Applied Ergonomics 45 (2014) 1157–1166

Contents lists available at ScienceDirect

Applied Ergonomics

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

Reliability and validity of the Home Care STAT (Safety Task Assessment Tool)

Ryan Olson [a,b,d,*], Brad Wipfli [a,1], Robert R. Wright [a,d,2], Layla Garrigues [e,3], Thuan Nguyen [b,4], Borja López de Castro [c,5]

[a] Oregon Health & Science University, Oregon Institute of Occupational Health Sciences, 3181 SW Sam Jackson Park Rd., L606, Portland, OR 97239-3098, USA
[b] Oregon Health & Science University, Department of Public Health and Preventive Medicine, 3181 SW Sam Jackson Park Road, CB 669, Portland, OR 97239-3098, USA
[c] University of Valencia, Department of Social Psychology, Avda. Blasco Ibáñez 21, 46010 Valencia, Spain
[d] Portland State University, Department of Psychology, P.O. Box 751, Portland, OR 97207-0751, USA
[e] Oregon Health & Science University, School of Nursing, 3455 SW US Veterans Hospital Rd., Portland, OR 97239-2941, USA

A R T I C L E I N F O

Article history:
Received 29 August 2013
Accepted 10 February 2014

Keywords:
Home care workers
Hazard exposure
Self-assessment

A B S T R A C T

Home care workers are a priority population for ergonomic assessment and intervention, but research on caregivers' exposures to hazards is limited. The current project evaluated the reliability and validity of an ergonomic self-assessment tool called Home Care STAT (Safety Task Assessment Tool). Participants ($N = 23$) completed a background survey followed by 10–14 days of self-monitoring with the STAT. Results showed that the most frequent task was house cleaning, and that participants regularly performed dangerous manual client moving and transferring tasks. Researcher in-home observations of 14 workers (duration ≤ 2 h) demonstrated that workers' self-assessments were moderately reliable. Correlational and multi-level analyses of daily self-assessment data revealed that several task exposures were significantly related to daily fatigue and/or pain. Other associations have implications for Total Worker Health™; for example, daily stress was positively associated with both pain and consumption of high calorie snacks. Findings support the STAT as a reliable and potentially valid tool for measuring home care workers' exposures to physically demanding tasks.

1. Introduction

Home care workers (HCWs) are a priority population for ergonomic assessment and intervention. These workers, who may be classified as personal or home care aides (BLS, 2010), assist the elderly and people with disabilities with activities of daily living in residential settings. Home care occupations have been characterized as offering low pay, few benefits, unstable employment, limited training and equipment, irregular work hours, and overall poor working conditions (Askew and Walker, 2008; Galinsky et al., 2001; Hayashi et al., 1994). The job is physically demanding, and HCWs often perform dangerous tasks, such as moving and transferring clients, without the help of co-workers or assistive devices. HCWs employed by private agencies may benefit from some supervision and safety training. However, independent HCWs are isolated and dependent on their clients, who are generally untrained private citizens, to provide safe working conditions. Not surprisingly, HCWs have elevated musculoskeletal problems relative to other occupational groups (Jellema et al., 2002; Ono et al., 1995; Simon et al., 2004; Tuomi et al., 1991; Zeytinoglu et al., 2000), and a lost-time injury rate that is 3–4 times higher than the US average (Bureau of Labor Statistics [BLS], 2010). The need for safety interventions is magnified by projections that HCW employment will grow by 70% by 2020 (BLS, 2012) to meet the needs of the growing elderly population (Administration on Aging, 2011).

* Corresponding author. Oregon Institute of Occupational Health Sciences, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd., L606, Portland, OR 97239-3098, USA. Tel.: +1 503 494 2501; fax: +1 503 494 4278.
E-mail addresses: olsonry@ohsu.edu (R. Olson), wipflib@ohsu.edu (B. Wipfli), wrightro@byui.edu (R.R. Wright), garrigue@ohsu.edu (L. Garrigues), nguythua@ohsu.edu (T. Nguyen), b.lopezdecastro@universitalent.com (B. López de Castro).
[1] Tel.: +1 503 494 2532.
[2] Tel.: +1 208 496 4085.
[3] Tel.: +1 360 600 5205.
[4] Tel.: +1 503 494 3697.
[5] Tel.: +34 151 5928 4094.

http://dx.doi.org/10.1016/j.apergo.2014.02.004
0003-6870/

1.1. Measuring hazards and exposures in home care

Despite the need for assessment and intervention research with HCWs, there is a relative paucity of relevant safety studies. Prior research with HCWs has focused on measuring injury prevalence and worker perceptions about the causes of stress and strain (Baron and Habes, 2004; Kim et al., 2010; Meyer and Muntaner, 1999; Myers et al., 1993; Wipfli et al., 2012). This research shows that HCWs are at greatest risk of experiencing musculoskeletal injuries to the low back, shoulders, and knees, and that lost time injuries are likely to be triggered by manual material handling and client lifts/transfers (Ono et al., 1995; Hess et al., 2007; Nelson et al., 2003; Wipfli et al., 2012). Other physically demanding tasks that may generate unusual postures and contribute to injuries include changing bedding and assisting clients with bathing and toileting (BLS, 2012). Workers' perceptions about injury hazards are aligned with injury data. HCWs in two independent focus group studies identified several tasks as highly physically demanding or causing high concern for injury: unassisted client lifting and transferring; supporting clients while walking or catching while falling; push/pull/lift wheelchair; bathing; moving household objects; and cleaning bathrooms, floors, and kitchens (Baron and Habes, 2004; Wipfli et al., 2012).

While certain caregiving tasks have been implicated as hazardous, we know little about actual task exposure levels among HCWs. Direct physical or observational measurement is ideal, but is rare among isolated working populations due to the effort and time required (Olson et al., 2009). If we are to gain a greater understanding of exposures and intervention needs among HCWs, observational methods for measuring worker postures and task exposures (e.g., Rapid Entire Body Assessment; Hignett and McAtamney, 2000) must be supplemented with validated worker self-assessment methods.

Ergonomic self-assessment research has predominantly focused on the reliability, bias, and validity of one-time ergonomic self-assessment surveys or ratings (Barrero et al., 2009a, 2009b; Hollmann et al., 1999; Spielholz, 2006; Spielholz et al., 2001; Wiktorin et al., 1993), and using survey measures to model or predict objective outcomes (e.g., Trask et al., 2010). In general, this literature suggests that worker self-assessments can be systematically biased, but also that workers are able to reliably self-assess exposures to a range of events that are easily discriminated or intense, and that these self-ratings can predict safety-related criteria. As an example of systematic bias, Barrero et al. (2009b) found that workers overestimated the duration of physically demanding tasks and underestimated less physically demanding tasks, and that these biases were inflated among participants who completed the tasks in a discontinuous pattern. As an example of the validity of worker self-reports, Trask et al. (2010) found that workers reports of task exposures following their workshift explained 21–36% of the variance in electromyography (EMG) of lumbar muscles during the workshift (versus direct observation of work characteristics explaining 31–47% of variance in EMG activity).

