Construct validity and test-retest reliability of the revised Upper Extremity Work Demands (UEWD-R) Scale

Cavalini, Miriam A.; Berduszek, Redmar J.; van der Sluis, Corry K.

Published in:
OCCUPATIONAL AND ENVIRONMENTAL MEDICINE

DOI:
10.1136/oemed-2017-104370

Publication date:
2017

ORIGINAL ARTICLE

Construct validity and test–retest reliability of the revised Upper Extremity Work Demands (UEWD-R) Scale

Miriam A Cavalini, Redmar J Berduszek, Corry K van der Sluis

Department of Rehabilitation Medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Correspondence to
Redmar J Berduszek, University of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine. Hanzeplein 1, 9713 GZ,Groningen, The Netherlands; r.j.berduszek@umcg.nl

Received 14 February 2017
Revised 19 May 2017
Accepted 30 May 2017
Published Online First
8 July 2017

ABSTRACT

Background The revised Upper Extremity Work Demand (UEWD-R) Scale is a six-item self-report questionnaire to measure the workload of the upper limbs. UEWD-R consists of a force/posture scale and a repetition scale. Psychometric properties are unknown so far.

Purpose Assess the construct validity and the test–retest reliability of UEWD-R.

Method Participants from different levels of physical work demands (based on the Dictionary of Occupational Titles (DOT) categories) were included. Construct validity was determined by testing 11 predefined hypotheses regarding UEWD-R related to other constructs, including a workplace observation using the Rapid Upper Limb Assessment (RULA). Correlations between these measures were calculated using Spearman correlation coefficients. Test–retest reliability was determined using the intraclass correlation coefficient (ICC) for agreement. The smallest detectable change (SDC) was calculated.

Results Fifty-four participants participated (63% men, mean age 39.4 years). The four DOT categories were equally represented. Nine out of 11 predefined correlations were confirmed (82%), indicating good construct validity. Strong expected correlations of UEWD-R-total versus RULA-C (r=0.69) and UEWD-R-repetition versus RULA-muscle (r=0.12) were not confirmed. The test–retest reliability was good (ICC agreement=0.79). The SDC was 4.85.

Conclusion Construct validity and the test–retest reliability of UEWD-R were good. UEWD-R can be used to evaluate the workload of the upper extremities. However, further research is advised to assess the validity of the UEWD-R not only by testing associations with RULA,but also with other observational measures.

What this paper adds

- There is no gold standard to evaluate the workload of the upper extremities. The Upper Extremity Work Demand Scale (UEWD) measured a good reliability, but the criterion validity came short.
- The construct validity and the test–retest reliability of revised UEWD (UEWD-R) were good.
- UEWD-R can be clinically used to evaluate the physical workload of the upper extremities.

INTRODUCTION

Work-related factors such as high intensity of exertion, awkward body postures, repetitive movements and a short recovery time increase the risk of developing work-related musculoskeletal disorders (WRMSDs).[1–4] Musculoskeletal complaints of the upper extremities are substantially contributing to WRMSDs.[1 5] With the increasing incidence of WRMSDs,[1] the reduction in work productivity and work absence[2 6–8] and the subsequent high healthcare costs,[2 9] a method is needed to assess the physical work demands of the upper extremities.

Several valid observational methods are available to assess the workload of the upper extremities, for example, Strain Index (SI),[3] Occupational Repetitive Actions (OCRA),[10] Threshold Limit Value for Hand Activity Level (TLV for HAL)[4] or Rapid Upper Limb Assessment (RULA).[11] However, there is no gold standard.[12] Moreover, observational methods need schooled observers and consume much time, which has implications for costs and feasibility. Limiting the number and duration of observations may reduce validity.[13]

Using questionnaires, a wider period of time can be studied, and all tasks of the job can be covered. Questionnaires could measure physical workloads that more accurately reflect the general workload.[13] Additionally, questionnaires are more efficient and cheaper to use than observations and, as such, can be used as screening methods and in larger studies.[7 13 14] However, questionnaires can be less accurate and less reliable than observational methods. Awareness of the body posture and movements is difficult, leading to over or underestimating the physical demands. Also, questions can be interpreted differently than how it was intended by the developers.[13] Moreover, the reproducibility of detailed questions (eg, postures of specific body regions and exerted force) is lower than of general questions.[13]

