, push, pull, or otherwise handle an object". Operationally, force data were collected if force was obvious and considered of importance by an ergonomist subjectively. For the four types of forceful exertions studied, conceptually, a forceful exertion was defined as a lifting force of $\geqslant 8.9$ Newton (N), a pushing/pulling force of $\geqslant 44.5$ N, a pinch griping force of $\geqslant 8.9$ N or a power griping force of $\geqslant 44.5$ N. These inclusion force values were chosen for consistency with Stetson et al. (1991) and hand force limits used in the repealed Washington State Ergonomics Rule (Washington State Department of Labor and Industries, 2000). The actual force level was not known until measured; therefore, the actual cut-point level depended on the subjective judgement of an individual ergonomist.

After the identification of forceful exertions, three different methods were used to quantify the pinch and power griping forces, and two methods were used to quantify the lifting and pushing/pulling forces. The three methods used to quantify pinch and power griping forces were: (1) estimation using the force-matching method (Bao and Silverstein, 2005), (2) observational ergonomist estimation, and (3) worker self-reports. The two methods used to quantify lifting and pushing/pulling forces were: (1) measurement using a force gauge, and (2) observational ergonomist estimation.

All measured forces were included in analyses of the present study, including forces that were actually lower than the conceptually defined levels.

The force-matching method (Bao and Silverstein, 2005) used to estimate pinch and power griping hand forces was done by asking a subject to recreate the amount of force he/she used in performing an identified work element on a force dynamometer (Flaghouse hydraulic hand dynamometer, Flaghouse, Inc., Hasbrouck Heights, NJ 07604, USA, or FEI hydraulic pinch gauge, Irvington, NY 10533, USA) using similar hand/wrist postures immediately after performing the actual element on the job. Keeping similar hand/wrist postures were important to obtain a reliable force estimation (Bao and Silverstein, 2005), for example, if the task was performed with a 3-finger pinch grip, the force matching was also done using a 3-finger pinch grip with similar grip span and hand orientation. This process was repeated three times, and the median of the three was used in the analysis. For the second method, observational ergonomist estimation, ergonomists rated the hand forces using a linear visual analog rating scale of 0 (none) to 10 (greatest imaginable) (Latko, 1997). The third method of pinch and power grip quantification, worker self-report, asked subjects to rate the levels of corresponding hand forces of the various task performance events where forceful exertions were identified using the Borg CR-10 scale with verbal anchors such as "nothing at all" for a Borg scale of 0, "strong (heavy)" for 5 and "extremely strong or almost maximal" for 10 (Borg, 1990).

Lifting force was measured by obtaining the weight of the lifted or lowered object. This was typically done by measuring the object weight using a force gauge (Chatillon, CSD 200 Strength dynamometer, Ametek, Largo, FL 33773, USA). In some instances, the weight information was already documented (provided by manufacturers, and are available to the company) and this information was used as the lifting force. Pushing/pulling force was also measured using a force gauge (Chatillon, CSD 200 Strength dynamometer, Ametek, Largo, FL 33773, USA). This was done by attaching the force gauge to the handle or object and simulating the pushing/pulling activity. For practical purposes, no distinction was made between pushing and pulling forces, although they may have different physiological impacts on the body. Additionally, both lifting force and pushing/pulling force were also estimated after observation by one of the three ergonomists using the same linear visual analog rating scale as used for the grip forces (Latko, 1997). Ergonomists performed the estimations without the knowledge of the object weights and force measurements (i.e. prior to taking the weight and force measurements with instrument). For practical reasons, workers were not asked to rate the lifting and pushing/pulling forces so to minimize the interference of task performances of the subjects.

Other details of the complete data collection and analysis procedures used in the large epidemiological study can be found in related publications (Bao et al., 2006b, c).

2.3. Data analysis and statistics

In order to study the ability of different force quantification methods in detecting exposure differences between jobs, a contrast of physical exposure among the selected jobs was created. This was done by categorizing the 733 jobs (individual subjects) into five different groups according to the selected characteristics of occupations defined in the revised dictionary of occupational titles (US Department of Labor, 1993)

