



HAL
open science

Differences between risk situations identified using a self-reported questionnaire and an observational method

Mohsen Zare, Maxime Norval, Julie Bodin, Yves Roquelaure

► To cite this version:

Mohsen Zare, Maxime Norval, Julie Bodin, Yves Roquelaure. Differences between risk situations identified using a self-reported questionnaire and an observational method. *WORK*, 2021, 68 (3), pp.759-769. 10.3233/WOR-203409 . hal-03169687

HAL Id: hal-03169687

<https://hal.science/hal-03169687>

Submitted on 31 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title page

Differences between Risk Situations Identified Using a Self-Reported Questionnaire and an Observational Method

Running Title

Risk Situations Identified by Questionnaire and Observation

Mohsen Zare^{1,2}

Maxime Norval²

Julie Bodin²

Yves Roquelaure²

¹ERCOS Group (pole), Laboratory of ELLIAD-EA4661, UTBM-University of Bourgogne Franche-Comté, France

²Univ Angers, CHU Angers, Univ Rennes, Inserm, Ehesp, Irset (Institut de recherche en santé, environnement et travail - Research Institute for Environmental and Occupational Health), UMR_S 1085, F-49000 Angers, France

Corresponding Author: Mohsen Zare

Email: mohsen.zare@utbm.fr Tel: +33787722141

Title page

Differences between Risk Situations Identified Using a Self-Reported Questionnaire and an Observational Method

Running Title

Risk Situations Identified by Questionnaire and Observation

BACKGROUND: This study investigates Work-related musculoskeletal disorders risk estimation by frequently as used as ergonomic methods in the field.

OBJECTIVE: The objective was: To identify the difference in risk estimation by an in-house observational method and a self-reported questionnaire, and to evaluate the complementary aspects of these methods.

METHODS: A sample of 15 operators who worked on the assembly workstations was selected from a truck manufacturing plant. The risk assessment of these workstations (28 scenarios) was performed by the observational method and the self-reported questionnaire. The agreement between both methods to identify risk situations was measured with the weighted Kappa coefficient.

RESULTS: The observational method and the self-reported questionnaire deployed on the same activity estimated different risk situations.

CONCLUSIONS: This analysis does not reveal that one tool is more powerful than the other one, but shows the probability of different risk estimation. The complementary effect of each method might be considered for further investigation concerning musculoskeletal risk factors.

KEYWORDS: risk estimation, musculoskeletal disorders, assessment tool, automotive assembly tasks

1. Introduction

Work-related musculoskeletal disorders (MSD) are multi-factorial, as they are caused by several factors (1). In 2017, MSD represented 87% of recognized occupational diseases. They are also the leading cause of lost working days due to absenteeism, with more than ten million working days lost in 2015 (2). Evaluating job characteristics constitutes a significant challenge in identifying the levels of exposure to MSD risk. Practitioners must be able to evaluate MSD risk with valid and reliable assessments (3). Several methods have been developed to assess the exposure to risk factors, including direct measurements, observational methods, questionnaires, and interviews (4,5). However, engineers and ergonomists need reliable and valid data, taking into account the variation and diversity of the job and the individual.

The results of risk estimation could encounter various possible biases related to data collection methods. Previous research has shown that direct measurement methods provide more reliable data than observation or self-reported questionnaires (6). However, direct measurement methods are time-consuming and require various support and skills (4,5,7). Two methods are commonly used to obtain ergonomic data on workers' activities: observational methods and self-reported questionnaires (4). Several previous studies have used observational methods to assess MSD risks (8,9). Paper-based observational methods, such as Rapid Upper Limb Assessment (RULA), Occupational Repetitive Action (OCRA), Rapid Entire Body Assessment (REBA) and Quick Exposure Check (QEC) tools (10-13) are used to address physical risk factors. Moreover, several large industrial companies have developed their in-house observational methods to identify risk factors that are specific to their sectors. Automotive industries, such as Volvo Car Corporation (VVC), Peugeot-Citroen (PSA), SCANIA and General Motors, have developed their own in-house method for their

ergonomics programs (14,15). Furthermore, large companies such as Fiat, Bosch, and Volkswagen used European Assessment Worksheet (EAWS) tool (16). Eliasson et al. (2019) reported that knowledge about risk assessment tools is low, and no systematic approach is defined for using these tools (17).

Self-reported questionnaire mostly used in epidemiological studies allows estimating the exposure to MSD risk by operators (18,19). Several studies have reported poor-to-moderate validity of the self-reported questionnaire compared to the direct measurement/observational method (20,21). However, Descatha et al. (2009) reported that self-reported questionnaires were more reliable and sensitive than observational methods (22). Stock et al. (2005) and Barrero et al. (2009) noted that current studies of self-reported questionnaires could not demonstrate the validity of self-reported exposure methods due to study design limitations (19,23). These studies have differed regarding the methodologies, the study population (age, gender, and education), and the questions so that their unanimous conclusions cannot be generalized (23).

Regarding the diversity of the conclusions about risk assessment tools, this study was designed to investigate the complementary aspects and the risk estimation difference between an in-house observational method and self-reported questionnaires in the automotive industry.

