JCQ Scale Reliability and Responsiveness to Changes in Manufacturing Process

Angelo d'Errico, MD,[1,2]* Laura Punnett, ScD,[1] Judith E. Gold, ScD,[1,3] and Rebecca Gore, PhD[1]

Background *The job content questionnaire (JCQ) was administered to automobile manufacturing workers in two interviews, 5 years apart. Between the two interviews, the company introduced substantial changes in production technology in some production areas. The aims were: (1) to describe the impact of these changes on self-reported psychosocial exposures, and (2) to examine test–retest reliability of the JCQ scales, taking into account changes in job assignment and, for a subset of workers, physical ergonomic exposures as assessed through field observations.*
Methods *The study population included 790 subjects at the first and 519 at the second interview, of whom 387 were present in both. Differences in demand and control scores between interviews were analyzed by Wilcoxon matched-pairs signed-rank test. Test–retest reliability of these scales was evaluated by the intraclass correlation coefficient (ICC) and the Spearman's ρ coefficient.*
Results *The introduction of more automated technology produced an overall increase in job control but did not decrease psychological demand. The reliability of the control scale was low overall but increased to an acceptable level among workers who had not changed job. The demand scale had high reliability only among workers whose physical ergonomic exposures were similar on both survey occasions.*
Conclusions *These results show that 5-year test–retest reliability of self-reported psychosocial exposures is adequate among workers whose job assignment and ergonomic exposures have remained stable over time.* Am. J. Ind. Med. 51:138–147, 2008. *© 2007 Wiley-Liss, Inc.*

KEY WORDS: psychosocial factors; stress; JCQ; work; automobile manufacturing; epidemiology; reliability; ergonomics

INTRODUCTION

One of the most widely used instruments to assess self-reported exposure to psychological stressors in the workplace is the job content questionnaire (JCQ), designed by Karasek [1985]. The underlying theoretical construct is the demand-control model, which postulates that a work environment characterized by high psychological demand and low decision latitude, defined as a low level of both autonomy and use of skills, exposes workers to a condition of high strain. Many epidemiologic studies using this model found subjects in the high strain group to be at increased risk of cardiovascular diseases [Schnall et al., 1994; Kristensen, 1996; Hemingway and Marmot, 1999], mental disorders [Braun and Hollander, 1988; Niedhammer et al., 1998; Bourbonnais et al., 1999; Stansfeld et al., 1999; Mausner-Dorsch and Eaton, 2000], musculoskeletal symptoms [Ariens et al., 2001; Kerr et al., 2001; Kivimaki et al., 2001; Nahit et al., 2001; Torp et al., 2001], and low back injuries [Myers et al., 1999].

[1]Department of Work Environment, University of Massachusetts Lowell, Lowell, Massachusetts
[2]Epidemiology Unit, Piedmont Region, Grugliasco (TO), Italy
[3]Department of Public Health, Temple University, Philadelphia, Pennsylvania
Contract grant sponsor: U.S. National Institute for Occupational Safety and Health; Contract grant number: R01-OH03514.
*Correspondence to: Dr. Angelo d'Errico, Epidemiologic Unit, Piedmont Region, Via Sabaudia 164, Grugliasco (TO) 10095, Italy. E-mail: angelo.derrico@epi.piemonte.it

Accepted 11 October 2007
DOI 10.1002/ajim.20545. Published online in Wiley InterScience (www.interscience.wiley.com)

The assessment of work-related stress by means of self-reports has been criticized by Kasl [1978] and others [Kristensen, 1996] on the grounds that associations between health status and exposure observed in cross-sectional studies may be due to differential reporting of workplace stressors by health status. However, evidence from longitudinal studies has demonstrated the ability of the demand-control model to predict the occurrence of diverse adverse health effects, such as cardiovascular diseases [Belkic et al., 2004], psychological disorders [Van der Doef and Maes, 1999; de Lange et al., 2003] and upper extremity musculoskeletal symptoms [Bongers et al., 2002]. Objective methods for assessing exposure may seem preferable, as they are expected to provide clearer identification of the environmental conditions on which to conduct interventions [Frese and Zapf, 1988; Kasl, 1998]. Nevertheless, the objective assessment of stressors appears limited by the construct of stress itself, which involves a cognitive appraisal of the stressors that is not accessible to the assessor [Lazarus and Folkman, 1986], in addition to general considerations related to the representativeness of the observed time and work processes.

Accurate exposure assessment is essential to the correct estimation of the risk of disease associated with an exposure. Test–retest reliability of a self-reported exposure is a measure of accuracy in the assessment of an exposure, which is mainly related to its precision, although other sources of measurement error may contribute to reducing it; the consequence of a loss in precision is non-differential exposure misclassification, which is expected to decrease the strength of the observed outcome-exposure association. According to the few published papers, the JCQ test–retest reliability appears to be high over the short-term [Escriba-Aguir et al., 2001], and moderate in the medium [Williams et al., 2001; Sale and Kerr, 2002; Cheng et al., 2003] to long-term [Landsbergis et al., 2002]. These data are mainly from countries other than the US and for a limited number of economic sectors and job titles.

