Agronomy Research **15**(3), 680–686, 2017

The search for practical and reliable observational or technical risk assessment methods to be used in prevention of musculoskeletal disorders

M. Forsman

Institute of Environmental Medicine, Karolinska Institutet, SE-171 77 Stockholm, Sweden
Correspondence: mikael.forsman@ki.se

Abstract. Work-related musculoskeletal disorders (WRMSDs) are still frequent, inducing very large costs for companies and societies all over the world. Ergonomists work to prevent these disorders and to make organisations sustainable. In their work it is important to identify risks in a reliable way, to prioritise risks, and then to perform interventions (participatory interventions have shown to more often be successful), so that the risks and the disorders may be reduced. Risks are most often assessed by observation. Two projects are described. In the first project the inter-observer reliability of six observational methods was found to be low in risk assessments concerning repetition, movements and postures. Also the inter-method reliability was often low, i.e. when the same work is assessed with different methods different risk estimates are often obtained. In the second described project, easy-to-use methods for measurements of postures and movements were developed and validated. Hence, there are now validated technical methods that are easy to use. But, today's inexpensive electronic devices should be utilized to a higher degree, in developing tools, together with practitioners, that are attractive, easy and time efficient to use, and which should increase the reliability in risk assessments of work tasks and jobs.

Key words: Biomechanical exposure, Observation, Direct measurements, Validity, Reliability, Usability.

INTRODUCTION

Work-related musculoskeletal disorders (WMSDs) are still frequent, inducing very large costs for organisations and societies all over the world. Factors in the physical workload, such as excessive and/or prolonged muscular load, repetitive work and work in awkward and constrained postures, are known risk factors for developing WMSDs in the neck/shoulder region and in arms and hands (European Agency for Safety and Health at Work, 2010).

In order to identify risk occupations, jobs and tasks, for interventions, as well as while planning new jobs and work stations, and to facilitate evaluations of interventions in terms of decreased exposure to risk factors, there is a need for valid, reliable and useful methods for risk assessment of biomechanical exposure.

After assessing and prioritise risks, by the level of acuteness and their seriousness versus the cost of the solutions, interventions should be performed (participatory interventions have shown to more often be successful; Westgaard & Winkel, 1997), so that the risks and the disorders may be reduced.

Many researchers and company ergonomists have worked with risk assessments and many methods, especially observational methods, have been proposed in the literature and are used by practitioners and researchers (Takala et al., 2010; Eliasson & Nyman, 2013; Kalkis et al., 2014; Kong et al., 2017). It is still a popular and interesting field which is indicated by the fact that the observational methods review article (Takala et al., 2010), that was the result of an Nordic collaboration led by Dr. Esa-Pekka Takala at the Finnish Institute of Occupational Health, have been cited about 100 times in other scientific publications. In that review, the authors identified 30 eligible observational methods. Of these, 19 had been compared to one or more methods. Intra- and inter-observer repeatability was reported for 7 and 17 methods, respectively. The methods are generally constructed based on epidemiologic findings, but their ability to predict future MSD (predictive validity) is rarely studied.

In a study comparing observations and inclinometer measurements, Trask et al. (2014) concluded: 'Since observations were biased, inclinometers consistently outperformed observation when both bias and precision were included in statistical performance'. Moreover, dynamic work is best quantified with technical measurements. The general opinion about technical measurements have been, at least until now, that they are time consuming, require expensive equipment, demand technical knowledge to perform, and are therefore not suitable for actors in the work environment field, such as the occupational health services. However, there are now, because of the development of electronics components, a number of low-cost (about $100–$200) technical devices available, i.e. accelerometers, that may be used to monitoring human motions (Korshøj et al., 2014; Skotte et al., 2014; Dahlqvist et al., 2016). These devices are accelerometers with integrated data loggers. There are also inertial measurement units (IMUs) that in addition to a three-axial accelerometer include a gyroscope and a magnetometer.

The overall aim of this research effort is to test existing methods, and to design and validate new methods for risk assessment of biomechanical exposure. In this article two just carried out research projects are described, and the future is discussed, in the field of ergonomics risk assessment.

MATERIALS AND METHODS

Observational methods – reliability, validity, and usability

Although there are many methods available, often ergonomists in the field, for different reasons, e.g. lack of time and/or lack of knowledge of adequate methods, use their own knowledge and experience, when performing risk assessments. In the first of the two described projects, six selected risk assessment methods were evaluated concerning their reliability, validity, and usability (Forsman et al., 2015). The reliability of assessments that ergonomists performed based on their own experience and knowledge, without any specific method, was also investigated.