While one-time ergonomic self-assessment surveys are useful, the data produced do not provide information about variability in task exposures that occur over time. In this regard, ecological momentary assessment (EMA) methods may provide enhanced ergonomic self-assessment data. In EMA participants are cued by alarms or events in their environment to self-monitor events, behaviors, or psychological states as they occur over many days or weeks. EMA (Bolger et al., 2003; Green et al., 2006) reduces retrospective bias and improves the reliability and validity of self-report data (Stone et al., 1998). However, EMA remains a self-report method, and while computerized EMA programs time stamp each self-monitoring instance, the accuracy of a participant's isolated self-recordings about their behavior may be difficult to assess. While we are not aware of any previous EMA research with HCWs, the approach has been applied in similar working populations such as nurses (Johnston et al., 2006). Most relevant to the current project, two prior studies have evaluated the reliability of EMA to measure safety-related task exposures.

Unge et al. (2005) used EMA to measure task exposures among office workers and hospital cleaners for 10 days, while experimenters observed several concurrent days. Workers produced reliable self-assessment data for most primary work tasks, including exposure to visual display unit work, which is a risk factor for musculoskeletal problems. In a study with a lone working population, Olson et al. (2009) evaluated truck drivers' self-assessments of non-driving work tasks and postures relative to video-based observations. Drivers were highly reliable at self-recording hazards that were environmental, relatively frequent, and easy to discriminate, such as receiving assistance from a forklift operator. Drivers were less reliable at self-assessing postural exposures, such as severe trunk flexion or rotation. These two studies demonstrate that workers can produce moderately reliable measures of work exposures using EMA methods, especially frequent and easily discriminated events. The research also demonstrates the practical value of EMA for measuring exposures among isolated and dispersed populations.

To address gaps in research and practice and extend EMA findings to HCWs, we created and evaluated an ergonomic self-assessment tool named Home Care STAT (Safety Task Assessment Tool). The STAT was programmed into a handheld computer, and a daily diary style study was designed to: 1) measure daily task exposures in a sample of HCWs, 2) evaluate the reliability of the STAT by comparing worker self-assessments to data from researcher observations, and 3) evaluate the criterion-related validity of the STAT by conducting correlational and multi-level analyses of associations between person-level and daily work task factors with psychophysical symptoms and lifestyle behaviors.

2. Material and methods

2.1. Participants

All participants worked as independent contractors for clients who qualified for publicly funded in-home care services in Oregon. Participants were recruited from training courses offered by the Oregon Home Care Commission and through SEIU Local 503 communications, and were required to work at least two days per week as a HCW and reside within the study metropolitan area. A total of 23 participants consented and completed all study procedures. The final sample was predominantly female ($n = 22$) with a mean age of 47.5 yrs ($SD = 12.7$), a mean body mass index of 29.9 ($SD = 7.4$), and was 57% minority, including African American ($n = 5$), Asian ($n = 2$), Native American ($n = 2$), Hispanic/Latino ($n = 2$), and other ($n = 2$). Participants worked an average of 34.1 h per week ($SD = 17.8$) and had been working as HCWs for a mean of 6.2 years ($SD = 6.2$). The majority of participants were single (divorced, separated or never married; 52%) and had an average of 13.2 years of education ($SD = 2.4$).

2.2. Study design and observation methods

The study employed a repeated measures daily diary design, where repeated daily measures were nested within individual participants. Participants first completed a demographic and background survey and received brief training with the STAT. Participants were then asked to use the STAT for the next 10–14 days to self-record task exposures during workdays, and a range of symptoms and behaviors every evening (including non-workdays). This design structure, with 23 participants × 10–14 days, possesses reasonable potency for accurately assessing variance components at person and daily measure levels of analysis (Hox, 2010; chapter 12). Study participation concluded with either an in-home observation of the participant at work (dependent upon client consent; $n = 14$ of 23), or a debriefing meeting with researchers. Observations were conducted by the first three authors, who were involved in STAT development and intimately familiar with variable definitions. As additional training, each researcher watched 2 h of safety training videos for HCWs in order to become familiar with and observe many repetitions of common home care tasks before going into the field. Each observation lasted until the participant had completed 10 work tasks, or until the observation duration reached 2 h. During the observation participants continued to complete their STAT self-assessments during natural breaks (which typically occurred during transitions between tasks), while the observer independently recorded the occurrence and duration of work tasks *in vivo*. Participants knew that researchers were observing their work, but were not aware that the primary purpose was to estimate the reliability of their self-observations. All measures of inter-rater agreement in the study were comparisons of a single researcher with a single worker during in-home observations.

2.3. Demographic and background measures

In addition to demographics reported in section 2.1, the demographic and background survey included measures of self-reported height and weight, concern for personal injury during work tasks; and self-reported injuries in the past 12 months. Participants also reported physical symptoms (3 items from the Physical Symptoms Inventory; Spector and Jex, 1998), life satisfaction (4-item Satisfaction with Life Scale; Diener et al., 1985), 12-month and past 7 day physical pain and discomfort (9-item Nordic Musculoskeletal Questionnaire: Kuorinka et al., 1987); fatigue (8-items from the Swedish Occupational Fatigue Inventory; Ahsberg, 2000); perceived stress (4-item Perceived Stress Scale; Cohen et al., 1983); positive and negative affect (8-item PANAS; Watson et al., 1988); sleep adequacy (1-item; Buxton et al., 2009); exercise (7-item International Physical Activity Questionnaire [iPAQ]; Ainsworth et al., 2000); and dietary habits (4-items: sugary snacks, sugary drinks, fast food; as described in Buxton et al., 2009; plus a fruit and vegetable item created for the current study).

Some demographic and background measures were adapted for the current study. Specifically, fatigue was assessed using two of the four subscales from the Swedish Occupational Fatigue Inventory (lack of energy, physical discomfort). Exercise was measured using the short version of the iPAQ (Ainsworth et al., 2000), however, the time-anchor period was changed from the past 7 days to the past 4 weeks to be consistent with the reference time period for all other behavioral scales in the survey. Finally, we created an item to measure fruit and vegetable consumption using the same scaling as the 3-item survey of high sugar and high fat food consumption (as described in Buxton et al., 2009; 1 = never to 10 = 5 or more times per day).

2.4. Daily Home Care STAT measures

The STAT was developed to help workers self-monitor their exposures to demanding and/or dangerous work tasks, along with daily physical symptoms, health behaviors, and mood states. To accomplish these measurement goals the STAT included two question sets: a work task survey and an evening survey. Question sets and branching schemes were programmed using survey development software (SurveyToGo, Dooblo Ltd., Kfar Sava, Israel) for administration on handheld computers (HP iPAQ 110, Hewlett–Packard Company, Palo Alto, CA). Before implementation the STAT was pilot tested with three HCWs and revised based on their feedback.

2.4.1. Work task survey measures

The work task survey focused on physically demanding and hazardous tasks identified in previous research (Baron and Habes, 2004; Kim et al., 2010; Meyer and Muntaner, 1999; Myers et al., 1993; Wipfli et al., 2012) and was completed throughout each workday during natural breaks in between work tasks (event contingent). The work task survey was accessible at any time during the day, and participants were instructed to complete it after each work task as possible. If this frequency was not possible, or if the participant forgot to make an entry, they were encouraged to complete the survey at least once every 2 h while at work. Once the work task survey was opened, workers were asked to report the occurrence, duration, and/or frequency of seven categories of home care tasks: driving to/for client, assisting client mobility, bathing/toileting, dressing, client lifting/transferring, house cleaning, and lifting/moving objects, as well as an "other" category to capture additional tasks. Two of the seven task categories included subcategories. Assisting client mobility was further specified as pushing or pulling in a wheelchair, or assisting with walking. The objects lifting/moving category was further specified by weight of objects handled using intervals from a prior ergonomic self-assessment survey (Wiktorin et al., 1993): light weight (<15 lbs), medium weight (15–35 lbs), and heavy weight (>35 lbs). Example objects of each weight category were provided to help workers make accurate estimations (e.g., light weight object examples were grocery bags or a loaded laundry basket). For each task a worker selected, the STAT presented further questions about the duration of the task in minutes, frequency counts if appropriate, the environmental context