- *Sedentary work* (*S*): Exerting up to 10 lb of force occasionally or a negligible amount of force to lift, carry, push, pull or move objects including the human body, involves sitting most of the time, but may include walking or standing for brief periods (122 workers or 16.6% of the study population),
- *Light work* (L): Exerting up to 20 lb of force occasionally or up to 10 lb of force frequently or negligible amount of force constantly to move objects, walking/standing for significant times, or sitting most of the time while using arm/leg controls, working at production rate while constantly pushing or pulling materials even thought the weight of the materials is negligible (210 workers or 29.1% of study population),
- *Medium work* (*M*): Exerting 20–50 lb of force occasionally, or 10–25 lb of force frequently, or up to 10 lb to constantly move objects (272 workers or 37.1% of the study population),
- *Heavy work* (H): Exerting 50–100 lb of force occasionally, or 25–50 lb of force frequently, or 10–20 lbs of force constantly to move objects (112 workers or 15.3% of the study population),
- *Very heavy work* (*V*): Exerting more than 100 lb of force occasionally, or more than 50 lb of force frequently, or more than 20 lb of force to constantly move objects (17 workers or 2.3% of the study population).

This categorization was done by an analyst who reviewed the jobs and compared to the above job categorization definitions. Descriptive statistics for different force estimation results were calculated by the job categories.

All analyses were conducted for all force measurement methods of the four types of forceful exertions. Analyses of variance (ANOVAs) were performed to compare the means of each of the force measurement methods by job category with a significance level of $p=0.01$. The Scheffe's multiple-comparison procedure was used in the post hoc analyses to compare the means of forceful exertions between job categories.

In order to study the degree of correlations between the different force measurement methods, correlation analyses were performed between pairs of forceful exertion measurements for each of the four types of forces when different measurement methods were used. Pearson product–moment correlation coefficients were also computed for each of the forces. All statistical analyses were performed using the SAS statistical program (V. 9E, SAS Institute, Cary, NC, USA).

3. Results

Five hundred seventy-nine (579) subjects among the 733 subjects had forceful exertion measurements during the course of the 3-year study period. Due to job and work process changes, exposures changed for some workers during the study. Forces were remeasured when forceful exertions or levels were thought to have been changed. All together, 753 forceful exertions profiles were measured among the 733 subjects. A forceful exertion profile contained one or more forceful exertion measures that occurred as part of a defined task in a subject's job. Some workers applied several different forces (different types and/or different levels). If a worker performed more than one task in a job, and all tasks required forceful exertions, the worker would have more than one forceful exertion profile. In total, 2482 forceful exertions were measured for the 579 subjects who had forceful exertions. In general, all three ergonomists measured many forceful exertions that were lower than the conceptually defined minimal inclusion levels (9.9% of the measurements were lower than the inclusion levels of the 4 different forces).

Table 1 shows the distribution of subject assignment to the three ergonomists. It could be seen that slightly more subjects were assigned to ergonomist A. Within the subjects assigned to each of the ergonomists, more subjects assigned to Ergonomist A were considered having forceful exertion measurements compared to

Table 1
Number (percent) of subject assignments/forceful exertion measurements among three ergonomists.

	Ergonomist A	Ergonomist B	Ergonomist C
Distribution by subjects assignments (all)	399 (39%)	294 (29%)	321 (32%)
Distribution by subjects assignments (with force)	315 (42%)	169 (22%)	269 (36%)
Distribution by forceful exertion measurements	1183 (48%)	456 (18%)	843 (34%)

the other two ergonomists. Ergonomist A perceived more exertions to be forceful and important, and therefore took more measurements than the other ergonomists.

Fig. 1 shows average forces by job categories using the different force measurement methods. Both the measured lifting force and ergonomists' estimated methods were able to show significant differences between the average lifting forces of the different job categories ($p<0.001$, Fig. 1a), with the highest average lifting force in *very heavy work* (V) job category, and least lifting force in *sedentary work* (S) job category. The post hoc tests showed both methods had similar abilities to detect differences between different jobs, with slightly higher *F* value for the measured force method (Table 2).

Both the measured pushing/pulling force and the ergonomists' estimated methods were able to detect significant differences between the average pushing/pulling forces of the different job categories ($p<0.001$, Fig. 1b), with job category V and H having significantly higher pushing/pulling force than the job category L. Again the measured method had slightly higher *F* value than the ergonomist estimated method (Table 2). The post hoc tests showed significant differences between jobs in the H and M categories, and between H and L categories, using the direct push/pull measurement method (Table 2). However, ergonomists' estimated push/pull forces were different between jobs in the V and L categories, and between the H and L categories (Table 2).

Force-matched pinch grip forces of job category H were significantly higher than those of job categories M, L and S ($p<0.001$, Fig. 1c, Table 2) after excluding job category V in the analysis due to having only one subject in that category. Similarly significant pinch forces differences were found between H and M job categories, and between H and L categories (Fig. 1c and Table 2). No statistically significant differences in pinch grip forces were found between any job categories ($p>0.01$) when they were self-reported by the subjects (Fig. 1c and Table 2).