Despite many studies, there still is no valid questionnaire that evaluates the actual workload of the upper extremities.[7] For this reason, the Upper Extremity Work Demand (UEWD) Scale was developed.[6] The physical work demand items of the upper extremities were extracted from the Dutch musculoskeletal questionnaire[14] to establish the UEWD. A previous study concerning UEWD measured a good reliability, but the criterion validity came short.[15] UEWD was proposed to be adapted to a revised version. Factor analysis distinguished two

To cite: Cavalini MA, Berduszek RJ, van der Sluis CK. *Occup Environ Med* 2017;**74**:763–768.

factors, namely the force exertion/awkward posture factor and the repetition factor, resulting in two subscales, 'force/posture subscale' and 'repetition subscale'. Furthermore, two questions were deleted because of a high correlation with other questions, and one question was rephrased.[15] This revised (R) version, from now on called UEWD-R, requires further testing of reliability and validity.[15] Therefore, the aim of this study is to assess the construct validity and the test–retest reliability of UEWD-R.

METHODS

Participants

A convenience sample of workers from a university hospital, a general hospital, an orthopaedic workshop and a cardboard factory was recruited between October and December 2016. They were included if they were 18 years of age or older, currently working and had sufficient command of the Dutch language. Based on the Dictionary of Occupational Titles (DOT), employees with various physical demands were approached to participate in the study. The intention was to include an equal number of participants from the different categories of the DOT.[6] The DOT categorises jobs according to workload from 'sedentary' (DOT 1) to 'very heavy' (DOT 5). Only workers from DOTs 1–4 could be included, because workers who are categorised as DOT 5 were not available. In the Netherlands, employees are not allowed to carry weights above 25 kg, which is one of the criteria of DOT 5.

Procedure

To assess construct validity, participants filled in a questionnaire, including UEWD-R, RAND 36-item Health Survey (RAND-36) physical functioning subscale, Maastricht Upper Extremity Questionnaire (MUEQ) posture subscale and information on current work ability (0: 'completely unable to work' to 10: 'work ability at its best') and general health perception (0: 'poor' to 5: 'excellent'). After the questionnaires were completed, a video-recording was performed at the participants' workplace during execution of their regular work tasks. Video recordings were used to compute RULA scores and lasted about 30 min, depending on the work tasks (longer if after 30 min a task with another physical demand would start). Objects that were lifted or pushed during the execution of the work tasks were measured with a weighing scale. When weighing was not possible, an estimation was made (less than 2 kg, 2–10 kg or more than 10 kg, conform RULA-force score). Lastly, it was asked whether the observation of the work tasks was representative for the participants' regular job activities. The videos were analysed by the first author (MC). Blinding was guaranteed as much as possible by first scoring RULA and subsequently entering the data of UEWD-R into the database.

Test–retest reliability was determined by completing UEWD-R twice. To reduce recall bias and avoid substantial change in the work tasks, the period in between both measurements lasted 1–3 weeks.[16] If there was no response after 7 days, a reminder was sent.

For both construct validity and test–retest reliability a sample size of at least 50 participants is advised as a sufficient sample size.[16]

Upper Extremity Work Demand-Revised

UEWD was originally developed by Opsteegh *et al*[6] and revised by Jacobs *et al*.[15] UEWD-R is a six-item questionnaire enquiring after the physical effort that participants exert with the upper extremities during their work. It is divided into two subscales, UEWD-R-force/posture and UEWD-R-repetitive (appendix 1). Each question is answered on a 4-point Likert scale, where 1 reflects 'rarely or never' and 4 'almost always'. The scores of the six questions are added up, where a higher score represents a higher workload of the upper extremities. Scores of the UEWD-R-total can range from 6 to 24.