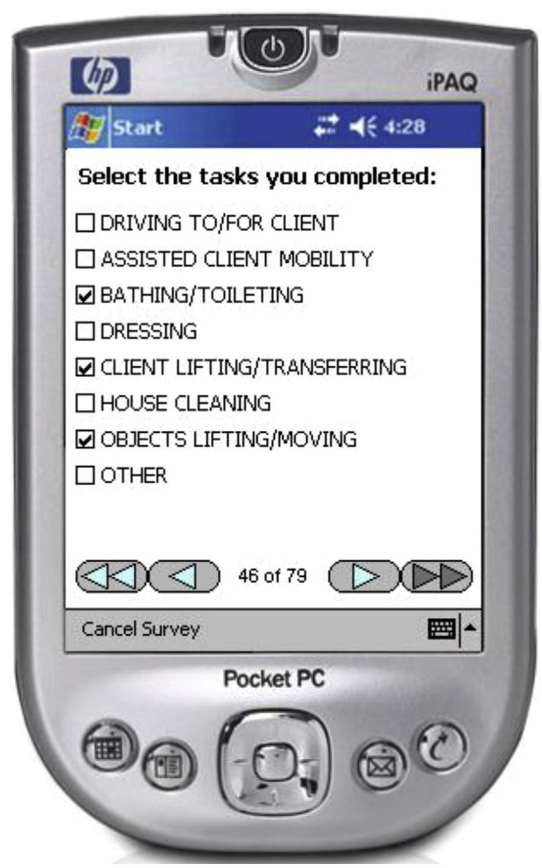

Fig. 1. Home Care STAT: Work task survey screen.

(e.g., bathing location such as in a shower or a bathtub, client transfer type such as bed to wheelchair/walker), and tools used (e.g., vacuum, long-handled mop). Question branching minimized participant effort by skipping questions about tasks workers did not report completing. See Fig. 1 to view the home screen for the work task survey component of the Home Care STAT.

2.4.2. Evening survey measures

The evening survey asked workers to rate or record a variety of psychophysical and behavioral variables when cued by a self-selected alarm time set between dinner and bedtime (signal contingent). The evening survey contained daily adaptations of psychophysical and behavioral constructs from the demographics and background survey. Daily adapted measures in the evening survey that were used for analyses in the current study included time working, sleep adequacy (Buxton et al., 2009), morning pain/discomfort, the Swedish Occupational Fatigue Inventory (Ahsberg, 2000), Perceived Stress (Cohen et al., 1983), Positive/Negative Affect scale (Watson et al., 1988), fruit and vegetable servings, three items about high fat/high sugar foods consumption, and minutes of exercise. In adaptations longer or more complex items were replaced with simplified items anchored at the daily time point. For example, musculoskeletal symptoms analyzed were a single item where participants rated their daily level of discomfort/pain upon awakening that morning on a 10-point severity scale. Morning pain was selected instead of evening pain because it was hypothesized that sore muscles or minor strains would be most strongly felt on the following day after an exposure or event. Another example of a daily adaptation was the measure of exercise, where the 4-item iPAQ (Ainsworth et al., 2000) was replaced with a single item asking workers to report daily minutes of moderate exercise, which was defined as “activity causing harder than normal breathing, such as brisk walking, biking, or sports.”

2.5. Analysis

2.5.1. Descriptives

To describe the study participants, task exposure levels, and daily symptoms and lifestyle behaviors, we computed means, standard deviations, and frequency distributions from measures in the demographics and background survey and daily STAT measures.

2.5.2. Reliability

The reliability of workers’ self-monitored STAT data was evaluated by comparing researcher and worker measurements during in-home observations. Comparisons included 1) contrasts of mean recorded task durations, 2) computation of percent agreement and kappa coefficients for dichotomous (yes/no) task occurrences (Cohen, 1960), and 3) computation of agreement Intra-Class Correlations (ICCs) for minutes and count measures. For occurrence/non-occurrence measures, two kappa coefficients were calculated. The first evaluated occurrence agreement for all the questions in the work task survey portion of the STAT (all eight task categories plus all branching sub-questions), and the second coefficient evaluated only occurrence agreement for the eight main task categories. Due to low occurrence of handling medium and heavy weight objects recorded by both observers and participants, durations for moving light, medium, and heavy objects were collapsed into a single aggregate ‘moving objects’ variable (for reliability analyses only).

2.5.3. Validity

To evaluate STAT criterion-related validity, we conducted both correlational and multi-level analyses. We implemented correlational analyses as a first step to explore criterion-related validity and inform more focused multi-level models. Subsequent focused multi-level models were then used to test person-level, day-level, and cross-level predictors of daily outcomes. For correlational analyses with daily data we used a recently developed hidden correlation technique that corrects for non-observable measurement error due to variations between occasions (i.e., day) in repeated measures (Nguyen and Jiang, 2011). This type of correlation maximizes the use of available data (the formula includes both between and within subject variance), yet captures the correlation among repeated observations within the same subject. Multilevel regression was performed with HLM software (v6.0; Raudenbush et al., 2004) to examine relationships among person- and daily-level factors. Models were designed to evaluate how person-level factors and daily-level work exposures related to daily physical and psychological symptoms (fatigue, pain, affect, and stress). Level 1 models tested relationships between daily predictor variables (within-person) and daily outcome variables. All Level 1 predictor variables were person mean-centered and slopes were allowed to vary. Level 2 models tested relationships between person-level predictors (between-person) and daily (within-person) outcome variables. Level 2 predictors were grand mean-centered and treated as fixed effects.

3. Results

3.1. Descriptive results

3.1.1. Demographics and background descriptives

Most participants ($n = 17$ of 23) regularly performed client lifting or transfer tasks. Tasks causing the most concern for personal injury (10 = highest concern for injury) were objects lifting/moving ($M = 3.9$, $SD = 3.2$), housecleaning ($M = 3.8$, $SD = 3.7$), and assisting client mobility (i.e., wheelchair, walking; $M = 3.4$, $SD = 2.9$). Eighty-six percent of workers reported musculoskeletal symptoms during the past 12 months and fifty-five percent reported experiencing symptoms during the past seven days. Thirty-four percent reported that symptoms caused them to modify or miss work in the past 12 months. During the same time period workers reported experiencing an average of 4.0 ($SD = 15.3$) minor work-related injuries (no lost work time), and an average of .2 ($SD = .5$) lost work time injuries. See Table 1 for additional baseline descriptive statistics of psychophysical and lifestyle behavior variables.

3.1.2. Daily STAT descriptives

HCWs contributed 242 participant days of STAT data (157 workdays, 85 non-workdays). Workers had an 85% compliance rate for completing the evening survey (206 completed out of 242 expected). Compliance for the work task survey was not computed because the daily frequency of submissions was free to vary. However, workers submitted work task surveys for only a portion of every workday (average 6.6 h worked reported in evening surveys vs. average ≈ 2 h total task times recorded). Visual inspection of time stamps for submitted surveys showed on only 6 occasions did a participant submit a work task survey after submitting an evening survey, which suggests that participants generally adhered to the protocol of submitting work task surveys during the day rather than completing them retroactively after being prompted by an alarm to complete an evening survey. No significant relationships between demographic variables and number of missing evening surveys were found. Participants recorded the most work time performing house cleaning tasks, followed by bathing/toileting, and dressing (see Table 2). Workers performed an average of 1.5 client lifts or transfers per workday ($SD = 3.3$). Across the 88 study days when pain was reported, the three most frequent body regions with pain were lower back (66 participant days), shoulders (49

Table 1
Descriptive statistics for psychophysical variables and lifestyle behaviors from the demographics and background survey.