There were no statistically significant differences between average power grip forces using any of the three different measurement methods ($p>0.01$, Fig. 1d) after exclusion of the only subject in job category V.

Pearson correlation coefficients between the different force quantification methods varied between 0.28 and 0.71, depending on the type of force (Table 3). Good correlations (0.71 and 0.65) were found between directly measured and ergonomists' estimated lifting and pushing/pulling forces (Table 3). Between-method correlations of pinch and power grip forces (0.28–0.45) were fair but not as good as that of lifting and pushing/pulling forces (Table 3).

4. Discussion

In general, determination of forceful exertions included in a job evaluation is a matter of subjective judgment. Although there was a conceptual definition of force levels as inclusion criteria in the present study, the force was still not known until measured. Results showed that the ergonomists were generally on the conservative side and measured many forceful exertions that were lower than the respective minimal inclusion levels. However, the results also

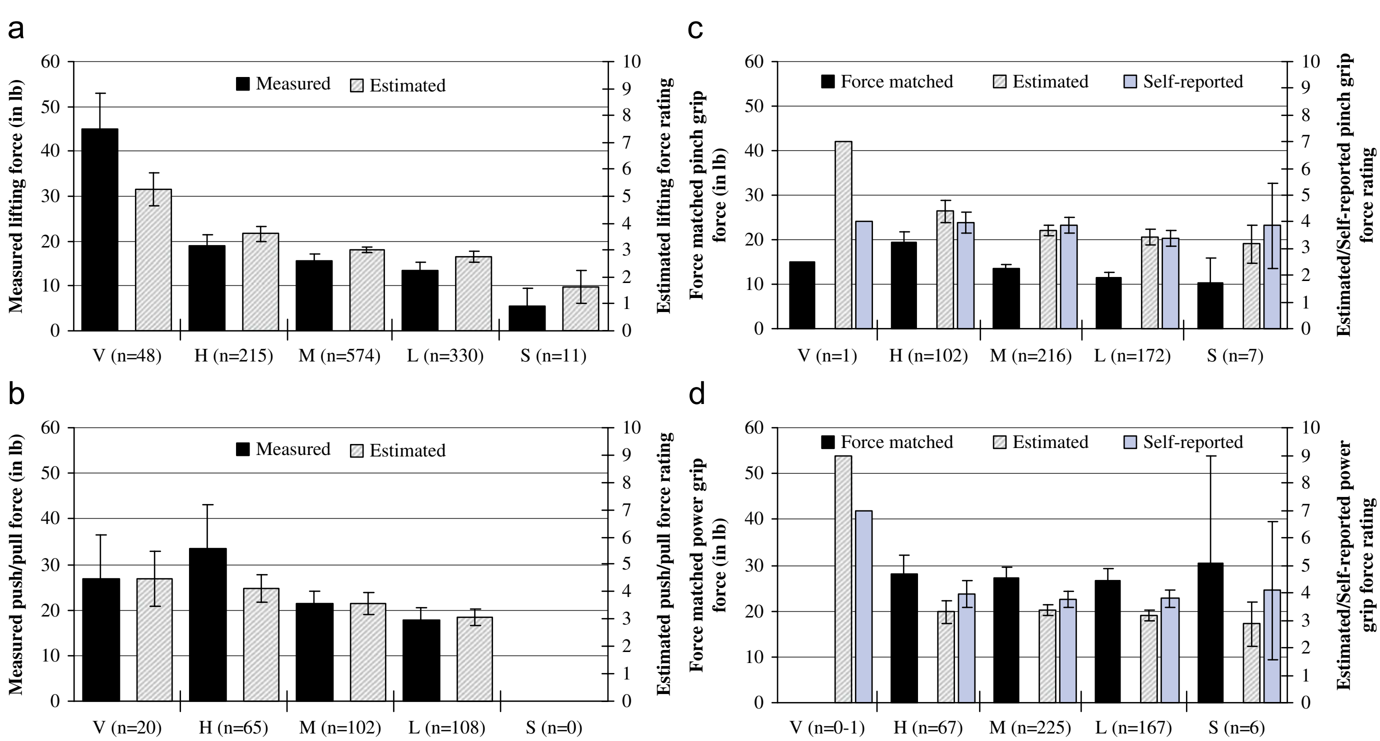

Fig. 1. Average (and 95% confidence interval) forceful exertion levels quantified by different methods by different job categories, *n*—number of subjects, V—very heavy work, H—heavy work, M—medium work, L—light work, and S—sedentary work.