The initial purpose of this study was to evaluate the test–retest reliability of the JCQ in a cohort of U.S. automobile manufacturing workers who were administered the JCQ short version in two interviews conducted 5 years apart. During the period between interviews, however, the plants where these workers were employed underwent important changes in production technology and work organization. A decrease in work pace, physical effort and other self-reported ergonomic exposures was observed in this cohort during the period under consideration [d'Errico et al., 2007]. These changes in physical exposures are an indication of the nature and magnitude of the changes in the production process, which were expected to have had an effect on the stability of exposure to both ergonomic and psychosocial hazards. In fact, numerous organizational and technological features of the workplace are believed to determine both physical and psychosocial exposures at the individual level, although empirical evidence of these relationships is rather limited [Punnett et al., 2004b]. Therefore, the first aim of this study was to describe the impact of the changes in technology and work organization on self-reported psychosocial exposures in the workplace; secondly, to estimate the test–retest reliability of psychosocial exposures in subgroups of workers who had experienced different degrees of change in working conditions, as assessed through a questionnaire and ergonomic field observations.

MATERIALS AND METHODS

Data Collection

This study is part of a prospective epidemiologic study on work-related upper extremity musculoskeletal disorders among automobile manufacturing workers [Punnett, 1998; Punnett et al., 2004a; d'Errico et al., 2007]. In 1992–1993 (survey occasion "T_0"), 613 production workers in an engine assembly plant and 670 employed in a stamping plant were interviewed using a standardized questionnaire to collect baseline information, which included past work history, current ergonomic exposures at work, upper extremity symptoms, relevant medical conditions and demographic data.

One year later during 1993–1994, 790 workers still employed in these plants were re-interviewed [Punnett et al., 2004a]. On this occasion, "T_1", the interview also included the JCQ short version, consisting of five items on quantitative job demands and nine items on decision latitude (job control) [Karasek et al., 1998]. In 1998, the same information was collected again in a second follow-up interview, "T_2", on 519 workers of whom 387 had been interviewed 5 years earlier.

At T_1, workers' department assignments were obtained from personnel rosters and confirmed by supervisors. During 1993–1994 (overlapping T_0 and T_1), field observations of work postures, manual material handling and other job-related ergonomic features were recorded with a checklist, for all jobs except those for which no fixed work cycle could be identified, such as forklift drivers and tool room personnel [Punnett, 1998]. At T_2, department and job assignments were obtained from personnel rosters and observations were conducted again, this time including all jobs [Gold et al., 2006]. A few workers had been transferred to a new engine plant; they were interviewed but job observations were not feasible.

Based on the field observations, a score on a composite index of upper extremity ergonomic exposure was assigned to each worker. A series of dichotomous variables were constructed, as described below, from the observed ratings on posture (three items), contact stress, vibration (two items), force and repetition. The composite index score was obtained by summing the values on these ergonomic exposure variables, yielding a value between 0 and 24, which is similar to the psychophysical exposure index constructed from a set of self-reported ergonomic exposures [Punnett, 1998].

The presence of the following exposures was rated as a "3" and its absence as a "0": awkward upper arm/shoulder posture (flexion or extension), awkward forearm/elbow posture (pronation, supination, or elbow extension), awkward wrist posture (extension, flexion, or ulnar deviation); contact pressure to the hand; segmental vibration transmitted to the hand, and whole body vibration. Force exposure was rated as a "3" if the object held in a power grasp exceeded ten pounds, or two pounds in a pinch grasp; otherwise, it was rated as "0". The repetition rating at T_2 was determined from the frequency of hand/arm pauses, rated on a Latko scale [Latko et al., 1997]; whereas at T_1 a "3" was assigned if the worker was observed to perform 1,800 or more cycles per shift (30 s cycle), and less repetitive jobs were rated as "0". A worker's final repetition rating was the sum of the repetition ratings for each daily task weighted by its duration in a given working day. For observations with missing hand/arm pause ratings and for unobserved tasks, the average repetition rating for that particular task or department was used.

This study was approved by the University of Massachusetts at Lowell Institutional Review Board.

Statistical Analysis

Standard JCQ algorithms were used to compute the overall scores on the demand and control scales, and the internal consistency of each scale was evaluated using Cronbach's alpha coefficient. Among those workers interviewed twice, the stability of self-reported exposure to psychosocial factors between the first and the second follow-up interview was assessed using the Wilcoxon matched-pairs signed-rank test on the change scores ($T_2 - T_1$). Stability was examined in the whole cohort, by plant, and by changes in job title and department. The difference in mean change scores by job title change (yes, no) was evaluated using the Mann–Whitney test, and the presence of a trend in change scores across ordered categories of department change (none, minor or major) was assessed by a non-parametric test for trend [Cuzick, 1985].

Job histories were reviewed manually to determine whether each worker was in the same job at T_2. Three groups were defined: workers who were still in the same job ($n = 166$), those who had changed job ($n = 162$), and those about whom the investigators were uncertain ($n = 58$). Changes in plant, production department and exact job title were assessed independently of each other; however, changes in department and job title were significantly correlated (Spearman $\rho = 0.37$, $P < 0.001$), with about 70% of the workers reporting either that both department and job had changed, or neither. Differences in the proportion of workers who had changed job by plant and department were assessed using the Chi-square test.