Maastricht Upper Extremity Questionnaire

MUEQ is a 95-item questionnaire about the potential risk factors for WRMSDs of the upper extremities, containing the following aspects: work station, posture during work, quality of break time, job demands, job control, social support, quality of the work environment and presence of complaints of the upper extremities.[17] The questions have to be answered on a 5-point Likert scale, where 1 reflects 'always' and 5 reflects 'never'. MUEQ has proven to have sufficient validity and reliability with computer office workers in different countries. The intraclass correlation coefficient (ICC) calculated for the separate subscales differed between 0.52 and 0.94.[17–20] The seven-item posture subscale was selected, which is further divided into an 'awkward body posture' and 'head and body posture' subscale (Cronbach's alpha: 0.71 and 0.54, respectively).[21]

RAND-36

RAND-36 is a 36-item reliable and valid questionnaire about physical, mental and social health and is used worldwide to measure health-related quality of life.[22] We used the 10-item subscale 'physical functioning' of the valid Dutch translation.[22] Each question can be answered with 'not limited at all' (1), 'slightly limited' (2) or 'severely limited' (3).

Rapid Upper Limb Assessment

RULA is a valid and reliable observational method for the upper extremities that was developed to investigate risk factors for the WRMSDs of the upper extremities.[12] RULA has proven to be a good screening tool, with an intraobserver repeatability of 91.7%[23] and an interobserver repeatability of 94.6%.[24] RULA is divided into subscores RULA-C and RULA-D. These subscores are computed from posture, muscle and force scores of the upper and lower arm and the wrist (RULA-C) or neck, trunk and legs (RULA-D), respectively. Force and muscle scores can also be used as separate outcomes. The combination of RULA-C and RULA-D leads to RULA-grand score. RULA-grand score is a risk score that indicates if a change in the work conditions should be made.

Predefined hypotheses

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) group recommended to create predefined hypotheses to calculate the construct validity.[16] Eleven predefined hypotheses (table 1) were constructed to estimate the strength of the correlation between UEWD-R-total, both subscales and MUEQ-posture, RAND-36-physical functioning, RULA-C, RULA-D, RULA-force and RULA-muscle. Correlations were defined as weak (0.0–0.3), low (0.3–0.5), moderate (0.5–0.7) or high (0.7–1.0).

The predefined hypotheses were based on the construct of the instruments involved. UEWD-R-total and RULA-C both concern repetition, force and posture of the upper extremities, whereas RULA-D concerns comparable items but body parts that are not covered in the UEWD-R-total. The construct of MUEQ-posture is also comparable with UEWD-R-total, but the questions are more generally asked and not specific for the upper extremities.

Table 1 Spearman correlations between UEWD-R, RULA, MUEQ and RAND-36

	RULA-C	RULA-D	RULA-muscle	RULA-force	MUEQ-posture	RAND-36-physical functioning
UEWD-R-total	**0.69***	**0.65***	**–**	**–**	**0.30****	**0.16**
UEWD-R-force/posture	0.67*	0.63*	–	0.74*	0.292**	–
UEWD-R-repetition	0.44*	0.40*	0.12	–	0.290**	–
Predefined hypotheses about the correlations						
	Reference test		Confirmed when the correlation is			Confirmed
1. UEWD-R-total versus	RULA-C		High			No
2. UEWD-R-total versus	RULA-C		Stronger than UEWD-R-total versus RULA-D			Yes
3. UEWD-R-total versus	RULA-C		Stronger than UEWD-R-total versus MUEQ-posture			Yes
4. UEWD-R-total versus	MUEQ-posture		Low to moderate			Yes
5. UEWD-R-total versus	MUEQ-posture		Stronger than UEWD-R-total versus RAND-36-physical functioning			Yes
6. UEWD-R-total versus	RAND-36-physical functioning		Weak to low			Yes
7. UEWD-R-force/posture versus	RULA-C		Stronger than UEWD-R-repetition versus RULA-C			Yes
8. UEWD-R-force/posture versus	RULA-D		Stronger than UEWD-R-repetition versus RULA-D			Yes
9. UEWD-R-force/posture versus	RULA-force		High			Yes
10. UEWD-R-force/posture versus	MUEQ-posture		Stronger than UEWD-R-repetition versus MUEQ-posture			Yes
11. UEWD-R-repetition versus	RULA-muscle		High			No

*Significant correlation p<0.01 (two tailed) **significant correlation p<0.05 (two tailed).
MUEQ, Maastricht Upper Extremity Questionnaire; RULA, rapid upper limb assessment; RULA-muscle, muscle score for RULA-A; RULA-force, force score for RULA-A; UEWD-R, Revised Upper Extremity Work Demands Scale.