2. Methods

2.1. Data collection

2.1.1. Observational method

An in-house ergonomic observational method (24) with a video recording was used to analyze the selected workstations. This observational method evaluates 20 risk criteria, which are grouped into four categories, including repetitiveness, working posture,

manual handling, and energy consumption (Table 1). The weights of objects, the magnitude of forces (using a dynamometer), and handle diameters (using caliper) were measured and recorded. Manual handling and lifting of loads with two hands were studied in more detail using the National Institute of Occupational Safety and Health (NIOSH) equation (25). The results were classified according to a color-coding scale. The green level indicated an acceptable situation, with minimal MSD risk. The yellow level indicated a moderate risk situation, which needs to be improved in the future, and the red and double red level corresponded to high-risk situations, which must be modified as soon as possible. The observational method had two high-risk categories (red and double red) that we merged both of them as a red category to facilitate the study analysis. After studying each criterion of the observational method for each workstation, the number of green, yellow and red criteria determined the final color of that workstation (Table 2). This color-coded method is based on Swedish guidelines, and it has been used in other observational methods, particularly in the automotive industry (14,26). Regarding the daily rotations of the various operators to all workstations in an Improvement Groups (IG), we developed a color-coded method representing the risk level for each criterion of an IG (including several workstations). This method is based on the logic of color attribution to one workstation (Table 2). The five thresholds were thus defined to determine the final color (green, yellow and red) for a criterion in an IG (Table 3). All threshold definitions were rounded down, and the most severe color decided the final color of each criterion for the IG.

2.1.2. Self-reported questionnaire

A self-reported questionnaire was used to evaluate the operators' perceptions of the physical exposure of their jobs. Several ergonomic epidemiological studies in France have used this tool to evaluate physical exposure (18,27). This tool is composed of

questions designed to identify the potential physical risk factors for MSD. It was developed using the European consensus criteria document for the evaluation of MSD (28). The questions concern repetition, neck, shoulder, wrist/hand and back postures, material handling, and force/effort for the entire body and wrists (Table 1). The response scale for each question has four levels: Never/Rarely/Often/Always. As shown in Table 1, several questions were asked to assess one ergonomic criterion. To ensure a single answer for each criterion, we combined the responses of several questions. If, for example, the answer to any of the three questions was 'Always,' then the final answer was 'Always.' If the answer to one of the three questions was 'Often,' then the final answer was 'Often,' otherwise, it was 'Never/Rarely.'

2.2. Data collection

This study was performed from September 2012 to August 2013. Analysis using the observational method was conducted at the first stage from September 2012 to March 2013 by an ergonomist worked in the industry. Twelve operators included in the observational phase were again assessed in different scenarios based on the video recordings in August 2013 (Table 4). Twenty-nine scenarios were evaluated (various truck models) for three IGs. Questionnaires were distributed in July 2013 on a Friday to allow operators to complete them carefully over the weekend, and they were collected on Monday, ensuring a high response rate. Fifteen operators responded to all questions, and they were included in the final analysis (Table 4). The institutional review board approved the study and informed consent was obtained from all subjects.

An ergonomist with the help of an assistant observed and video recorded all the scenarios. The ergonomist analyzed the workstation by the observational method by viewing work in person. The video recordings were also used to review and revise the assessments.

2.3. Comparison criteria

We selected 11 criteria from the observational method and the self-reported questionnaire for comparison (Table 1). When several questions existed for one criterion in the questionnaire, they were regrouped by a statistical method based on the frequency of occurrences of the exposure level (rare, often, and always) to achieve a final evaluation. The “manual handling of loads with two hands” and “one hand” criteria were studied at two different levels, as the questions on the self-reported questionnaire concern manual handling of various loads and do not specify whether these loads are handled with one or two hands. Therefore, two subgroups were defined for these two questionnaire items: manual handling with two hands allowing analysis of loads weighing from 10 to 25 kg and loads weighing more than 25 kg, and manual handling with one hand allowing analysis of loads weighing from 1 to 4 kg and >4 kg.

The criteria for effort/force of arms and the effort/force of the entire body corresponded to the same questions in the self-reported questionnaire. The various compared criteria from the questionnaire and observational method are presented and defined in Table 1.

2.4. Statistical analysis

We used descriptive statistics to summarize the percentage of risk factors in different IGs. The responses to the questionnaire (exposure to risk factors of different workstations in one IG) were compared with the results of the observational method for the IG. Categories of ‘Never/Rarely’ from the questionnaire and ‘green’ from the observational method were considered to be low risk. ‘Often’ from the questionnaire and “yellow” from the observation were a moderate risk, and ‘Always’ from the questionnaire and ‘red’ from the observation were high risk. The agreement between the

criteria of the two methods was assessed using the weighted Kappa coefficient (29). The Kappa coefficient interpretation is presented in Table 5 (30). The unit of comparison between both methods was the operator.

3. Results

3.1. Observational method

Table 6 presents the results of the ergonomic risk assessment for three IGs (IG₁, IG₂, and IG₃), according to the observational method (the results of various workstations are shown in the Appendix).

Whole-body work and back, neck, shoulder, and wrist postures were the main risk factors identified in IG₁. Awkward wrist posture was reported for all workstations. Exposure to risk factors, such as one-handed manual handling and surface area for pressure, was low in IG₁ (Table 6).