Table 2
Detected differences between paired job categories based on Scheffe's multiple-comparison procedure when different measurement methods were used.

a. Lifting force by two methods

Measured lifting force (DF=1177, *F*=38.59, *P*<0.0001)

Job	V	H	M	L	S
V (48)		***	***	***	***
H (215)				***	
M (574)					
L (330)					
S (11)					

Ergonomist estimated lifting force (DF=1181, *F*=27.92, *P*<0.0001)

Job	V	H	M	L	S
V (49)		***	***	***	***
H (215)			***	***	***
M (576)					
L (331)					
S (11)					

b. *Push/pull force by two methods*

Measured push/pull force (DF=294, *F*=7.71, *P*<0.0001)

Job	V	H	M	L	S
V (20)					
H (65)			***	***	
M (102)					
L (108)					
S (0)					

Ergonomist estimated push/pull force (DF=319, *F*=6.92, *P*=0.0002)

Job	V	H	M	L	S
V (21)				***	
H (69)				***	
M (112)					
L (118)					
S (0)					

c. *Pinch grip force by three methods*

Force-matched pinch grip force (DF=497, *F*=13.76, *P*<0.0001)

Job	V	H	M	L	S
V (1)					
H (102)			***	***	
M (216)					
L (172)					
S (7)					

Ergonomist estimated pinch grip force (DF=505, *F*=8.35, *P*<0.0001)

Job	V	H	M	L	S
V (1)					
H (103)			***	***	
M (217)					
L (178)					
S (7)					

Worker self-reported pinch grip force (DF=502, *F*=2.00, *P*=0.0937)

Job	V	H	M	L	S
V (1)					
H (103)					
M (216)					
L (176)					
S (7)					

d. *Power grip force by three methods*

Force-matched power grip force (DF=465, *F*=0.19, *P*=0.9009)

Job	V	H	M	L	S
V (0)					
H (67)					
M (225)					
L (168)					
S (6)					

Ergonomist estimated power grip force (DF=472, *F*=4.41, *P*=0.0017)

Job	V	H	M	L	S
V (1)		***	***	***	***
H (70)					
M (227)					
L (168)					
S (7)					

Worker self-reported power grip force (DF=465, *F*=0.71, *P*=0.5821)

Job	V	H	M	L	S
V (1)					
H (67)					
M (225)					
L (167)					
S (6)					

*** Significant difference at $p<0.05$, blank: no significant difference, number in the parentheses: number of subjects in the job category for the force measurement method, DF: total degree of freedom, *F*: *F* value, and *P*: probability >*F*).

showed that this determination varied between individual ergonomists. Some may have been more conservative than the others. For example, Ergonomist A considered more exertions as forceful exertions and took corresponding measurements (Table 1). It is unknown what the ultimate difference in assessment of physiological impact is when including every single forceful exertion and how low the minimal inclusion levels should be. This could be an important issue for ergonomics practitioners in their job evaluation practices and may need to be further studied to set more objective guidelines. Since the present study only had three ergonomists, the finding of judgment differences between ergonomists was empirical. Further systematic study on factors affecting the determination of forceful exertions may be needed in future studies in order to help developing more objective guidelines in forceful exertion measurement.

Table 3
Pearson correlation coefficients between the different measurement methods (number of subjects in parentheses, $p<0.0001$ for all analyses).

	Lifting	Pushing/ pulling	Pinch gripping	Power gripping
Measured/force matched vs. estimated	0.71 (1177)	0.65 (295)	0.38 (498)	0.37 (466)
Force matched vs. self-reported	N/A	N/A	0.28 (495)	0.36 (463)
Estimated vs. self-reported	N/A	N/A	0.40 (503)	0.45 (466)

Different quantification methods of forceful exertions seemed to be able to detect differences related to forceful exertion levels as shown in Fig. 1 and Table 2. This was better with lifting and pushing/pulling forces, showing that both measurements with force gauges and estimates by ergonomists were able to detect the differences of exertion levels between different job categories, with some slight variations between the methods possibly due to random errors (Fig. 1a and b, and Table 2). The force-matching and ergonomists' estimated methods used for the pinch force estimation seemed to be more sensitive than the subjects' self-reported methods, as they were able to better detect differences between job categories. With regards to the methods used for the power grip force estimation, none were able to find any significant differences between job categories. It should be noted that the present job categorization was mainly based on manual material handling activities and used simply as a way to create a contrast in physical exertions among the jobs. It is not clear (and it may well be possible) if the lack of significant differences was because there were no actual differences in power grip forces between the job categories or because the methods used were not able to detect the differences between power grip forces of the different job categories. Since the present study was a field-based research, variations of forceful exertion of various types were not controlled by researchers. A well-designed laboratory study may provide further detailed information about the different forceful exertion quantification methods.

