

Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: www.tandfonline.com/journals/terg20

Analytical comparison of selected international tools for the risk assessment of upper-limbs biomechanical overload in repetitive tasks

D. Colombini, E. Occhipinti, R. Fox, D. Moresco & M. Candoli

To cite this article: D. Colombini, E. Occhipinti, R. Fox, D. Moresco & M. Candoli (26 Feb 2026): Analytical comparison of selected international tools for the risk assessment of upper-limbs biomechanical overload in repetitive tasks, Ergonomics, DOI: 10.1080/00140139.2026.2618597

To link to this article: https://doi.org/10.1080/00140139.2026.2618597

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 26 Feb 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=terg20

ERGONOMICS
https://doi.org/10.1080/00140139.2026.2618597

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Analytical comparison of selected international tools for the risk assessment of upper-limbs biomechanical overload in repetitive tasks

D. Colombini[a], E. Occhipinti[a], R. Fox[b], D. Moresco[a] and M. Candoli[a]

[a]Scientific Association EPMIES (Ergonomics of Posture and Movement International Ergonomics School), Milan, Italy; [b]Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan, USA

ABSTRACT

This study compared eight upper-extremity risk assessment tools applied to five different manual tasks. The aim is to provide users with insight into how the tools compare and to contribute to the ongoing revision of the ISO 11228-3 Standard. Eight upper limbs risk assessment tools were compared (ACGIH® HA, DUET, EAWS4, HARM, OCRA index, OCRA checklist, RSI, and RULA). The predictive or concurrent validity of each tool was described and considered. Tool merits and limitations, such as underestimating high action frequencies, static postures, or upper limb areas to study, were described. The ability of the tools to distinguish risk levels was examined. All tools differentiated the risk level of the five tasks, except HARM and RULA, which showed similar risk levels for all tasks. Suggestions are provided to prospective users for choosing tools appropriate to their application, with an emphasis on predictive validity to ensure more reliable risk assessment results.

Practitioner Summary: The objective of this paper is to provide guidance on the selection of the most suitable tools. The authors compared eight tools (ACGIH® HA, DUET, EAWS4, HARM, OCRA index, OCRA checklist, RSI, and RULA) by applying them to five representative tasks. A tool with good predictivity is more likely to provide reliable results in terms of risk assessment.

ARTICLE HISTORY
Received 10 May 2025
Accepted 10 January 2026

KEYWORDS
Work-related musculoskeletal disorders of the upper limbs (UL-WMSDs); risk assessment; biomechanical overload, comparison of methods; predictive validity

1. Introduction and aim

Work-related musculoskeletal disorders of the upper limbs (UL-WMSDs) are prevalent in the working population, with approximately 60% of European workers reporting different musculoskeletal disorders in the last 12 months (EU-OSHA (European Agency for Safety and Health at Work)) 2019).

Due to the high prevalence of UL-WMSDs, risk-reduction should begin with risk screening and assessment to guide prevention efforts.

1.1. General framework for risk assessment

According to a *Consensus Document*, which was prepared and published by the International Ergonomics Association (IEA), the Technical Committee on Musculoskeletal Disorders, and the International Commission on Occupational Health (ICOH) (Colombini et al. 2006), at least four main risk factors should be analysed to obtain a complete risk assessment of the occupational biomechanical overload of the upper limbs. These four risk factors are repetitiveness, forces, awkward postures and movements, and a lack of proper recovery periods. These factors should be assessed as a function of time (mainly considering their respective durations). In addition to these factors, others grouped under the term 'additional risk factors' should be considered. These include mechanical factors (e.g. vibrations and localised mechanical compressions), environmental factors (e.g. exposure to cold), and organisational factors (e.g. work pace determined by machinery). Separately describing each risk factor helps identify specific requirements, early preventive actions, and assess each factor's contribution to overall exposure. From this viewpoint, the numerical or categorical classification of results can help risk management (Lind and Rose 2016).

1.2. Brief review of previous comparative studies

Currently, many observation-based tools are used to assess hazards related to repetitive movements and

CONTACT Colombini Daniela daniela.colombini50@gmail.com Occupational Medical Doctor, Euro Erg, President of Scientific Association EPMIES (Ergonomics of Posture and Movement International Ergonomics School), Cadore 2, 20135 Milan, Italy

Supplemental data for this article can be accessed online at https://doi.org/10.1080/00140139.2026.2618597.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

exertion of the upper limbs (Eliasson, Lind, and Nyman 2019; Rose et al. 2020). A non-exhaustive list can be found in IEA TC on the MSD (Musculo Skeletal Disorders) website (IEA 2024).

There are also studies that compare the tools and their output metrics. It is beyond the scope of the present study to refer analytically to all the various tools.

Several studies have indicated that the results from different observation-based risk assessment tools often diverge (Chiasson et al. 2012; Nyman et al. 2023, Valentim et al. 2025). This was synthesised in a recent review (Joshi and Deshpande 2019) which included a comparison of 18 evaluation tools found in 39 papers applied in a variety of industrial sectors. This work summarises these comparative studies, reporting, in percentage, the number of comparative studies available for each of the 18 tools: REBA, Rapid entire body assessment tool: 22%; RULA, Rapid upper limbs assessment: 21%; SI, Strain Index: 12%; OWAS, Ovako Working Posture Analysing System: 11%; OCRA, Occupational repetitive actions: 10%; QEC, Quick exposure check: 7%. Joshi and Deshpande (2019) concluded that the number of available publications related to such comparisons is limited. Ergonomic assessment tools have been applied in a variety of sectors. The results of comparative studies have demonstrated that there is variation in the outcomes in most cases. The cause of this deviation was unidentified. Nyman et al. (2023) reached a similar conclusion about the agreement between the classified outputs of commonly used observational tools by comparing six tools for risk assessment of manual repetitive work.

Although these studies are valuable, there remains a lack of in-depth analysis regarding the differences among tool outputs, particularly in the context of repetitive tasks, highlighting the necessity for further investigation.

1.3. Aim

The aim of this study is to compare all the characteristics of eight, specifically selected, observational-based, risk assessment tools for evaluating repetitive work through the analysis of the differences in risk categories/classifications generated by these tools when applied to five different representative manual repetitive tasks. The purpose is also to make an applicative contribution to ISO 11228-3 Standard: Ergonomics – Manual handling – Part 3. Repetitive movements and exertions of the upper limbs, currently under revision, which offers a list of the most suitable tools for studying biomechanical overload of the upper limbs.

Numerous papers compare various methods; however, most of the comparisons are based on bibliographic reviews.

The utility and the originality of this study is that it compares the results obtained by these risk assessment tools (eight tools of which seven are part of the main tools proposed in this ISO 11228-3 draft Standard, and an additional tool chosen because it is widely used) when they are simultaneously applied to the study of the same representative manual repetitive tasks (five tasks, different in terms of biomechanical overload level for the upper limbs, from low to very high, especially in terms of frequency or strength demands).

2. The selected risk assessment tools for evaluation of manual repetitive tasks

To identify potential tools for risk assessment of repetitive work of the upper limbs, a two-step process was applied:

a. First, ergonomic experts of the ISO working group (ISO/TC 159/SC 3/WG 4 - Human physical strength: manual handling and force limits) reviewed standards and research literature to find tools that specifically targeted repetitive work of the upper limbs. The tools presented in 11228–3:2007 (ISO, 2007) and ISO Technical Report 12295 (ISO, 2014) were considered, together with other tools (explicitly targeting repetitive work of the upper limbs) from literature reviews (e.g. Beliveau et al. 2022; Chiasson et al. 2012; Eliasson, Lind, and Nyman 2019; Graben et al. 2022; Joshi and Deshpande 2019; Lind et al. 2013, 2021; Lowe, Dempsey, and Jones 2019; Nyman et al. 2023; Oakman et al. 2022; Rose et al. 2020; Takala et al. 2010) and ergonomic experts.
b. Second, based on a consensus procedure, the expert group discussed and iteratively developed a list of characteristics that had to be met for the tool to be included as a real risk assessment tool (Table 1). The list includes the following:

Table 1. Suggested characteristics of tools for evaluation of upper limb repetitive movements.

- to consider the relevant risk factors for UL-WMSDs including force, posture, repetition, frequency, recovery periods, movements, duration, or derivatives thereof (e.g. duty cycle)
- to consider the different areas of the upper limbs, (i.e. fingers, hands, wrists, forearms, elbows, and shoulders). when studying awkward postures
- easy to use not only for ergonomists and similar experts but also for non-ergonomists experts such as managers and safety representatives.
- not requiring costly or complex measurement instruments or equipment
- including procedures for multitask analysis of jobs rotations
- having predictive validity or, at least, concurrent validity

- Considerations regarding the main UL-WMSDs risk factors as documented in a previous Consensus Document (Colombini et al. 2006), and usability related aspects related to risk assessment of UL-WMSDs risk factors. These aspects include ensuring that the tool can assess more than just one body part, that it is usable by ergonomists as well as managers and safety representatives, and that it provides a straightforward method for analysing multiple tasks, such as in job rotation.
- Consideration about the tool *predictive validity* (i.e. demonstrated relationship with UL-WMSDs within an adult working population in at least one epidemiological study, either longitudinal, prospective, retrospective, or a large cross-sectional study, and having clearly defined health outcomes) or at least presence of a *concurrent validity* (i.e. the output metrics of the tool exhibit a high level of agreement with another tool that has established predictive validity), (Takala et al. 2010).

Seven risk assessment tools met the inclusion criteria and were included in this study. To these, the authors added the Rapid Upper Limbs Assessment (RULA) tool, (McAtamney and Corlett 1993), despite not meeting the inclusion criteria, given its vast application worldwide in the evaluation of repetitive exertions and movements. The eight risk assessment tools considered in this study are:

- ACGIH® Hand Activity TLV (Threshold Limit Value)- (revised ACGIH® HA), (ACGIH® (American Conference of Governmental Industrial Hygienists)) 2022),
- Ergonomic Assessment Worksheet-Section 4 (EAWS4), (Schaub et al. 2013),
- Distal Upper Extremity Tool (DUET) (Gallagher et al. 2018; Bandekar et al. 2021),
- Hand-Arm Risk Assessment method (HARM) (Douwes and De Kraker 2014),
- Occupational Repetitive Actions Checklist (OCRA ck) (Colombini et al. 2011; Colombini and Occhipinti 2017; ISO, 2014; Occhipinti and Colombini 2004),
- Occupational Repetitive Actions Index (OCRA in), (Colombini and Occhipinti 2017; Occhipinti 1998),
- Revised Strain Index (RSI), (Garg, Moore, and Kapellusch 2017, 2017b),
- Rapid Upper Limb Assessment (RULA), (McAtamney and Corlett 1993).

All these methods appear in the literature to have already been validated; therefore, this paper does not intend to proceed with further validation studies.

In Appendix A, at the end of the paper, the main characteristics of the eight selected tools are reported. In this final section, we thought it useful to first provide a descriptive summary of the eight selected tools, then to focus on a more detailed analysis of the differences presented within each of the risk factors considered.

In the different subsections of the Appendix, the following topics are covered: a brief description of the main characteristics of the selected tools is provided along with the relevant bibliography; the estimated 'weight' that each tool assigns to different risk factors when deciding its own final risk score; details about the criteria used for evaluating the single risk factors as frequency (of actions/exertions/movements), and static posture. Table 2 summarises the principal risk factors (Colombini et al. 2006) addressed by each tool.

Considering now the predictive characteristic of the selected tools, Table 3 reports details about their predictive or concurrent validity. The results are, when available, primarily expressed in terms of odds ratios (OR), which provide a measure of the strength of association and the trends between level of exposure and the occurrence of UL-WMSDs.

Table 2. Risk factors considered by the different tools.

	ACGIH® HA	DUET	EAWS4	HARM	OCRA ck	OCRA in	RSI	RULA
Frequency	x	x	x	x	x	x	x	x
Force	x	x	x	x	x	x	x	x
Shoulder Posture	–	–	x	x	x	x	–	x
Elbow Posture	–	x	x	x	x	x	–	x
Hand/wrist Posture	–	x	x	x	x	x	x	x
Kind of grasp	–	–	x	–	x	x	–	–
Stereotypy	–	–	+/–	–	x	x	–	–
Additional risk factors	–	–	x	x	x	x	–	–
Lack of recovery periods	–	–	x	x	x	x	–	–
Net repetitive work duration	>4 h	–	x	x	x	x	x	–
Multitask analysis / jobs rotations	x	x	x		x	x	x	–

Table 3. Predictivity features of selected tools with supporting studies.