The results for IG₂ showed high-risk exposure for the wrist and shoulder. The categories repetitiveness and manual handling with two hands were low, while back and neck posture, manual handling with one hand, and whole-body force/effort were moderate. The final risk evaluations for back, neck, shoulder, and wrist postures and whole-body force/effort for this IG were high (red). Awkward body posture was observed at most workstations in IG₃ (see the Appendix). Wrist posture and manual handling with two hands were red at many workstations, while repetition and surface area for pressure were green.

3.2. Self-reported questionnaire

Analysis of the self-reported questionnaires for the three IGs showed that 13 operators (87%) identified awkward back postures as being present often at their work positions.

Repetitiveness and awkward whole-body work postures were identified as being “Often” present for 5 (33%) and 9 (60%) operators, respectively. Furthermore, 10 operators (67%) reported “Always” for exposure to repetitiveness (Table 6).

All operators in IG₁ and IG₂ reported that they were often exposed to awkward back postures. More than half of the operators in IG₁ reported that they were always exposed to manual handling, awkward wrist postures, and excessive effort/force of the body. For IG₃, force and effort of the whole body were often or always present. The majority of operators reported “Often” for exposure to different risk factors (see the Appendix).

3.3. The difference of risk estimation between the observational method and self-reported questionnaires

Table 7 presents the results of a comparison of the data derived from the observational method and the self-reported questionnaire for the three IGs. Both tools identified several risk factors, while the results for certain factors differed considerably, according to the method of analysis, particularly for the back (weighted Kappa coefficient = -0.29), shoulder, neck, and wrist postures and repetitiveness (weighted Kappa coefficient = 0). For most criteria, the risk estimation of the two tools was similar in the moderate risk range but different for extreme situations (high risk and no risk). The agreement between both methods for whole-body effort/force, as well as for effort of palms of hands, was better than that for the other criteria (weighted Kappa coefficient = 0.07 to 0.09). Handling criteria (component size) and two-handed manual lifting imply a low agreement between the operators’ estimations and the ergonomist’s assessments in the material handling criteria (Kappa factor = -0.05 and -0.03); however, the percentage of agreement was 53% between two methods for the one-handed manual lifting criterion.

4. Discussion

This study compared the difference of risk estimation obtained with two risk assessment tools: the in-house observational method and the self-reported questionnaire. The agreement between these tools was investigated for the identification of physical risk factors in a truck assembly plant. This study showed that the observational method and the self-reported questionnaire do not represent a similar risk estimation, and they provided contradictory risk estimation for the analysis of certain physical risk factors. Several studies have performed this type of comparison and have reached different conclusions. Descatha et al. (2009) concluded that the results of self-reported questionnaires differed from those of the observational method and that the self-reported questionnaire was a better predictor of the incidence of future MSDs (22). The study by Spielholz et al. (2001) and Chiasson et al. (2015) reported that the operators' perceptions were different from the results of the reference methods (observation and direct measurement), and the self-reported questionnaire was unreliable (21,31). Three review articles have reported a low-to-moderate agreement between self-reported questionnaires and observational methods (19,23,32). The observational method is based on the expertise of an ergonomist. The observation grid and the ergonomist's knowledge guide the evaluation (32). In contrast, data from the self-reported questionnaire were essentially based on the operators' perceptions concerning their job, and it might differ in terms of the perceived muscular fatigue and musculoskeletal pain (31,32).

These two methods estimated repetitiveness, differently, as the observation tool revealed a low level of exposure, while the self-reported questionnaire identified repetitiveness as a commonly present risk factor. Other studies have also reported poor agreement for repeated movements evaluated by questionnaires and other reference

methods (34). Lowe and Krieg (2009) reported low accuracy of the counted number of repetitive motions by observation analysis compared to direct measurements, although the severity of repetitive motions was accurately estimated (33).

The results concerning whole-body working, neck, and back postures varied considerably, as the self-reported questionnaires revealed a lower risk than the observational method. Burdorf and Laan (1991) reported that operators might underestimate the trunk postures adopted at work (35). Takala et al. (2010) and Stock et al. (2005) reported that micro-postures (the neck and wrists and trunk rotations) are difficult for observers and operators to determine (5,19). Lowe (20014) reported that the ergonomists might underestimate posture duration severity because of discrete intervals of angular position (36).

Observational analysis of shoulder postures revealed high exposure in all groups, while only 13% of the operators identified these postures as being constraining on the self-reported questionnaire. The results for wrist postures presented a similar tendency, with higher sensitivity for the observational method and less definitive results for the self-reported questionnaire. Operators' reports in our study often underestimated the postures adopted and tended to focus on the pain experienced at a particular point in time. Some types of pain are experienced in the context of overexposure, and it is only at this time that the operator becomes aware of the posture adopted (31). Previous studies have reported inconsistent results regarding the presence of MSDs and estimation of the risk exposure by operators. Balogh et al. (2004) and Chiasson et al. (2015) reported the impact of MSDs on the overestimation of the exposure (31,37), while Burdorf and Laan (1991), and Toomingas et al. (1997) found no relationship between MSDs symptoms and overexposure reporting (35,38).

The self-reported questionnaire and the observational method did not provide consistent results for the handling characteristics. Few studies have compared the results of various analysis tools for this type of criterion. Nevertheless, Stock et al. (2005) reported a moderate correlation between the results of such tools for the handling of large objects (19).