The above findings seem to lend some support to the idea that one might be able to use different methods to estimate forceful exertions in ergonomics job evaluations, as suggested by some of the commonly used assessment methods such as the SI (Moore and Garg, 1995) and ACGIH HAL TLV (ACGIH, 2001). This may hold as long as they are sensitive enough to detect exposure differences between jobs. However, the results also showed that the correlations between the measurements using different methods were not strong. This is especially true for pinch and power grip hand forces, where the Pearson correlation coefficients were only in the 0.28–0.45 range for the pinch and power grip forces (Table 3). The results of the Pearson correlation analysis of the present study confirmed the findings of our previous study (Bao et al., 2006), which was based on a much smaller dataset. The lack of strong correlations may suggest that, although different forceful exertion estimation methods may be used in job evaluations, they may not be interchangeable. Other factors such as work technique, hand/wrist postures and worker's individual characteristics may have influenced the surrogate measures of the forceful exertions. Different weights and/or cut-points may need to be used in job evaluations when different force quantification methods are used. Different methods have different characteristics of validity, reliability and sensitivity, and may estimate different aspects of forceful exertions. It is necessary to have a discussion on the most commonly used force measurement methods with relation to the findings of the present study. General discussions of validity, reliability and sensitivity of different exposure measurement methods have also been discussed by Winkel and Mathiassen (1994).

Although direct measurement of handgrip forces with instrumentation has been considered the most reliable and valid gold standard among hand force measurement methods (Spielholz et al., 2001). Studies have shown that the between-subject reliability of direct hand force measurement method, though not poor, was not perfect. Reliability coefficient of 0.77 was reported by Koppelaar and Wells (2005). This may be due to the actual variation of performing the same task with different work techniques between different operators, resulting in different actual forces application. It is also very difficult for most practitioners to instrument this type of measurement in real workplace settings in order to conduct direct handgrip force measurements. In addition to any reliability and versatility issues associated with direct measurement, the validity of the measurement is also dependent on how the handles and tools are instrumented. Different instrumentation may measure different aspects of a griping activity. Direct measurement of lifting and pushing/pulling forces is relatively easy. This can be done by measuring the weight of objects with a weight scale or a force gauge as used in the present study, as long as conditions permit performing the measurements. Of course, the accuracy is relative, as many measurements are performed in relatively slow speed so the dynamic force component of lifting or pushing/pulling is not as readily considered.

The measurement of muscle activities using electromyography (EMG) may be considered a more relevant measure of hand/arm exposure, as tasks with hand force may also involve other exposure modifiers such as awkward postures and repetitive motions. However, this is not a direct measure of the applied hand force, but an indirect estimate of the hand force. Correlations between EMG measurements from the forearm muscles and directly measured hand force have been found to be moderate (Koppelaar and Wells, 2005). It has been well-known that the variability of EMG measurement can be large between different subjects (Mathiassen et al., 1995), partially due to the variability inherent in the present technology (e.g. electrode placement) and partially due to the use of different muscles between individuals for the same tasks (e.g. one may use different work techniques in performing a task, assume different hand/wrist postures, or hold a tool in different ways). Koppelaar and Wells (2005) found that the between-participant reliability coefficients of EMG measurements ranged between 0.44 and 0.81, depending on the forearm and hand muscles. EMG measurement from the hand and forearm usually requires the measurement of multiple muscles. It is difficult to interpret the EMG levels from multiple muscle groups. Practically, it is also both time consuming and equipment intensive to measure EMG in most worksites. This technique is difficult to be used in large-scale epidemiological studies such as the present study. This technique is also unlikely to be widely used by ergonomics practitioners in job evaluations given the cost and complexity of the currently available technologies.