Case numbers of data base	Risk areas	OR - central value (low-upper)	Predictive or concurrent validity studies
ACGIH® HA	≥AL - ≤TLV (3.6–5.6)	1.16 (0.73–1.85)	**Predictive validity** All the MSD of distal UE. Validation studies for CTS clinically diagnosed (symptoms and EMG)) using the revised HA TLV (after 2018) (Yung et al., 2019) CTS and tendinosis (Harris et al., 2011; Bonfiglioli et al., 2013; Kapellusch et al, 2002) in studies prior to 2018.
No.4321	> TLV (> 5.6)	1.99 (1.28–3.10)	
DUET		1.62 (1.32–2.09)	**Predictive validity:** only based on cross-sectional data.
No.772			Statistically significant association with distal upper extremity outcomes (injury/pain), dose-response relationship. (Gallagher et al. 2018). Statistically
			significant association between DUET Log CD and physician-diagnosed CTS (Bandekar et al. 2021).
	Concurrent validity:*Fair* agreement (Cohen's kappa = 0.299) with RSI based on risk category, and strong correlation (Spearman's rank correlation coefficient = 0.933, $p < 0.001$) for risk factor exposure and output metrics, (Jorgensen et al., 2024).		
EAWS4			**Predictive validity:** No studies with direct predictivity were identified.
	Concurrent validity. The Author of the tool claims a total concurrent validity with OCRA Index based on a preliminary study in the automobile sector (45 workstations) comparing EAWS4 with OCRA tool (Lavatelli et al., 2012): the correlation between the respective scores was high (in the work only OCRA results in the green/yellow risk category are compared with EAWS4). However, results of a more recent large comparison of OCRA Checklist and EAWS4 in the automobile sector (about 300 workstations), (Occhipinti and Ghibaudo, 2019) indicated only a partial concurrence.		
HARM	≥ 50	4.57 (1.34–15.57)	**Predictive validity** Self-reported neck and upper limb symptoms in a follow-up study involving a small group of workers (Douwes et al, 2014).
No.88			
OCRA Index	2.3–3.5	2.16 (1.45–3.23)	**Predictive validity:** Prediction function: PA (Prevalence of affected subjects by one or more UL- WMSDs) = 2.39 (± 0.14) x OCRA INDEX, (adjusted $R^2 = 0.92$).
No. 5372	3.6–4.5	3.74 (2.38–5.89)	
	4.6–9	5.30 (3.63–7.78)	Predictivity for all the UL-WMSDs clinically diagnosed (Occhipinti and Colombini, 2007; Nicoletti et al., 2008).
OCRA Ck	7.6–11.0	2.18 (1.69–2.8)	**Predictive validity:** Prediction function: PA (Prevalence of affected subjects by one or more UL-WMSDs) = 0.742 (± 0.055) x OCRA Ck Score. (adjusted $R^2 = 0.86$).
N0.11734	11.1–16.0	2.77 (2.25–3.41)	
	16.1–22.55	4.36 (3.63–5.24)	Predictivity for all the UL-WMSDs clinically diagnosed (Occhipinti and Colombini, 2021).
RSI	8.6–15.0	1.42 (0.96–2.09)	**Predictive validity:** All the MSD of distal UL. Validation study for Carpal Tunnel Syndrome (CTS) clinically diagnosed (symptoms and EMG), (Kapellusch et al., 2021).
No.1372	> 15	1.79 (1.19–2.69)	
RULA	The relationship of the individual RULA body part scores to the development of pain or discomfort is statistically significant for the neck and lower arm scores ($p < 0.01$) and not significant for the trunk, upper arm or wrist scores (McAtamney L, Corlett N., 1993). **Concurrent validity:** no identified studies; **Predictive validity:** no identified studies.		

The data reported in Tables 2 and 3 are very useful and important for users in deciding which tool to use. Starting from the intrinsic characteristics of the tools (Table 2), it results useful to highlight and underline the degree of depth with which these tools address the study of the biomechanical overload of the upper limbs. We can see how some tools address the study of the risk factors, indicated in the table, more completely than others.

The OCRA tools are the most comprehensive, together with EAWS4. The OCRA tools have also a high predictive validity, supported by a numerically impressive database (OCRA in No. 5372 and OCRA ck No. 11734). It seems important to point out that the larger the database that supports a high level of predictive validity, the higher the significance of the results, the more reliable the chosen tool in defining the achievable risk level. EAWS4, while analysing all the different risk factors, does not present data for supporting its predictive validity, having only a partial concurrent validity versus OCRA tools.

Other tools present a high predictive validity for wrist/hand pathologies (ACGIH® HA, RSI, DUET) but offer the possibility of studying only the distal part of the upper limbs. They are not completely exhaustive regarding the study of the job organisation (exposure times, pauses), data needed for better modulating the final risk score.

OCRA tools will be chosen as the reference tool for the comparison with the other tools, due to the following main characteristics: response to all the risk factors necessary for the estimation of biomechanical overload (including the additional factors as cold temperature, vibration, golpes etc.) and high predictive validity.

3. Methods

No peculiar ethical approval was required. The study was carried out with compliance with the Declaration of Helsinki, and the participants performing the tasks provided informed consent. In any case, because some

of the authors of this paper were the subjects involved in carrying out the 5 tasks described, they have given their full consent to the inclusion of the images and videos in the paper itself.

3.1. Selected tasks for evaluation

To compare the risk scores of the eight risk assessment tools, five distinct repetitive tasks were selected for the analysis. Since most of the selected tools do not estimate awkward shoulder postures, to compare the evaluation results, tasks were selected in which the shoulders operated in a suitable posture.

Each task was designed to illustrate varying degrees of operational involvement and musculoskeletal commitment, representative of common industrial jobs. The tasks were analysed through video recordings (see Supplementary Material), which allowed a detailed examination of each task by analysing one work cycle. To ensure participant privacy and ethical compliance, all tasks were reconstructed in a controlled laboratory environment, performed by some of the authors of this paper, avoiding complications associated with using video footage from actual work settings.

The five repetitive tasks chosen for this study were selected based on their relevance to industrial work and their potential to cause musculoskeletal disorders. The tasks are summarised in Table 4a, which provides a brief description of each task, along with the specific operational characteristics.

All the potential risks presented are summarised in Table 4b. The tasks are represented in a video clip capturing only one complete work cycle to facilitate the analysis by reducing the complexity of the risk assessment due to inter-cycle exposure variations. This approach allows for controlled comparison across various assessment tools.

The tasks were performed by the authors themselves. They verbally consented to participate in the evaluation. No other external participants were involved in performing the tasks. As a preliminary activity necessary to fully classify the tasks, the participants were interviewed to obtain information on their perceived level of exertion using the Borg CR10 scale,

Table 4a. Description of the repetitive work tasks that were analysed.

Task		Given description	Right upper limb	Left upper limb
1		Very high frequency of very similar actions (one take/position every second), very low force (0,5 in Borg scale CR10) and pinch grip type (pinch or precision grip), lack of variation in gestures. Rhythm self-paced; Cycle time: 60 sec. (a deck of cards). Additional notes: The left hand is in static posture (in pinch) during the full work cycle. (*see* Movie TASK 1).	**Frequency** very high **Force** Very low **Kind of grasp** pinch	**Frequency** Static posture 100% of time **Force** Very low **Kind of grasp** pinch
2		Manual assembly. For the last two screwing of the piece, during its screwing the worker judged the force exertion as *moderate* (3 on Borg CR10 scale). Cycle time: 11.6 sec. (*see* Movie TASK 2).	**Frequency** Medium/high **Force** Low, moderate for short time **Kind of grasp** pinch	**Frequency** low **Force** low **Kind of grasp** pinch
3		The work consists of completing the assembly of two components and then packing them. When screwing the two components together the worker judged the force exertion as *light* (*2* on Borg CR10 scale). Cycle time: 27 sec. (*see* Movie TASK 3).	**Frequency** medium **Force** low **Kind of grasp** pinch	**Frequency** low **Force** low **Kind of grasp** pinch
4		Preparation of an electric cable: the different copper wires, covered in plastic, placed inside the cable, are prepared: first cutting the coloured plastic surface that covers them, then tearing off this coloured plastic surface, exposing the internal copper part. When tearing the plastic coating, the worker judged the force exertion as *strong* (*5* on Borg CR10 scale (5% of cycle time). Cycle time: 45 sec. (*see* Movie TASK 4).	**Frequency** Static posture 60% of time **Force** strong for almost 5% of the time **Kind of grasp** Power grip	**Frequency** Static posture 60% of time **Force** low **Kind of grasp** pinch
5		Screwing with an electric screwdriver, components on an appliance placed on an assembly line. Rhythm imposed by the machine. For the placement of the second screw, the worker judged the force exertion as *moderate* (*3* on Borg CR10 scale). Cycle time: 18 sec. (*see* Movie TASK 5).	**Frequency** Static posture 100% of time **Force** Low, moderate for short time **Kind of grasp** Power grip	**Frequency** low **Force** Very low **Kind of grasp** pinch

Table 4b. Summary of the main potential risk factors present in the tasks analysed.

Right limb
- Task 1: very high frequency, fingers in 'pinch' and very light force,
- Task 2: medium/high frequency, fingers in 'pinch' and very light force,
- Task 3: medium/low frequency, fingers in 'pinch' and very light force,
- Task 4: presents force peaks for at least 5% of the time, fingers in 'pinch',
- Task 5: static posture in 'power grip' for 100% of the time with negligible force,

Left limb. To better understand the illustrated trend, it is worth recalling that,
- Task 1: presence of static posture with fingers in 'pinch' for the whole time,
- Task 2 and Task 3: low frequency, with very light force and fingers in 'pinch',
- Task 4: static posture for the entire time, with light force and fingers in 'pinch',
- Task 5 medium/low frequency with fingers in 'pinch' grip.

Table 5. Description of work schedule.

	8-hour shift	4-hour shift
Shift start-end	8.00 am to 5 pm	8.00 am to 12 am
Lunch time (out of the shift duration)	60 min	absent
Shift duration	480 min	240 min
Net duration of repetitive task	410 min	210 min
Auxiliary tasks: cleaning, etc.	40 min	20 min
Breaks	2 of 15 min, morning and afternoon	1 of 10 min.

a validated tool for assessing perceived physical exertion (Borg 1998).

3.2. Risk evaluation of selected tasks

The five repetitive tasks were evaluated using all the eight ergonomic assessment tools. Each work task was assessed by a group of two to three experts, including the authors of this article, who were experienced in using these specific tools. The assessment process involved individual evaluations, followed by a group discussion to reach a consensus on the final ratings for each task. In general, the individual ratings for each task were equal or very close to the group final rating. When the scores were assigned to a risk category (for each tool), according to classification criteria reported in Appendix A, the inter-rater reliability resulted in an almost perfect agreement (Cohen K: > 0.8 – 1).

To support consistency across the experimental conditions, the net durations of the shift as well as the duration and distribution of breaks were predetermined and kept constants for all five tasks. This standardisation ensured that the effects of the shift length and break distribution could be accurately assessed without confounding variables. Risk assessments were performed for both 8-hour and 4-hour shifts (see Table 5), allowing for a comparative analysis of how shift duration may affect risk exposure. The assessment was conducted separately for both the right and left upper limbs, resulting in 20 independent risk assessments (ten for full-time shifts and ten for part-time shifts).

3.3. Data processing and analysis for comparison

Each risk assessment tool generates its own unique numerical risk score, which allows the classification of the risk into different risk categories (see Appendix A). Given that the risk categories scores differed for the different tools, a harmonisation process was necessary to facilitate and allow comparisons. To this end, the following steps were applied:

a. *Risk Score Calculation*. For each of the five tasks, the risk score was calculated separately for the right and left limbs. The risk score reflects the assessed risk level for each task.

- The *risk scores range* defined by each tool are reported in Table 6 (part a). The Table illustrates in detail, for each tool being analysed, the numerical values describing the respective risk areas. These values are also complemented by the colours, from green to purple, as specifically defined for each tool by their authors, thus indicating the progression of the exposure level with respect to the numerical parameter represented. The colours represent: no increased risk (green), very low risk or borderline (yellow), presence of low risk (orange), medium/high risk (red), and very high risk (purple). Notably, only the OCRA index and OCRA ck, EAWS use purple to distinguish between high and very high risks. All the tools analysed define the yellow colour as low risk or borderline, except the ACGIH® HA method, which instead defines yellow as presence of significant risk. When graphically displaying the results, in having to compare the evaluation results of the different tools, this choice could have created interpretative confusion. For this reason, the ACGIH® HA intermediate scores (between AL and TLV) are assigned orange: the beginning of a risk band. These colours, specific to each tool in defining the risk category, will be maintained even when the results of the risk assessment will be presented using harmonised risk values, in percentage, represented by histograms. From Table 6 (part a) two values are extracted: the Minimum Risk and Maximum Risk score,

Table 6. The values of scores describing the risk areas as proposed for the various tools by their authors and the minimum and maximum risk scores.