Both tools identified the whole-body force/effort risk factor. A slight agreement was observed between the two methods. The force/effort required by a task was measured by a dynamometer and reported as the results of the observational method, while the questionnaire provided the general perceptions of the operators. Based on practice and experience, operators might identify their exposure to the effort. However, an operator may become so accustomed to the working conditions that he/she no longer accurately perceives the effort involved in performing an activity. Working habits, individual experience, and perceptions are essential elements in identifying high-risk exposure, as the level of sensitivity of an operator can result in different responses to the same situation (31). Other studies have reported a good agreement between these tools for the whole-body force/effort criterion (19). However, a low agreement was reported in four studies that compared assessments of push/pull forces using questionnaires and observational methods (32).

The proportions of the agreement for manual handling of loads with two hands and one hand were moderate. Despite certain limitations concerning the analysis of these criteria, our results align with those reported by Stock et al. (2005), who demonstrated poor agreement, particularly for questions about the number of hours/working day spent lifting or carrying loads (19).

Several explanations can be proposed for the different results obtained with these two methods. First, the low Kappa coefficient that was, generally, observed

between the compared criteria might not necessarily be related to the disagreement between the methods. A highly agreed estimation may receive a low Kappa coefficient because other factors, such as limited variability in the distribution of exposure in the categories influenced this coefficient (39). We used in this study a weighted Kappa coefficient that is more adapted for the data with limited variability in the distribution of exposure in the categories.

The limitations of our study such as the reliability of the observational methods and questionnaire (question formulation, response scale, etc.), small sample size, and the respondents' pain/fatigue, were the sources of errors in this experimentation. In the current study, pictograms were used to represent degrees of flexion/extension of each body segment in the questionnaire (Table 1), while the categorical limit was used for observational methods. One limitation may be that we did not compare identical variables. Nevertheless, providing categorical limits in the questionnaire might be a source of error, as the operators might incorrectly estimate the degree of flexion/extension on a numerical scale (36,40,41). The pictogram considered the workers' mental representations of the workload and provided meaningful measures. Another limitation was the small sample size, but the real setting constraints make it impossible to include more participants. The bias induced by the time interval was negligible as we had a short time interval between two measurement methods. We also revised and modified the results of the observational method using videos recorded less than two months after the operators completed the questionnaires. Few variations occurred in the work situations over this short time interval.

5. Conclusion

This study compares the difference in risk estimation by two frequently used methods in the field. The probability of risk estimation differed between the

observational method and the self-reported questionnaire. This difference could enable us to recognize the positions and roles of these tools in the analysis of MSD risk factors and might highlight the complementary aspect of each method. This study did not analyze the validity of these two methods or demonstrate the superiority of any method. However, the findings raise several questions concerning the level of risk estimation using the two frequently used ergonomic methods in the field. We propose to extend this comparison to other tools used in risk assessment, such as interviews and the direct measurement method, which would provide more information on the validity and the use of each method during risk assessment in the workplace.

Acknowledgments

The authors would like to thank all of the people and institutions that contributed to this study, [Edited for Review Process] This study was financially supported by the [Edited for Review Process]

Conflict of interest

None conflict of interest to report for this study.

References

1. Aptel M, Vézina N. “What model for understanding and preventing MSDs ? For a comprehensive and dynamic approach.” [Quels Modèles Pour Comprendre et Prévenir Les TMS? Pour Une Approche Holistique et Dynamique.] Paper presented at the second French congress about MSDs, Montréal, Canada, 2008;1–27.
2. L'Assurance Maladie - Risques professionnels. Paris: Rapport Annuel; 2017. Available from: <https://bit.ly/39U7P3w>
3. Van der Beek AJ, Mathiassen SE, Windhorst J, Burdorf A. An evaluation of methods assessing the physical demands of manual lifting in scaffolding. *Applied ergonomics*. 2005;36(2):213–222.
4. David GC. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occup Med (Lond)*. 1 mai 2005;55(3):190-9.
5. Takala E-P, Pehkonen I, Forsman M, Hansson G-AAke, Mathiassen SE, Neumann WP, et al. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scandinavian journal of work, environment & health*. 2010;36(1):3–24.
6. Koch M, Lunde L-K, Gjulem T, Knardahl S, Veiersted KB. Validity of Questionnaire and Representativeness of Objective Methods for Measurements of Mechanical Exposures in Construction and Health Care Work. *PLOS ONE*. 2016;11(9):e0162881.
7. Zare M, Biau S, Brunet R, Roquelaure Y. Comparison of three methods for evaluation of work postures in a truck assembly plant. *Ergonomics*. 2017;60(11):1551-63.
8. Dempsey PG, McGorry RW, Maynard WS. A survey of tools and methods used by certified professional ergonomists. *Applied Ergonomics*. 2005;36(4):489-503.
9. Habibi E, Zare M, Haghi A, Habibi P, Hassanzadeh A, others. Assessment of physical risk factors among artisans using occupational repetitive actions and Nordic questionnaire. *International Journal of Environmental Health Engineering*. 2013;2(1):14.
10. McAtamney L, Corlett EN. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied ergonomics*. 1993;24(2):91–99.
11. Occhipinti E. OCRA: a concise index for the assessment of exposure to repetitive movements of the upper limbs. *Ergonomics*. 1998;41(9):1290–1311.
12. Hignett S, McAtamney L. Rapid Entire Body Assessment (REBA). *Applied Ergonomics*. 2000;31(2):201-5.
13. David G, Woods V, Li G, Buckle P. The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors for work-related musculoskeletal disorders. *Applied ergonomics*. 2008;39(1):57–69.

