Archives of Occupational Health | Volume 3 | Issue 1 | January 2019 | 259-67.

Original

Inter-Rater Reliability of Ergonomic Risk Assessment Methods

Atefeh Siahi Ahangar[1], Sahebeh Ghanbari[2], Majid Hajibabaei[3*], Mahnaz Saremi[4], Narges Azadi[1], Fereshteh Jahani[1], Sanaz Karim Pour[1], Moslem Abedini[1], Hassan Mohammadpour[5]

[1] MSc of Occupational Hygiene, School of Health , Students Research Committee, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran• [2]health products safety research center, Qazvine university of medical science , Qazvine, Iran • [3]PhD student of Occupational Hygiene, Department of Occupational Health Engineering, Students Research Committee, School of Public Health and safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran• [4]Assistant professor, Department of Ergonomics, School of Public Health and safety, Shahid Beheshti University of Medical Science, Tehran, Iran• [5]MSC of Occupational Hygiene, Department of Occupational Health Engineering, Students Research Committee, School of Public Health and safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran• *Corresponding author: Majid Hajibabaei, Email: mhajibabaei2050@sbmu.ac.ir, Tel: +98-0912-4134720

Abstract

Background: Musculoskeletal disorders are one of the most common occupational diseases, and in recent years, several methods have been developed to evaluate risk factors for these types of disorders. **Methods:** In this cross-sectional study, 40 tasks in small industries including carpentry, turning, welding, loading and unloading, and sewing were recorded with a video camera and in the second stage, the postures were reviewed and evaluated by six raters. In total, forty of the worst and most frequent postures were analyzed by self-raters and then, the same risk levels were determined for the six methods and analyzed with correlation and Kappa agreement coefficient tests using SPSS (version 19), and then they were compared with each other using the Intraclass correlation coefficient (ICC). **Results:** The results revealed the importance of Kappa Coefficient in which it shows the risk level of different method and specified pair method: OCRA/SI =0.25, OCRA/HAL=0.2, SI/HAL= 0.32, SI/ RULA= 0.33, REBA/OCRA = 0.4, QEC/SI= 0.27, QEC/ RULA= 0.23Inter-rater Reliability of the methods was found as follow:ICCOCRA=0.3, ICCSI= 0.67, ICCHAL= 0.8, ICCRULA= 0.85, ICCREBA=0.8, ICCQEC=0.972.**Conclusions:** The results showed that there was no complete agreement among the methods. This agreement among methods is evaluated from poor to good (0.2-0.4). The ICC showed high reliability in the methods except in the OCRA method.

Keywords: Risk assessment, Reliability, Agreement of Methods, Posture.

Introduction

The advancement of technology has led to change in the situations and working conditions of the industrial workers.[1,2] Today, many of the workers have aligned themselves with the inappropriate work environment.[3,4] Work-related musculoskeletal disorders are one of the most common occupational diseases and the cause of many disabilities in developing industrial countries.[4-7] According to studies conducted from 1992 to 2010, these disorders account for 29% to 35% of absenteeism among workers.[8] Musculoskeletal disorders computed %31 (356,910 cases) for all workers of the total cases in

Citation: Siahi Ahangar A, ghanbari S, Hajibabaei M, Saremi M, Azadi N, Jahani F, et al. **Inter-Rater Reliability of Ergonomic Risk Assessment Methods.** Archives of Occupational Health. 2019; 3(1): 259-67.

Article History: Received: 20 September 2018; Revised: 21 December 2018; Accepted: 27 December 2018

2015.[9] Moreover, the cost of these disorders is estimated to be between 0.5% - 2% of gross national product in Europe and 3.4% in Canada.[10,11]