Part a

	Negligile	Very low or borderline	Low risk	Medium/High	Very high
OCRA in	<2.3	2.3–3.5	3.6–4.5	4.6–9	≥9.1
OCRA ck	<7.6	7.6–11	11.1–16	16.1–22.5	≥22.6
ACGIH® HA	<3.6 TLV		≥3.6 TLV ≤5.6	5.7–7.7	≥7.8
RSI	<8.5	8.5–10	10,1–15	15.1–24	≥25
HARM	<30		30–50	>50	≥70
EAWS4	=<25	25.1–50	50.1–79		≥80
DUET	<25%	25%–40%	41–50%	50.1%–79%	≥80%
RULA	1–2	3–4	5–6		≥7

Part b

	ACGIH®HA	DUET	EAWS4	HARM	OCRA in	OCRA ck	RSI	RULA
Minimum risk level	3.6	25%	50.1	30	2.3	7.6	8.5	5
Maximum risk level	7.8	80%	80	70	9,1	22,6	25	7

essential values for the subsequent harmonisation process:

- *Minimum Risk score* is here defined as the lowest score at which the author of each tool establishes the start of the risk (yellow area, very low o borderline risk, when present, or orange area when the yellow area is absent). When the score is lower than the Minimum Risk Score (no risk, green area), the result in percentage would have a negative sign. To avoid unnecessary negative values when there is no risk, a 0% default score is assigned.
- *Maximum Risk score* is defined as the score indicative of a high level of exposure (but not the highest possible score achievable) indicated for each tool. As an example, in the 'OCRA ck' the start of a very high level of exposure is 22.6 (but not the highest possible score achievable), which in the method coincides with the beginning of the purple band. When score values exceed the pre-established Maximum risk score results are capped at 100%.

The Minimum and Maximum risk score values are reported in Table 6, part b.

b. *Harmonisation process.* Since each tool uses different numerical reference risk scores, we convert all results into percentage to harmonise them for comparison across tasks.

For the harmonisation of the results, the following methodology is used: *Min-Max normalisation*. Each assessed exposure risk score for each task, is transformed into percentage using the following equation:

$$(Risk\ score - Minimum\ risk\ score) / (Maximum\ risk\ score - Minimum\ risk\ score). \quad (1)$$

To illustrate the normalisation methodology outlined in this section, an example is provided. Having a measured value of 18 with 'OCRA ck', since the *Minimum risk scores* denoting the risk for this method is equal to 11.1, and the *Maximum risk score* equal to 22,6, the result in percentage presents positive sign and is equal to 73%: (18 − 7.6)/(22,6-7.6).

Conversely, when the score evaluated is lower than the *Minimum risk score* (green area with no risk) the normalisation would yield a negative value.

c. *Result presentation*. In the results section, the harmonised percentage, calculated using *Min-Max normalisation,* are presented by histograms.

As previously underlined, to simplify the representation, when the score is lower than the *Minimum Risk Score (no risk)* a 0% default score is assigned, avoiding negative signs to denote the absence of risk; when score values exceed the pre-established *Maximum risk score* results are capped at 100%.

Initially, the results are paired by tool, followed by a pairing based on task, thereby facilitating meaningful comparisons. Each histogram is colour-coded (green, yellow, orange, red, or purple) to indicate its risk category, which is based on the actual risk value calculated and assigned by each tool during evaluation.

As remarked in section 2, the results obtained with the OCRA methods will be used as a benchmark in comparison with the results of other methods.

To enhance the clarity of the results, the comparative analysis is presented with the tools arranged in the following order:

- tools that offer predictive validity and analyse all risk factors and all segments of the upper limb, (both OCRA in and OCRA ck),

- tools with predictive validity with analysis only of the distal upper limb and/or only part of the various risk factors (ACGIH® HA, RSI, HARM),
- tools with predictive and concurrent validity with analysis only of the distal upper limbs (DUET),
- tools with only concurrent validity (EAWS4),
- tools that do not offer predictive or concurrent validity for the upper limbs (RULA).

The final harmonised risk indices (expressed in percentages), relating to the results of the risk assessment, are always provided, for all the tools and for the individual tasks analysed, separately for the right and left limb.

4. Results

4.1. Analytical risk scores and their classification

Table 7 shows the results derived by applying the eight risk assessment tools to the five work tasks during both the 8- and 4-hour shifts. This table presents the risk scores (and risk categories) obtained for both the right and left limbs using the colour-coded system to show different levels of risk as defined by each risk assessment tool.

Figure 1 shows a comparison of the risk score variations as a function of the shift duration for the different tools, limitedly to the right upper limb. The values represent the difference between 8-hour and 4–hour risk score, given as a percentage on the 8-hour risk score.

When the net task duration was reduced from 410 min. to 210 min. (i.e. the 8-hour shift and the 4-hour shift), the ACGIH® HA and RULA tools did not reflect a decrease in risk scores. In contrast, the OCRA index consistently showed a reduction in risk scores, averaging approximately 60% across all five tasks. The OCRA checklist and EAWS4 exhibited a risk score reduction of approximately 40%. The RSI tool shows a 23% decrease, whereas DUET and HARM show reductions of 14% and 57%, respectively.

4.2. Comparison of results of the 8 tools for each task based on min-max normalisation

In the following. additional comparison of tools, limited to the 8-hour shift, is reported. Since different

Table 7. Risk scores *(current values)* obtained from the eight proposed evaluation tools related to the 5 tasks under analysis, for 8-hr shift and 4-hr shift.

			OCRA in		OCRA ck		ACGIH® HA		RSI		HARM		EAWS4		DUET		RULA	
Tasks	Hr shift	Tasks	right	left	right	left	right	left	right	left	right	left	right	left	right	left	right	left
1	8	1	14.4	7.1	25.3	19.4	4.5	4,5	4.6	3.7	40.8	35	54.8	7.8	30.1	8.4	3	3
	4		7.94	3.6	16.8	13.2	4.5	4.5	3.5	2.8	17.5	15	31.4	4.5	21	6.7	3	3
2	8	2	6.1	3.2	19.6	6.9	6.49	2.6	18.3	1.9	52.2	49.9	38.4	18.0	42	19.9	4	4
	4		2.3	0.8	12.1	4.6	6.49	2.6	13.9	1.5	22.5	21.3	26.9	11.2	36.8	16.6	4	4
3	8	3	3.4	3.7	10.1	5.1	4.37	2.5	5.0	0.9	46.7	52.5	29.6	19.4	27	20	4	4
	4		1.3	1.3	6	3	4.37	2.5	3.8	0.9	20	22.5	18.4	12.2	23	16.7	4	4
4	8	4	33.6	5.4	31.6	15.8	7.6	3.6	23.45	7	49.6	40.8	51.6	16.7	56.5	15.9	3	3
	4		24.3	3.8	18	9.4	7.6	3.6	18.7	3.4	21.3	17.5	31	15	51.1	13.2	3	4
5	8	5	3.9	2.5	10.7	10.7	6.12	2.96	9.6	1.2	46.7	43.8	27.4	41.9	32.6	20	3	4
	4		0.7	1.3	6.7	66.7	6.12	2.96	7.4	0.9	20	18.8	11.5	26.1	28.1	16.8	3	4

The colours respect those proposed by each evaluation tool: no increased risk (green); very low risk or borderlines (yellow); presence of low risk (orange); medium/high risk (dark red); very high risk (purple).

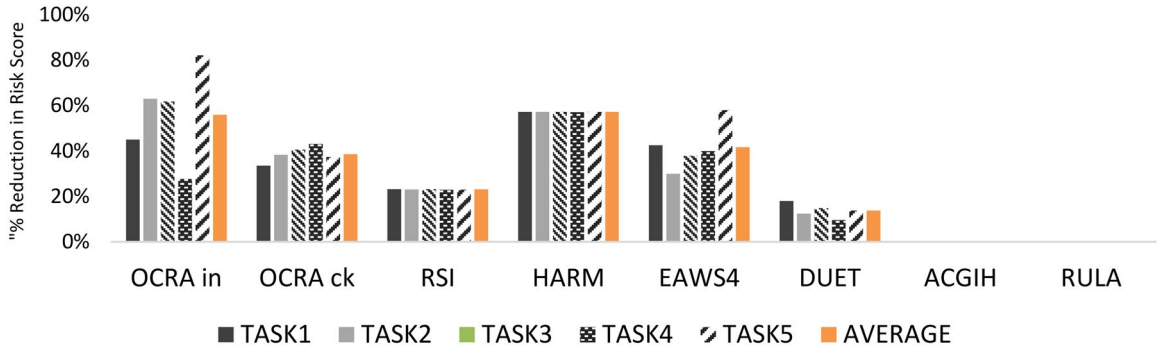


Figure 1. Differences between the 8-hour shift risk score and the 4-hour shift risk score as a percentage on the 8-hours risk score. Evaluation, using the eight tools and the five tasks, only for the right upper limb.

tools produce very different numerical risk indices, the comparison is made based on normalised results in percentage (obtained with the *Min-Max normalisation*). The authors intentionally selected assessment and presentation methods that are straightforward to understand: histograms instead of mathematical models and formulas. To point out risk level, colours are selected according to Table 6.

The results, obtained with eight tools, for each task (separately for right and left), are reported in Figure 2(a–e), in parallel, for all the tools within each task. In detail:

a. *Task 1 (*Figure 2(a)*: see* Movie TASK 1):

- *Task 1 right limb* (highest frequency of technical actions: 120 technical actions per minute! tight

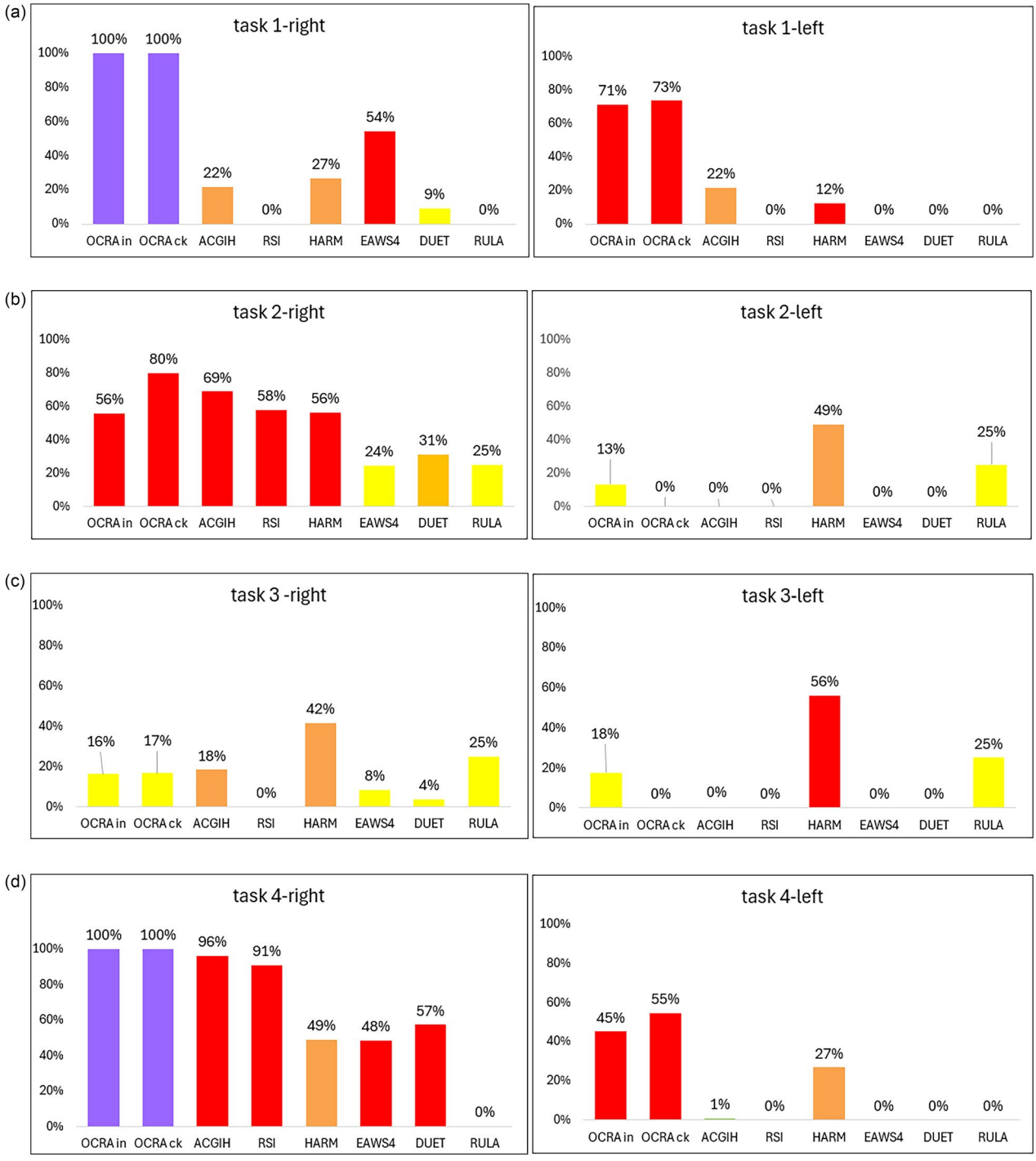


Figure 2. (a) Task 1, right and left upper limbs: risk harmonised (in percentage). (b) Task 2, right and left upper limbs: risk harmonised (in percentage). (c) Task 3, right and left upper limbs: risk harmonised (in percentage). (d) Task 4 right and left upper limbs risk harmonised (in percentage). (e) Task 5, right and left upper limbs risk harmonised (in percentage) using Min-Max normalisation.