Variable[a]	Mean (SD)	Scale description
Occupational fatigue	19.43 (11.43)	Sum of 7-pt frequency scale, 0 = never, 3 = sometimes, 6 = very often (highest possible fatigue = 48)
Perceived stress	3.08 (.31)	5-pt frequency scale, 1 = never, 3 = sometimes, 5 = very often (highest possible stress = 5)
Negative affect	2.23 (.83)	5-pt agreement scale, 1 = not at all, 3 = moderately, 5 = extremely (worst negative affect = 5)
Positive affect	3.37 (.80)	5-pt agreement scale, 1 = not at all, 3 = moderately, 5 = extremely (best positive affect = 5)
Daily fruit/vegetable servings[b]	2.47 (1.38)	Weighted 10 point frequency intervals scale (see note below for computation)
High calorie food/drink consumption[b,c]	1.88 (2.36)	Weighted 10 point frequency intervals scale (see note below for computation)
Sleep adequacy	3.43 (1.20)	5-pt frequency scale, 1 = never, 3 = sometimes, 5 = very often (best sleep adequacy = 5)

Note.
[a] All psychophysical and lifestyle behavior scales asked workers to reflect on the past month when answering questions.
[b] Fruit and vegetable servings and high calorie food/drink consumption were self-reported on a scale of 10 different frequency intervals. We translated interval ratings to daily servings/counts using the following rating system (0 = 0 (Never in the past 4 weeks); 1 = .04 (less than once in the past 4 weeks); 2 = .07 (1–3 times in the past 4 weeks); 3 = .21 (1–2 times a week in the past 4 weeks); 4 = .5 (3–4 times a week in the past 4 weeks); 5 = .79; (5–6 times a week in the past 4 weeks); 6 = 1 (once a day in the past 4 weeks); 7 = 2 (2 times a day in the past 4 weeks); 8 = 3.5 (3–4 times a day in the past 4 weeks); 9 = 5 (5 or more a day in the past 4 weeks).
[c] High calorie food/drink consumption does not translate perfectly into servings per day because of mixed item anchoring (servings vs. meals); instead, the statistic reflects a summed count of instances of consumption of sugary drinks, sugary snacks, and fast food. Exercise questions were asked in the Demographics and Background survey using the 7-item iPAQ scale, but were excluded from analyses in the current paper due to observed measurement errors. Multiple participants provided extreme answers and appeared to misunderstand or make errors on the iPAQ. To illustrate, participants reported an average of 99.0 min ($SD = 86.1$) of daily physical activity in the Demographics and Background survey (2855 MET-min per week; $SD = 3002$), compared to only 17.7 min ($SD = 20.0$) of exercise per day in daily STAT evening surveys.

participant days) and legs/knees (46 participant days). Tasks that most frequently caused workers concern for personal injury were housecleaning (44 participant days), client lifting/transferring (32 participant days), lifting/moving objects (30 participant days), and assisting client mobility (30 participant days).

With regard to daily health behaviors, participants reported an average sleep sufficiency rating of 3.7 (1 = no, 5 = yes; $SD = 1.4$); an average of 17.7 min of exercise ($SD = 20.0$); and eating an average of 2.9 servings of fruit and vegetables ($SD = 2.2$), 1.3 servings of sugary snacks ($SD = 1.6$), 1.3 servings of sugary drinks ($SD = 1.9$), and 0.3 fast food meals ($SD = 0.6$). In subsequent correlational analyses sugary snacks/drinks and fast food meals were combined into a single summed "high calorie food/drink" measure. Exploratory contrasts between workdays and non-workdays for evening survey items showed that workdays were associated with significantly higher fatigue ($t = 2.32$, $p < .05$) and lower positive affect ($t = -3.60$, $p < .01$), while non-workdays were associated with higher consumption of high calorie food/drinks ($t = -2.60$, $p < .05$).

3.2. Reliability results

Mean differences between participants and observers in recorded task durations varied from 1.8 to 12.6 min (see Table 3). To illustrate, participants recorded an average of 38.2 min of house cleaning during observation periods, which was 6.1 min more than the average of 32.3 min observed by researchers. In general, participants overestimated task durations relative to researcher measurements, with the exception of the 'other' category where participants recorded an average of 4.3 fewer minutes spent performing these unspecified tasks (see Table 3). Percent agreement for the occurrence of all events in the work task survey ($n = 1204$ events, according to total occurrences recorded by participants) was 88.7%, while the kappa coefficient, which corrects for chance agreement, was $\kappa = .59$ (Landis and Koch, 1977). Similarly, percent agreement for the occurrence of the eight main task categories ($n = 98$ events) was 79.6%, with $\kappa = .58$. The overall agreement ICC for participants vs. observers in minutes recorded across all tasks was .73. ICCs for agreement for minutes recorded in specific tasks are reported in Table 3, but should be interpreted cautiously given the infrequent occurrence of many tasks (and thus, limited data points for comparison). The overall agreement ICCs for participants vs. observers in recording count variables was .46 for the number of transfers performed and .74 for the number of objects moved.

3.3. Validity results

Hidden correlations showed several significant ($p < .05$) associations between work exposures and daily symptoms. Time spent driving and dressing a client were positively associated with daily fatigue ($r = .44$ and .41, respectively), and time spent moving heavy objects was associated with pain the next morning ($r = .99$).[6] Certain work tasks were also associated with various mood and lifestyle behaviors. Daily client transfer count was significantly

Table 2
Reported time allocation and frequency counts for work tasks across all workdays.

	Mean minutes	Total % time	Mean daily frequency
House cleaning	47.49	39.1%	NA
Moving objects	13.33	10.9%	2.46
Bathing/toileting	11.03	9.1%	NA
Assisting with mobility (walking, wheelchair)	9.14	7.6%	6.88
Dressing	8.45	7.0%	NA
Driving for client	7.11	5.9%	NA
Client lifting/transferring	3.77	3.1%	1.48
Other tasks	21.11	17.4%	NA
Total	121.43	100.0%	NA

[6] It is theoretically and empirically established that hidden correlations typically reveal higher and more accurate associations between variables than standard Pearson correlations for repeated measures datasets (Nguyen and Jiang, 2011). Moreover, the Pearson correlation is inappropriate for application with repeated measures datasets due to the dependency of observations within each subject. However, our observed .99 hidden correlation between moving heavy objects and pain is abnormally, and perhaps impossibly, high. In contrast, the Pearson correlation between within subject means for these two variables was .50. We therefore inspected and plotted the data for patterns that may have inflated the estimated magnitude of the hidden correlation in this instance. Fourteen of the twenty two subjects recorded zero time moving heavy objects, and five of those subjects who reported no exposure to moving heavy objects also reported zero ratings of daily pain. The hidden correlation was developed under the assumption of smooth data (e.g., data are normally or uniformly distributed). However, the data in the aforementioned variables are sparse due to many missing values and zero-inflation observations. These data patterns appear to be the driving empirical factor behind the abnormally high hidden correlation coefficient in this case. Therefore, while the hidden correlation is an empirical innovation for detecting significant associations in repeated measures datasets, researchers must inspect data for patterns in variance that may generate biased estimates of association in this paradigm.