Estimation through observation by trained ergonomists or analysts is a commonly used method to quantify forceful exertions among ergonomics practitioners and researchers. This can be done either on-site in a workplace (Armstrong et al., 1982; Baty et al., 1986), as we used in the present study, or afterwards from video recordings of jobs (Spielholz et al., 2001). The latter allows for a more detailed and reproducible evaluation, but makes it difficult to estimate events when the video camera was in a poor position or when the element was not captured. In contrast, the analyst who is directly observing might be able to adjust his/her position to get the best view of an exertion, and obtain as much information as possible to help estimate the forceful exertion. However, the analyst may not have enough time to deal with large amount of information generated when many different activities occurring simultaneously at a fast pace. When observers are required to estimate forceful exertions in a task, they may sometimes have details about the task, such as weight of an object handled and/or the amount of forces required to perform the task, or they may know nothing about the task except what they see from the video. Koppelaar and Wells (2005) found that observation with information about the task could significantly improve the reliability of the estimation. Observational methods are relatively easy to perform and generally less intrusive to the workers than other methods. Our results showed that observational methods could quite well detect differences of lifting and pushing/pulling forces between different job categories, although the direct measurement method might be slightly more sensitive as shown with higher F values (Table 2). The observational methods also seemed to be more sensitive than the self-reporting technique in detecting differences of pinch grip force levels between the different job categories (Fig. 1c and Table 2).

Estimation of forceful exertion through workers' self-reports is based on the perception of exertion effort by the workers (Wiktorin et al., 1996). The Borg scale was used in the present study, although other types of visual analogue scales (VAS) could have been used. This method is probably the most cost-effective and feasible way to obtain force data in large population studies (Hjelm et al., 1995). Spielholz et al. (2001) concluded, however, that self-reports are a very imprecise exposure assessment method when assessing hand force and argued that they should not be used solely for any scientific assessment of exposure to physical risk factors such as hand force. Results of the present study on the pinch grip force estimation (Fig. 1c) may also suggest that this method might be the least sensitive one among the three used in the present study to detect pinch grip force differences between jobs.

The force-matching method used for pinch and power grip force estimation seems to be welcomed by many ergonomics consultants and practitioners. This method is basically a psychophysical approach, which relies on the memory of subjects when applying a hand force during a task and replicating the same amount of force on a dynamometer. Therefore, this method is actually a subjective indirect hand force measurement method. This method has been assumed to be appropriate for hand force estimation and is discussed in various textbooks and ergonomic manuals (Bohr, 1998; O'Callaghan and Switzer-McIntyre, 1995; Selan, 1994). Recently, this method has been studied by several researchers (Bao and Silverstein, 2005; Casey et al., 2002; King and Finet, 2004; Koppelaar and Wells, 2005; McGorry et al., 2004). In general, it has been shown that this method has some level of inaccuracy. King and Finet (2004) found significant differences between actual and perceived grip force performance using three different levels of forces (10, 30 and 50 lb). However, the highest average differences were just 1.9, 2.4 and 4.5 lb for the three force levels, respectively. In a previous study by Bao and Silverstein (2005), it was found that the estimated hand forces of both pinch and power grips only slightly overestimated the actual forces (9.7 vs. 8.9 N for pinch grip and 44.6 vs. 44.3 N for power grip). Casey et al. (2002) and McGorry et al. (2004) noticed that the accuracy of force estimation varied greatly from individual to individual. McGorry et al. (2004) found greater discrepancy between the actual and estimated grip forces at the lower force levels (80.8% at 44.5 N force level) and less discrepancy at higher force levels (18.9% on average for force levels of 89, 133.5, and 178 N). This finding is important as the accuracy of estimation at low force levels may have less consequence in an ergonomic evaluation, as it represents a relatively low exposure. Koppelaar and Wells (2005) found a correlation coefficient of 0.58 between hand force estimate using force-matching and direct measurement. They concluded that the direct measurement and force-matching methods seem to cluster as 'gripping force' and the EMG, force rating from observation, and self-reporting methods seem to cluster as the 'overall' response to the task. This is similar to what was found in the present study. The different measurement methods may have quantified different aspects of forceful exertions (i.e. feeling of exertion on the localized areas of the hand vs. overall feeling of the exertion). Therefore, the correlations between results obtained by the different methods were not strong (Table 3).

It has also been reported that the reliability of the force-matching method is not very satisfactory. Large between-subject variations have been reported using force-matching in the previous studies (Bao and Silverstein, 2005; Casey et al., 2002; King and Finet, 2004; Koppelaar and Wells, 2005; McGorry et al., 2004). Bao and Silverstein (2005) found large between-subject coefficients of variations of up to 72% (pressing staple remover using pinch grips). Although, this may partially be attributable to the differences in force that subjects might actually apply when performing the same tasks, and using different muscles due to variations in work techniques and hand/wrist postures, the differences in force estimation might still be large. The large between-subject variation may be improved to some extent by using proper and consistent estimation procedures and instructions. In many situations, subjects may receive only simple instructions for the force-matching estimation such as "grip the handle with the same force you just used while you performed the task" (Casey et al., 2002). Bao and Silverstein (2005) found that with improved instructions the between-subject coefficients of variation were greatly reduced. Although, the reliability of the force-matching method was not satisfactory (between-subject coefficient of variance or CV=0.29), Koppelaar and Wells (2005) reported that force-matching method is not worse than EMG, worker self-report, and estimates from observation (between subject CV=0.33–0.66), and it was just slightly worse than the direct measurement method using force gauges (between subject CV=0.26). Results from the present study may also suggest that the force-matching method may be more sensitive than at least the workers' self-reports, as the force-matching method was better able to detect differences between the different job categories, and slightly better than the observational method as shown having a higher F value (Fig. 1c and Table 2).