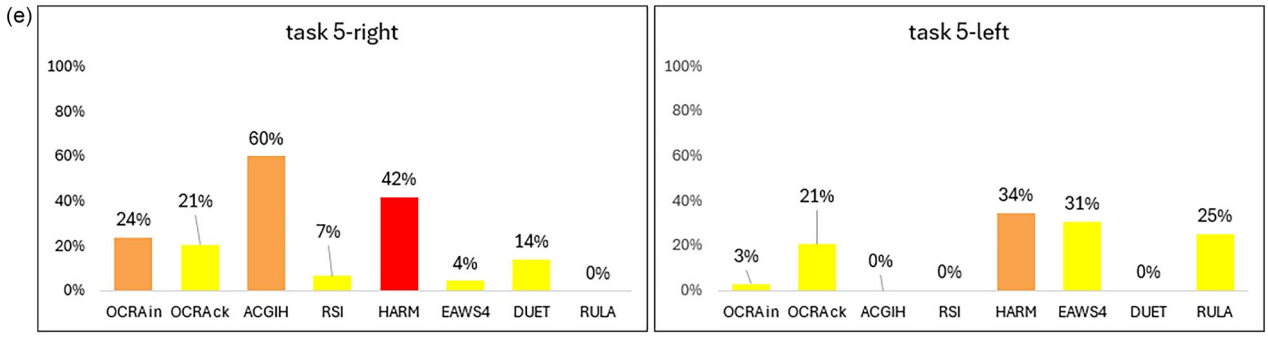


Figure 2. Continued.

pinch, absolute lack of variation (stereotypy), and force practically absent). OCRA tools, in these and similar situations, assigns maximum risk levels (100%, purple); EAWS4 reports values in red areas (medium risk); ACGIH® HA, HARM report values showing similar risk categories, all at low risk in orange area. For ACGIH® HA and RSI, when very high frequency and low force is present, the 'duty cycle equation' does not work perfectly: to obtain a score the suggested scale estimators are used). RSI, DUET and RULA offer lower values, in the green (0%) or in yellow risk categories.

- *Task 1, left limb*, (work with prevalent static posture with fingers in tight pinch 100% of the time). Differences in risk estimation are noted among different tools. In RSI, EAWS4, DUET and RULA all scores are in the green risk category (0%); OCRA tools settle in red area (medium risk); ACGIH® HA results in a low-risk category (orange); RSI, EAWS4, DUET, and RULA (all in green risk category: 0%) seem to give less evaluative weight to static postures in 'grasping in tight pinch'.

b. *Task 2 (*Figure 2(b)**:** *see* Movie TASK 2):

- In *Task 2 right limb* (medium/high frequency with 60 technical actions for minutes, no force, fingers in tight pinch), all the results indicate the presence of risk (red risk category, DUET is in orange), with EAWS4 and RULA, in borderline/yellow risk category.
- In *Task 2 left limb (*low frequency and low force*)* the results are similar (all in yellow areas). HARM is an exception, appearing in the orange risk category, where other tools do not signal a significative presence of risk.

c. *Task 3 (*Figure 2(c)**:** *see* Movie TASK 3):

- In *Task 3 right limb* (medium/low frequency with 35–40 technical actions for minutes, no force, fingers in tight pinch), the data are again more similar for almost all methods, indicating no significant risk category (green or yellow risk category). Only the results for ACGIH® HA and HARM are an exception, in low-risk category (orange, low risk).
- In *Task 3 left limb (*low frequency and low force*)* the data are again similar for almost all methods, appearing in no significant risk category (green or yellow risk category). The presence of risk appears in HARM (red risk category).

d. *Task 4* (Figure 2(d)**:** *see* Movie TASK 4):

- *Task 4 right limb* (action frequency low/medium with 38 technical actions per minute), presence of five force peaks with score 5 in CR10 Borg scale, with duration exceeding 5% of cycle time. Given the force peaks, almost all the evaluations appear at risk, even if with different intensities: OCRA tools all in the purple risk category, ACGIH® HA, RSI, EAWS4, DUET, appear all in red risk category, HARM in orange risk category. Only RULA appears in green category (0%, absence of risk).
- *Task 4 left limb* (prevalent static posture with fingers in tight pinch 100% of the time). Differences in risk estimation are noted. OCRA tools appear in the red risk category, ACGIH® HA and HARM in orange risk category. The other tools do not indicate presence of risk: RSI, EAWS4, DUET and RULA (in green risk category).

e. *Task 5 (*Figure 2(e)**:** *see* Movie TASK 5):

- In *Task 5 right limb* (hand is in a static posture 100% of time with low strength but the grip type is, unlike the other tasks, in power grip, therefore in a more favourable grip compared to the tight pinch). Only 'OCRA in', ACGIH® HA, HARM signal the presence of significant risk.

The other tools are all in yellow or green *areas*.

- In *Task 5 left limb (*low frequency and low force*)* most of the tool's result in green or yellow risk category. Only HARM appears in low-risk category, (orange).

4.3. Comparison of results of the 5 tasks for each tool using percentage based on Min-Max normalisation

The same results are now compared differently: for each tool the histograms representing the risks obtained for the 5 tasks, are presented in parallel, allowing to identify a trend in the results, depending on the intrinsic risk weight that each tool gives to the specific risks of each task, differing for biomechanic risk characteristics, as previously described in Table 4b.

a. 'OCRA in' and OCRA ck tools.
Figure 3 reports a comparison of results relating to the five tasks for the right and left limbs. The trend of the results obtained is also highlighted by a dotted line. Trend lines appear similar for the OCRA tools.

b. HARM and RULA tools.
Figure 4 reports the results for HARM and RULA tools. For HARM, the results indicate the presence of risk at a constant rate for all tasks (orange areas and only one red area) both on the right and left. For RULA, the results indicate almost no risk, with a constant rate for all tasks (green or yellow areas) both on the right and left. The trends of the results within HARM and RULA, when paired with OCRA, clearly mark completely different trends.

c. ACGIH® HA, RSI and DUET tools
Figure 5 reports the results for ACGIH® HA, RSI and DUET tools. Similar trends with the OCRA tools are noted except for situations with a high frequency of actions, where, both in ACGIH® HA and in RSI there are problems of 'out of scale' when the action frequencies are very high (task 1, right side). ACGIH® HA, RSI and DUET indicate low risk in the presence of static postures without the use of force (tasks 1 and 4 on the left), remarkably differing from OCRA tools.
It is also of interest to compare directly the results of DUET and RSI, since DUET defines itself as superimposable to RSI in terms of results. The comparison of the RSI and DUET results confirms the concurrent validity.

d. EAWS4
Figure 6 reports the results for EAWS4 to be compared with the OCRA tools. In a single publication, (Lavatelli, Schaub, and Caragnano 2012) EAWS4 is defined as fully superimposable to OCRA ck in terms of results, speaking about total *concurrent validity* (see also Table 3**)**. The trends shown in Figure 6 help to

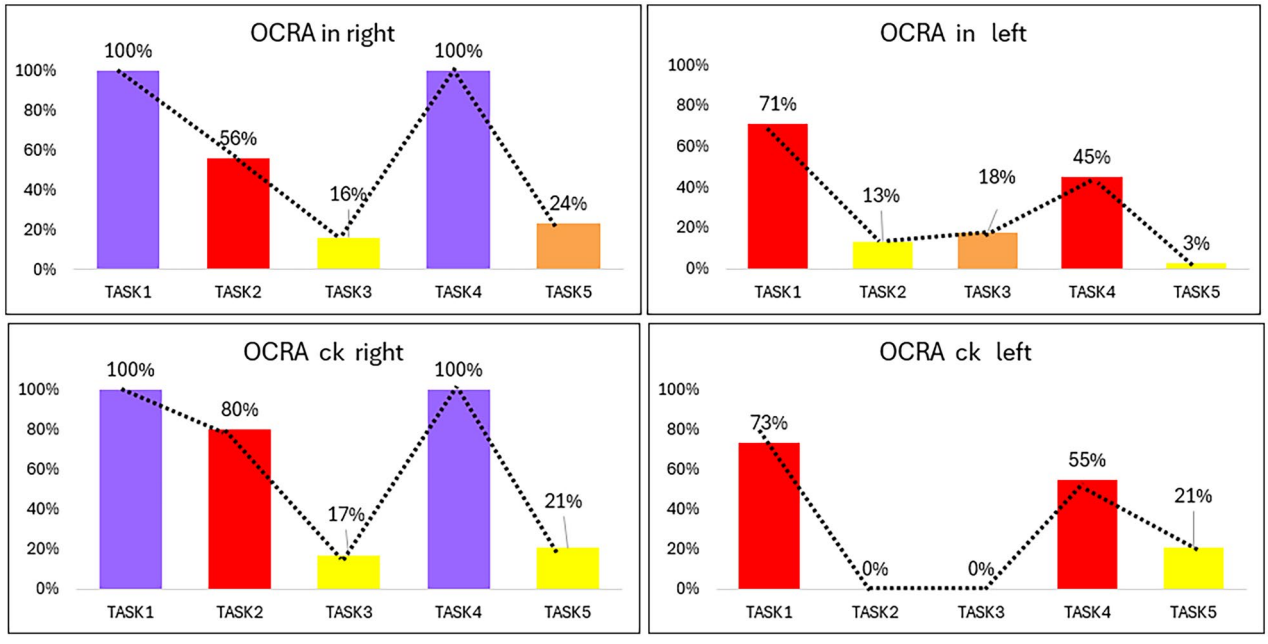


Figure 3. OCRA tools: harmonisation with Min-*Max normalisation*. Comparison of final risk levels (in percentages) for the 5 tasks.



Figure 4. HARM and RULA tools: harmonisation with Min-*Max normalisation*. Comparison of final risk levels (in percentages) for the 5 tasks.