Various evaluation techniques, such as RULA, REBA, QEC, SI, OCRA, HAL, OWAS, LUBA, PATH, etc. are considered as observational methods. These methods are used as the most common methods for evaluating working postures due to their ease of use, low cost, and availability.[12] The priority in calculating and recording physical exposures during the performance of individuals for risk assessment using these methods is very important. Raters can apply their opinions in evaluations according to their perceptions and interpretation of the process. These methods have evolved over time, and on the other hand, various researchers have achieved other methods by adding additional options to each method. In total, the ultimate goal of all these methods can be a comprehensive and complete evaluation of the activities of individuals in the workplace and investigation of the existing risk factors to prioritize and intervene in the workplace. Therefore, an important challenge in these methods is the correct use of the methods and awareness of the amount of agreement and correlation between the evaluation methods and inter-rater reliability of the method rates.[13,14] If a valid method is not selected, a good evaluation will not be done, and prioritizations, controls, and possible interventions will also be questioned. An analyzer should select the method based on the applicability, validity, reliability, and purpose of the evaluation. Various semi-quantitative or observational risk assessment methods are usually used instead of each other based on various criteria such as cost, time, goal, etc. This has led several researchers to investigate the correlation between methods and their reliability and obtain different results. [15-21] The results of this study can predict whether these methods can replace each other or that each of them has its own existential philosophy and they can be used for specific purposes in evaluations. This study also shows if the Inter-Rater Reliability of any method is acceptable.

The purposes of this study were to investigate the agreement between six ergonomic risk assessment methods that are common in Iran, including RULA, REBA, QEC, SI, OCRA, and HAL, and to measure the reliability of these methods by various raters. These methods are common in most industries and work environment in Iran.

Methods

1. Provide information on tasks

This study tried to use the usual methods utilized to evaluate posture in Iran and the experts apply them in the industries and work environment. The observation and sampling methods were also carried out using common and simple techniques by the same process that experts and raters perform at operational levels. Therefore, forty tasks were selected from the jobs of small industries, including carpentry, turning, welding, loading and unloading, and sewing, and the information on the physical exposures of individuals in these tasks was collected. The work cycle was video-recorded by Samsung Galaxy S4 camera, a 13-megapixel camera that is capable of capturing images at a resolution of 1080 x 1920 pixels. Video-recording was done at a suitable distance and from the place of the work and at different angles of the task. The work cycle in each task was then checked by the leader of the research team and the worst and most frequent postures were selected according to the posted videos. Photos taken from posture along with videos were transferred to the raters.

2. Rater

Six students of the Master of Science in occupational hygiene engineering were employed to evaluate these tasks. Each student, in addition to their specialized training during education, also got further information on expected evaluations in various ways. All videos were copied and shared to the raters. First, the videos that were related to a particular job and task were well assessed by the leader of the research team. Then each video related to the intended job was provided to the

raters. The worst and most frequent posture were selected according to the opinion of raters, and they were evaluated using the six methods, including RULA, REBA, QEC, SI, OCRA and HAL, based on the specific guidelines which are available for each method and available software that users and experts apply in the industry. A total of 40 postures were extracted from all of these tasks, and they were assessed by six raters with the six methods. In total, 1440 samples were analyzed.

3. Exposure assessment methods

3.1. Quick Exposure Check (QEC)

The QEC is a quick and practical method for evaluating musculoskeletal disorders that evaluate the exposure of the four areas of the body, including the back, shoulder-arm, wrist-hand, and neck.[21] In this study, QEC2003 software was used.

3.2. Rapid Upper Limb Assessment method (RULA)

The RULA is an observational method that, in addition to the posture, also takes into account the force and related motions[22] .In this study, Ergointelligence software was used that is a Canadian company's product. This software is completely designed based on procedures of the method evaluation, and it has been approved and used by various sources [23,24]

3.3. Rapid Entire Body Assessment method (REBA)

In REBA, the human body is divided into two groups. Group A includes trunk, neck, and legs, and group B includes upper and lower arms and wrists. Other items including the load handled, couplings with the load, and physical activity are individually scored, and ultimately, the final score for the corresponding posture is calculated.[25] Ergointelligence software was used in this study.