RAND-36-physical functioning measures physical restrictions of activities in daily life, which is a different construct than UEWD-R-total.

The constructs of RULA-C, RULA-D and MUEQ-posture are mostly based on the exerted forces and postures during work and less on repetition. Therefore, their constructs are more comparable with UEWD-R-force/posture than to UEWD-R-repetition. Also, RULA-muscle is based on repetitive movements, as is UEWD-R-repetition. RULA-force is based on the exerted force, as is UEWD-R-force/posture.

Statistical analysis

The statistical analyses were performed with IBM SPSS Statistics for Windows, V.22.0.

Construct validity

Spearman correlation coefficients were used to determine if the predefined hypotheses could be confirmed, because the questionnaires and RULA scores were considered as ordinal data. If more than 75% of the predefined hypotheses are confirmed, it can be concluded that the construct validity is good.[16]

Reproducibility

Agreement

The limits of agreement were presented in a Bland-Altman plot. These limits were put at the mean of the difference between the repeated measures±1.96 × the SD of the difference.[25]

The SE of measurement (SEM) was calculated (SEM=SD difference x √2). Herewith, the smallest detectable change (SDC) was calculated to see if the agreement of UEWD-R-total was sufficient (SDC=1.96 x √2 x SEM). When the SDC is smaller than the minimal important change (MIC), the agreement is sufficient.[26]

Test–retest reliability

For the test–retest reliability, the ICC for absolute agreement was calculated. This measure includes the error for comparing two different measurements. An ICC agreement >0.70 was regarded to reflect a sufficient test–retest reliability.[26]

RESULTS

Demographic characteristics

A total of 54 participants were included (table 2); they were equally distributed over the four DOT categories (table 3).

Construct validity measurements were completed for 50 participants. Four participants (all with a profession classified in DOT 4) did not return the second questionnaire. Therefore, we recruited four additional participants in the DOT 4 category, who fulfilled UEWD-R twice.

Construct validity

Nine (82%) of the 11 predefined hypotheses were confirmed (table 1). Confirmation of more than 75% was

Table 2 Participant characteristics

	Total (n=54)	Construct validity (n=50)	Test–retest reliability (n=50)
Age (years), *mean (SD)*	39.4 (13.0)	40.6 (12.8)	40 (13.0)
Gender, *n (%)*			
Male	34 (63)	30 (60)	30 (60)
Female	20 (37)	20 (40)	20 (40)
Dominant hand, *n (%)*			
Left	9 (17)	9 (18)	9 (18)
Right	45 (83)	41 (82)	41 (82)
General health perception, *n (%)*			
Excellent	13 (24)	10 (20)	13 (26)
Very good	23 (43)	22 (44)	21 (42)
Good	17 (31)	17 (34)	15 (30)
Moderate	1 (2)	1 (2)	1 (2)
Poor	0 (0)	0 (0)	0 (0)
Work ability perception, *mean (SD)*	8.4 (1.1)	8.4 (1.1)	8.4 (1.1)

Table 3 Participants per DOT category (n=54)

DOT category (n)	DOT 1 sedentary work (14)	DOT 2 light work (12)	DOT 3 medium work (13)	DOT 4 heavy work (15)
Profession (n)	Researcher (8) Secretary (6)	Isolation supervisor (1) Physical therapist assistant (1) Prosthetist (7) Catering assistant (1) Speech therapist (2)	Isolation contributor (2) Painter (2) Physical therapist (3) Manual therapist (1) Kitchen helper (3) Cook (2)	Stucco worker (1) Construction worker (6) Machine operator (4) Order picker (4)*

*Only contributed to the test–retest reliability investigation.
DOT, Dictionary of Occupational Titles.