Table 3
A contrast of worker and researcher recordings of task exposures in minutes during in-home observations.

	Workers completing task	Worker			Researcher			Difference (worker-researcher)			ICC
Task	n	Mean	Median	Min-max	Mean	Median	Min-max	Mean	Median	Min-max	
Assist wheelchair	2	3.0	2.0	(2–5)	.67	0	(0–2)	2.33	2.0	(2–3)	.43
Assist walking	4	4.8	2.0	(0–20)	3.0	1.0	(0–13)	1.8	1.0	(−1 to 7)	.88
Bathing/toileting	4	24.8	30.0	(2–60)	12.2	9.0	(0–31)	12.6	21.0	(−7 to 48)	.43
Dressing	4	12.0	11.5	(5–20)	5.5	1.5	(0–19)	6.5	10.0	(1–15)	.72
Client lifting/transferring	4	11.0	3.0	(1–30)	8.6	5.0	(0–33)	2.4	−2.0	(−13 to 30)	.49
House cleaning	13	38.4	23.0	(0–140)	32.3	20.0	(1–119)	6.1	3.0	(−55 to 71)	.68
Moving objects	9	11.9	0	(0–45)	6.8	1.0	(0–27)	5.1	−1.0	(−5 to 35)	.31
Other	11	8.9	3.0	(0–55)	13.2	9.0	(0–52)	−4.3	−6.0	(−42 to 55)	−.12

Note: All variables are minutes of exposure recorded for the 14 workers who were observed in their clients' homes. Minimum and maximum values for each variable (except for ICCs) were rounded to the nearest whole number.

associated with daily fruit and vegetable consumption. Additionally, several "symptom" factors varied together significantly ($p < .05$); for example, fatigue was associated with increased levels of stress, negative affect, and pain; and positive affect was positively associated with sleep sufficiency. Sleep sufficiency was also significantly related to increased levels of fruit and vegetable consumption (see Table 4). As with all correlation matrices that test a large number of relationships, the probability of type I errors is inflated in our analyses. We did not apply any family-wise corrections for significance tests of correlations, such as the false discovery rate correction (Benjamini and Hochberg, 1995, 2000), because they have not yet been adapted for hidden correlations. Moreover, the function of the hidden correlation matrix was to explore the data and inform subsequent construction of multilevel models rather than to draw firm conclusions about associations. All this considered, correlations on their own should be interpreted with caution due to the probability of type I errors in the matrix.

Multilevel regression analyses revealed that several STAT daily work task exposure measures were significant predictors of daily outcomes. Time spent dressing the client, lifting/moving medium objects (time), and heavy household objects (count) were predictive of higher levels of lagged pain the next morning (see Table 5). In addition, lifting/moving medium (time) or heavy (count) household objects was significantly predictive of daily negative affect. Lifting/moving heavy objects (both time and count) significantly predicted daily stress. In person-level regression models (Level 2), BMI and age were significant predictors of some daily symptoms (see Table 5). Our examination of cross-level interactions between significant work task exposures and person characteristics produced one significant relationship. Specifically, the interaction between increasing person-level BMI and increasing day-level time spent moving medium objects was associated with increased levels of daily negative affect ($p = .0002$).

4. Discussion

The aim of the current research was to evaluate the Home Care STAT as an ergonomic self-assessment tool for HCWs. We conducted a daily diary study with the STAT to: 1) measure daily task exposures in a sample of HCWs, 2) evaluate the reliability of the STAT by comparing worker self-assessments to researcher observations, and 3) evaluate the criterion-related validity of the STAT through correlational and multi-level analyses. Results confirm that HCWs are a population in need of health and safety interventions, and that the STAT is an acceptably reliable and valid tool for safety self-assessment in home care.

4.1. Strengths and limitations

Strengths of the current study include the use of an empirically developed safety self-assessment tool; daily self-report measures

Table 4
Hidden correlations among daily variables using grand-mean (sample) centred data.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Driving time	1																	
2. Wheelchair time	−.03	1																
3. Bathing time	.16	.62	1															
4. Dressing time	.02	.63†	.84*	1														
5. Transfer count	−.07	.52*	.52*	.45*	1													
6. Transfer Time	−.34†	.94	.69*	.71*	.41*	1												
7. Houseclean time	.88*	−.05	−.05	−.10	−.50†	−.16	1											
8. Light objects time	.57	.02	−.01	−.01	−.19	−.03	.75	1										
9. Med objects time	.03	.82	.70*	.60*	.19	.77	.20	.36	1									
10. Hvy objects time	.31	.27	.41	.36	.02	.46	.59*	.84†	.98	1								
11. Fatigue	.44*	.43	.11	.41*	−.01	.19	.11	.34	.31	.57	1							
12. Positive affect	−.41	.17*	−.07	.42	−.06	.01	−.72	−.43	.00	−.57	−.15	1						
13. Negative affect	−.09	.24	−.05	.56†	−.11	.22	−.41†	−.14	.12	−.09	.81*	.19	1					
14. Lagged sleep	−.30	.53	.24	.58	.16	.25*	−.28	.14	−.18*	.33	.14	.66*	.20	1				
15. Stress	−.46	.09	−.09	.37	−.17	.30	−.49	−.40	.22	−.59	.54*	−.35	.74*	−.07*	1			
16. High cal food/drink sum	−.08	.06	.08	.22†	−.04	.02	−.43	−.10	.08	−.14	.44	.20†	.52*	.41	.53*	1		
17. Lagged pain	.32	.34†	.15	.31	.39	.31	.28	.55	.30	.99*	.64*	−.34	.21	−.14	−.08	.33	1	
18. Exercise (min)	−.44†	.77	.15	.14	.07	.60	−.15	−.14	.81	.23	.12	.24†	.08	.37	.28	.25	−.03	1
19. fruit/veg (serv)	.20	.41	.33	−.09	.39*	.39	−.15	.18	.20	.41†	−.03	.02	−.03	.24*	−.27	.05	.01	.24

Note: Daily work task survey items (exposures) are above the dark double line; Evening survey (symptom) items are below the dark double line. While the current table correlations were performed on grand mean centred variables, means and standard deviations for these same variables in Tables 2 and 3 were computed from uncentred data. †$p < .10$, *$p < .05$.

Table 5
Multilevel regression results predicting daily outcomes from daily work task exposures.