3.4. Strain Index (SI)

In the SI method, the combined effect of six variables affecting the development of skeletal disorders is examined. Three task variables such as intensity of exertion, Hand/ wrist posture, and speed of work are observationally estimated and other task variables including duration of exertion, efforts per minute, and duration of task per day are measured. The task variables are categorized using the corresponding tables. Each of the task variables takes a coefficient, and the SI index is calculated from their multiplication.[26] The Strain Index 22 software was used in this study. All parameters were estimated based on qualitative tables and observation according to the method described by Moore and Garg.[27]

3.5. Hand Activity Level (HAL)

In the HAL-ACGIH method, two indicators that are hand activity level (HAL) and Normalized Peak Force (NPF) of the hand are used to assess the risk factors for developing skeletal disorders in hand, wrist, or forearm. The hand activity level is determined based on the exertion frequency and rest time within the range of 0 to 100. The Normalized Peak Force is set on a scale of 0–10, and finally the number from each indicator in the TLV graph is obtained. Their intersection point specifies the exposure limit for individuals.[28]

3.6. Occupational Repetitive Actions (OCRA)

The OCRA method is based on the correlation between the daily number of actual technical actions performed by the upper limbs in repetitive tasks and the corresponding number of recommended technical actions. In addition to a relatively precise analysis of various risk factors, such as repetitiveness of movements, force, posture, effective and additional factors, and recovery period distribution, this index can also be useful for further preventive measures.[26,27] Ergointelligence software was used in this study.[29]

4. Risk assessment

Table 1. Risk level classification of the ergonomic risk assessment methods

Risk Assessment	Level 1	Level 2	Level 3
QEC-GENERAL	<%40	%40-%70	>%70
RULA	1-2	3-6	7
REBA	1	2-7	8-15
SI	0-3	3.1-7	>7
ACGIH- HAL	<0.56	0.56-0.78	>0.78
OCRA INDEX	<1	1.1-4	>4

The risk assessment was in accordance with those described by previous studies.

The results of all methods were defined as three risk levels according to Table 1[30,31] . As shown in Table 1, the corrective action levels are defined according to each method protocol.

Therefore, to compare the agreement among the methods, each corrective action level was considered to three agreed levels according to Table1.

5. Statistical analysis

5.1. Correlation and agreement analysis between methods

This analysis was performed based on the classified information of each method in three risk levels according to Table 1. The Spearman and Cohen's Kappa correlation coefficients were used to evaluate the agreement among the methods. Spearman coefficient expresses the strength of the relationship between categorized methods and the Kappa coefficient eliminates the probable chance agreement and shows the actual agreement.[32] According to the Landis-Koch benchmark scale in 1977, the Kappa result is interpreted as follows: values ≤0 indicating no agreement and 0.01–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and ≥0.81 an almost perfect agreement.[32] Each rater has evaluated 40 postures of various tasks in the 6 methods .Then, the risk levels of each method in these 40 postures have been compared together.

5.2. Inter-Rater Reliability analysis between raters

This analysis was performed based on the raw scores from each rater's evaluation using the Intraclass correlation coefficient (ICC), as proposed by Shrout and Fleiss in 1976. Confidence intervals with standard errors ($\alpha = 0.05$) were also computed using SPSS software version 21.[33] The reliability of the method based on the Intraclass correlation coefficient (ICC) is interpreted as follows: values less than 0.4 as poor, between 0.4 and 0.75 as moderate and greater than 0.75 as an excellent reliability.[34] In the reliability analysis, the results of each method have been analyzed among six raters and their inter-rater reliability has been measured. For example, 40 postures were evaluated by the raters using RULA method, and the results of these six raters were compared together.

Results

1. Comparison of risk levels

Table 2 presents the risk level distribution in various methods. As shown in Table 2, approximately 60.8 percent of postures are at a dangerous level in the OCRA index method, and the lowest percentage of risk among postures has been evaluated in the QEC-General method. The REBA and HAL methods report approximately the same percentage of the risk level (50.4 and 51.7 percent, respectively) and this similarity exists in the rest of the risk levels.

2. Agreement and correlation between methods

Table 3 shows the paired correlation of methods. The correlation of each method has been evaluated based on the specified risk levels. The Spearman and Kappa correlation coefficients have been used in this table, and the results of previous studies related to each method have also been presented in separate columns. The results of some studies have been reported using the Pearson correlation coefficient, which has also presented in the same way. As shown in Table 3, RULA-REBA possesses a good correlation of 0.808 and REBA-HAL has a weak correlation of 0.12.