14. Törnström L, Amprazis J, Christmansson M, Eklund J. A corporate workplace model for ergonomic assessments and improvements. *Applied ergonomics*. 2008;39(2):219–228.
15. Hägg GM. Corporate initiatives in ergonomics—an introduction. *Applied Ergonomics*. 2003;34(1):3–15.
16. Schaub K, Caragnano G, Britzke B, Bruder R. The European assembly worksheet. *Theoretical Issues in Ergonomics Science*. 2013;14(6):616–639.
17. Eliasson K, Lind CM, Nyman T. Factors influencing ergonomists' use of observation-based risk-assessment tools. *Work*. 2019;64(1):93–106.
18. Roquelaure Y, Ha C, Leclerc A, Touranchet A, Sauteron M, Melchior M, et al. Epidemiologic surveillance of upper-extremity musculoskeletal disorders in the working population. *Arthritis Care & Research*. 2006;55(5):765–778.
19. Stock SR, Fernandes R, Delisle A, Vézina N. Reproducibility and validity of workers' self-reports of physical work demands. *Scandinavian journal of work, environment & health*. 2005;409–437.
20. Trask C, Teschke K, Morrison J, Village J, Johnson P, Koehoorn M. Using observation and self-report to predict mean, 90th percentile, and cumulative low back muscle activity in heavy industry workers. *Annals of occupational hygiene*. 2010;54(5):595–606.
21. Spielholz P, Silverstein B, Morgan M, Checkoway H, Kaufman J. Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors. *Ergonomics*. 2001;44(6):588–613.
22. Descatha A, Roquelaure Y, Caroly S, Evanoff B, Cyr D, Mariel J, et al. Self-administered questionnaire and direct observation by checklist: comparing two methods for physical exposure surveillance in a highly repetitive tasks plant. *Applied ergonomics*. 2009;40(2):194–198.
23. Barrero LH, Katz JN, Dennerlein JT. Validity of self-reported mechanical demands for occupational epidemiologic research of musculoskeletal disorders. *Scandinavian journal of work, environment & health*. 2009;35(4):245.
24. Zare M, Malinge-Oudenot A, Höglund R, Biau S, Roquelaure Y. Evaluation of ergonomic physical risk factors in a truck manufacturing plant: case study in SCANIA Production Angers. *Industrial health*. 2016;54(2):163–176.
25. Waters TR, Putz-Anderson V, Garg A, Fine LJ. Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics*. 1993;36(7):749–776.
26. Berlin C, Örtengren R, Lämckull D, Hanson L. Corporate-internal vs. national standard—A comparison study of two ergonomics evaluation procedures used in automotive manufacturing. *International Journal of Industrial Ergonomics*. 2009;39(6):940–946.

27. Descatha A, Roquelaure Y, Chastang JF, Evanoff B, Melchior M, Mariot C, et al. Validity of Nordic-style questionnaires in the surveillance of upper-limb work-related musculoskeletal disorders. *Scand J Work Environ Health*. 2007;33(1):58-65.
28. Sluiter JK, Rest KM, Frings-Dresen MH. Criteria document for evaluating the work-relatedness of upper-extremity musculoskeletal disorders. *Scandinavian journal of work, environment & health*. 2001;27(Suppl 1): 1–102.
29. Tang W, HU J, Zhang H, Wu P, He H. Kappa coefficient: a popular measure of rater agreement. *Shanghai Arch Psychiatry*. 2015;27(1):62-7.
30. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37(5):360-3.
31. Chiasson M-È, Imbeau D, Major J, Aubry K, Delisle A. Influence of musculoskeletal pain on workers' ergonomic risk-factor assessments. *Appl Ergon*. 2015;49:1-7.
32. Barriera-Viruet H, Sobeih TM, Daraiseh N, Salem S. Questionnaires vs observational and direct measurements: a systematic review. *Theoretical Issues in Ergonomics Science*. 2006;7(3):261–284.
33. Lowe BD, Krieg EF. Relationships between observational estimates and physical measurements of upper limb activity. *Ergonomics*. 2009;52(5):569-83.
34. Juul-Kristensen B, Hansson G-AA, Fallentin N, Andersen JH, Ekdahl C. Assessment of work postures and movements using a video-based observation method and direct technical measurements. *Applied ergonomics*. 2001;32(5):517–524.
35. Burdorf A, Laan J. Comparison of methods for the assessment of postural load on the back. *Scandinavian journal of work, environment & health*. 1991;17(6):425–429.
36. Lowe BD, Weir P, Andrews D. Observation-based Posture Assessment: Review of Current Practice and Recommendations for Improvement. U.S. DHHS (NIOSH) Publication 2014-131. Cincinnati, OH: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.
37. Balogh I, Ørbaek P, Ohlsson K, Nordander C, Unge J, Winkel J, et al. Self-assessed and directly measured occupational physical activities—influence of musculoskeletal complaints, age and gender. *Applied ergonomics*. 2004;35(1):49–56.
38. Toomingas A, Alfredsson L, Kilbom Å. Possible bias from rating behavior when subjects rate both exposure and outcome. *Scandinavian Journal of Work, Environment & Health*. 1997;23(5):370-7.
39. Hansson G-A, Balogh I, Byström JU, Ohlsson K, Nordander C, Asterland P, et al. Questionnaire versus direct technical measurements in assessing postures and