Daily level 1 predictors	Pain[a]	Fatigue	Negative affect	Stress
	ICC = .34	ICC = .32	ICC = .22	ICC = .30
Driving time	.01 (.003)*	.01 (.025)	.01 (.006)*	−.01 (.005)
Wheelchair time	.02 (.026)	.05 (.027)†	.01 (.005)	.00 (.011)
Dressing time	.03 (.013)*	.03 (.079)	.00 (.009)	.00 (.005)
Light objects time	−.01 (.008)	.03 (.124)	−.01 (.007)	−.03 (.005)**
Medium objects count	.18 (.214)	.57 (.680)	.02 (.013)†	.00 (.030)
Medium objects time	.03 (.013)*	.03 (.040)	.01 (.002)**	.00 (.005)
Heavy objects count	.10 (.044)*	.53 (.548)	.07 (.026)*	.24 (.071)**
Heavy objects time	.00 (.017)	.02 (.021)	.01 (.005)	.03 (.006)***
Individual level 2 predictors				
BMI	.23 (.058)***	.68 (.255)*	.22 (.123)†	.08 (.096)
Age	.05 (.024)†	−.03 (.123)	−.06 (.076)	−.08 (.035)*

Note: BMI = Body Mass Index. Values reported are the unstandardized within-person partial regression slopes (fixed effects) and robust standard errors are reported in parentheses. All level 1 variables were group-mean centred and all level 2 variables were grand-mean centred. Degrees of freedom varied from 20 to 22 across all models. Only those work task exposure variables that yielded significant relationships are reported (time spent assisting clients walking, bathing/toileting, client lifting/transferring, and house cleaning as well as the number of client lifts/transfers and light objects moving were all nonsignificant predictors of outcome variables). BMI and age were selected due to strong correlations with outcome variables.
†$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.
[a] Outcome variable was lagged in models.

that were digitally time stamped; in-home observations to estimate the reliability of worker self-assessments; and the use of new correlational techniques to evaluate work task exposures as predictors of daily outcomes. In addition, the study design allowed for multi-level modeling of relationships between person characteristics, work exposures, and evening symptoms. We also measured lifestyle behaviors, which allowed us to explore correlations between safety and health factors.

Limitations of the project suggest opportunities for future research and practice, and include the somewhat small convenience sample, omission of one-time survey measures of task exposures, timing of the daily pain measure, a lack of inter-rater reliability assessment between research observers, potential participant reactivity to research observers, and a variety of biases inherent to self-reported data. These several limitations and their potential implications are discussed below.

With regard to the nature and size of our sample, a convenience sample like ours means that participants may not have been representative of the population at large, and thus findings may not generalize to other HCWs more broadly. As for sample size, statistical modeling demonstrates that multi-level studies should produce increasingly accurate estimates of between and within person variances as the number of level 2 observations approach 30 (Hox, 2010). In this light, our sample size of 23 participants × 10 days (for a total of 242 participant days) was sufficient for producing reasonably precise estimates of these variance components. With regard to our ability to detect effects, multi-level power analysis, and power analysis for hidden correlations, are both developing areas. Most multi-level power calculators are tailored to experimental rather than descriptive designs like ours, and the exploratory nature of our study meant that we did not possess a priori effect size estimates for relationships tested in our models. In sum, we did not design the study based on power for exploratory tests of STAT validity, but instead, designed it to produce reasonable estimates of variance components in task exposures and evening symptoms among HCWs. Limitations related to sample size and selection method may be addressed through future replication studies and/or through random selection of study participants from all available members of a sub-population.

Measurement limitations may also be addressed in future research. It was a limitation that the study did not include a global one-time survey self-assessment of daily work task exposures that could be compared to parallel daily exposure measures collected with the STAT. It would be valuable for future researchers to contrast the accuracy of both types of self-reported task exposures so that the added value of EMA methods can be determined. Also, we may have observed larger or different effects of exposures on pain if we had analyzed post-shift instead of morning pain. In past studies pain is lower and less variable in the morning compared to right after the end of a shift (See Amick et al., 2003). As in similar studies of worker self-assessments (e.g., Unge et al., 2005), for practical reasons we did not evaluate inter-rater reliability between researcher observers, which means that the degree of calibration among researcher observers is unknown. Observations were conducted in client's private homes, and we opted to send only a single research observer to maximize client acceptance of observations, and also to minimize crowding and obstructed views inside small rooms (e.g., bathrooms). However, future research would be enhanced by the inclusion of inter-rater reliability assessments among researcher observers during training and/or in the field.

While in-home observations were a strong point of the study, participants may have reacted to researcher presence by temporarily increasing the accuracy of their self-monitoring, or by performing different work tasks than normal. While Olson et al. (2009) demonstrated that truck drivers' behavior was not reactive to the onset of a self-monitoring procedure during video camera observations, an in-person observation is more intrusive than a camera. If participants were reactive, this may have resulted in inflated reliability estimates for the STAT, or in workers performing unrepresentative work on those study days. While the 10–14 day sample period protects against the problem of unrepresentative descriptive STAT results, future studies could protect further against inflated reliability by increasing the number of researcher observation days per participant to allow for habituation to the procedure.

4.2. Discussion of descriptive findings

Descriptive findings from both retrospective surveys and daily STAT measures of exposures were generally aligned with existing evidence that HCWs are a population in need of health and safety interventions. Our small convenience sample was predominantly female, obese, had a high prevalence of musculoskeletal symptoms, and reported low levels of fruit and vegetable consumption. With regard to task exposures and physical symptoms, housekeeping consumed about 40% of caregivers' self-reported work task time.

Given the significant time spent housekeeping, interventions to reduce ergonomic demands and awkward postures during housekeeping could generate meaningful reductions in musculoskeletal symptoms for HCWs. However, exposures to less frequent but more physically demanding tasks remain important. For example, in the current study manual handling of household objects was associated with daily pain, and in other studies manual handling of objects and assisting client mobility were associated with lost time injuries (Baron and Habes, 2004; Kim et al., 2010; Meyer and Muntaner, 1999; Myers et al., 1993; Wipfli et al., 2012). With regard to dietary findings, low fruit and vegetable consumption is associated with increased risk of stroke, heart disease, cancer, and mortality (Hung et al., 2004; Neuhouser et al., 2003). However, we note that our fruit and vegetable measure was a single question created for the current study, and may not provide an estimate as valid as those produced by more sophisticated measures (e.g., Thompson et al., 2002).

4.3. Discussion of reliability findings

In our analyses of in-home observations, kappa coefficients for occurrence/non-occurrence of events between workers' and observers' measurements approached .60, which exceeds traditional standards for acceptable reliability (acceptable kappa coefficient $\kappa > .40$; Landis and Koch, 1977). And, agreement ICCs for task exposure minutes and counts exceeded .70. Mean differences in minutes between workers' and observers' recordings varied in absolute value from 1.8 to 12.6 min. For surveillance purposes, these mean differences seem tolerable for producing actionable assessments of HCWs task exposures. Moreover, these mean differences are comparable to previous diary style ergonomic self-assessment data collected with nurses and office workers (Unge et al., 2005). These findings suggest that, when participant and observer data are pooled and contrasted at the sample level, the STAT is moderately reliable for measuring HCWs' task exposures.

While STAT reliability during in-home observations was acceptable, variance in reliability across measures may inform future research and practice. In general, workers overestimated task durations, but this tendency was magnified for the potentially unpleasant task of bathing and toileting. This elevated discrepancy was partly explained by one worker who recorded bathing and toileting time when the researcher observer recorded no bathing and toileting time, which likely occurred prior to the researcher's arrival, but was recorded by the worker during the in-home observation. Alternatively, the researcher may have missed the occurrence of the task. However, the overestimation of time spent performing physically demanding tasks such as bathing/toileting and dressing clients is consistent with systematic bias observed in prior research (e.g., Barrero et al., 2009b). Such biases could result in overestimation of task exposures in physically demanding tasks. Also, if the magnitude of a certain bias varied greatly across workers, then this error variance might reduce statistical power for detecting actual cross-level associations between worker characteristics and daily level exposure–symptom associations. Testing for biases such as these in the future is important because ergonomic self-assessment inherently focuses on physically demanding work. A simple method that may attenuate such bias would be to build timers into computerized tools, or provide workers with a manual timer rather than having them rely on clocks for self-monitoring. A variety of off-the-shelf time tracking applications for smart phones could be used to measure time-based exposures (e.g., Time Tracker Pro application for iPhone/iPod Touch).