Test–retest reliability of demand and control scales between interviews was evaluated by the intraclass correlation coefficient (ICC) and Spearman's correlation coefficient (ρ), treating the scores as continuous. ICC is a measure of agreement that has been shown to reproduce the weighted kappa coefficient perfectly when Fleiss–Cohen weights are used [Fleiss, 1981]. ICC was computed using random-effects one-way ANOVA, which assumes homogeneous variance within groups [Armstrong et al., 1992]. We also assessed agreement with Spearman ρ to evaluate possible distortions of ICC estimates. Agreement was evaluated separately for the two dimensions of job control, skill use, and decision authority, in order to identify potential sources of low reliability. Test–retest reliability was not computed using Cohen's kappa coefficient on the dichotomized scales, as scores had a normal-like distribution with high density at the middle values of the scales. This could have artificially reduced agreement, because of the increased likelihood that workers with similar scores measured on two different occasions would be assigned to different exposure categories. Test–retest reliability was analyzed after stratifying by job and department change status because agreement was expected to be higher among workers who had experienced no such changes. Workers who reported no job or department changes might have actually undergone changes in job content and/or tasks, owing to the technological changes introduced between surveys (described below). Therefore, agreement was also examined in the subgroup of workers ($n = 130$) who had been observed at both T_1 and T_2, permitting calculation of the ergonomic exposure index based on observations. We analyzed test–retest reliability separately for workers whose score was in the same versus different quartile in both interviews.

All analyses were performed using the statistical software Stata, version 8 (Stata Corporation, College Station, TX).

RESULTS

From 1993 to 1998, the overall work flow in both the stamping and engine operations remained organized around a moving production line while the production technology became more automated [Gold et al., 2006]. For example, manual part transfer between pairs of stamping presses was replaced by automated transfer. There was more job rotation at T_2, although no formal teamwork or manufacturing cells had been introduced.

In the overall cohort, 42% of workers changed job between interviews and 43% were in the same job at T_2 (Table I). The proportion of workers still in the same job at T_2

TABLE I. Change in Job Title Between T_1 and T_2, Stratified by Plant and Department at T_2; 387 U.S. Auto Manufacturing Workers, Detroit, 1993–1998

	Job change, no		Job change, yes		Job change, uncertain		
	n	%	n	%	n	%	Total, N
Overall—all workers	166	42.9	163	42.1	58	15.0	387
Stamping plant (SP)	111	49.8	79	35.4	33	14.8	223
Engine plant (EP)	54	37.0	68	46.6	24	16.4	146
Engine plant—new	1	5.6	16	88.9	1	5.6	18
Department groups							
Non-routine (EP)	14	17.0	41	50.0	27	32.9	82
Machining (EP)	40	40.4	50	50.6	9	9.1	99
Assembly/subassembly (EP)	12	44.4	13	48.2	2	7.4	27
Large press (SP)	46	79.3	9	15.6	3	5.2	58
Small press (SP)	8	53.3	3	20.0	4	26.7	15
Assembly (SP)	46	43.4	47	44.3	13	12.3	106

was markedly higher in the stamping than the engine plant ($P = 0.016$) and was also unevenly distributed by department ($P < 0.001$), with the least amount of change in the press departments of the stamping plant. More than 40% of workers changed production department between interviews (data not shown); this was similar in the two plants, but the highest proportion was among those found to be working in the non-routine department at T_2.

Internal consistency of the full decision latitude scale was quite high and similar in the two interviews ($\alpha = 0.76$ and 0.77 at T_1 and T_2, respectively); it was higher for the skill use subscale ($\alpha = 0.69$ and 0.70) than for the decision authority subscale ($\alpha = 0.47$ and 0.62), particularly at T_1. On both occasions, internal consistency was lower for the demand scale ($\alpha = 0.45$ and 0.53 at T_1 and T_2, respectively) than for job control.

Responsiveness of Psychosocial Exposures to Manufacturing Changes

The average scores on the control and demand scales among workers surveyed only at T_1 were not significantly different from those of participants in both surveys ($P = 0.20$ and $P = 0.25$, respectively; Wilcoxon rank-sum test), indicating that the study population was not an extreme subgroup with respect to the level of psychosocial exposures at T_1. Overall, the mean score on the psychological demand scale was relatively stable between the two interviews (Table II), as well as by plant and department (data not shown), except for an increase in the assembly department of the engine plant. The dispersion of the demand scores was smaller at T_2 compared to the previous interview 5 years earlier.

The average job demands value among workers who had not changed department was quite stable, whereas a slight decrease was found among those reporting minor or major changes ($P = 0.07$ for trend in the mean change score by department change status). Demand increased among workers still in the same job and decreased among those who had changed job. The difference in the mean change scores between these two groups was statistically significant ($P = 0.002$).

The mean value on the control scale increased significantly overall; the increase was larger in the engine plant (Table II). Job control increased in all departments (data not shown), except for a moderate reduction in the press departments of the stamping plant; the largest increase was in the non-routine department of the engine plant, where the change was significant ($P = 0.03$). Control increased slightly among workers who had not changed department; the larger increases observed among those who reported minor or major changes were statistically significant, as was the trend in the mean change score by department change status ($P < 0.01$). Among workers who had not changed job, the mean score on the control scale was practically unaltered, whereas a highly significant increase was found among workers who had. The mean change scores in these two groups were statistically different ($P = 0.017$). The mean scores on the individual scale items at T_1 and T_2 indicate a slight reduction in unfavorable working conditions on all items of control and on most of demand (Table III).

For both demand and control, workers in the lowest exposure quartile at T_1 reported the highest increase in demand and control at T_2, whereas those in the highest quartile reported a decrease, regardless of job change status (Table IV). For both scales, the mean change score was smaller for workers who were in the middle quartiles of exposure at T_1. The larger mean change scores at both the lower and higher levels of exposure at T_1 demonstrate the presence of regression toward the mean (RTM).