Table 2. Distribution of risk levels in the various methods (percent)

Risk levels	QEC-General	REBA	RULA	SI	ACGIH- HAL	OCRA index
First Level (Safe)	43.30	0.80	31.30	32.10	39.60	6.20
Second Level (Moderate)	40.40	48.80	32.90	27.90	8.80	32.90
Third Level (Dangerous)	16.30	50.40	35.80	40.00	51.70	60.80

Table 3. Paired correlation of the methods

Paired methods	Spearman correlation coefficient (SP)	*P*-VALUE	The results of previous studies	Kappa coefficient	*P*- VALUE	The results of previous studies
QEC-RULA	0.46	P<0.001	0.37P#	0.23	*P*<0.001	-
QEC-REBA	0.25	*P*<0.001	0.89+++SP, 0.35P#	0.04	*P*>0.05	-
QEC-SI	0.45	*P*<0.001	0.17P#	0.27	*P*<0.001	-
QEC-OCRA	0.39	*P*<0.001	0.56spΔ, 0.03P#	0.08	*P*<0.05	-
QEC-HAL	0.40	*P*<0.001	0.01P#	0.24	*P*<0.001	-
RULA-REBA	0.81	*P*<0.001	0.67P#	0.25	*P*<0.001	-
SI-REBA	0.36	*P*<0.001	-	0.18	*P*<0.001	-
OCRA-REBA	0.47	*P*<0.001	-	0.40	*P*<0.001	-
HAL-REBA	0.12	*P*>0.05	-	0.07	*P*<0.05	-
SI-RULA	0.50	*P*<0.001	-	0.33	*P*<0.001	0.11-
OCRA-RULA	0.60	*P*<0.001	-	0.18	*P*<0.001	-
HAL-RULA	0.20	*P*<0.001	-	0.15	*P*<0.001	-
SI-HAL	0.41	*P*<0.001	0.77##SP, 0.69P#, 0.48++, 0.73sp+	0.32	*P*<0.001	0.45+, 0.33++
SI-OCRA	0.53	*P*<0.001	0.75sp+, 0.32P#, 0.52##SP	0.25	*P*<0.001	0.55+
OCRA-HAL	0.38	*P*<0.001	0.74sp+, 0.42##SP, 0.16*#	0.20	*P*<0.001	0.52+

P Pearson correlation
Results of Chiasson et al.
Results of Serranheira et al
Δ Results of Joseph et al.
+ Results of Mohammadian et al.
++ Results of Spielholz et al.
+++ Results of Motamedzade et al.
- Results of Drinkaus et al.

Table 4. Reliability and validity between raters in various methods

Method	QEC-General	REBA	RULA	SI	ACGIH-HAL	OCRA index
ICC	0.97	0.80	0.85	0.67	0.80	0.30
%95 CI	0.95-0.98	0.69-0.88	0.77-0.91	0.48-0.8	0.69-0.88	(-0.09)-0.59
PVALUE	0.001	0.001	0.001	0.001	0.001	0.06
results of previous studies	0.82-0.90a	-	0.5-0.7b	0.56c, 0.59g, 0.43i	0.71-0.79k, 0.69g	-

ICC: Intraclass correlation coefficient
%95 CI: 95% confidence interval
a Comper et al.
b Dockrell et al.
c Stephens et al.
g Paulsen et al.
i Stevens et al.
k Ebersole et al.

3. Inter-Rater Reliability and validity among raters

In this study, six raters evaluated the postures using six methods. The inter-rater reliability among the raters was analyzed based on the raw scores of each raters using the Intraclass correlation coefficient (ICC). A total of 240 tasks has been evaluated so that, the contributions of each rater were 40 different postures. The results of these analyzes are presented in Table 4. As shown in table 4, the highest reliability is associated with the QEC method (0.972), and the lowest reliability is related to the OCRA index (0.3). This table also presents the results of previous studies.