The six selected methods were:

1. Occupational Repetitive Actions checklist (OCRA; Takala et al., 2010),
2. Quick Exposure Checklist (QEC; Takala et al., 2010),
3. Strain Index (SI; Takala et al., 2010),
4. Assessment of Repetitive Tasks (ART; www.hse.gov.uk/msd/uld/art/index.htm),
5. Hand Arm Risk-assessment Method (HARM; Douwes & de Kraker, 2009),

6. Repetitive work model by the Swedish Work Environment Authority (SWEA, 2012).

Ten video-recorded (3-6 minutes) work tasks were included: 2 supermarket work tasks, meat cutting and packing, engine assembly, hairdressing, 2 cleaning tasks and 2 post sorting tasks. For each work task, data of the work task length (between 2–7 hours per workday), pause schedules, handled weights and physical factors, as well as the employees' ratings of force exertion, work demands and control were given.

Twelve experienced ergonomists made assessments of the ten work tasks in their own pace. The ergonomists were all educated physiotherapists, who were recruited from occupational health services (OHS) and through social media posts to members of the Swedish Ergonomist and Human Factors Society. To be included in the study they should be employed by OHS (or equivalent) and have at least one year of work experience with risk assessments. Firstly they did assessments without using any specific method, as over-all risk, and specified for eight body regions into: high risk (red), moderate risk (yellow) and or low risk (green) (Eliasson et al., 2017). Then, they used the six methods twice, with at least four weeks between the occasions. Before the first assessment, the ergonomists were trained in each method. The videos could be paused or repeated as needed. The assessment times were registered, and the ergonomists were given an evaluation questionnaire on completion of each of the methods.

As an alternative for predictive validity, consensus assessments were carried out by three experts (Kjellberg et al., 2015). They first did own assessments, carefully, in accordance with the manual of each method, and then agreed in the group. Three months later, they again, together, repeated the assessments, in the reversed method order. The experts' assessments were used as a gold standard for concurrent validity of the ergonomists' ratings, and for inter-method comparisons.

To take the agreement due to chance into account, the linearly weighted Kappa coefficient, K_w was the parameter primarily chosen for inter- and intra-observer reliability and validity. The K_w was firstly computed pairwise for all pairs, and then averaged in the way recommended by Davies & Fleiss (1982), Hallgren (2012).

New easy-to-use technical measurements of postures and movements

The second project of the two projects described here was set up to (1) together with actors at the OHS develop easy-to-use methods of technical measurements of postures and movements during work, and which automatically provides informative charts and graphs; (2) validate the new methods against previously validated methods, which today are used by researchers, and (3) test the new methods concerning the time required and their usability for actors within the OHS (Dahlqvist et al., 2016; Yang et al., 2017).

The project included one method for full-day measurements, with a quality equal to similar research methods. The high data quality makes these simplified measurements fully comparable with those reported in the literature. It consists of small accelerometer-devices with integrated USB-memories. They measure postures and movements of head, back and both upper arms. The method includes a simple protocol where you note the start- and end times for work and breaks. After the measurement, the devices are connected to a computer and the noted times are used in a program that computes and presents the workload in figures and tables. An 'even easier' method for shorter measurements was developed as an application for iPhone/iPod. It is called

ErgoArmMeter and measures the arm elevations. Directly after a measurement it shows statistical parameters of angles and angular velocities (Yang et al., 2017).

The new methods were validated, i.e. they were compared with previously validated methods, in static postures, during standardised movements (moving the arm back and forth in different paces), and during simulated work tasks.

RESULTS AND DISCUSSION

Observational methods – reliability, validity, and usability

For sole observation without any specific method, the average inter-observer, weighted Kappa, K_w, for the over-all risk was 0.32, i.e. the agreement above what could be expected by random was 32%. The intra-observer ditto was 0.41. The corresponding K_w for 8 body-part-ratings were in average 0.21, and 0.35.

The K_w of the inter-observer reliability for over-all risk in three levels were in OCRA 0.37, QEC 0.54, HARM 0.65, and SWEA 0.28. The K_w for specific body parts were, in QEC, 0.44 (shoulder), 0.49 (back), 0.67 (shoulder), 0.86 (neck), SI 0.47 (hand), ART 0.58 (left side) and 0.65 (right side). In the SWEA model, the K_w was below 0.4 for all five questions.

The relatively high reliability found for HARM ($K_w = 0.65$) was at first thought to depend on the clear pictures (photographs) in the scoring sheets showing neutral and awkward postures for wrists, shoulders and neck. But when observing the computed separate K_w for each rated HARM item, all K_w for items concerning repetition, movements and postures were below 0.3, i.e. at the same level as those without any specific method. Hence, the clear pictures did not seem to make it easier to rate postures and movements. The relatively high reliability of HARM instead seems to depend on the data of the work task that were given to the ergonomists (see methods above), for example the task length which in HARM has a high impact on the resulting estimated risk level. Throughout the methods, the K_w was generally the lowest for ratings of body postures.