TABLE II. Mean and Standard Deviation of Psychological Demand and Job Control Scores Reported at T_1 and T_2 (Overall, by Plant, Change in Job Title and Change in Department)—U.S. Auto Manufacturing Workers

		T_1			T_2				
Exposure	**Group**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**	**Interviewed at T_1 and T_2, n**	**Change T_1 to T_2, *P*-value***
Demand	Overall—all workers	372	28.9	4.2	386	28.8	3.4	371	0.69
	Stamping plant	219	28.9	4.1	223	28.8	3.4	217	0.61
	Engine plant	153	28.9	4.3	145	28.8	3.5	132	1.00
	Engine plant—new	—	—	—	18	29.1	4.1		
	Job change—no	166	28.4	4.1	166	28.9	3.1	166	0.17
	Job change—yes	162	29.4	4.1	162	28.5	3.4	161	0.13
	Job change—uncertain	44	29.2	4.6	58	29.3	4.2	44	0.14
	Department change—none	207	28.8	4.2	216	29.0	3.8	207	0.30
	Department change—minor	100	28.6	3.8	101	28.2	3.3	99	0.66
	Department change –major	65	30.0	4.5	69	29.0	3.7	65	0.24
Control	Overall—all workers	370	60.8	9.4	385	62.7	9.4	368	**0.01**
	Stamping plant	217	60.9	8.5	222	61.9	9.0	214	0.14
	Engine plant	153	60.7	10.5	145	63.8	9.1	132	0.06
	Engine plant—new	—	—	—	18	63.7	14.6		
	Job change—no	166	62.3	9.4	165	62.1	9.5	165	0.67
	Job change—yes	160	59.2	9.4	162	62.7	9.4	159	**0.02**
	Job change—uncertain	44	61.0	8.0	58	64.5	8.8	44	0.08
	Department change—none	206	61.2	9.1	215	61.7	9.6	205	0.76
	Department change—minor	100	61.3	8.7	101	63.7	8.7	99	**0.04**
	Department change—major	64	58.6	10.9	69	64.4	9.4	64	**0.01**

*Computed by Wilcoxon matched-pairs signed-rank test.
Bold face values signify $P < 0.05$.

TABLE III. Mean and Standard Deviation of Single Items Scores Related to Demand and Control Reported at T_1 (372 Subject) and T_2 (386 Subjects)—U.S. Auto Manufacturing Workers

	T_1		T_2	
JCQ items	**Mean**	**s.d.**	**Mean**	**s.d.**
Control				
My job requires that I learn new things	2.59	0.67	2.74	0.61
My job involves a lot of repetitive work	3.19	0.58	2.95	0.52
My job requires me to be creative	2.47	0.69	2.53	0.64
My job allows me to make a lot of decisions on my own	2.66	0.70	2.74	0.67
My job requires a high level of skill	2.39	0.65	2.42	0.66
On my job, I have very little freedom to decide how I do my work	2.33	0.63	2.32	0.64
I get to do a variety of different things on my job	2.59	0.63	2.72	0.65
I have a lot of say about what happens on my job	2.65	0.66	2.70	0.66
I have an opportunity to develop my own special abilities	2.58	0.66	2.65	0.63
Demand				
My job requires working very fast	2.67	0.68	2.51	0.63
My job requires working very hard	2.53	0.62	2.49	0.62
I am not asked to do an excessive amount of work	2.68	0.61	2.62	0.58
I have enough time to get the job done	2.95	0.42	2.89	0.43
Some demands I face at work are in conflict with other demands at work	2.70	0.58	2.52	0.63

TABLE IV. Change in Psychological Demand and Job Control Scores Between T_1 and T_2 by Quartile of Demand and Control at T_1, and Change in Job Title at T_2—U.S. Auto Manufacturing Workers

	Job change, no		Job change, yes		Job change, uncertain	
Exposure quartile at T_1	n (%)	Change score, mean (SD)	n (%)	Change score, mean (SD)	n (%)	Change score, mean (SD)
Demand						
1 (low)	59 (54)	3.27 (3.4)	38 (35)	3.55 (4.3)	12 (11)	3.17 (4.0)
2	50 (49)	0.44 (2.4)	41 (40)	0.22 (2.5)	11 (11)	2.0 (4.47)
3	36 (37)	−1.11 (3.0)	50 (52)	−2.1 (3.3)	11 (11)	0.64 (2.4)
4 (high)	21 (33)	−4.24 (4.8)	32 (51)	−5.81 (4.6)	10 (16)	−6.1 (8.3)
Total	166 (45)	0.52 (4.1)	161 (43)	−0.91 (4.9)	44 (12)	0.14 (6.1)
Control						
1 (low)	36 (39)	5.22 (9.3)	46 (49)	12.91 (13.8)	11 (12)	13.1 (12.9)
2	46 (38)	0.13 (7.4)	60 (49)	2.80 (9.4)	16 (13)	2.5 (6.8)
3	44 (50)	−1.05 (7.4)	33 (38)	−2.72 (7.9)	11 (13)	3.09 (5.1)
4 (high)	39 (60)	−4.46 (7.8)	20 (31)	−4.90 (4.2)	6 (9)	−5.0 (6.9)
Total	165 (45)	−0.16 (8.5)	159 (43)	3.61 (12.0)	44 (12)	4.27 (10.0)

Test–Retest Reliability of Psychosocial Exposures

Test–retest reliability of the demand scale, as assessed by the stability of scores at the individual level between T_1 and T_2, was low overall (Table V). As expected, scores were more stable among workers who were in the same job in both surveys, compared to those who had changed job. Similarly, the scores among workers who reported minor or major department changes between surveys were less stable than those among workers who had no department change. Stability of the demand scale was, however, low in absolute value even among workers who had not changed job or department.