Discussion

1. Risk assessment and agreement between methods

The purpose of this study was to compare the correlation and agreement of risk levels in the most widely used and most important posture assessment methods for various tasks. Several studies have been performed on this subject with various approaches

and characteristics and a complete agreement between these methods has not been reported. The findings of this study shows that the correlation coefficients do not have a strong agreement and as shown in Table 3, the Kappa agreement coefficient is in the range of 0.2-0.4 in most methods which, according to the criteria set out in Section 2.5.1, indicates that there is a fair agreement between methods. The agreement is even slight between some methods (QEC-REBA, HAL-REBA, QEC-OCRA). The results of Chiasson et al. 's Studies in 2012 are very close to the findings of this study.[30] Furthermore, the results of a study by Serranheira et al. in 2008 provides a quite close and very good agreement with the methods of OCRA-HAL and SI-OCRA with the results of this study, but there is no correlation between RULA and other methods and there was more agreement between the two methods of SI and HAL (RSP = 0.77).[35] Drinkaus et al. obtained slight agreement between RULA and SI in 2003 (K = 0.1) and in the present study, this agreement is evaluated as fair based on the Kappa coefficient (0.33).[36] The study of Mohammadian et al. in 2013 does not match with the results of this study, and the agreement class is evaluated as moderate, based on the criteria of Kappa coefficients, while the results of the present study shows fair agreement.[37] The results of the agreement between QEC-REBA in the present study are much closer to the study of Chiasson et al. and they are very different from the study by Motamedzade et al.[38] Mohammadian and Motamedzade have suggested that the reason for the difference in the results of their studies compared to other studies was the homogeneity of the industries studied. However, this research is more similar to the industries of Chiasson and Saranian studies, where the heterogeneity of the industry is evident. Joseph et al., in 2015, described the weak correlation of OCRA-QEC methods.[39] Many other studies have also evaluated the agreement between methods, but most of them have considered the percentage agreement approach. For example, Jones and Kumar in 2010 found a good correlation between OCRA, RULA and REBA methods.[40] Since the purpose of this study was to compare the agreement coefficients such as Kappa and Spearman, it has been devoted to studies which have analyzed their results with this approach. The results of this study and previous studies show that none of the ergonomic evaluation methods can fully agree on each other and that a method cannot be considered as an alternative to other methods. These methods are often performed with a screening approach, and It would be better to choose the most appropriate method by considering a series of items such as the purpose of the evaluation, time, sample size, speed, method complexity, simplicity of implementation, facility in learning of the method, etc. Preferably, two methods should be used for screening. For example, simple and rapid methods such as QEC, RULA, REBA, ACGIH HAL should be used for general screening; and more precise numerical methods such as OCRA, strain index should be applied to identify complex and difficult jobs and tasks. [17,41]The results showed that the output risk in each method is strongly dependent on exposure conditions such as activity, status, repetition, etc. For example, the repetition in the OCRA index method is the most important parameter that greatly changes the output risk.[30,31] The results of this study indicated that there is no complete agreement between the methods.

2. Inter-Rater Reliability and validity between raters

In this study, the reliability of each of the RULA, REBA, QEC, SI, OCRA, and HAL methods was investigated by six raters. Investigation of the validity of each method will enhance the credibility of the evaluation using that method. In total, accreditation methods are one of the priorities in the evaluation process. The results of each six raters in each method were examined using the intraclass correlation