While the current study did not explicitly test whether STAT measures of work task exposures were more accurate or representative than a one-time worker self-assessment survey, we expect that EMA methods like the STAT would help workers produce more accurate information because self-recordings occur in close temporal proximity to the events being assessed. This logical probability and the initial reliability of the STAT should encourage researchers and practitioners to use the STAT to measure task exposures among HCWs. Daily process measurement methods also open the door for studying more complex and multi-level relationships between worker characteristics, daily work tasks, physical and psychological states, and health behaviors.

4.4. Discussion of validity findings

Hidden correlations and multi-level models showed that several task exposures measured by the STAT had significant associations with daily pain, including driving time, dressing time, and moving objects. If our assumption that daily pain is a risk factor for long-term injury and disability is correct, these findings support the STAT as a valid ergonomic exposure assessment tool. In support of this assumption, the tasks associated with daily symptoms in the current study have also been associated with workers' concern for injury and lost time injuries in previous research (Baron and Habes, 2004; Wipfli et al., 2012). In addition to effects of work task exposures on daily symptoms, person-level factors of BMI and age were associated with daily symptoms. We also observed one significant cross-level interaction between BMI and the daily-level association between moving medium objects and negative affect. These findings suggest that some person-level factors may make HCWs more susceptible to certain safety and health hazards. Our analyses also identified associations between daily symptoms and lifestyle behaviors. For example, in multi-level models measures of manual material handling significantly increased daily negative affect and stress. In hidden correlations, these symptom variables (negative affect, stress) were also significantly associated with high calorie food/drinks. These findings suggest that home care task exposures may generate physical or psychological symptoms that have indirect negative effects on healthful behaviors. Taken together, these findings suggest that the STAT measures task exposures that are related to Total Worker Health™ (the integration of injury prevention with health promotion; NIOSH, 2012).

4.5. Discussion of the usability and future dissemination of the STAT

The Home Care STAT, and other hand-held computer apps used for worker self-assessments, will be most likely to be adopted and have an impact in the home care industry if they are easy to use. On this theme, most workers reported few problems using the device, and found the touch screen survey interface to be very user friendly. The version of the STAT used in the current study was programmed with SurveyToGo software, which can be purchased off the shelf and has a great visual interface for programming logic for a large array of question types. However, the software is actually designed to be used by researchers to collect data via interviews. This means that a login is required when the app is launched, and if a participant accidentally logs themselves out, they have to enter a username and password provided by researchers. We also found the data export to be problematic for questions where a user could select multiple options. Rather than reserving a unique column for each possible answer, the software only created a data column for items selected. This meant that each exported dataset had to be manipulated extensively before data analysis. So, the application was easy to program and easy for workers to use (unless they logged themselves out), but it did require a high level of expertise to analyze exported data. Future versions of the software, or alternative smart phone applications, may enhance the STAT for dissemination and impact in the home care industry.

5. Conclusions

The current project was designed to evaluate a new safety self-assessment tool for HCWs. Background surveys and workers' self-assessments confirm that HCWs are a priority population for ergonomic assessment and intervention research. The predominantly female sample was older, obese, experienced a high prevalence of musculoskeletal symptoms, and reported regularly performing dangerous and demanding tasks such as client transfers. In-home observations demonstrated that workers produced moderately reliable self-assessments of both task occurrence and duration using the STAT. The daily diary design and an evening symptom survey showed that exposures to tasks such as dressing clients and moving/lifting household objects were related to daily levels of pain, occupational fatigue, negative affect, and stress. We also observed significant hidden correlations between postulated outcomes of negative affect and stress with health factors, suggesting that work exposures in home care may have indirect effects on worker health behaviors. Limitations in the current study suggest areas for future research and practice, such as providing workers with devices with built-in timers to enhance the completeness and reliability of self-assessments. Altogether, our findings support the STAT as a reliable and valid exposure assessment tool for HCWs that researchers and practitioners are encouraged to use to measure HCWs' task exposures, evaluate risk, and guide intervention selection.

Funding

Notes: This project was supported with funding from the National Institute for Occupational Safety & Health (grant # 5 T42 OH008433-03, Noah Seixas, PI; and grant # 1 U19 OH010154-01, Kent Anger, PI). The funding source had no involvement in designing and conducting the study or in preparing and submitting the manuscript. The authors declare no conflicts of interest.

Acknowledgments

The authors would like to acknowledge Joanne Lees and David Meier for their contributions to formative research and data analysis, and the SEIU Local 503 and the Oregon Home Care Commission for guidance and logistical support.