Reproducibility of the control scale was also low, although higher than the demand scale (Table VI); similar values of ICC and Spearman's ρ were found for the control scale and its two subscales. Stratification by job change status revealed that the scores among workers who had changed job between interviews were less stable than among those who had not, for whom reproducibility of control was good. Similarly, workers who had experienced minor or major department changes had less stable scores than those who reported no changes.

In the subgroup of workers for whom an ergonomic exposure index score was assigned in both surveys (Table VII), moderate/good demand stability was found among workers who were in the same quartile of exposure in both surveys ($n = 40$), whereas it was very low among those who changed quartile ($n = 90$). Unexpectedly, control stability was lower among workers whose ergonomic exposure score was in the same quartile in both surveys compared to those whose score was not. Only half of the

TABLE V. Test–Retest Reliability of JCQ Psychological Demand Scale Between T_1 and T_2—U.S. Auto Manufacturing Workers

Group	ICC	ICC LCL 95%	ICC UCL 95%	Spearman's ρ coefficient
Overall—all workers	0.23	0.13	0.33	0.30
Job change—no	0.36	0.23	0.50	0.38
Job change—yes	0.16	0.01	0.31	0.20
Job change—uncertain	0.04	0.00	0.35	0.33
Department change—none	0.34	0.22	0.47	0.39
Department change—minor	0.02	0.00	0.22	0.16
Department change—major	0.13	0.00	0.37	0.20

TABLE VI. Test–Retest Reliability of JCQ Job Control Scale Between T_1 and T_2—U.S. Auto Manufacturing Workers

Group	ICC	ICC LCL 95%	ICC UCL 95%	Spearman's ρ coefficient
Overall—all workers	0.36	0.27	0.44	0.40
Skill use subscale—all workers	0.29	0.19	0.38	0.38
Decision authority subscale—all workers	0.35	0.26	0.44	0.35
Job change—no	0.60	0.50	0.70	0.59
Job change—yes	0.14	0.00	0.29	0.23
Job change—uncertain	0.22	0.00	0.51	0.32
Department change—none	0.50	0.40	0.60	0.51
Department change—minor	0.40	0.23	0.56	0.40
Department change—major	0.00	0.00	0.25	0.15

TABLE VII. Test–Retest Reliability of Psychological Demand and Job Control Scales by Change in Quartile on the Ergonomic Exposure Index Between T_1 and T_2—U.S. Auto Manufacturing Workers

Group	ICC	ICC LCL 95%	ICC UCL 95%	Spearman's ρ coefficient
Demand—all workers (n = 130)	0.26	0.10	0.42	0.38
Demand—change in quartile (n = 90)	0.15	0.00	0.35	0.26
Demand—no change in quartile (n = 40)	0.64	0.45	0.82	0.70
Control—all workers (n = 130)	0.46	0.32	0.59	0.43
Control—change in quartile (n = 90)	0.48	0.32	0.64	0.45
Control—no change in quartile (n = 40)	0.35	0.08	0.62	0.34

workers in the same quartile on the exposure index, however, were still in the same job at T_2; for these workers (n = 20), reliability of the control scale was closer to that observed in the whole cohort with no job change (ICC = 0.48, 95% CI: 0.14–0.82; Spearman's ρ = 0.59).

DISCUSSION

In these automotive manufacturing operations, the introduction of new production technology resulted in substantial changes in job and department assignments during the 5-year period between interviews, especially for workers who were in the engine plant at T_2. The technological and organizational changes did not substantially modify psychological demand in the whole cohort or within plants or departments, but overall the jobs became less routinized [Gold et al., 2006] and there was an increase in perceived job control. The decrease in demand and increase in control scores from T_1 to T_2 were greater among workers changing job title or department than among those who did not. As these workers were more likely to have been affected by the organizational and technological changes introduced between the surveys, it seems probable that the small observed decrease in exposure to psychosocial stressors was caused by these changes.

In this data set, internal consistency of the job control scale was high and in close agreement with the literature [Brisson et al., 1998; Karasek et al., 1998; Pelfrene et al., 2001; Niedhammer, 2002; Sale and Kerr, 2002]. Internal consistency of the job demand axis in both surveys was lower than that observed in six major international studies, in which Cronbach's alpha values ranged from 0.51 to 0.72 on the five-item quantitative demand scale (mean alpha = 0.63) [Karasek et al., 1998]. However, lower consistency has been reported by others [MacDonald et al., 2001; Williams et al., 2001; Cheng et al., 2003; Li et al., 2004]. The low internal reliability demonstrated by the demands scale does not automatically imply low validity of the exposure assessment, but it likely reflects the number of the exposure dimensions that this scale is intended to capture. The trade-off between content validity (measuring all relevant types of job demands) and internal reliability cannot be optimized on the basis of these data; we merely note that consistency of the scale is about the same in this occupational cohort as in others, strengthening the generalizability of our findings.

The present study is the first to assess test–retest reliability of the JCQ short version over a period of more than 1 year. Its major strengths are the availability of detailed information on job assignments at the time of each survey and a measure of ergonomic exposure independent of worker self-assessment, albeit limited to a subset of the cohort. The evaluation of the reliability of self-reported psychosocial exposures among workers with relatively constant job content and physical demands was of particular value, as levels of work-related physical and psychosocial exposures tend to be correlated in different working populations [Johansson and Nonas, 1994; Ingelgard et al., 1996; MacDonald et al., 2001].