coefficient (ICC), which is currently one of the best indicators of reliability measurement.[32] As shown in Table 4, the results demonstrated that QEC (p=0.972), REBA (p=0.8), RULA (p=0.85) and HAL (p=0.8) have excellent reliability according to the interpretation criteria of reliability which was set out in Section 2.5.2. In 2005, David et al. examined the QEC method with ordinary raters and expert raters, and they reported that the Pearson correlation coefficient was 0.79 between the raters.[42] David et al. also evaluated the validity of the QEC method in 2008 and they suggested that this method is valid.[21] In 2012, Comper et al. expressed the range of reliability to be 0.82-0.9 for the QEC method [43]. In a study by Dockrell et al. in 2012, the reliability of the RULA method in individuals working computers was evaluated, and moderate reliability (p=0.5-0.7) was reported. The results of the study indicated that the REBA method also has excellent reliability, but no reliable studies have been done in this field.[44] Stevens et al. (2006) examined the repeatability of the SI method, and the intraclass correlation coefficient of 0.56 was obtained, which was a coefficient for individual assessments; the ICC has been upgraded to 0.82 for team evaluations.[33,45] Paulsen et al. (2015) and Stevens et al. (2004) also expressed the inter-rater reliability of the SI method to be 0.59 and 0.43, respectively which is consistent with the present study and the reliability of the SI method is evaluated as poor to moderate.[45] Paulsen et al., In 2014, examined the reliability of the HAL method and the mean Pearson correlation coefficient of 0.69 was obtained.[31] Ebersole (p= 0.71), Spielholz (p= 0.4) and Takala et al. reported that the reliability coefficient of the HAL method was poor to moderate.[17,46,47] No study reported the reliability for the OCRA index method, but there are several studies for the OCRA Checklist that report the reliability of 0.8 for this method.[17] Typically, a checklist is used for simple screening and OCRA index is applied for deep work analyses.[41-48] As shown in Table 4, the ICC of the OCRA index is 0.3, which indicates that the reliability of this method is poor. Repetition factor is very sensitive in this method, which causes the raters to have a a large number of errors and also reduces repeatability. According to the results of this research and previous studies, six considered methods can be evaluated to have a relatively good reliability. However, the OCRA index method has some specific complexities that makes its repeatability questionable, and more studies are needed in this field. Overall, it seems that the reliability of the methods is decreased when the focus is on repetitive activities. No matter how much the systematic methods design is precise, they do not have a 100 percent reliability for some reasons, which are as follows:

1. Workers demonstrate different movements, postures, and techniques in the same jobs.
2. Activity time can also affect the movements and work postures in the task.
3. The study on the posture and the movements of the small parts of the body, which affect repeatability needs to be more precise.
4. The interpretations that raters make about movements and postures are varied (48).

Conclusion

The results of the present study showed that neither of the two methods are in complete agreement. Because these methods have completed each other for the sake of evolution or they are designed for specific purposes, for example, to evaluate small parts of the body such as wrists. Therefore, a relative agreement is expected between them, but there is no complete agreement between them based on the current study and other similar studies. In this regard, different factors must be taken into account for the use of these methods. So, although the ultimate goal of selecting these methods is to evaluate the postures and estimate the risk factors to prioritize and intervene in postures, each method can be selected specifically in different conditions. For example, if we want to make a

general assessment of the hands, arms, and shoulders, then QEC or REBA methods are more appropriate than SI or OCRA INDEX, but if we want to measure the risk of carpal tunnel syndrome, the SI method seems to be more appropriate. The results of the current study and similar studies confirm this fact.[21] The objectives of the evaluation, the assessment time, the cost of performance method, method complexity, simplicity of implementation, facility in learning of the method, etc., can be other criteria for choosing the methods.[41] Except for selecting a suitable method, selecting a method that has good repeatability and reliability is also very important. The design of systematic methods enables the methods to be as reliable and valid as possible, but the upgrading of sampling systems (advanced video-recording equipment), work experience, skill creation and training needed to learn the method can enhance the validity of the method.[41,48] In the end, all studies have evaluated the validity of the methods as moderate to good, but it is recommended that, as far as possible, multiple methods and even several evaluations be used in repetitive tasks to achieve valid and acceptable results.

Conflict of interest

The authors did not report any contradiction of interests.

Acknowledgments

The authors would like to thank the student research committee at Shahid Beheshti University of Medical Sciences for the financial grants of this study (1394/2659).Also, The authors wish to thank all the people of the various industries who have co-ordinated the work