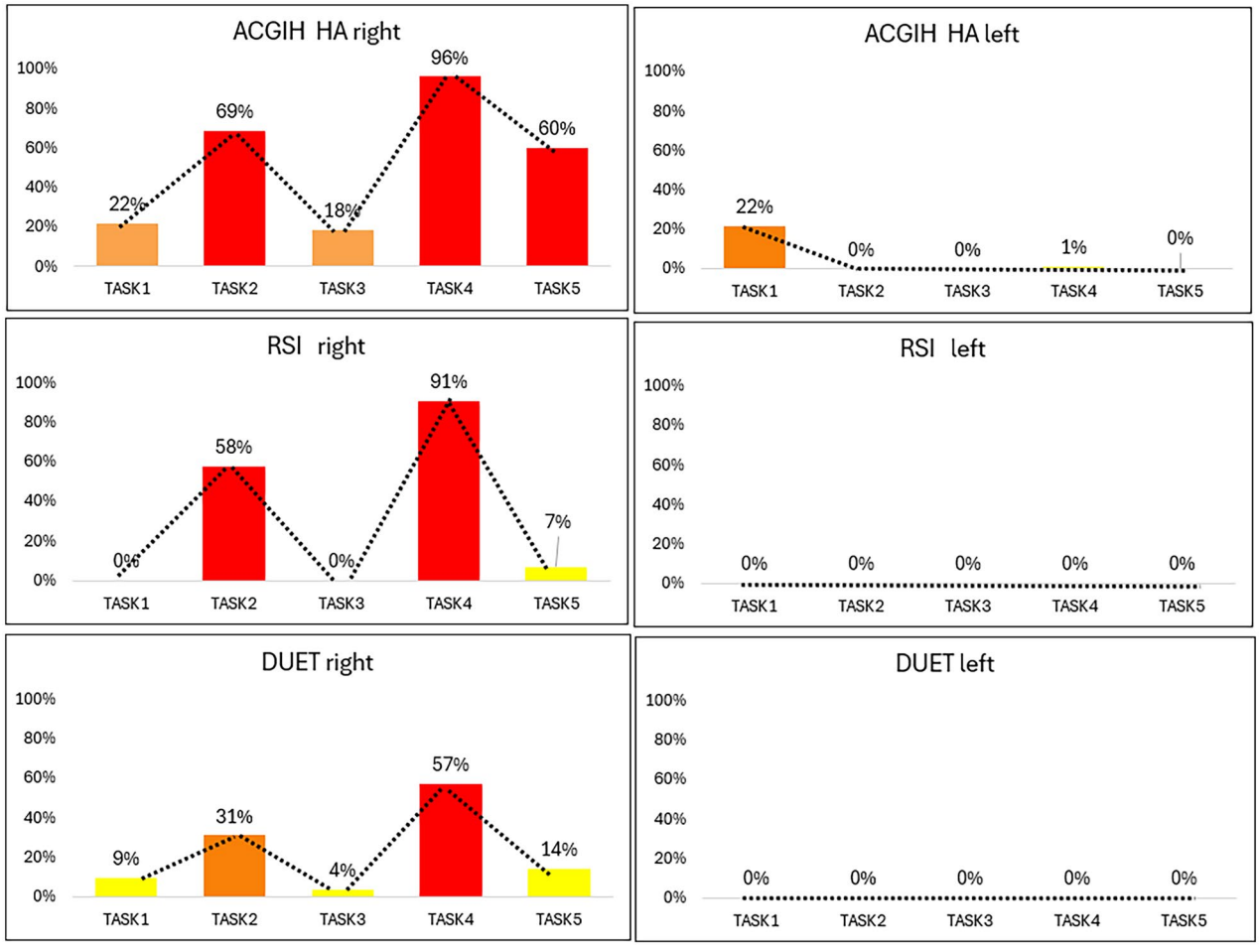


Figure 5. ACGIH HA, RSI and DUET tools: harmonisation with Min-*Max normalisation*. Comparison of final risk levels (in percentages) for the 5 tasks.

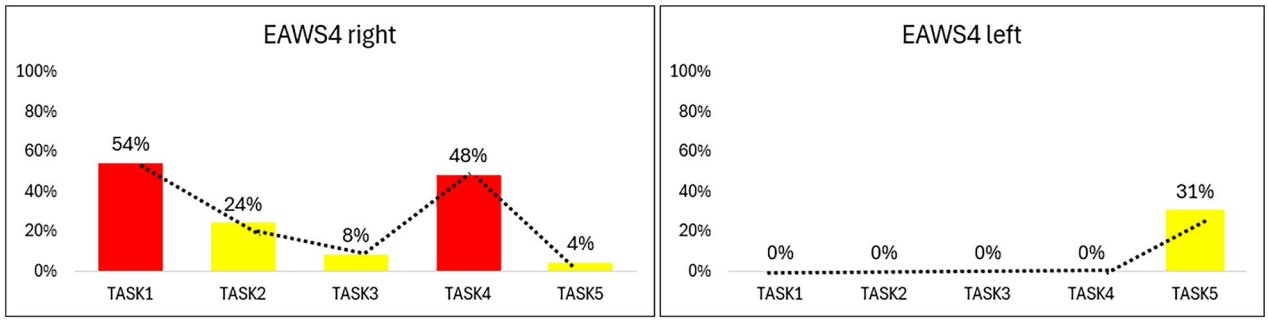


Figure 6. EAWS4: harmonisation with Min-*Max normalisation*. Comparison of final risk levels (in percentages) for the 5 tasks.

Table 8. Comparison of the exposure risk assessment results (in % of the own intrinsic maximum scores) between OCRA tools and EAWS4.

	Right upper limb (% on max)					Left upper limb (% on max)				
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 1	Task 2	Task 3	Task 4	Task 5
OCRA in	100	67	37	100	43	78	35	41	59	27
OCRA ck	100	87	45	100	48	86	31	23	70	45
EAWS4	69	38,3	29,6	51,6	27,4	10	18,0	19,4	21	32
Lack of agreement		x				x		x	x	

highlight whether there is a complete agreement of the results between the EAWS4 tool versus OCRA: the *concurrent validity* appears to be only partial. Even in this case, the static postures are rated differently (see task 1 and 4 on the left arm).

In the present study, as noted more analytically in Table 8, the agreement of the measured risk scores and the respective risk categories between the EAWS4 and OCRA tools can be considered only partial. In four of the ten evaluations (in the 8-hours shift duration, especially for the left hand), EAWS4 offers lower risk estimation in comparison with the OCRA tools results (Index and Checklist). In summary, concordance of the results is seen here only when both EAWS4 and OCRA denote the absence of risk or when the risk is very high.

5. Discussion

There are many larger comparative studies of tools in literature, but these are essentially comparative studies of bibliographic reviews. Here, we aim to bring the results of an application experience to be able to compare the results using real results in analysing the same tasks.

Given the purpose of this study, the authors deliberately chose assessment and exposition of the results techniques that are easily understandable. This involves the use of histograms rather than mathematical models and formulas.

Because there are very different numerical risk indices for different tools, to be able to compare them, the original expedient of normalising the results in % for each tool. This solution has helped in the comparison and interpretation of results.

In consideration of the important and multiple results that emerged, we proceed to divide the discussion into sub-topics.

5.1. Comparison of the risk scores: strengths and limitations of the eight tools studied

We would like to underline that the five tasks have been selected to represent different typologies and levels of risk (Table 4b), with the aim of verifying the response of the different tools to each condition.

- The OCRA in and OCRA ck tools show consistent results. The OCRA tools appear to adequately represent variations in risk levels, both in the presence of elevated level of action frequency (Task 1, right) and in the presence of peaks in force (Task 4, right). When observing the trend, sensitivity to variations in action frequency (Tasks 2 and 3 on the right and left) and grip type was also noted, differentiating scores for close grip and power grip. The OCRA method also increases the risk attributed to static postures when these are present for prolonged periods.
  The OCRA tools are compliant with predictive validity criteria for all UL-WMSDs, based on a

very large database of exposure and clinical outcomes. The OCRA tools, differently from most of the other tools, analyse all the different risk factors indicated as determinants of biomechanical overload and for each area of the upper limb separately (Colombini et al. 2006). According to Paulsen et al. (2015), the completeness, precision, and comprehensibility with which the evaluation criteria are defined make the results of inter-rater reliability tests excellent.

Based on the above considerations, OCRA tools are taken, in the present study, as the reference for comparison with the other tools.

- The HARM tool proposes similar risk scores for all tasks, both on the right and left, mainly in the orange risk category (low risk level) and only in two cases red (medium risk level), remarkably differing from OCRA results. The HARM tool has yielded results that do not seem to adequately read the different levels of risk presented by the different tasks, having shown constant results always around the cut-off between low and moderate risk for all tasks, as well as between the right and left limbs. Given these findings, its use cannot be recommended.
- RULA tool provides identical risk values, all in the yellow risk category (in practice, borderline risk), remarkably differing from OCRA tools' results. Regarding the examined tasks RULA shows an underestimation of the risk with respect to OCRA and to the other tools. This underestimation is easily interpretable because this tool was created mainly to describe postures (gives the greatest weight, more than 60%, to postures to obtain its final scores). In the tasks analysed, the postures were not problematic. In RULA, the evaluation of frequency and force is too synthetic and not analytical as the in other tools and frequency and force scores have only a limited weight on the final score. The tool is not modulated by exposure time RULA has been included in the study because of its extensive use as a tool for the complete study of the risk of biomechanical overload of the upper limbs. The results confirm that the tool is unsuitable for the study of overall biomechanical overload. It remains useful for initial and rapid screening of posture description. RULA, as HARM, does not appear to assess the biomechanical overload sufficiently coherent with the variations present in the 5 tasks analysed here.
- ACGIH® HA and RSI tools yield similar trends to OCRA tools in evaluating the present tasks. However, both tools struggle with 'out of scale' issues at very high frequency (as in task 1 on the right) especially in the absence of force. Furthermore, both methods provide low-risk evaluation parameters in the presence of static postures without the use of force (tasks 1 and 4 on the left). They do not study either the elbow or the shoulder. Exertions/repetitions of the actions may need clearer definitions for both tools to avoid mistakes in interpretation and counting.
- The DUET tool results have shown similarity with the RSI tool results, confirming the assumed concurrent validity (Jorgensen, Martinez, and Hakansso 2024).
- EAWS4 and OCRA, regarding the risk factors considered, are the most comprehensive and similar (Table 2)

  Both methods offer, unlike other tools, well-defined (even if different) criteria for the analysis of the action frequency: technical actions for OCRA methods (partially derived from MTM); the real actions for EAWS4 (derived directly from MTM-UAS). The EAWS4 has not yet been supported by studies on its predictive validity. Furthermore, the estimated risk level for the cases here examined has shown only a partial agreement with the OCRA tool results.

  The partial differences, found in the final risk evaluation, between the two methods, are, however, technically predictable, due to the existing evaluation differences and the weighting of the assigned risk scores (currently not supported by predictive studies). In summary, we can say that these partial differences are essentially due to:

  - the 'fingers in pinch hand score', which in EAWS4 is assigned only in the presence of a certain level of strength,
  - the partial study of stereotypy (extreme lack of variation in actions and postures).

  In automotive assembly, where work is performed using a power grip and in the absence of stereotypy, the values may be similar. Conversely, in companies with fine assembly (hand in tight pinch and/or with stereotypy), the differences are evident: in OCRA, the results would always be higher.

5.2. Primary criteria for choosing tools: risk assessment factors studied and their predictive characteristics

The criteria for choosing which method to use to assess the risk of biomechanical overload of upper

limbs, whether based on the tools analysed here or any other available ones, are, in order of priority: first priority: presence of high predictive validity; second priority: completeness and ability to comprehensively assess the various risk factors causing biomechanical overload, both physical and organisational.

a. First priority: presence of high predictive validity. Risk assessment is a scientific qualitative-quantitative evaluation of the probability, individual or collective, of an adverse health effect in the present case of UL-WMSDs. The more a risk assessment tool can classify exposure with good predictability of related clinical outcomes, the more certainty it provides regarding the consequences of its analysis. The predictive characteristics of a tool should be considered by future users as the primary criterion for choosing a method for studying biomechanical overload of the upper limbs. Summarising the characteristics regarding the predictive validity of each tool:

- the OCRA tools (both index and checklist) are compliant with predictive validity criteria for all UL-WMSDs and are also based on a very large database of exposure and clinical outcomes.
- RSI and ACGIH® HA tools present predictive validity, especially versus CTS (carpal tunnel syndrome), as also confirmed by studies following their respective updates in 2016 and 2018.
- For HARM, limited predictive validity was reported, being based only on symptoms and on a limited number of observations.
- For DUET data for predictive validities are available, limited to the distal segment of the upper limbs. DUET has also concurrent validity towards RSI.
- no predictive or concurrent validity studies have been found for the RULA tool, which is still widely used for the study of biomechanical overload.
- EAWS4 presents only concurrent validity (concurrent predictivity. The authors of the tool reported a preliminary study in the automobile sector (45 workstations) comparing EAWS4 with OCRA Index (Lavatelli, Schaub, and Caragnano 2012) indicating a high correlation between scores. In this study only the OCRA results in the green/yellow risk category were compared with those of EAWS4.

  Moreover, results in more recent large, paired applications of the OCRA Checklist and EAWS4 in the Italian automobile sector (approximately 300 workstations), showed only partial agreement between EAWS4 and OCRA. For these reasons, the concurrent validity of the EAWS4 is to be considered as partial, as appears also in the results of the present study.

  Therefore, for the purpose of a risk assessment, a procedure with the combinations of the two tools can be suggested, as it was adopted by the largest Italian automotive producer, by using EAWS4 tool in the task design phase as a basic screening tool for risk assessment and OCRA tool for a second level analysis. This procedure can be applied when some characteristics of the task (so called 'cut-off', working condition indicating where the EAWS4 tool has to be integrated by OCRA tools, since EAWS4 would provide a risk underestimation) regarding frequency, hand grasp, force and others emerge from the initial screening (Occhipinti and Ghibaudo 2019).

  The most commonly 'cut-off' are represented by the following scenarios:

- the hand is NOT in a power grip (tight pinch or similar) for more than 20%-25% of the time,
- the elbow exceeds the shoulder for at least 10% of the time,
- when the real actions of EAWS4 are more than 25 per minute.

  In the assessments repeated according to the cut-offs cited here, the risk value obtained with EAWS4 determines an underestimation of the risk, compared to OCRA risk results, in 65-70% cases.