achieved, indicating a good construct validity. We did not find the expected high correlation (>0.70) between UEWD-R-total and RULA-C (r=0.69) nor did we find a high correlation between UEWD-R-repetition and RULA-muscle (r=0.12). The mean duration of the observations was 30.5 (SD 4.6) minutes. For four participants, the observation was terminated prematurely (after 18, 20, 27 and 28 min, respectively), due to a low camera battery (n=2) and refusal of recording of patients treated by two physical therapists. All participants indicated that the observed work tasks represented their general work tasks. However, eight participants, for example, physical therapists and constructions workers, explained that in their job they had varying work tasks, signifying alternating physical demands that were not recorded. Therefore, the observation did not reflect their entire workload. A repeated analysis with exclusion of these eight participants showed comparable correlation coefficients.

Reproducibility

The UEWD-R was completed for the second time after an average of 12.7 (SD 7.5) days, range 7–41 days. Five participants replied later than the preferred 3 weeks.

Agreement

The limits of agreement (–4.15 and 5.59), as shown in the Bland-Altman plot (figure 1), showed a random distribution, with one outlier. The SEM was 1.75 and the SDC was 4.85.

Test–retest reliability

The ICC agreement was 0.79 (95% CI 0.66 to 0.88), which signifies a good test–retest reliability of the UEWD-R.

DISCUSSION

A good construct validity and test–retest reliability of UEWD-R were found. These findings suggest that UEWD-R can be used to screen the physical work demands of the upper extremities and, with that, offer an alternative for more resource demanding observational methods.

Construct validity

Previously Jacobs *et al*[15] calculated the criterion validity of UEWD using direct observations of upper limb work demands. These authors found a low correlation (r=0.44) and concluded that direct

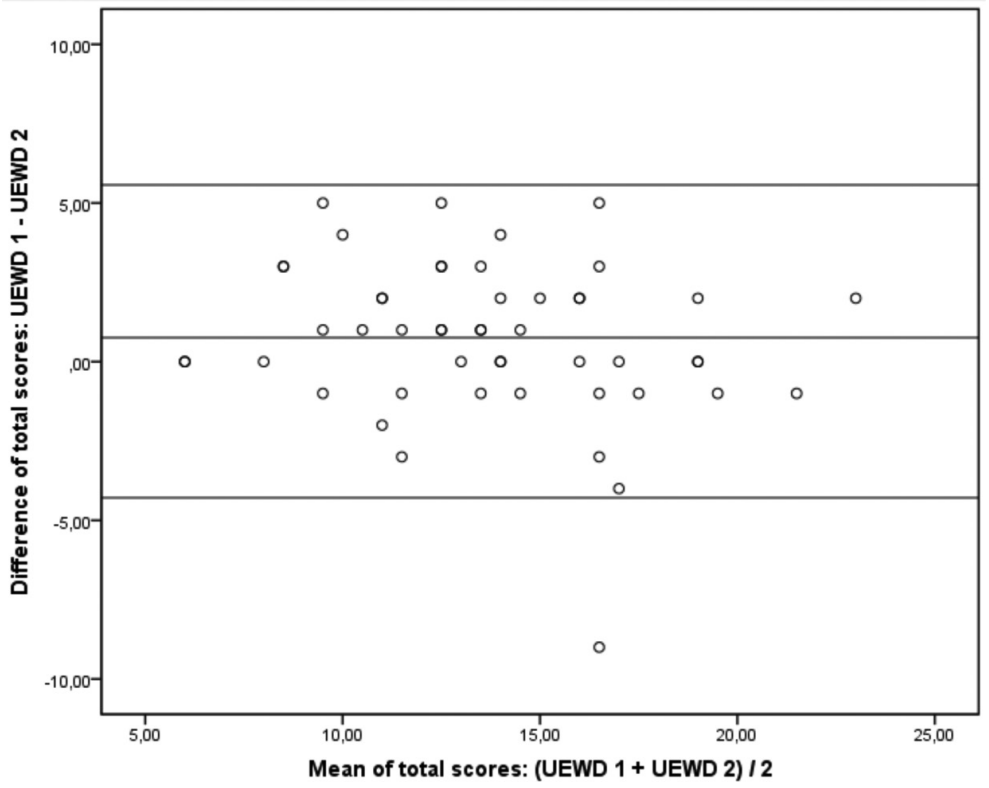

Figure 1 Bland-Altman plot. Difference between the sum scores of first and second applications of UEWD-R against the mean sums scores of first and second applications of UEWD-R. UEWD-R, Revised Upper Extremity Work Demands.