In conclusion, measurement of forceful exertion in job evaluations is both important and challenging. In practice, the determination of whether a forceful exertion is important for inclusion in the measurement is very subjective. More objective inclusion criteria may need to be developed. Different methods are available to quantify the exertions in order to distinguish different exertion levels between jobs, though with varied sensitivity. Direct measurement of lifting and pushing/pulling forces seems to be slightly more sensitive than ergonomists' estimates, and the force matching of grip forces may be more sensitive than workers' self-reports. Ergonomists estimations through observation seem to be a good alternative for measuring forceful exertions. Lifting and pushing/pulling forces are easier to measure compared with pinch and power grip forces. Correlations between results of forceful exertions measured with different methods were not strong. This indicates that different methods may measure different aspects of forceful exertions. It can be surmised that different force measurement methods may not necessarily be used interchangeably in job evaluation tools. Different weights and/or cut-points may need to be developed when different methods are used to quantify the force component of a job evaluation and exposure assessment.

Acknowledgments

We acknowledge the important contributions of Ruby Irving, Benjamin Hamilton, Cindy Orr, Jessica Keller and Larry Taing in assisting with the data collection in the field, and of Caroline Smith in coordinating the fieldwork. This research was funded in part by the US National Institute for Occupational Safety and Health (OH1007316) and the Washington State Department of Labor and Industries.

References

ACGIH, 2001. Hand Activity Level. TLVs and BEIs—Threshold Limit Values for Chemical Substances and Physical Agents. ACGIH, Cincinnati, OH, p. 110–112.

Armstrong, T.J., Foulke, J.A., Joseph, B.S., Goldstein, S.A., 1982. Investigation of cumulative trauma disorders in a poultry processing plant. American Industrial Hygiene Association, Journal 43, 103–116.

Bao, S., Silverstein, B., 2005. Estimation of hand force in ergonomic job evaluations. Ergonomics 48, 288–301.

Bao, S., Howard, N., Spielholz, P., Silverstein, B., 2006a. Quantifying repetitive hand activity for epidemiological research on musculoskeletal disorders—part II: comparison of different methods of measuring force level and repetitiveness. Ergonomics 49, 381–392.

Bao, S., Silverstein, B., Howard, N., Spielholz, P., 2006b. The Washington state SHARP approach to exposure assessment. In: Marras, W.S., Karwowski, W. (Eds.), Fundamentals and Assessment Tools for Occupational Ergonomics. Taylor & Francis Group, Boca Raton, FL, USA, pp. 44-1–44-22.

Bao, S., Spielholz, P., Howard, N., Silverstein, B., 2006c. Quantifying repetitive hand activity for epidemiological research on musculoskeletal disorders—part I: individual exposure assessment. Ergonomics 49, 361–380.

Baty, D., Buckle, P.N., Stubbs, D.A., 1986. Posture recording by direct observation, questionnaire assessment and instrumentation: a comparison based on a recent field study. In: Corlett, N., Wilson, J. (Eds.), The Ergonomics of Working Postures. Taylor & Francis, London, p. 293.

Bohr, P., 1998. Work analysis. In: King, P. (Ed.), Sourcebook of Occupational Rehabilitation. Plenum Press, New York, pp. 229–246.

Borg, G., 1990. Psychophysical scaling with applications in physical work and the perception of exertion. Scandinavian Journal of Work, Environment, and Health 16 (Suppl. 1), 55–58.

Burdorf, A., van der Beek, A., 1999. Exposure assessment strategies for work-related risk factors for musculoskeletal disorders. Scandinavian Journal of Work, Environment, and Health 25, 25–30.

Casey, J.S., McGorry, R.W., Dempsey, P.G., 2002. Getting a grip on grip force estimates: a valuable tool for ergonomic evaluations. Professional Safety, 18–24.