  To conclude this part of the discussion, the best predictive validity is offered by the OCRA tools for UL-WMSDs which concern all areas of the upper limb, both proximal and distal and by RSI and ACGIH® HA for UL-WMSDs limited to pathologies of the most distal district (wrist/hand).

b. Second priority: completeness and ability to comprehensively assess the various risk factors causing biomechanical overload, both physical and organisational. Among the tools with demonstrated predictive validity, it is immediately evident that only the OCRA methods assess the entire upper limb. The other tools, which focus solely on the distal segments, can provide an exhaustive risk assessment **only** when the task performed does not involve the elbow or shoulder.

Regarding organisational factors, it is worth noting that the organisational analysis proposed by the OCRA tools is both comprehensive and highly detailed. It examines net repetitive work time as well as the duration and distribution of breaks factors that substantially influence the final risk level.

RSI also provides a thorough organisational analysis, but as previously mentioned, its scope is limited to the wrist and hand. The remaining tools offer little to no assessment of organisational aspects. In fact, taking up the results reported in Figure 1, we note that having performed risk assessments for both 8-hour and 4-hour training durations, ACGH HA and RULA do not indicate any difference in exposure risk.

5.3. Strengths and limitations of the study

Among the strengths of the study, two points appear relevant:

a. The selection criteria of the tools considered (apart from RULA) were oriented to include those tools that allow a 'real' risk assessment (estimation of the induced health effects) through a verification of their predictive (or concurrent) validity. These criteria have already been adopted by the ergonomics experts of the working group which is revising and updating Standard ISO 11228-3.
b. It is likely the first time that risk assessment tools (for UL-WMSDs) have been compared by directly applying them to the same tasks, which has allowed us to highlight procedures, analogies, and differences, as well as the results with which the different tools analyse simple but common repetitive manual tasks.

A possible limitation of this study is that it involved only a small number of repetitive tasks. The selected tasks are representative of many industrial situations and asymmetric tasks, in which the two upper limbs work differently. An extension of the study to a greater range of tasks would certainly have been desirable, although it is not feasible in the present context and within the space limits offered by a paper.

6. Conclusions

Some of the most well-known risk assessment tools for biomechanical overload of the upper limbs were considered in this study. They were applied to five tasks chosen to represent different biomechanical commitments, from low to very high, and for the involvement of different risk factors. The results of the risk analyses were compared.

Similar results were obtained using some tools, but not for all. This is not unexpected because the tools use different criteria for choosing the factors to be assessed and different criteria for assigning the weight to the different factors, in determining the respective final values. This study has tried to highlight these aspects in the assessment strategies for the benefit of potential users and for a more conscious choice of the tool(s) to apply to their specific applicative needs. Which criteria can be suggested to users to address the choice of tool to use is an important question. We think that the comparative results offered by this study may have achieved the objective of offering a comprehensive answer. Risk assessment tools with predictive validity are the most relevant for evaluating exposure and prioritising risk-reduction measures.

The more a risk assessment tool can classify exposure with good predictability of related clinical outcomes, the more certainty it provides regarding the consequences of its analysis. This includes the expected burden of diseases in the exposed population and necessary measures (technical, organisational, and training) to prevent them. Additionally, this aspect can support cost-benefit analyses, including legal considerations in defining occupational diseases across different countries.

Acknowledgements

a. Lind C.M (Unit of Occupational Medicine, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden).
b. Eliasson K (Department of Medical Sciences, Occupational and Environmental Medicine, Uppsala University, Uppsala).
c. Johansson P. (Department of Medical Sciences, Occupational and Environmental Medicine, Uppsala University, Uppsala, Sweden; Department of Occupational and Environmental Medicine, Uppsala University Hospital, Uppsala, Sweden).
d. Alberto Pigin (EPMIES Scientific Association) Milan-Italy.

Many thanks for the essential collaboration provided for the analysis and application of the tools presented. Colombini Daniela: the corresponding author (daniela.colombini50@gmail.com). Substantive intellectual contributions to a paper given by all the paper authors: Colombini D.[1], Occhipinti E.[1] Fox R.[2], Moresco D.[1], Candoli M.[1]

- Substantial contributions to the conception and design of the work; the acquisition, analysis, interpretation of data for the work; AND
- Drafting the work, reviewing it critically for important intellectual content; AND
- Final approval of the version to be published; AND
- Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

All authors agree to be accountable for all aspects of the work.

Author contributions

CRediT: **D. Colombini**: Conceptualization, Data curation, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing; **E. Occhipinti**: Conceptualization, Data curation, Methodology, Supervision, Writing – original draft, Writing – review & editing; **R. Fox**: Conceptualization, Data curation, Formal analysis, Methodology, Supervision; **D. Moresco**: Data curation, Investigation, Software, Supervision; **M. Candoli**: Data curation, Investigation, Software, Supervision.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

D. Colombini http://orcid.org/0000-0001-6318-268X
E. Occhipinti http://orcid.org/0000-0002-1969-0510
R. Fox http://orcid.org/0000-0003-4486-425X
D. Moresco http://orcid.org/0009-0009-6274-4519
M. Candoli http://orcid.org/0009-0007-9876-9756

References

ACGIH® (American Conference of Governmental Industrial Hygienists). 2022. *TLVs® and BEIs®. Threshold Limit Value (TLV) for Hand Activity*, 187–191.

Bandekar, A. J., R. Sesek, M. Schall, Jr, R. Huangfu, D. Bani Hani, and S. Gallagher. 2021. September. "Validation of Fatigue Failure Risk Assessment Tools against Physician-Diagnosed Outcomes." In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 65, 710–714. Los Angeles, CA: SAGE Publications. doi:10.1177/1071181321651194.

Beliveau, Peter Jh., Heather Johnston, Dwayne Van Eerd, and Steven L. Fischer. 2022. "Musculoskeletal Disorder Risk Assessment Tool Use: A Canadian Perspective." *Applied Ergonomics* 102: 103740. doi:10.1016/j.apergo.2022.103740.

Bonfiglioli, Roberta, Stefano Mattioli, Thomas J. Armstrong, Francesca Graziosi, Francesco Marinelli, Andrea Farioli, and Francesco S. Violante. 2013. "Validation of the ACGIH TLV for Hand Activity in the OCTOPUS Cohort: A Two-Year Longitudinal Study of the Carpal Tunnel Syndrome." *Scandinavian Journal of Work, Environment & Health* 39 (2): 155–163. doi:10.5271/sjweh.3312.

Borg, G.A.V. 1998. *Borg's Perceived Exertion and Pain Scales*. Champaign (USA): Human Kinetics.

CEN. 2008. *EN 1005-3 Safety of Machinery Human Physical Performance Part 3: Recommended force Limits for Machinery Operation*.

Chiasson, M.T., D. Imbeau, K. Aubry, and A. Delisle. 2012. "Comparing the Results of Eight Methods Used to Evaluate Risk Factors Associated with Musculoskeletal Disorders." *International Journal of Industrial Ergonomics*. 42 (5): 478–488. doi:10.1016/j.ergon.2012.07.003.

Colombini, D., and E. Occhipinti. 2017. *Risk Analysis and Management of Repetitive Actions: A Guide for Applying OCRA System (Occupational Repetitive Actions)*. London , New York: CRC Press, Taylor and Francis Group.

Colombini, D., E. Occhipinti, M. Cerbai, N. Battevi, and M. Placci. 2011. "Updating of Application Procedures and Criteria for OCRA Checklist." *Med.Lav* 102: 1–39.

Colombini, D., E. Occhipinti, R. Fox, E. Alvarez, and A. Orlei Michaloski. 2021. "The Multi Task General Exposure Index (MultiGEI): an Original Model for Analyzing Biomechanical Risk Factors in Multitask Jobs. featuring Weekly, Monthly and Annual Macro-Cycles." *International Journal of Industrial Ergonomics*. 86: 103212. doi:10.1016/j.ergon.2021.103212.

Colombini, D., N. Delleman, N. Fallentin, A. Kilbom, and A. Grieco. 2006. "Exposure Assessment of Upper Limb Repetitive Movements: A Consensus Document.". In *International Encyclopedia of Ergonomics and Human Factors*, edited by W. Karwowski. 2nd ed., 55–71. London, New York: Taylor and Francis.

Douwes, M., and H. De Kraker. 2014. "Development of a Non-Expert Risk Assessment Method for Hand-Arm Related Tasks (HARM)." *International Journal of Industrial Ergonomics*. 44 (2): 316–327. doi:10.1016/j.ergon.2013.09.002.

Douwes, M., M. Boocock, P. Coenen, S. van den Heuvel, and T. Bosch. 2014. "Predictive Validity of the Hand Arm Risk Assessment Method (HARM)." *International Journal of Industrial Ergonomics*. 44 (2): 328–334. doi:10.1016/j.ergon.2013.09.003.

Eliasson, K., C. M. Lind, and T. Nyman. 2019. "Factors Influencing Ergonomists' Use of Observation-Based Risk-Assessment Tools." *Work (Reading, Mass.)* 64 (1): 93–106. doi:10.3233/WOR-192972.

EU-OSHA (European Agency for Safety and Health at Work). 2019., *Work-Related Musculoskeletal Disorders: Prevalence, Costs, and Demographics in the EU*. Luxembourg: Publications Office of Europe. Union.

Gallagher, S., M.C. Schall, Jr., R.F. Sesek, and R. Huangfu. 2018. "An Upper Extremity Risk Assessment Tool Based on Material Fatigue Failure Theory: The Distal Upper Extremity Tool (DUET)." *Human Factors* 60 (8): 1146–1162. doi:10.1177/0018720818789319.

Garg, A., J. S. Moore, and J. M. Kapellusch. 2017a. "The Revised Strain Index: An Improved Upper Extremity Exposure Assessment Model." *Ergonomics* 60 (7): 912–922. doi:10.1080/00140139.2016.1237678.

Garg, A., J. S. Moore, and J. M. Kapellusch. 2017b. "The Composite Strain Index (COSI) and Cumulative Strain Index (CUSI): Methodologies for Quantifying Biomechanical Stressors for Complex Tasks and Job Rotation Using the Revised Strain Index." *Ergonomics* 60 (8): 1033–1041. doi:10.1080/00140139.2016.1246675.

Graben, P.R., M.C. Schall, Jr., S. Gallagher, R. Sesek, and Y. Acosta-Sojo. 2022. "Reliability Analysis of Observation-Based Exposure Assessment Tools for the Upper Extremities: A Systematic Review." *International Journal of Environmental Research and Public Health* 19 (17): 10595. doi:10.3390/ijerph191710595.

Harris, Carisa, Ellen A. Eisen, Robert Goldberg, Niklas Krause, and David Rempel. 2011. "Workplace and Individual Factors in Wrist Tendinosis among Blue-Collar Workers: The San Francisco Study." *Scandinavian Journal of Work, Environment & Health* 37 (2): 85–98. doi:10.5271/sjweh.3147.

IEA. 2024. "Musculoskeletal Disorder (MSD) Risk Assessment Tools." Accessed December 12, 2024.

ISO. 2000. *ISO 11226. Ergonomics—Evaluation of Static Working Postures*. Geneva (CH): International Organization for Standardization.

ISO. 2007. *ISO 11228- 3: Ergonomics—Manual Handling—Part 3. Handling of Low Loads at High Frequency*. Geneva (CH): International Organization for Standardization.

ISO. 2014. *ISO TR 12295. Ergonomics—Application Document for International Standards on Manual Handling (ISO 11228-1, ISO 11228-2 and ISO 11228-3) and Evaluation of Static Working Postures (ISO 11226)*. Geneva (CH): International Organization for Standardization.

ISO. 2021. *ISO TR 23076. Ergonomics—Recovery Model for Cyclical Industrial Work*. Geneva (CH): International Organization for Standardization.

Jorgensen, M.J., A. Martinez, and N.A. Hakansso. 2024. "Comparison of Multi-Task Ergonomic Assessment Methods to Assess the Risk of Upper Extremity and Lower Back Musculoskeletal Disorders." *Applied Ergonomics* 119: 104313. doi:10.1016/j.apergo.2024.104313.

Joshi, M., and V. Deshpande. 2019. "A Systematic Review of Comparative Studies on Ergonomic Assessment Techniques." *International Journal of Industrial Ergonomics*. 74: 102865. doi:10.1016/j.ergon.2019.102865.

Kapellusch, J.M., F. Gerr, E.J. Malloy., et al. 2002. "Exposure-Response Relationships for the ACGIH TLV for Hand Activity Level: results from a Pooled Data Study of the Carpal Tunnel. syndrome." *Scandinavian Journal of Work, Environment & Health*. 40 (2014): 610–620. doi:10.1080/15459624.2017.1401709.