observation appeared not to be the finest method. As a result, we improved the study design by using RULA as a valid observational method to evaluate the workload of the upper extremities.[11] Different observational methods were considered for the current study: RULA, OCRA, SI and, TLV for HAL. SI and TLV for HAL comprise assessments for the distal upper extremities, whereas OCRA and RULA take both the distal and proximal parts of the upper extremities into account. Different studies have compared these observational methods. Two studies showed that SI estimates a higher risk for WRMSDs than TLV for HAL, but they did not define which one had the best risk estimation.[27 28] Another study showed that SI and RULA were only weakly correlated.[29] Furthermore, a review showed that OCRA correlated moderately with SI and TLV for HAL.[12] To conclude, there is no gold standard for the observation of the workload on the upper extremities.[12] Since SI and TLV for HAL only concern the distal upper extremities and OCRA concerns irrelevant additional aspects, we decided RULA seemed to measure most closely the same construct as UEWD-R. With confirmation of 82% of the predefined hypotheses, the construct validity of UEWD-R appeared to be good.[16] With UEWD-R, we have developed a screening method for the physical workload of the upper extremities. When a high physical workload is measured by UEWD-R, more specific assessments can be used, for example, TLV for HAL or SI.

We expected a strong correlation between UEWD-R-total and RULA-C because of their corresponding constructs. Both measurements evaluate awkward postures and repetitive movements of the upper extremities and in both measurements exertion of forces are classified. Although the correlation coefficient was statistically significant and a high correlation was nearly reached, this hypothesis had to be rejected ($r=0.69$ instead of $r>0.70$). A short observation of work task execution may not give a full representation of all job tasks. As such, the correlation with UEWD-R-total, where a judgement of the general workload is asked, might be lower than expected. Although the hypothesis was rejected, the correlation was strong enough to imply that UEWD-R-total indeed measures the physical workload of the upper extremities. The finding that UEWD-R-total correlated stronger with RULA-C than with MUEQ-posture and RAND-36-physical restriction underlines this assumption.

UEWD-R-repetition and RULA-muscle did also not correlate as presumed. UEWD-R-repetition only consists of questions on repetitive movements, where RULA-muscle also takes other movements into account. When a posture is statically held and no repetitive movements are made, RULA-muscle score is high, whereas UEWD-R-repetition score is low. This might well explain the much weaker correlation than expected and in hindsight we should have defined this hypothesis differently.

As predicted, both subscales of UEWD-R correlated higher with RULA-C than with RULA-D. RULA-C evaluates the upper and the lower arm, whereas RULA-D evaluates the neck, trunk and legs. However, the differences were small (force/posture $r=0.67$ vs $r=0.63$ and repetition $r=0.44$ vs $r=0.40$). During the observations we noticed that when a high force is exerted by the upper extremities, frequently the neck, trunk and sometimes also the legs were burdened. We did not considered this in our hypotheses beforehand, and this may explain the comparable correlations between UEWD-R and its both subscales versus RULA-C and RULA-D.

Reproducibility

With an ICC of 0.79, the test–retest reliability is above the recommended 0.70 and therefore regarded to be good.[26] The limits of agreement ranged from −4.15 to +5.59 (UEWD-R-total ranges: 6–24), implying that these differences might possibly be due to the measurement error.[26] The SDC of 4.85 signifies that a score difference of more than 4.85 is required to identify a true difference in UEWD-R-total score. The ICC, limits of agreement and SDC are conform the previously calculated measured values for UEWD.[15] Clinical relevance can be determined when the MIC for UEWD-R is defined.