Five-year test–retest reliability of both the demand and control scales was low overall. However, the reliability of the decision latitude scale increased to moderate/good among workers in the same job at the time of both surveys. More surprisingly, it was low among workers who were in the same quartile on the (physical) ergonomic exposure index. Although the latter analysis was based on few subjects, this finding would indicate that the control dimension is more closely related to the overall work organization, in this case apparently represented by job title, rather than to the level of physical demand. Therefore, we can conclude that test–retest reliability of job control was good over the long-term among workers with the same job assignment over time, regardless of the level of physical demand.

On the other hand, intra-individual stability of quantitative demand scores was low among workers in the same job, whereas it increased to moderate/good among workers who were in the same quartile on the ergonomic exposure index in both surveys. This would suggest that the level of "psychological" demands is more closely associated with physical demands than with overall work organization, at least in these two factories. It appears that when physical demands are stable, the reliability of self-reported ‘‘psychological’’ demands is satisfactory, even over a longer time interval.

The similarity of the mean scores on the demand and control scales at T_1 for the study population and the workers surveyed only at T_1 demonstrates that the study population was not a subgroup of workers with extreme values of the exposures at T_1. Therefore, the effect of RTM on the mean changes in exposure between the surveys will likely have been minimal [Bland and Altman, 1994a,b]. On the other hand, the strong RTM effect observed when workers were divided into quartiles with respects to their scores on both the demand and control scales at T_1, suggests that the true scores of workers with extreme values at T_1 were actually closer to the population mean. This RTM effect demonstrates the presence of large within-subject variation and low precision in the assessment of exposure [Barnett et al., 2005], which is likely the most important source of lower test–retest reliability in this population. Multiple measurements at baseline would increase the precision in the assessment of exposures to workplace psychological stressors, as would be true for quantitative measurements of other occupational risk factors.

In general, test–retest reliability of the JCQ over 1 month or less has been moderate or good. Over the short-term, both the demand and control scales were highly reliable among Spanish hospital nurses [Escriba-Aguir et al., 2001]. Similar results were obtained by Salerno et al. [2001], who evaluated the reproducibility of the JCQ control dimension among American keyboard operators.

Over longer retest periods, reproducibility of the JCQ scales is less consistent. Low test–retest reliability of the high strain scale ($r = 0.33$) after 11 weeks was reported among sewing machine operators in a Canadian study [Williams et al., 2001]. In contrast, good reliability over a 3-month interval was observed for both the demand ($r = 0.62$) and control scales ($r = 0.73$ for skill discretion, $r = 0.64$ for decision authority) in a population-based study of Taiwanese workers [Cheng et al., 2003]. An epidemiologic study of hospital workers in Ontario [Sale and Kerr, 2002] found good 1-year agreement for decision latitude ($ICC = 0.68$) and moderate for psychological demand ($ICC = 0.57$). The only evidence on the reproducibility of the JCQ over a longer time interval comes from a study in the US [Landsbergis et al., 2002] that reported moderate/good 3-year test–retest reliability ($r = 0.64$) for both the demand and control scales among healthy full-time employees resident in New York City.

Reproducibility of exposure to psychosocial factors has also been evaluated for questionnaires other than the JCQ, in most cases showing good agreement. In a Spanish study on hospital nursing staff [Escriba et al., 1999], the Nursing Stress Scale, which measures exposure to 34 workplace stressors, was re-administered after a 15-day interval with moderate reproducibility ($ICC = 0.49$). Female healthcare workers in Norway [Vasseljen et al., 2001] showed much higher reliability over 4 weeks for job control and job pressure ($ICC = 0.85$ and 0.81, respectively). Short-term test–retest reliability (1–3 weeks) of a composite measure of work-related stress, based on six psychosocial items, was high ($k = 0.61$) in a sample of carpenters employed in the construction industry in the US [Booth-Jones et al., 1998]. Salerno et al. [2001] also found high 3-week reproducibility ($ICC = 0.88$) of the Perceived Stress Scale [Cohen et al., 1983] in keyboard operators.

With a few exceptions, these studies demonstrate acceptable test–retest reliability of the JCQ and other questionnaires in assessing exposure to psychosocial factors. The overall good reliability observed in these studies conducted in various countries and economic sectors provides evidence of the ability of workers to recognize and accurately report psychosocial exposures in the workplace. Our results suggest that when assessing test–retest agreement of self-reported psychosocial exposures, it is important to ascertain and take into account changes in work organization and job content between interviews. Changes in working conditions occurring between interviews may be responsible for the low reproducibility in some of the above-cited studies.

Although almost a decade has elapsed since the T_2 survey was carried out, we do not believe that this reduces the generalizability of our findings to present-day workplaces. In fact, the technology introduced in these plants in the 1990s was among the most advanced available at the time, being characterized by semi-automated processes, often sequential, requiring on average limited manual intervention of moderate physical effort by the operators. It is very likely that production machinery and work processes similar to those examined in this study are still in widespread use in this and other mechanical industries. Moreover, the ongoing process of technology transfer to smaller enterprises and to industries in developing countries means that this technology will not become obsolete in the near future.

CONCLUSIONS

The introduction of more automated technology in the car manufacturing plants examined in the present study produced an increase in self-reported decision latitude but did not affect perceived quantitative job demands. The results demonstrate that 5-year test–retest reliability of psychosocial exposures is adequate when assessing workers with stable job assignments and ergonomic exposures.