Kapellusch, Jay M., Stephen S. Bao, Elizabeth J. Malloy, Matthew S. Thiese, Andrew S. Merryweather, and Kurt T. Hegmann. 2021. "Validation of the Revised Strain Index for Predicting Risk of Incident Carpal Tunnel Syndrome in a Prospective Cohort." *Ergonomics* 64 (11): 1369–1378. doi:10.1080/00140139.2021.1940306.

Lavatelli, I., K.H. Schaub, and G. Caragnano. 2012. "Correlations between EAWS and OCRA Index for Repetitive Loads on the Upper Limbs in Automobile Manufacturing Industries." *Work (Reading, Mass.)* 41 (Suppl. 1): 4436–4444. doi:10.3233/WOR-2012-0743-4436.

Lind, C.M, and L.M. Rose. 2016. "Shifting to Proactive Risk Management: Risk Communication Using the RAMP Tool." *Agron Res* 14: 513–524.

Lind, C.M., E. Vidlund, K. Lorén, and M. Wersäll. 2021. "Observation-Based Tools for Assessment of Risk for Musculoskeletal Disorders. European Agency for Safety and Health at Work, OSHwiki." Accessed May 10, 2024. https://oshwiki.osha.europa.eu/en/themes/observation-based-tools-assessment-risk-musculoskeletal-disorders.

Lind, C.M., F. Abtahi, and M. Forsman. 2023. "Wearable Motion Capture Devices for the Prevention of Work-Related Musculoskeletal Disorders in Ergonomics-An Overview of Current Applications, Challenges, and Future Opportunities." *Sensors (Basel, Switzerland)* 23 (9): 4259. doi:10.3390/s23094259.

Lowe, B.D., P.G. Dempsey, and E.M. Jones. 2019. "Ergonomics Assessment Methods Used by Ergonomics Professionals." *Applied Ergonomics* 81: 102882. doi:10.1016/j.apergo.2019.102882.

McAtamney, L., and E.N. Corlett. 1993. "RULA: A Survey Method for the Investigation of Work-Related Upper Limb Disorders." *Applied Ergonomics* 24 (2): 91–99. doi:10.1016/0003-6870(93)90080-s.

Moore, J.S, and A. Garg. 1995. "The Strain Index: A Proposed Method to Analyze Jobs for Risk of Distal Upper Extremity Disorders." *American Industrial Hygiene Association Journal* 56 (5): 443–458. doi:10.1080/15428119591016863.

Nicoletti, S., D. Consonni, M. Carino, G. Di Leone, G. Trani, N. Battevi, Daniela Colombini, and L. Ambrosi. 2008. "Patologia Da Sovraccarico Biomeccanico Degli Arti Superiori: studio Retrospettivo Di Coorte in Tre Grandi Aziende Del Mobile Imbottito." *La Medicina Del Lavoro* 99 (4): 281–296. doi:10.23749/mdl.v99i4.1684.

Nyman, Teresia, Ida-Märta Rhén, Peter J. Johansson, Kristina Eliasson, Katarina Kjellberg, Per Lindberg, Xuelong Fan, and Mikael Forsman. 2023. "Reliability and Validity of Six Selected Observational Methods for Risk Assessment of Hand Intensive and Repetitive Work." *International Journal of Environmental Research and Public Health* 20 (8): 5505. doi:10.3390/ijerph20085505.

Oakman, J., V. Weale, N. Kinsman, H. Nguyen, and R. Stuckey. 2022. "Workplace Physical and Psychosocial Hazards: A Systematic Review of Evidence Informed Hazard Identification Tools." *Applied Ergonomics* 100: 103614. doi:10.1016/j.apergo.2021.103614.

Occhipinti, E. 1998. "OCRA, a Concise Index for the Assessment of Exposure to Repetitive Movements of the Upper Limbs." *Ergonomics* 41 (9): 1290–1311. doi:10.1080/001401398186315.

Occhipinti, E., and D. Colombini. 2007. "Updating Reference Values and Predictive Models of the OCRA Method in the Risk Assessment of Work-Related Musculoskeletal Disorders of the Upper Limbs." *Ergonomics* 50 (11): 1727–1739. doi:10.1080/00140130701674331.

Occhipinti, E., and D. Colombini 2004. "The Occupational Repetitive Action (OCRA) Methods: OCRA Index and OCRA Checklist.". In *Handbook of Human Factors and Ergonomics Methods, Chapter 15, 15/1–15/14*, edited by N. Stanton, et al. London, New York: CRC Press.

Occhipinti, E., and D. Colombini 2021. "Validation of the OCRA Checklist Score as Predictive of the Occurrence of UL-WMSDs in Workers Exposed to Manual Repetitive Tasks.". In *Proceedings IEA 2021, LNNS*, edited by N. L. Black., et al, 794–780 Springer Nature Switzerland. doi:10.1007/978-3-030-74611-7_108.

Occhipinti, E., and L. Ghibaudo 2019. "Development of a Risk Assessment Procedure for Upper Limbs Based on Combined Use of EAWS 4th Section and OCRA High

Precision Checklist.". In *IEA 2018, Aisc 820*, edited by S. Bagnara., et al, 793–799. London, New York.
Paulsen, Robert, Tommaso Gallu, David Gilkey, Raoul Reiser, Lelia Murgia, and John Rosecrance. 2015. "The Inter-Rater Reliability of Strain Index and OCRA Checklist Task Assessments in Cheese Processing." *Applied Ergonomics* 51: 199–204. doi:10.1016/j.apergo.2015.04.019.
Rose, L. M., J. Eklund, L. Nord Nilsson, L. Barman, and C. M. Lind. 2020. "The RAMP Package for MSD Risk Management in Manual handling - A Freely Accessible Tool, with Website and Training Courses." *Applied Ergonomics* 86: 103101. doi:10.1016/j.apergo.2020.103101.
Schaub, K., G. Caragnano, B. Britzke, and R. Bruder. 2013. "The European Assembly Worksheet." *Theoretical Issues in Ergonomics Science* 14 (6): 616–639. doi:10.1080/1463922X.2012.678283.
Takala, Esa-Pekka, Irmeli Pehkonen, Mikael Forsman, Gert-Åke Hansson, Svend Erik Mathiassen, W Patrick Neumann, Gisela Sjøgaard, Kaj Bo Veiersted, Rolf H. Westgaard, and Jörgen Winkel. 2010. "Systematic Evaluation of Observational Methods Assessing Biomechanical Exposures at Work." *Scandinavian Journal of Work, Environment & Health* 36 (1): 3–24. doi:10.5271/sjweh.2876.
Valentim, DP., MLC. Comper, L. Sandy Medeiros Rodrigues Cirino, PR. da Silva, M. Padilha Alonso Gomes, A. Martins da Silva, and RS. Padula. 2025. "Observational Methods for the Analysis of Biomechanical Exposure in the Workplace: A Systematic Review." *Ergonomics* 68 (10): 1561–1582. Epub ahead of print. PMID: 39611722.) doi:10.1080/00140139.2024.2427864.
Yung, Marcus, Ann Marie Dale, Jay Kapellusch, Stephen Bao, Carisa Harris-Adamson, Alysha R. Meyers, Kurt T. Hegmann, David Rempel, and Bradley A. Evanoff. 2019. "Modeling the Effect of the 2018 Revised ACGIH Hand Activity Threshold Limit Value (TLV) at Reducing Risk for Carpal Tunnel Syndrome." *Journal of Occupational and Environmental Hygiene* 16 (9): 628–633. doi:10.1080/15459624.2019.1640366.

Appendix

In this final section, we thought it useful to first provide a descriptive summary of the eight selected tools, then focusing on a more detailed analysis of the differences presented within each of the risk factors considered.

A1. Main characteristics of the tools and weight of different risk factors considered

Table A2 illustrates the weight in % that the intrinsic scores propose for each of the risk factors considered.

It is clear how different methods assign different weights to different risk factors. For example, as expected, the RULA method assigns the greatest weight to postures, while the ACGIH® HA method, studying only two risk factors, equally distributes the weight.

A2. Definition of the working actions in the different tools and an applicative example

Generic working actions are defined differently by the different tools.

Table A3 presents the definitions and explanations.

To clarify the notable differences in the choice of measurement units for the frequency estimation, a useful comparative reading of the working actions is proposed in Table A.4 (as an example, the working actions present in task 3, only for the right limb, are considered) to better understand the meaning of each of the different 'frequency measurement unit' used by the tools here analysed.

Given that the *technical actions,* in OCRA Index and OCRA ck., seem to be the most detailed in describing the work performed in an operating cycle, along with DUET, the other four tools' procedures for estimating the frequency are illustrated by aggregating the *technical actions*.

RULA does not count action numbers but only defines, in the example, that the actions are more than 4 per minute.

A3. The evaluation criteria of static postures used by the different evaluation tools

Static action is defined as the maintenance of an action for a consecutive time >4 seconds) (ISO 2000). The evaluation criteria for static postures are listed in Table A.5.

A4. The evaluation criteria for force used by the different evaluation tools

Table A6 reports the procedures, and the criteria used for the evaluation of force by the different tools.

Table A1. Reports on the main characteristic of the eight tools.

ACGIH® Hand Activity TLV-revised (ACGIH® HA) (ACGIH® (American Conference of Governmental Industrial Hygienists)) 2022) It is a method for monotask handwork, lasting 4 or more hours per shift, based on the analysis of frequency of exertions or actions (in relation to duty cycle) and of peak force. Other main factors are only generically considered. The tool is based on the joint assessment of average hand activity and normalised peak hand force, and it includes a threshold limit value (TLV) and a (lower) Action Limit (AL) for which general controls are recommended. It can be applied to manufacturing industries and computer-related jobs where there is a possibility for UL-WMSDs. Rotations may be considered with a time-weighted average approach.

The original 2001 Hand Activity TLV has demonstrated a predictive validity in several epidemiological studies especially regarding the carpal tunnel syndrome (CTS) and distal extremities tendinosis (Bonfiglioli et al. 2013; Harris et al. 2011; Kapellusch et al. 2002). However, most of these studies suggested to lower the TLV and the AL for a better protection of exposed workers. The ACGIH reconsidered the limits in 2018. After the publication of the 'revised' ACGIH® Hand Activity TLV, one relevant epidemiological study, regarding the CTS outcome confirmed the predictive validity of this tool and will be considered in the present study (Yung et al. 2019)

The Distal Upper Extremity Tool (DUET) (Gallagher et al. 2018). It is based on Fatigue Failure Theory, and the DUET score is based on the rated intensity of the exertion (based on the OMNI-RES scale) for the task and the number of repetitions of the task during the workday. DUET calculates the cumulative damage of the task or tasks over a working day. The tool approach does not consider working posture, duty cycle or how repetitions of the task are distributed over the course of the workday. DUET is designed to assess the likelihood of distal upper extremity outcomes due to repeated stress associated with hand-intensive tasks. DUET can assess irregular and multiple tasks performed within a day. Multiple task analysis is intrinsic to the method given that it considers a 'cumulative dose' during the shift independently of the number of tasks performed. A statistically significant dose-response relationship with distal upper extremity outcomes (injury and pain) was proven in a revision of an earlier cross-sectional study including 772 participants. In a more recent study (Bandekar et al. 2021) a sstatistically significant association between DUET Log CD and physician-diagnosed CTS was also demonstrated.

Regarding the concurrent validity, a recent study (Jorgensen, Martinez, and Hakansso 2024) related DUET with RSI that showed a strong correlation where increasing risk factor exposure resulted in increasing outputs for both methods. However, RSI and DUET demonstrated only a fair agreement (Cohen's Kappa) in how the two methods classified outputs into three risk categories (high, moderate, or low) when assessing the same tasks.

Ergonomic Assessment Worksheet Section 4 (EAWS4) (Schaub et al. 2013). Tool for analysis of cyclical industrial work that covers four risk areas: whole body postures, action forces, manual material handling and, in section 4, repetitive motions of the upper limbs, related to Methods-Time Measurement (MTM) systems. In the design of EAWS section 4, that is specifically considered in this paper, the OCRA method served as the foundational reference; in fact, EAWS4 studies the same risk factors as the OCRA method, but with different scores and criteria. The main differences regard (ISO. 2021): frequency of actions as it counts the Real Actions, based on the element definitions in MTM-UAS (Methods-Time Measurement Universal Analysis System); the integrated scoring of frequency, force, and type of hand grip, that are evaluated together; the scoring for lack of variation that only consider cycle time and not stereotypy of postures and movements.

Task rotations may be considered with a time-weighted average approach.

No studies evaluating direct predictivity were found but only studies of concurrent predictivity.