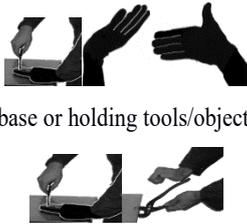
movements of the head, upper back, arms and hands. *Scandinavian journal of work, environment & health*. 2001;27(1):30–40.

40. Lowe BD. Accuracy and validity of observational estimates of shoulder and elbow posture. *Applied Ergonomics*. 2004;35(2):159-71.
41. Lowe BD. Accuracy and validity of observational estimates of wrist and forearm posture. *Ergonomics*. 2004;47(5):527-54.

Accepted Manuscript

Table 1: Explanation of criteria evaluated by the observational method and the self-reported questionnaire. The similar sub-criteria for each method were regrouped as the exposure criteria.

Observational analysis data (indicators: green, yellow, red)		Self-reported questionnaire analysis data (never, rarely/often/always)*
Repetitiveness	Number of repeated movements per hour. < 150 rep/hour Green 150-300 rep/hour Yellow > 300 rep/hour Red	Very precise repeated movements. Same movement for reasons of procedure or quality. Repetition of the same action more than twice a minute.
Whole body work postures	Normal working position: Standing/walking/sitting Green Uncomfortable/twisted position while standing/sitting Yellow Lying, kneeling, squatting, reclining on one side or back, standing on one leg Red	Squatting or kneeling position  Twisting or uncomfortable posture
Back posture	Fixed back posture while working: 0 – 20° bending forward Green 20 – 45° bending forward/ 20° - 45° sideways/rotation Yellow > 45° bending forward or > 45° sideways/ rotation or bending backward Red	Regular or prolonged anterior flexion or to one side.
Neck posture	Fixed neck posture while working: 0-20° bending forward Green 20-45° bending forward or 20-30° sideways/ rotation Yellow > 45° bending forward or > 30° sideways/ rotation or bending backwards Red	Regular or prolonged anterior flexion movements of the head.  Regular or prolonged posterior flexion movements of the head. 
Shoulder posture	Fixed shoulder / arm posture while working: < 45° upper arm lifting Green 45°-90°upper arm lifting Yellow > 90° upper arm lifting Red	Regular or prolonged working with one or both arms above the shoulder level.  Holding the hand behind the trunk Regular or prolonged working with one or both arms abducted 

Wrist posture	Wrist working posture:		Full prono-supination movements or extreme wrist bending posture	
	Neutral wrist	Green		
	Non-neutral wrist	Red		
	<ul style="list-style-type: none"> ➤ 30° bending upwards ➤ 45° bending downwards ➤ > 10° bending sideways 			
				
Effort of palms of hands	Pressure zone area > 1 kg (> 1 kg)		Pressing with palm base	
	$A \geq 1.7 \text{ cm}^2$ or $A \geq 7 \text{ cm}^2$	Green		
	$A < 1.7 \text{ cm}^2$ or $A < 7 \text{ cm}^2$	Red		
				
Force and effort of the whole body	Force of whole-body pushing/traction		Pushing or pulling.	
	Initial force (starting)	Continuous		
	< 100 N	< 50 N		Green
	100-150 N	50-110 N		Yellow
	> 150 N	> 110 N		Red
				
			Manual handling of very large and cumbersome objects with the arms outstretched.	
				
			Manual handling of objects that are difficult to hold, unstable or without handles	
				
Handling	Size of component handled (Size (mm) = Length + Height + Width)		Handling very large and cumbersome objects with arms outstretched.	
	< 1000 mm	Green		
	1000-2000 mm	Yellow		
	> 2000 mm	Red		
				
Two-handed manual handling of loads (lifting with two hands)	Evaluation by revised NIOSH equation method		Manual handling of a load weighing 1 to 10 kg Manual handling of a load weighing 10 to 25 kg Manual handling of a load weighing more than 25 kg	
	< 10 Nm	Green		
	10-35 Nm	Yellow		
	> 35 Nm	Red		
One-handed manual handling of loads (lifting with one hand)	Lifting with one hand:		Manual handling of a load weighing 1 to 10 kg Handling of a tool or an object weighing 1 to 4 kg	
	< 2 kg	Green		
	2-5 kg	Yellow		
	> 5 kg	Red		

*Several questions for one criterion regrouped based on the frequency of occurrence of the exposure level (rare, often, and always) to achieve a final evaluation

Table 2: Prioritization of risk factors by the observational method and the NIOSH equation method for a workstation

Methods	Evaluation Criteria	Green**	Yellow**	Red**
Ergonomic Standard method (SES)	Number of Yellows*	0-8	9-16	≥ 17
	Number of Reds	0-6	7-9	≥ 10
	Number of Yellows + Reds	0-16	-	≥ 17
NIOSH Lifting Equation	Lifting Index	<1	1 -1.6	>1.6

*The worst color dictates the final evaluation of the work position

**Final color based on the number of criteria evaluated Yellow and Red.

Accepted Manuscript

Table 3: Final color definition for each criterion for the improvement group (IG: including several workstations). This method is developed based on the color-coded method explained in Table 2.