ACKNOWLEDGMENTS

This research was supported by the National Joint Committee on Health and Safety jointly sponsored by the

Company and the United Automobile Workers (UAW), and by Grant Number R01-OH03514 from the National Institute for Occupational Safety and Health (NIOSH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIOSH or other funding agencies. We thank Ms. Tony King and Ms. Carolyn Wynn, UAW Ergonomic Trainers, as well as other employees of the car manufacturer for their key roles in facilitating data collection in 1998. Participation of the individual workers is gratefully acknowledged. Job observations were carried out by Andrew Bigelow at T_0/T_1 and at T_2 by Judith Gold, Jung-Soon Park, Jamie Tessler, Jenny Ro, and Gary Davis. Jatin Soneta and Devang Shah performed extensive data entry, coding and cleaning. Dr. Robert Karasek offered invaluable suggestions for refining the approach to these analyses.

REFERENCES

Ariens GA, van Mechelen W, Bongers PM, Bouter LM, van der Wal G. 2001. Psychosocial risk factors for neck pain: A systematic review. Am J Ind Med 39:180–193.

Armstrong BK, White E, Saracci R. 1992. Principles of exposure measurement in epidemiology. Oxford—New York: Oxford University Press. p 351.

Barnett AG, van der Plos JC, Dobson AJ. 2005. Regression to the mean: What it is and how to deal with it. Int J Epidemiol 34:215–220.

Belkic KL, Landsbergis PA, Schnall PL, Baker D. 2004. Is job strain a major source of cardiovascular disease risk? Scand J Work Environ Health 30:85–128.

Bland JM, Altman DG. 1994a. Regression toward the mean. Br Med J 308:1499.

Bland JM, Altman DG. 1994b. Some examples of regression toward the mean. Br Med J 309:780.

Bongers PM, Kremer AM, ter Laak J. 2002. Are psychosocial factors, risk factors for symptoms and signs of the shoulder, elbow, or hand/wrist?: A review of the epidemiological literature. Am J Ind Med 41: 315–3342.

Booth-Jones AD, Lemasters GK, Succop P, Atterbury MR, Bhattacharya A. 1998. Reliability of questionnaire information measuring musculoskeletal symptoms and work histories. Am Ind Hyg Assoc J 59:20–24.

Bourbonnais R, Comeau M, Vezina M, Dion G. 1999. Job strain, psychological distress, and burnout in nurses. Am J Ind Med 34:20–28.

Braun S, Hollander RB. 1988. Work and depression among women in the Federal Republic of Germany. Women Health 14:3–26.

Brisson C, Blanchette C, Guimont C, Dion G, Moisan J, Vezina M. 1998. Reliability and validity of the French version of the 18-item Karasek Job Content Questionnaire. Work Stress 12:322–336.

Cheng Y, Luh WM, Guo YL. 2003. Reliability and validity of the Chinese version of the Job Content Questionnaire in Taiwanese workers. Int J Behav Med 10:15–30.

Cohen J, Kamarck T, Mermelstein R. 1983. A global measure of perceived stress. J Health Soc Behav 24:385–396.

Cuzick J. 1985. A Wilcoxon-type test for trend. Stat Med 4:87–90.

de Lange AH, Taris TW, Kompier MA, Houtman IL, Bongers PM. 2003. The very best of the millennium: longitudinal research and the demand-control-(support) model. J Occup Health Psychol 8:282–305.

d'Errico A, Gore R, Gold J, Park JS, Punnett L. 2007. Medium- and long-term reproducibility of self-reported exposure to workplace ergonomic factors in a cohort of auto workers. Appl Ergon 38:167–175.

Escriba V, Mas R, Cardenas M, Perez S. 1999. Validación de la escala de estresores laborales en personal de enfermería: «the nursing stress scale». Gac Sanit 13:191–200.

Escriba-Aguir V, Mas Pons R, Flores Reus E. 2001. Validacion del Job Content Questionnaire en personal de infermeria hospitalario. Gac Sanit 15:142–149.

Fleiss JL. 1981. Statistical methods for rates and proportions. New York: Wiley. 321 pp.

Frese M, Zapf D. 1988. Methodological issues in the study of work stress: Objective vs. subjective measurement of work stress and the question of longitudinal studies. In: Cooper CL, Payne RL, editors. Causes, coping and consequences of stress at work. Chichester, England: Wiley. p 375–411.

Gold JE, Park J-S, Punnett L. 2006. Work routinization and implications for ergonomic exposure assessment. Ergonomics 49:12–27.

Hemingway H, Marmot M. 1999. Evidence based cardiology: Psychosocial factors in the aetiology and prognosis of coronary heart disease. Systematic review of prospective cohort studies. Br Med J 318: 1460–1467.

Ingelgard A, Karlsson H, Nonas K, Ortengren R. 1996. Psychosocial and physical work environment factors at three workplaces dealing with material handling. Int J Ind Ergon 17:209–220.

Johansson JA, Nonas K. 1994. Psychosocial and physical working conditions and associated musculoskeletal symptoms among operators in five plants using arc welding in robot stations. Int J Hum Factor Man 4:191–204.

Karasek RA. 1985. Job Content Questionnaire and user's guide. Los Angeles/Lowell, MA: University of Southern California/University of Massachusetts Lowell.

Karasek RA, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. 1998. The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. J Occup Health Psychol 3:322–355.