In a preliminary study in the automobile sector (45 workstations) comparing EAWS4 with OCRA Index (Lavatelli, Schaub, and Caragnano 2012), only OCRA results in the green/yellow risk category were compared with EAWS4. Moreover, results in more recent, large, paired application comparison of OCRA Checklist and EAWS4 in the automobile sector (about 300 workstations), (Occhipinti and Ghibaudo 2019), showed an only partial agreement of EAWS4 versus OCRA that confirm the presence of a partial concurrent validity versus OCRA method tools.

Hand Arm Risk assessment Method (**HARM)** (Douwes and De Kraker 2014). Semi-detailed tool for determining the risk of arm, neck or shoulder complaints when performing a single task that involve the use of the hands or arms. The tool considers: task duration, breaks, force, posture, repetition, additional factors (i.e. vibrations). It does not directly support a multiple task analysis where separate analysis of single tasks is combined. It may be useful for risk surveillance in a variety of repetitive work situations in most industrial and service sectors.

Predictive validity has been partially demonstrated only for self-reported Musculo-skeletal symptoms in single study with a small sized study involving a 3-year follow-up of 88 cases (Douwes et al. 2014).

OCRA index (Colombini and Occhipinti 2017; Occhipinti 1998) It is a detailed tool that considers the following risk factors: frequency of technical actions, repetitiveness/lack of variation, awkward postures, force, additional factors, lack of recovery periods, duration of repetitive task. Applicable also to multitask repetitive jobs. Exposure/risk level is classified in five zones (green; yellow; red respectively low-medium and purple when extremely high). The final OCRA Index defines the overall risk of all UL-WMSDs based on predictive models derived from large (5278 subjects) cross-sectional and longitudinal epidemiological studies (Nicoletti et al. 2008; Occhipinti and Colombini 2007). Task rotations considered by two approaches: 'Time Weighted Average' (TWA) for frequent rotations and 'Incremental' (MultiGei) for infrequent rotations (Colombini et al. 2021)

OCRA checklist (OCRA ck) (Colombini et al. 2011; Colombini and Occhipinti 2017; ISO, 2014; Occhipinti and Colombini 2004). It is a tool that considers, in a simplified way, the same risk factors as the OCRA index. Exposure/risk level is classified in five zones too. OCRA Checklist has three main variants with different levels of detail and fields of application (from simpler to more complex): a) OCRA mini checklist; b) OCRA Checklist classic; c) OCRA Checklist high precision. Variants a) and b) are applicable also to non-industrial jobs. The OCRA Mini checklist is the first screening tool. The OCRA Checklist high precision is more complex and should be used only in industrial jobs. They are all applicable to multitask repetitive jobs also for rotations over prolonged periods (weeks, months, year). Task rotations considered by two approaches: 'TWA' for frequent rotations and 'Incremental' (MultiGei) for infrequent rotations (Colombini et al. 2021). Exposure/risk level is classified in five zones (green, yellow, red respectively 'low', 'medium' and 'high'). The final score defines the overall risk of all UL-WMSDs based on forecasting models derived from large cross-sectional epidemiological studies involving more than 11,000 workers (Occhipinti and Colombini 2021).

Revised Strain Index (RSI). (Garg, Moore, and Kapellusch 2017a, 2017b). It is an analysis tool for identifying high-risk jobs by calculating an overall numerical score: the 'Strain Index'. The tool considers the following five risk factors: intensity and frequency of exertion, duration per exertion, hand/wrist posture, and duration of task per day. It is an evolution of the original proposal of the Strain Index (Moore and Garg 1995). With respect to the original SI, the RSI was also adapted to analyse single complex tasks with varying force, duration, and posture requirements among exertions (COSI approach) and multiple rotating tasks (CUSI approach). Both the COSI and the CUSI adopt an incremental approach in analysing multiple tasks or subtasks. The tool is intended for professional and ergonomic teams to predict the raised risk of developing a UL- WMSDs for only distal upper limbs. The original SI demonstrated a predictive validity in different studies for CTS, epicondylitis and hand/wrist tendinitis. The RSI has demonstrated a predictive validity for CTS in a recent longitudinal epidemiologic study involving 1372 workers in the manufacturing, service and healthcare working sectors (Kapellusch et al. 2021).

RULA (Rapid Upper Limb Assessment). The method was developed in 1993 (McAtamney. 1993) and provides a final score that is the level of postural load of the musculoskeletal system of a worker in relation to four action levels with a sign of urgency of intervention. The final score is obtained through the evaluation of three combined factors: the posture of the different areas of the body, the load or force exerted and the muscular activity (static posture or repetitive movements). The body segments considered are the arm, forearm, wrist (group A) and neck, trunk, legs (group B). Each of these body segments has a specific score decided by angular thresholds and is then combined with the factor load/force exerted and the muscular activity factor.

The method has four risk levels: 1–2 acceptable (green); 3–4 further investigation, changes may be necessary (yellow); 5–6 further investigation, changes soon (red); 7 + investigation and implement changes (purple).

RULA scores should be derived from snapshots or video recordings. To perform RULA, the analyst should select the most critical body postures or those that are performed for a long duration, or that are repetitive, or specifically known to be harmful. For this reason, the first step in RULA is to see the different activities performed by the workers. If the work cycle is very long, professionals often choose to take snapshots at regular time intervals.

Table A2. The weight (in %) of the different risk factors that determine the final risk index, in the different methods compared.

	OCRA index	OCRA ck	ACGIH® HA	RSI	HARM	EAWS4	DUET	RULA
Frequency (*exertion/repetition)	19%	11%	50%	67%	23%	56%	~0% to >90%	7,7%
Force (*exertion magnitude)	21%	36%	50%	28%	19%		~0% to >80%	14%
Hand/fingers posture	11%	9%	N/A	2%	N/A		N/A	
Total posture	20%	27%	N/A	2%	33%	13%	N/A	69%
Stereotypy (lack of variations)	6%	3%	N/A	N/A	N/A	6%	N/A	N/A
Additional risk factors	4%	7%	N/A	N/A	30%	13%	N/A	N/A
Lack of recovery period	9%	8%	N/A	N/A	N/A	13%	N/A	N/A

Table A3. Definition of the working actions in the different tools.

Tools	Name of working actions	Definition
OCRA index and ck	Technical actions	*Technical actions* are simple work elements that make up the task, such as *grasping*, *positioning*, *turning*, *pushing*, etc. They are not the individual movements of the hand, wrist, elbow or shoulder, but rather include the overall movement performed by one or more joint segments for performing the work elements. Technical actions may be *dynamic* (when they involve movement) or *static* (when workers 'hold' an object in their hand for more than 4 consecutive seconds).
ACGIH® HA	Exertions/ repetitions	Exertions (efforts per minute) Counting exertions can be confusing given the two different ways of assessing repetitiveness (HAL scale or duty cycle equation) and their relationship to force.: • Hand repetition or hand exertions are characterised using the 0 to 10 Hand Activity Level (HAL) scale (where 0 = completely idle and 10 = the greatest level of repetition imaginable) proposed by Latko (1997) and Latko et al. (1997) • HAL values (0–10) corresponding to different combinations of exertion frequency, period, and duty cycle when using the HAL duty cycle equation. • If using strength and force data to figure out the normalised peak force, forceful hand exertions were those requiring > 9 N pinch force or > 45 N power grip force. This equates to about 10% of female strength.
RSI	Effort per min. Exertions/ repetitions	An effort is defined as a direct application of force through the hand and usually occurs with prehension (e.g. holding an object or using a tool) but also includes direct applications of force with the hand such as typing, pressing controls or pushing objects. Efforts per minute, which is synonymous with frequency of exertion, is a measure of repetitiveness and is defined as the number of exertions per minute. The repetition in RSI is simply a count of the efforts exerted by the hand and fingers and there are no stipulations as far as minimum force exertions. The RSI instructions concerning intensity (no change from the original SI) of exertion are in some respects unclear. They include either taking a Borg scale assessment or finding the %MVC. While the HAL TLV is directed towards the peak force exertion, the RSI does not appear to have such a stipulation. If exertions of varying forces are present on a task, does the analyst take a weighted average of all or a peak force?
EAWS 4	Real actions	The exact definition of the RA is based on the movement definitions of the building blocks of MTM-UAS. Example: *get and place* an object to a specific destination = 1 real action.
HARM	Force exertions	No specific definition of force exertions or working actions can be traced in the supporting material. Postures are assessed as proportion of the tasks. Exertions are assessed according to their frequency and proportion within a task. Force exertions are exemplified as 'components within a task which involve the application of a force, for example: picking up a part; using an electric screwdriver to put in a screw; or placing a product in a container'.
DUET	Force exertions	A change in the amplitude or direction of a force.
RULA	*Actions*	No definitions of actions. When repeated 4x minutes, add the score 1

Table A4. A comparative reading of the working actions presented in Task 3, only for the right limb.

OCRA index and OCRA ck		RSI	ACGIH® HA	HARM	EAWS 4	DUET
Technical actions	No.	No. Exertions/ repetitions	No. Exertions/ repetitions	No. Exertions/ repetitions	No. Real actions (MTM-UAS terminology)	No. Repetitions
TAKES OBJECT 2	1	**1**	**1**	**1**	(AC2) **–1**	**1**
PLACE ON OBJECT 1	1					**1**
SCREW	1	**1**	**1**	**1**	(ZA)-**1**	**1**
TAKES BACK LEVER OF OBJECT 2	1				(AA1 + ZC1)-**2**	**1**
SCREW	1			**1**		**1**
TAKES BACK LEVER OF OBJECT 2	1				(ZC1)-**1**	**1**
SCREW	1			**1**		**1**
PUT BACK ASSEMBLED OBJECT IN BOX (object 1+ object 2	1				(AA2)-**1**	**1**
INSERTS hand into box 2, -TAKES instructions and releases instructions into box 1	2	**1**	**1**	**1**	(AC1) **1**	**2**
THREAD-TAKES same as before	2			**1**	(AC1)**1**	**2**
THREAD-TAKES same as before	2			**1**	(AC1)-**1**	**2**
TAKES BOX 2	1			**1**	(AA2) **1**	**1**
THROW BOX 2	1					**1**
TAKES LID box 1	1	**1**	**1**	**2**	(AA2 + PC1) **2**	**1**
ROTATE LID	1					**1**
PLACE LID on box 1	1					**1**
CLOSE LID	1				(ZC1) **1**	**1**
TAKES BOX 1 with lid	1				(AA1) **1**	**1**
Total number of actions	21	**4**	**14**	**10**	**14**	**21**

Table A5. The evaluation criteria of static postures used by the different evaluation tools.

Tool	Evaluation criteria
OCRA index	Since the risk index is obtained from the ratio between the total number of technical actions carried out in the shift and those recommended (reduced by multipliers representing the different risk factors), the static grip times are transformed into fictitious technical actions (60 sec in static = 45 fictitious technical actions)
OCRA ck	Static actions begin to be scored if they occupy at least 51% of the time in the cycle. When they occupy quite all the time (≥80%), they are assigned a score like for a frequency of dynamic actions equal to 40–50 actions per minute.
ACGIH® HA- RSI	In the presence of a static action that lasts the entire cycle, the duration of the task would occupy the entire time and the efforts per minute, or the duration of exertion would automatically be the maximum.
EAWS4	Calculate the static actions in sec/min (static time/cycle time*60) assigning different scores based on the strength level and the type of grip
HARM	The proportional duration of force exertions and non-neutral posture are assigned different scores that increase with increased duty cycle. There are, however, no additional penalties for sustained exertions or postures
DUET	Each exertion is accounted as one exertion without taking account to its duration.
RULA	If a mostly static posture is present (maintained for more than 10 minutes), a fixed score is assigned that denotes the presence of a static posture

Table A6. The procedures and criteria for evaluating force by the different tools.

Tool	Evaluation criteria
OCRA index and ck	Borg CR10. In the presence of peaks of strength ≥5 on the Borg CR10 scale (even for a duration of 5% or more) the risk scores are very high (critical).
ACGIH® HA	Borg scale CR10
RSI	Borg scale CR10 (evolution with respect to earlier SI)
EAWS4	Use force readings in Newtons (external force). For force exertion a table is being used in which the level of force, frequency of force exertion and duration of force exertion are combined.
HARM	• Force level in HARM is classified in 2 categories according to the mass of the object (g. or kg). and /or the estimated force exerted in Newtons (N):(extremely) low to average: weight <100 g to 1 kg; force <1 N to 10 N • somewhat high to high: weight: 1 kg to 6 kg; force: 10 N to 60 N
DUET	The OMNI-RES scale (here, transposed from ratings using the Borg CR10 scale)
RULA	If load >4.4 lbs (intermittent)= 0 (4.4 lbs = 2 kg) If load between 4.4 and 22 lbs (intermittent)= 1 (22 lbs = 10 kg) If load between 4.4 and 22 lbs (static or repetitive) = 2 If load > 22 lbs (repetitive and shocks) = 3