Strengths and weaknesses

The first strength of our study is that we measured the construct validity using predefined hypotheses. Different constructs were compared and expected correlations were determined, leading to accuracy. Second, we succeeded to include a sufficient number of participants from a diverse working population of both sexes and a broad age range. Therefore, generalisability of UEWD-R seems to be appropriate. Third, we used a valid observation method to measure the workload.[11] Lastly, all RULA scores were rated by the same researcher, preventing inter-rater variability. However, some limitations should be noted as well. The predefined hypotheses were based on the authors' insights, due to a lack of relevant information in the literature. Hence, the hypotheses could have been composed too strong or too weak. Furthermore, four observations lasted shorter than 30 min. However, the participants with the shorter observation times indicated that continuation of their work would not alter the content of their tasks. Therefore, we do not expect that these participants' results substantially influenced the outcomes.

Clinical relevance and future perspectives

Our results contribute to a valid and reliable questionnaire to evaluate the workload of the upper extremities. With UEWD-R, one of the risk factors of developing WRMSDs of the upper extremities can be evaluated easily and cost-effectively. Repetitive movements without assessment of great levels of force do not necessarily signify an increased risk for developing WRMSDs.[30] Therefore, further research should investigate if there is indeed an association between UEWD-R and the presence of WRMSDs of the upper extremities.

CONCLUSION

UEWD-R appeared to be the first valid and reliable questionnaire to evaluate the workload of the upper extremities. In contrast to UEWD, UEWD-R can be clinically used as a screening method for the physical workload of the upper extremities. However, further research is advised to assess the validity of UEWD-R not only by testing the association with RULA scales, but also with other observational measures. Also, the clinical value of UEWD-R in relation to WRMSDs of the upper extremities should be investigated.

Acknowledgements We would like to thank P.U. Dijkstra, MSc, PhD, for his help with the statistics.

Contributors MAC executed the study and conducted the statistical analysis. MAC wrote the first draft of the manuscript with further contributions from CKvdS and RJB. All authors interpreted the data, reviewed and edited the manuscript and approved the final version of the manuscript.

Competing interests None declared.

Patient consent Patient details have been anonymised.