Kasl SV. 1978. Epidemiological contributions to the study of work stress. In: Cooper CL, Payne RL, editors. Stress at work. New York: Wiley & Sons. p 3–38.

Kasl SV. 1998. Measuring job stressors and studying the health impact of the work environment: An epidemiologic commentary. J Occup Health Psychol 3:390–401.

Kerr MS, Frank JW, Shannon HS, Norman RWK, Wells RP, Neumann WP, Bombardier C. Ontario Universities Back Pain Study Group. 2001. Biomechanical and psychosocial risk factors for low back pain at work. Am J Public Health 91:1069–1075.

Kivimaki M, Vahtera J, Ferrie JE, Hemingway H, Pentti J. 2001. Organisational downsizing and musculoskeletal problems in employees: A prospective study. Occup Environ Med 58:811–817.

Kristensen TS. 1996. Job stress and cardiovascular disease: A theoretic critical review. J Occup Health Psychol 1:246–260.

Landsbergis PA, Schnall PL, Pickering TG, Schwartz JE. 2002. Validity and reliability of a work history questionnaire derived from the Job Content Questionnaire. J Occup Environ Med 44:1037–1047.

Latko WA, Armstrong TJ, Foulke JA, Herrin GD, Rabourn RA, Ulin SS. 1997. Development and evaluation of an observational method for assessing repetition in hand tasks. Am Ind Hyg Assoc J 58:278–285.

Lazarus RS, Folkman S. 1986. Cognitive theories of stress and the issue of circularity. In: Appley ML, Trumbull R, editors. Dynamics of stress. New York: Plenum. p 62–80.

Li J, Yang W, Liu P, Xu Z, Cho SI. 2004. Psychometric evaluation of the Chinese (mainland) version of Job Content Questionnaire: A study in university hospitals. Ind Health 42:260–267.

MacDonald LA, Karasek RA, Punnett L, Scharf T. 2001. Covariation between workplace physical and psychosocial stressors: Evidence and implications for occupational health research and prevention. Ergonomics 44:696–718.

Mausner-Dorsch H, Eaton WW. 2000. Psychosocial work environment and depression: Epidemiologic assessment of the demand-control model. Am J Public Health 90:1765–1770.

Myers AH, Baker SP, Li G, Smith GS, Wiker S, Liang KY, Johnson JV. 1999. Back injury in municipal workers: A case-control study. Am J Public Health 89:1036–1041.

Nahit ES, Macfarlane GJ, Pritchard CM, Cherry NM, Silman AJ. 2001. Short term influence of mechanical factors on regional muskoskeletal pain: A study of new workers from 12 occupational groups. Occup Environ Med 58:347–381.

Niedhammer I. 2002. Psychometric properties of the French version of the Karasek Job Content Questionnaire: A study of the scales of decision latitude, psychological demands, social support, and physical demands in the GAZEL cohort. Int Arch Occup Environ Health 75:129–144.

Niedhammer I, Goldberg M, Leclerc A, Bugel I, David S. 1998. Psychosocial factors at work and subsequent depressive symptoms in the Gazel cohort. Scand J Work Environ Health 24:197–205.

Pelfrene E, Vlerick P, Mak R, De Smet P, Kornitzer M, De Backer G. 2001. Scale reliability and validity of the Karasek 'Job Demand-Control-Support' model in the Belstress study. Work Stress 15:297–313.

Punnett L. 1998. Ergonomic stressors and upper extremity disorders in vehicle manufacturing: Cross sectional exposure-response trends. Occup Environ Med 55:414–420.

Punnett L, Gold J, Katz JN, Gore R, Wegman DH. 2004a. Ergonomic stressors and upper extremity musculoskeletal disorders in automobile manufacturing: A one-year follow-up study. Occup Environ Med 61: 668–674.

Punnett L, MacDonald L, Härenstam A, Warren N. 2004b. Work Organization: Contextual exposure assessment for occupational health epidemiology (invited keynote). Exposure assessment in epidemiology and practice, Utrecht, The Netherlands, June 2004.

Sale JE, Kerr MS. 2002. The psychometric properties of Karasek's demand and control scales within a single sector: Data from a large teaching hospital. Int Arch Occup Environ Health 75:145–152.

Salerno DF, Franzblau A, Armstrong TJ, Werner RA, Becker MP. 2001. Test–retest reliability of the Upper Extremity Questionnaire among keyboard operators. Am J Ind Med 40:655–666.

Schnall PL, Landsbergis PA, Baker D. 1994. Job strain and cardiovascular disease. Annu Rev Public Health 15:381–411.

Stansfeld SA, Fuhrer R, Shipley MJ, Marmot MG. 1999. Work characteristics predict psychiatric disorder: Prospective results from the Whitehall II Study. Occup Environ Med 56:302–307.

Torp S, Riise T, Moen BE. 2001. The impact of psychosocial work factors on musculoskeletal pain: A prospective study. J Occup Environ Med 43:120–126.

Van der Doef M, Maes S. 1999. The Job Demand-Control(-Support) model and psychological well-being: A review of 20 years of empirical research. Work Stress 13:87–114.

Vasseljen O, Holte KA, Westgaard RH. 2001. Shoulder and neck complaints in customer relations: Individual risk factors and perceived exposures at work. Ergonomics 44:355–372.

Williams RM, Sundelin G, Schmuck ML. 2001. Reliability of the Demand-Control Questionnaire for sewing machine operators. Work 16:71–75.