Reference

1. Cascio WF. Changes in workers, work, and organizations. Handbook of psychology. 2003;12:401-22.
2. Pheasant S, Haslegrave CM. Bodyspace: Anthropometry, ergonomics and the design of work. US: CRC Press; 2016.
3- Kennedy CA, Amick III BC, Dennerlein JT, Brewer S, Catli S, Williams R, et al. Systematic review of the role of occupational health and safety interventions in the prevention of upper extremity musculoskeletal symptoms, signs, disorders, injuries, claims and lost time. occupational rehabilitation. 2010;20(2):127-62.
4- Chung SH, Her JG, Ko T, Ko J, Kim H, Lee JS, et al. Work-related musculoskeletal disorders among Korean physical therapists. Physical therapy science. 2013;25(1):55-9.
5- Valipour Noroozi M, Hajibabaei M, Saki A, Memari Z. Prevalence of Musculoskeletal disorders among office workers. Jundishapur Journal of Health Sciences. 2015;7(1).
6- Karwowski W, Marras WS. Occupational ergonomics: engineering and administrative controls. US: CRC Press; 2003.
7- Mojadam M, Ebadi Z, Ghanbari S, Hajibabaei M. The prevalence of musculoskeletal disorders and assessment of body status among employees of Shahid Chamran University of Ahwaz. Archives of occupational health. 2018;2(4):240-4.
8- (AFL-CIO) AFoLaCoIO. Report on 'Death on the Job, the Toll of Neglect: a National and State-bystate Profile of Worker Safety and Health in the United States. US; 2012.
9- Bureau of Labor Statistics. Nonfatal Occupational Injuries and Illnesses Requiring Days Away From Work. US Department of Labor, Bureau of Labor Statistics Washington, DC; 2011.
10- Buckle PW, Devereux JJ. The nature of work-related neck and upper limb musculoskeletal disorders. Applied ergonomics. 2002;33(3):207-17.
11- Coyte PC, Asche CV, Croxford R, Chan B. The economic cost of musculoskeletal disorders in Canada. Arthritis & Rheumatism: official journal of the american college of rheumatology. 1998;11(5):315-25.
12- David GC. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. Occupational medicine. 2005;55(3):190-9.
13- Kanis H. Reliability and validity of findings in ergonomics research. Theoretical issues in ergonomics science. 2014;15(1):1-46.
14- Li G, Buckle P. Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods. Ergonomics. 1999;42(5):674-95.
15- Kilbom Å. Assessment of physical exposure in relation to work-related musculoskeletal disorders-what information can be obtained from systematic observations. Scandinavian journal of work, environment & health. 1994;20:30-45.
16- Takala EP, Pehkonen I, Forsman M, Hansson GÅ, 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.
17- Paulsen R, Gallu T, Gilkey D, Reiser II R, Murgia L, Rosecrance J. The inter-rater reliability of Strain Index and OCRA Checklist task assessments in cheese processing. Applied ergonomics. 2015;51:199-204.
18- Eliasson K, Palm P, Nyman T, Forsman M. Inter-and intra-observer reliability of risk assessment of repetitive work without an explicit method. Applied ergonomics. 2017;62:1-8.