Criteria	Threshold	Calculation method*	Final color of one criterion in an IG**
Number of yellow in an IG for one criterion	T_{1y}	$T_1 = \frac{N}{2} + 1$	Yellow
	T_{2y}	$T_2 = N + 1$	Red
Number of red in an IG for one criterion	T_{3r}	$T_3 = \frac{N-1}{3}$	Yellow
	T_{4r}	$T_4 = \frac{N}{2}$	Red
Number of yellow + red in an IG for one criterion	T_{5r+y}	$T_5 = 0.8 N$	Red

*N = the number of measurements (workstation) in an IG.

**The worst color dictates the final evaluation of the IG for a criterion

Table 4: The number of operators included in the study and the number of scenarios evaluated by the observational method for the workstations in the various Improvement Groups (IG). An ergonomist observed and assessed all the scenarios.

Improvement Group	Operators responding questionnaires	Operators observed*	Scenarios evaluated by observational method***
IG 1	6	4**	8
IG 2	5	5	12
IG 3	4	4	9
Total	15	13	29

*The worst results of each operator considered as a final evaluation of the workstation

** Two operators were excluded from the observation phase because of unavailability

***A video recording was performed for each scenario.

Accepted Manuscript

Table 5. Interpretation of the Kappa coefficient (30)

Kappa	Agreement
< 0	Poor agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

Accepted Manuscript

Table 6: Analysis of observational method and self-reported questionnaire of physical risk factors for Improvement Groups (IG) 1, 2, and 3 for 11 criteria of physical risk factors (see Appendix for details).

Item	Analysis of observational method ^a			Self-reported questionnaire (N=15)					
	Final color ^b			Never/Rarely ^b		Often		Always	
	IG ₁	IG ₂	IG ₃	n	%	n	%	n	%
Repetitiveness	Green	Green	Green	0	0	5	33	10	67
Whole-body work postures	Red	Yellow	Yellow	3	20	9	60	3	20
Back posture	Red	Red	Yellow	0	0	13	87	2	13
Neck posture	Red	Red	Red	1	7	11	73	3	20
Shoulder posture	Red	Red	Red	7	47	6	40	2	13
Wrist posture	Red	Red	Red	2	13	9	60	4	27
Effort of hands' palms (Surface area for pressure)	Green	Green	Green	10	67	4	27	1	6
Whole-body force/effort Handling (Component size)	Red	Red	Yellow	1	7	9	60	5	33
Manual handling	Red	Yellow	Red	10	67	5	33	0	0
Two-handed (NIOSH method) ^c				0	0	8	53	7	47
10- 25kg ^d	Red	Green	Red	13	87	2	13	0	0
>25 kg ^d				15	100	0	0	0	0
One-handed ^c	Yellow	Yellow	Yellow						
1-4kg ^d				8	53	7	47	0	0
>4kg ^d				12	80	3	20	0	0

^a Eight scenarios were evaluated at the Improvement Group 1, 12 scenarios at the Improvement Group 2, and 9 scenarios at the Improvement Group 3 (see the Appendix of this paper)

^b “Green” and “Never/Rarely” show low risk; “Yellow” and “Often” show moderate risk; “Red” and “Always” show high risk

^c The items of the observational method for two and one-handed manual handling evaluation

^d The questions of the self-reported questionnaire for manual handling evaluation

Table 7: Comparison between observational method and questionnaire with the calculation of the kappa coefficient and proportion of agreement

	Low (Never/ Rarely; Green)		Moderate (Often; Yellow)		High (Always; Red)		Proportion of agreement	Kappa Coefficient
	n	%	N	%	n	%		
Repetitiveness							0	0
Observational Method	15	100	0	0	0	0		
Self-reported Questionnaire	0	0	5	33	10	67		
Whole-body work posture							47	-0.04
Observational Method	0	0	9	60	6	40		
Self-reported Questionnaire	3	20	9	60	3	20		
Back posture							13	-0.29
Observational Method	0	0	4	27	11	73		
Self-reported Questionnaire	0	0	13	87	2	13		
Neck Posture							20	0
Observational Method	0	0	0	0	15	100		
Self-reported Questionnaire	1	7	11	73	3	20		
Shoulder Posture							13	0
Observational Method	0	0	0	0	15	100		
Self-reported Questionnaire	7	47	6	40	2	13		
Wrist Posture							27	0
Observational Method	0	0	0	0	15	100		
Self-reported Questionnaire	2	13	9	60	4	27		
Effort hands' palms (Surface area for pressure)							87	0.09
Observational Method	10	67	5	33	0	0		
Self-reported Questionnaire	10	67	4	27	1	6		
Whole-body force/effort							28	0.07
Observational Method	0	0	4	27	11	73		
Self-reported Questionnaire	1	7	9	60	5	33		
Handling (Component size)							7	-0.05
Observational Method	0	0	5	33	10	67		
Self-reported Questionnaire	10	67	5	33	0	0		
Two-handed manual handling								
Observational Method (NIOSH equation)	5	33	0	0	10	67		
Self-reported Questionnaire (handling >10kg)	13	86,7	2	13,3	0	0	26,7	-0,03
Self-reported Questionnaire (handling >25kg)	15	100	0	0	0	0	33,3	0
One-handed manual handling								
Observational Method	0	0	15	100	0	0		
Self-reported Questionnaire (handling 1 to 4kg)	0	0	8	53	7	47	53	0