References

Administration on Aging. Aging statistics. 2011. Available from:: http://www.aoa.gov/AoARoot/Aging_Statistics/index.aspx.
Ahsberg, E., 2000. Dimensions of fatigue in different working populations. Scand. J. Psychol. 41 (3), 231–241.
Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., O'Brien, W.L., Bassett Jr., D.R., Schmitz, K.H., Emplaincourt, P.O., Jacobs Jr., D.R., Leon, A.S., 2000. Compendium of physical activities: an update of activity codes and MET intensities. Med. Sci. Sports Exerc. 32 (9), S498–S516.
Amick III, B.C., Robertson, M.M., Bazzani, L., DeRango, K., Moore, A., Rooney, T., Harrist, R.H., 2003. Effect of office ergonomics intervention on reducing musculoskeletal symptoms. Spine 28 (24), 2706–2711.
Askew, R., Walker, J.T., 2008. Ergonomics for homecare providers. Home Healthc. Nurse 26 (7), 412–417.
Baron, A.D., Habes, D.J., 2004. Health hazard evaluation report #2001-0139-2930: Alameda county public authority for in-home support services. Available from: www.cdc.gov/niosh/hhe/reports/pdfs/2001-0139-2930.pdf.
Barrero, L.H., Katz, J.N., Dennerlein, J.T., 2009a. Validity of self-reported mechanical demands for occupational epidemiologic research of musculoskeletal disorders. Scand. J. Work Environ. Health 35 (4), 245–260.
Barrero, L.H., Katz, J.N., Perry, M.J., Krishnan, R., Ware, J.H., Dennerlein, J.T., 2009b. Work pattern causes bias in self-reported activity duration: a randomized study of mechanisms and implications for exposure assessment and epidemiology. Occup. Environ. Med. 66 (1), 38–44.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Methodol. 57, 289–300.
Benjamini, Y., Hochberg, Y., 2000. On the adaptive control of the false discovery rate in multiple testing with independent statistics. J. Educ. Behav. Stat. 25, 60–83.
Bolger, N., Davis, A., Rafaeli, E., 2003. Diary methods: capturing life as it is lived. Annu. Rev. Psychol. 54, 579–616.
Bureau of Labor Statistics U.S, 2010. Department of Labor. Nonfatal Occupational Injuries and Illnesses Requiring days Away from Work, 2009. Available from: http://bls.gov/iif/.
Bureau of Labor Statistics U.S, 2012. Department of Labor. Occupational Outlook Handbook, 2012–13 ed. Available from: http://www.bls.gov/ooh/healthcare/home-health-and-personal-care-aides.htm.
Buxton, O.M., Quintilliani, L.M., Yang, M.H., Ebbeling, C.B., Stoddard, A.M., Pereira, L.K., Sorensen, G., 2009. Association of sleep adequacy with more healthful food choices and positive workplace experiences among motor freight workers. Am. J. Public. Health 99 (S3), S636–S643.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psy. Meas. 20, 37–46.
Cohen, S., Kamarck, T., Mermelstein, R., 1983. A global measure of perceived stress. J. Health Soc. Behav. 24 (4), 385–396.
Diener, E., Emmons, R.A., Larsen, R.J., Griffin, S., 1985. The satisfaction with life scale. J. Person. Assess. 49 (1), 71–75.
Galinsky, T., Waters, T., Malit, B., 2001. Overexertion injuries in home health care workers and the need for ergonomics. Home Health Care Serv. Q. 20 (3), 57–73.
Green, A.S., Rafaeli, E., Bolger, N., Shrout, P.E., Reis, H.T., 2006. Paper or plastic? Data equivalence in paper and electronic diaries. Psychol. Meth. 11 (1), 87–105.
Hayashi, R., Gibson, J.W., Weatherly, R.A., 1994. Working conditions in home care: a survey of Washington state's home care workers. Home Health Care Serv. Q. 14 (4), 37–48.
Hess, J.A., Kincl, L.D., Mandeville, D.S., 2007. Three single-person manual patient transfer techniques for bed-to-wheelchair transfers. Home Healthc. Nurse 25 (9), 3–9.
Hignett, S., McAtamney, L., 2000. Rapid entire body assessment (REBA). Appl. Ergon. 31, 201–205.
Hollmann, S., Klimmer, F., Schmit, K.-H., Kylian, H., 1999. Validation of a questionnaire for assessing physical work load. Scand. J. Work Environ. Health 25, 105–114.
Hox, J.J., 2010. Multilevel Analysis. Techniques and Applications, 2nd ed. Routledge, New York.
Hung, H.-C., Joshipura, K.J., Jiang, R., Hu, F.B., Hunter, D., Smith-Warner, S.A., Colditz, G.A., Rosner, B., Spiegelman, D., Willett, W.C., 2004. Fruit and vegetable intake and risk of major chronic disease. J. Natl. Cancer Inst. 96 (21), 1577–1584.
Jellema, P., Bierma-Zeinstra, S.M.A., van Poppel, M.N.M., Bernsen, R.M.D., Koes, B.W., 2002. Feasibility of lumbar supports for home care workers with low back pain. Occup. Med. 52 (6), 317–323.
Johnston, D.W., Beedie, A., Jones, M.C., 2006. Using computerized ambulatory diaries for the assessment of job characteristics and work-related stress in nurses. Work Stress 20 (2), 163–172.
Kim, I., Geiger-Brown, J., Trinkoff, A., Muntaner, C., 2010. Physically demanding workloads and the risks of musculoskeletal disorders in homecare workers in the USA. Health Soc. Care Community 18 (5), 445–455.
Kuorinka, I., Jonsson, B., Kilbom, A., Vinterberg, H., Biering-Sorensen, F., Andersson, G., et al., 1987. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl. Ergon. 18 (3), 233–237.
Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.
Myers, A., Jensen, R.C., Nestor, D., Rattiner, J., 1993. Low back injuries among home health aides compared with hospital nursing aides. Home Health Care Serv. Q. 14, 149–155.
Meyer, J.D., Muntaner, C., 1999. Injuries in home health care workers: an analysis of occupational morbidity from a state compensation database. Am. J. Ind. Med. 35, 295–301.
Nelson, A., Fragala, G., Menzel, N., 2003. Myths and facts about back injuries in nursing. Am. J. Nurs. 103 (2), 32–40.
Neuhouser, M.L., Patterson, R.E., Thornquist, M.D., Omenn, G.S., King, I.B., Goodman, G.E., 2003. Cancer epidemiology. Biomarkers Prev. 12, 350–358.
NIOSH., 2012. Research Compendium: the NIOSH Total Worker Health™ Program: Seminal Research Papers 2012 (DHHS (NIOSH) Publication No. 2012-146). U.S. Department of Health and Human Services, Washington, DC. Public Health Service, Centers for Disease Control and Prevention Retrieved from. http://www.cdc.gov/niosh/docs/2012-146/.
Nguyen, T., Jiang, J., 2011. Simple estimation of hidden correlation in repeated measures. Stat. Med. [Internet]. http://dx.doi.org/10.1002/sim.4366. Available from:
Olson, R., Hahn, D.I., Buckert, A., 2009. Predictors of severe trunk postures among short-haul truck drivers during non-driving tasks: an exploratory investigation involving video-assessment and driver behavioral self-monitoring. Ergonomics 52, 707–722.
Ono, Y., Loagerstrom, M., Hagberg, M., Linden, A., Malker, B., 1995. Reports of work related musculoskeletal injury among home care service workers compared with nursery school workers and the general population of employed women in Sweden. Occup. Environ. Med. 52, 686–693.
Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., Congdon, R.T., 2004. HLM 6: Hierarchical Linear and Nonlinear Modeling. Scientific Software International, Lincolnwood, IL.
Simon, M., Tackenberg, P., Nienhaus, A., Estryn-Behar, M., Conway, P.M., Masselhorn, H.-M., 2004. Back or neck-pain-related disability of nursing staff in

hospitals, nursing homes and home care in seven countries: results from the European NEXT-Study. Int. J. Nurs. Stud. 45 (1), 24–34.

Spielholz, P., 2006. Calibrating Borg scale ratings of hand force exertion. Appl. Ergon. 37 (5), 615–618.

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.

Spector, P.E., Jex, S.M., 1998. Development of four self-report measures of job stressors and strain: interpersonal conflict at work scale, organizational constraints scale, quantitative workload inventory, and physical symptoms inventory. J. Occup. Health Psy. 3 (4), 356–367.

Stone, A.A., Schwartz, J.E., Neale, J.M., Shiffman, S., Marco, C.A., Hickcox, M., Paty, J., Porter, L.S., Cruise, L.J., 1998. A comparison of coping assessed by ecological momentary assessment and retrospective recall. J. Pers. Soc. Psychol. 74, 1670–1680.

Thompson, F.E., Subar, A.F., Smith, A.F., Midthune, D., Radimer, K.L., Kahle, L.L., Kipnis, V., 2002. Fruit and vegetable assessment: performance of 2 new short instruments and a food frequency questionnaire. J. Am. Diet. Assoc. 102, 1764–1772.

Trask, C., Teschke, K., Morrison, J., Village, J., Johnson, P., Koehoorn, M., 2010. Using observation and self-report to predict mean, 90^{th} percentile, and cumulative low back muscle activity in heavy industry workers. Ann. Occup. Hyg. 54 (5), 595–606.

Tuomi, K., Ilmarinen, J., Eskelinen, L., Jarvinen, E., Toikkanen, J., Klockars, M., 1991. Prevalence and incidence rates of diseases and work ability in different work categories of municipal occupations. Scand. J. Work Environ. Health 37 (6), 455–463.

Unge, J., Hansson, G., Ohlsson, K., Nordander, C., Axmon, A., Winkel, J., Skerfving, S., 2005. Validity of self-assessed reports of occurrence and duration of occupational tasks. Ergonomics 48, 12–24.

Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Person. Soc. Psychol. 47, 127–144.

Wiktorin, C., Karlqvist, L., Winkel, J., 1993. Validity of self-reported exposures to work postures and manual materials handling. Scand. J. Work Environ. Health 19, 208–214.

Wipfli, B., Olson, R., Wright, R.R., Garrigues, L., Lees, J., 2012. Characterizing hazards and injuries among home care workers. Home Healthc. Nurse 30, 387–393.

Zeytinoglu, I.U., Denton, M.A., Webb, S., Lian, J., 2000. Self-reported musculoskeletal disorders among visiting and office home care workers. Work Health 31 (2–3), 1–35.