Cook, T., Rosecrance, J., Zimmermann, C., Gerleman, D., Ludewig, P., 1998. Electromyographic analysis of a repetitive hand-gripping task. International Journal of Occupational Health & Safety Ergon 4, 185–200.

Fan, J., Silverstein, B., Viikari-Juntura, E., Bonauto, D., Bao, S., Howard, N., Spielholz, P., Smith, C., 2007. Work-related lateral epicondylitis: quantitative exposure–response relations with force and posture. In: Proceedings of the Sixth International Scientific Conference on Prevention of Work-related Musculoskeletal Disorders, Boston, MA, USA, p. 295.

Hjelm, E.W., Winkel, J., Nygard, C.H., Wiktorin, C., Karlqvist, L., 1995. Can cardiovascular load in ergonomic epidemiology be estimated by self-report? Stockholm MUSIC 1 study group. Journal of Occupational and Environmental Medicine 37, 1210–1217.

Hoozemans, M., van der Beek, A., Frings-Dresen, M., van Dijk, F., van der Woude, L., 1998. Pushing and pulling in relation to musculoskeletal disorders: a review of risk factors. Ergonomics 41, 757–781.

King, P.M., Finet, M., 2004. Determining the accuracy of the psychophysical approach to grip force measurement. Journal of Hand Therapy 17, 412–416.

Koppelaar, E., Wells, R., 2005. Comparison of measurement methods for quantifying hand force. Ergonomics 48, 983–1007.

Latko, W.A., 1997. Development and evaluation of an observational method for quantifying exposure to hand activity and other physical stressors in manual work. Ph.D. Dissertation, The University of Michigan, Ann Arbor.

Mathiassen, S.E., Winkel, J., Hagg, G.M., 1995. Normalization of surface EMG amplitude from the upper trapezius muscle in ergonomic studies—a review. Journal of Electromyography and Kinesiology 5, 197–226.

McGorry, R.W., Dempsey, P.G., Casey, J.S., 2004. The effect of force distribution and magnitude at the hand–tool interface on the accuracy of grip force estimates. Journal of Occupational Rehabilitation 14, 255–266.

Moore, J.S., Garg, A., 1995. The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. American Industrial Hygiene Association, Journal 56, 443–458.

National Institute for Occupational Safety and Health (NIOSH), 1997. Musculoskeletal disorders and workplace factors: a critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back. DHHS (NIOSH).

O'Callaghan, J., Switzer-McIntyre, S., 1995. Job analysis. In: Key, G.L. (Ed.), Industrial Therapy. Mosby, St. Louis, MO, pp. 85–109.

Selan, J.L., 1994. The Advanced Ergonomics Manual. Advanced Ergonomics, Inc., Dallas, TX.

Silverstein, B., Fan, J., Bonauto, D., Bao, S., Howard, N., Spielholz, P., Smith, C., Viikari-Juntura, E., 2007. Work related rotator cuff and carpal tunnel syndromes: quantitative exposure–response relationships with force and posture. In: Proceedings of the Sixth International Scientific Conference on Prevention of Work-related Musculoskeletal Disorders, Boston, MA, USA, p. 305.

Spielholz, P., Silverstein, B., Morgan, M., Checkoway, H., Kaufman, J., 2001. Comparison of self-report, video observation and direct measurement methods

for upper extremity musculoskeletal disorder physical risk factors. Ergonomics 44, 588–613.
Stetson, D.S., Keyserling, W.M., Silverstein, B.A., Leonard, J.A., 1991. Observational analysis of the hand and wrist: a pilot study. Application Occupational Environment Hygiene 6, 927–937.
US Department of Labor., 1993. Selected characteristics of occupations defined in the revised dictionary of occupational titles. Employment and Training Administration, Washington, DC.
van der Beek, A.J., Hoozemans, M.J.M., Frings-Dresen, M.H.W., Burdorf, A., 1999. Assessment of exposure to pushing and pulling in epidemiological field studies: an overview of methods, exposure measures, and measurement strategies. International Journal of Industrial Ergonomics 24, 417–429.
Washington State Department of Labor and Industries, 2000. General occupational health standards, Chapter 296-62 WAC, Part A-1, Ergonomics. Department of Labor and Industries, Olympia, WA, USA
Wiktorin, C., Selin, K., Ekenvall, L., Kilbom, A., Afredsson, L., 1996. Evaluation of perceived and self-reported manual forces exerted in occupational materials handling. Applied Ergonomics 27, 231–239.
Winkel, J., Mathiassen, S.E., 1994. Assessment of physical work load in epidemiologic studies: concepts, issues and operational considerations. Ergonomics 37, 979–988.