Appendix

In this study, several workstations in an improvement group (IG) evaluated by the observational method and the detailed results for each criterion are provided in Table 1. Table 2 provides the detailed results of the self-reported questionnaire for each IG separately.

Table 1: The percentage of physical risk factors identified by the observational method (11 criteria) for Improvement Groups (IG) 1, 2 and 3

Item	Improvement Group ₁ ^a							Improvement Group ₂							Improvement Group ₃						
	Green		Yellow		Red		Final color	Green		Yellow		Red		Final color	Green		Yellow		Red		Final color
	n	%	n	%	n	%		n	%	n	%	n	%		n	%	n	%	n	%	
Repetitiveness	4	50	4	50	0	0	Green	9	75	3	25	0	0	Green	9	100	0	0	0	0	Green
Whole-body work posture	0	0	3	38	5	62	Red	6	50	1	8	5	42	Yellow	3	33	3	33	3	34	Yellow
Back posture	0	0	4	50	4	50	Red	1	8	8	67	3	25	Red	2	22	5	56	2	22	Yellow
Neck posture	2	25	1	13	5	62	Red	1	5	9	75	2	10	Red	0	0	5	56	4	44	Red
Shoulder posture	2	25	1	13	5	62	Red	0	0	6	50	6	50	Red	1	11	7	78	1	11	Red
Wrist posture	0	0	0	0	8	100	Red	2	17	0	0	10	83	Red	1	11	0	0	8	89	Red
Effort of hands' palms	8	100	0	0	0	0	Green	9	75	0	0	3	25	Green	9	100	0	0	0	0	Green
Whole-body force/effort	0	0	4	50	4	50	Red	1	8	8	67	3	25	Red	3	33	3	33	3	34	Yellow
Handling (component size)	1	13	5	62	2	25	Red	8	67	1	8	3	25	Yellow	1	11	7	78	1	11	Red
Two handed manual handling (NIOSH method)	1	13	3	37	4	50	Red	11	92	0	0	1	8	Green	4	44	1	11	4	45	Red
One handed manual handling	1	13	7	87	0	0	Yellow	4	33	8	67	0	0	Yellow	3	33	4	45	2	22	Yellow

^aEight scenarios were evaluated at the Improvement Group 1, 12 scenarios at the Improvement Group 2 and 9 scenarios at the Improvement Group 3

Table 2: The percentage of physical risk factors (11 criteria) reported by the self-reported questionnaire in Improvement Groups (IG) 1, 2 and 3

Item	Improvement Group ₁ (N=6)						Improvement Group ₂ (N=5)						Improvement Group ₃ (N=4)						All of the subjects (N=15)					
	Never/ Rarely		Often		Always		Never/ Rarely		Often		Always		Never/ Rarely		Often		Always		Never/ Rarely		Often		Always	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Repetitiveness	0	0	2	33	4	67	0	0	2	40	3	60	0	0	1	25	3	75	0	0	5	33	10	67
Whole-body work posture	2	33	3	50	1	17	0	0	4	80	1	20	1	25	2	50	1	25	3	20	9	60	3	20
Back posture	0	0	6	100	0	0	0	0	5	100	0	0	0	0	2	50	2	50	0	0	13	87	2	13
Neck posture	1	17	3	50	2	33	0	0	4	80	1	20	0	0	4	100	0	0	1	7	11	73	3	20
Shoulder posture	3	50	1	17	2	33	2	40	3	60	0	0	2	50	2	50	0	0	7	47	6	40	2	13
Wrist posture	1	17	2	33	3	50	1	20	4	80	0	0	0	0	3	75	1	25	2	13	9	60	4	27
Effort of hand palms	3	50	2	33	1	17	3	60	2	40	0	0	4	100	0	0	0	0	10	67	4	27	1	6
Whole-body force/effort	1	17	2	33	3	50	0	0	4	80	1	20	0	0	3	75	1	25	1	7	9	60	5	33
Handling (Component size)	3	50	3	50	0	0	4	80	1	20	0	0	3	75	1	25	0	0	10	67	5	33	0	0
Manual handling																								
10- 25kg	5	83	1	17	0	0	4	80	1	20	0	0	4	100	0	0	0	0	13	87	2	13	0	0
>25 kg	6	100	0	0	0	0	5	100	0	0	0	0	4	100	0	0	0	0	15	100	0	0	0	0
1-4kg	2	33	4	67	0	0	3	60	2	40	0	0	3	75	1	25	0	0	8	53	7	47	0	0
>4kg	4	67	2	33	0	0	4	80	1	20	0	0	4	100	0	0	0	0	12	80	3	20	0	0

Accepted Manuscript