Ethics approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Schneider E, Irastorza X. *Work EA for S and H at, Copsey S. Work-related musculoskeletal disorders in the EU — Facts and figures. European risk observatory report*. Luxembourg: Pubications Office of the European Union, 2010. https://osha.europa.eu/en/tools-and-publications/publications/reports/TERO09009ENC. (Accessed 23 Sept 2017).
2 Waters TR. National efforts to identify research issues related to prevention of work-related musculoskeletal disorders. *J Electromyogr Kinesiol* 2004;14:7–12.
3 Moore JS, Garg A. The Strain Index: a proposed method to analyze jobs for risk of distal upper extremity disorders. *Am Ind Hyg Assoc J* 1995;56:443–58.
4 Latko WA, Armstrong TJ, Foulke JA, *et al*. Development and evaluation of an observational method for assessing repetition in hand tasks. *Am Ind Hyg Assoc J* 1997;58:278–85.
5 Van ED, Munhall C, Irvin E, *et al*. Effectiveness of workplace interventions in the prevention of upper extremity musculoskeletal disorders and symptoms: an update of the evidence. *Occup Env Med* 2016;73:62–70.
6 Opsteegh L, Soer R, Reinders-Messelink HA, *et al*. Validity of the dictionary of occupational titles for assessing upper extremity work demands. *PLoS One* 2010;5:e15158.
7 Dale AM, Strickland J, Gardner B, *et al*. Assessing agreement of self-reported and observed physical exposures of the upper extremity. *Int J Occup Environ Health* 2010;16:1–10.
8 Roy JS, Desmeules F, MacDermid JC. Psychometric properties of presenteeism scales for musculoskeletal disorders: a systematic review. *J Rehabil Med* 2011;43:23–31.
9 Bongers PM, Ijmker S, van den Heuvel S, *et al*. Epidemiology of work related neck and upper limb problems: psychosocial and personal risk factors (part I) and effective interventions from a bio behavioural perspective (part II). *J Occup Rehabil* 2006;16:272–95.
10 Occhipinti E. OCRA: a concise index for the assessment of exposure to repetitive movements of the upper limbs. *Ergonomics* 1998;41:1290–311.
11 McAtamney L, Nigel Corlett E, Corlett EN. RULA: a survey method for the investigation of work-related upper limb disorders. *Appl Ergon* 1993;24:91–9.
12 Takala EP, Pehkonen I, Forsman M, *et al*. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scand J Work Environ Health* 2010;36:3–24.
13 Stock SR, Fernandes R, Delisle A, *et al*. Reproducibility and validity of workers' self-reports of physical work demands. *Scand J Work Environ Health* 2005;31:409–37.
14 Hildebrandt VH, Bongers PM, van Dijk FJ, *et al*. Dutch musculoskeletal questionnaire: description and basic qualities. *Ergonomics* 2001;44:1038–55.
15 Jacobs NW, Berduszek RJ, Dijkstra PU, *et al*. Validity and reliability of the Upper Extremity Work Demands Scale. *J Occup Rehabil* 2016.
16 Terwee CB, Bot SD, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
17 Eltayeb SM, Staal JB, Hassan AA, *et al*. Complaints of the arm, neck and shoulder among computer office workers in Sudan: a prevalence study with validation of an arabic risk factors questionnaire. *Environ Health* 2008;7:33.
18 Bekiari EI, Lyrakos GN, Damigos D, *et al*. A validation study and psychometrical evaluation of the Maastricht Upper Extremity Questionnaire (MUEQ) for the Greek-speaking population. *J Musculoskelet Neuronal Interact* 2011;11:1108–7161.
19 Turci AM, Bevilaqua-Grossi D, Pinheiro CF, *et al*. The brazilian portuguese version of the revised Maastricht Upper Extremity Questionnaire (MUEQ-Br revised): translation, cross-cultural adaptation, reliability, and structural validation. *BMC Musculoskelet Disord* 2015;16:41.
20 Ranasinghe P, Perera YS, Lamabadusuriya DA, *et al*. Work-related complaints of arm, neck and shoulder among computer office workers in an asian country: prevalence and validation of a risk-factor questionnaire. *BMC Musculoskelet Disord* 2011;12:1471–2474.
21 Eltayeb S, Staal JB, Kennes J, *et al*. Prevalence of complaints of arm, neck and shoulder among computer office workers and psychometric evaluation of a risk factor questionnaire. *BMC Musculoskelet Disord* 2007;8:68.
22 VanderZee KI, Sanderman R, Heyink JW, *et al*. Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. *Int J Behav Med* 1996;3:104–22.
23 Kee D, Karwowski W. A comparison of three observational techniques for assessing postural loads in industry. *Int J Occup Saf Ergon* 2007;13:3–14.
24 Breen R, Pyper S, Rusk Y, *et al*. An investigation of children's posture and discomfort during computer use. *Ergonomics* 2007;50:1582–92.
25 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
26 de Vet HCW, Terwee CB, Mokkink LB, *et al*; *Measurement in medicine: a Practical Guide. Practical guides to biostatistics and epidemiology*. Cambridge: Cambridge University Press, 2011.
27 Spielholz P, Bao S, Howard N, *et al*. Reliability and validity assessment of the hand activity level threshold limit value and strain index using expert ratings of mono-task jobs. *J Occup Environ Hyg* 2008;5:250–7.
28 Bao S, Spielholz P, Howard N, *et al*. Quantifying repetitive hand activity for epidemiological research on musculoskeletal disorders--part I: individual exposure assessment. *Ergonomics* 2006;49:361–80.
29 Drinkaus P, Sesek R, Bloswick D, *et al*. Comparison of ergonomic risk assessment outputs from rapid Upper Limb Assessment and the Strain Index for tasks in automotive assembly plants. *Work* 2003;21:165–72.
30 Harris-Adamson C, Eisen EA, Kapellusch J, *et al*. Biomechanical risk factors for carpal tunnel syndrome: a pooled study of 2474 workers. *Occup Environ Med* 2015;72:33–41.

Construct validity and test–retest reliability of the revised Upper Extremity Work Demands (UEWD-R) Scale

Miriam A Cavalini, Redmar J Berduszek and Corry K van der Sluis

Occup Environ Med 2017 74: 763-768 originally published online July 8, 2017
doi: 10.1136/oemed-2017-104370