19- Rosecrance J, Paulsen R, Murgia L. Risk assessment of cheese processing tasks using the Strain Index and OCRA Checklist. International journal of industrial ergonomics. 2017;61:142-8.
20- Kong YK, Lee Sy, Lee KS, Kim DM. Comparisons of ergonomic evaluation tools (ALLA, RULA, REBA and OWAS) for farm work. International journal of occupational safety and ergonomics. 2018;24(2):218-23.
21- 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.
22- McAtamney L, Corlett EN. RULA: a survey method for the investigation of work-related upper limb disorders. Applied ergonomics. 1993;24(2):91-9.
23- Neumann P. Inventory of tools for ergonomic evaluation: Arbetslivsinstitutet, förlagstjänst; 2006.
24- Das SK, Mukhopadhyay S. Integrating ergonomics tools in physical therapy for musculoskeletal risk assessment and rehabilitation–a review. International Journal of Engineering & Scientific Research. 2014;2(10):136-55.
25- Hignett S, McAtamney L. Rapid entire body assessment (REBA). Applied ergonomics. 2000;31(2):201-5.
26- Snook SH, Ciriello VM. The design of manual handling tasks: revised tables of maximum acceptable weights and forces. Ergonomics. 1991;34(9):1197-213.
27- Steven Moore J, Garg A. The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. American Industrial Hygiene Association Journal. 1995;56(5):443-58.
28- Ebersole ML, Armstrong TJ, editors. Inter-rater reliability for hand activity level (HAL) and force metrics. [POSTER] at: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications Sage CA: Los Angeles, CA; 2002.
29- Occhipinti E, Colombini D. Proposal of a concise index for the evaluation of the exposure to repetitive movements of the upper extremity (OCRA index). La Medicina del lavoro. 1996;87(6):526-48.
30- Chiasson MÈ, Imbeau D, Aubry K, Delisle A. Comparing the results of eight methods used to evaluate risk factors associated with musculoskeletal disorders. International Journal of Industrial Ergonomics. 2012;42(5):478-88.
31- Paulsen R, Schwatka N, Gober J, Gilkey D, Anton D, Gerr F, et al. Inter-rater reliability of cyclic and non-cyclic task assessment using the hand activity level in appliance manufacturing. International journal of industrial ergonomics. 2014;44(1): 32-8.
32- Landis JR, Koch GG. The measurement of observer agreement for categorical data. biometrics. 1977;33(1):159-74.
33- Stephens JP, Vos GA, Stevens Jr EM, Moore JS. Test–retest repeatability of the Strain Index. Applied ergonomics. 2006;37(3):275-81.
34- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological bulletin. 1979;86(2):420.
35- Serranheira F, Uva A. WRULMSDs risk assessment: different tools, different results! What are we measuring? Medicina y Seguridad del Trabajo. Med Segur Trab. 2008;54(212):35-44.
36- Drinkaus P, Sesek R, Bloswick D, Bernard T, Walton B, Joseph B, et al. Comparison of ergonomic risk assessment outputs from Rapid Upper Limb Assessment and the Strain Index for tasks in automotive assembly plants. Work. 2003;21(2):165-72.
37- Mohammadian Mastanabad M, Motamedzade M, Faradmal J. Investigating the correlations of OCRA index, strain index and ACGIH HAL methods for assessing the risk of upper limb musculoskeletal disorders. Ergonomics. 2013;1(2):63-71.
38- Motamedzade M, Ashuri MR, Golmohammadi R, Mahjub H. Comparison of ergonomic risk assessment outputs from rapid entire body assessment and quick exposure check in an engine oil company. Journal of research in health sciences. 2011;11(1):26-32.
39- Joseph C, Imbeau D, Nastasia I. Measurement consistency among observational job analysis methods during an intervention study. International journal of occupational safety and ergonomics. 2011;17(2):139-46.
40- Jones T, Kumar S. Comparison of ergonomic risk assessment output in four sawmill jobs. International journal of occupational safety and ergonomics. 2010;16(1):105-11.
41- Eliasson K, Nyman T, Forsman M, editors. Usability of six observational risk assessment methods. [POSTER] at: Proceedings 19th Triennial Congress of the IEA; 2015Aug.9-14; Melbourne; 2015.
42- David G, Buckle P, Woods V. Further development of the usability and validity of the Quick Exposure Check (QEC). Bootle, United Kingdom: Health & Safety Executive; 2005.
43- Comper MLC, Costa LOP, Padula RS. Clinimetric properties of the brazilian-portuguese version of the quick exposure check (QEC). Brazilian journal of physical therapy. 2012;16(6):487-94.
44- Dockrell S, O'Grady E, Bennett K, Mullarkey C, Mc Connell R, Ruddy R, et al. An investigation of the reliability of Rapid Upper Limb Assessment (RULA) as a method of assessment of children's computing posture. Applied ergonomics. 2012;43(3):632-6.
45- Stevens EM, Vos GA, Stephens JP, Moore JS. Inter-rater reliability of the strain index. occupational and environmental hygiene. 2004;1(11):745-51.
46- Ebersole ML, Armstrong TJ. Analysis of an observational rating scale for repetition, posture, and force in selected manufacturing settings. Human factors. 2006;48(3):487-98.
47- Spielholz P, Bao S, Howard N, Silverstein B, Fan J, Smith C, et al. Reliability and validity assessment of the hand activity level threshold limit value and strain index using expert ratings of mono-task jobs. Occupational and environmental hygiene. 2008;5(4):250-7.
48- Eliasson K, Palm P, Nyman T, Forsman M. Inter-and intra-observer reliability of risk assessment of repetitive work without any explicit method. Applied ergonomics. 2016;62:1-8.