As expected, the intra-rater K_w was somewhat higher than the corresponding inter-rater K_w in all methods, and the validity K_w. correlated with the inter-rater K_w. The obtained risk levels varied considerably between the methods, the pairwise K_w ranged from 0.10 (HARM-QEC) to 0.74 (ART-OCRA) (Kjellberg et al., 2015).

The mean score of 8 usability ratings was the highest for ART and the lowest for OCRA. OCRA also had the longest average assessment time.

As shown above, and which is in agreement with previous findings, there is a considerable variation not only between ergonomists' assessments of risks levels for MSDs in the observation methods, but also between methods. However, since observation without the use of any specific method have a low, non-acceptable reliability, it is recommended to use one or more systematic methods, and to a larger degree combine observations with validated methods of direct measurements.

New easy-to-use technical measurements of postures and movements

The three different types of methods all showed similar results to the previously validated methods (i.e. they are comparable with the previously used technical research methods). One example of a comparison is shown in Fig. 1, where a high correlation between the optical system and the iPhone system for arm flexion postures is shown.

Each one of the here described three methods for measurements of postures and movements have been presented at conferences and in education of ergonomists. The methods, especially the iPhone application, have been used by physiotherapists/ergonomists for workplace improvements, and in master thesis projects. The new methods are easier to use. They are also easier to wear (less obstructive) for the workers than the cabled sensors of the previous research methods. Because the validity of the new methods, they are now also used in research projects. During the development process, we received feedback from the OHS, and we counteracted the weaknesses that we and the OHS discovered.

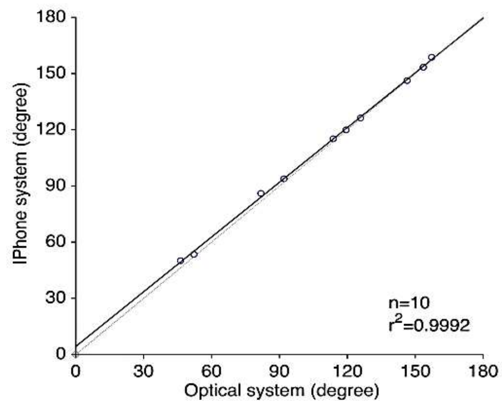

Figure 1. Upper arm inclination results from arm flexion validation experiment (Yang & Forsman, 2015).

The future of ergonomics risk assessments

A vision, which researchers in this field share, is that there should be practical tools available and that the ergonomists and other practitioners use them. Now, more and more methods are being developed, but there is a delay in the ergonomists' use of these methods. In Sweden the use of the QEC method (Takala et al., 2010) is increasing, also the use of HARM (Douwes & de Kraker, 2009) is becoming more frequently used. The methods are often presented as good if they cover many aspects, very little information is usually given about reliability issues. Also ergonomists, as well as other professionals, have their basic education and their usual way to work; it is hard to change the way you usually do things. A recommendation out of this research effort is to use systematic tools for risk assessments, at least the scientifically documented observational methods. When the results are presented from the part where no specific method was used, and a very low reliability in the assessments is revealed, ergonomists agree that systematic methods should be used. Systematic and direct measurement risk assessment methods should be included in the education programs for ergonomists.

Direct measurements are more reliable. But observational methods may cover more dimensions. A possible future is to use a combination, i.e. combined methods, where the dimensions of the lowest reliability in observational methods are replaced by technical, practical methods.

The new iPhone application (ErgoArmMeter; free to download from App Store) for upper arm posture and movement measurements is very easy to use, and may be a start for of ergonomists to a new way to work. The application only include upper arm, but other measurements, as e.g. wrist postures and movements should be possible to include, by utilizing gyroscopes, accelerometers, and magnetometers on external inertial measurement units (IMUs). These applications would then include measurements of repetitively, estimated by measured angular velocity. They could also include checklists for input of e.g. forces (measured or observed), additional factors such as auto-control (job-control), temperature, rest-schedules and vibrations.

A development of an easy-to-use method including postures, movements, and handled tum-grip forces, have been started by a consortium with Swedish universities and companies (e.g. KTH Royal Institute of Technology, Karolinska Institutet, Chalmers, Universities of Borås and Gävle; Scania, Volvo and Hultafors workwear). The project is called 'Smart textiles', and the plan is to integrate sensors in textiles. Measurements will be made easy, since analyses and interpretation will be carried out automatically.

More reference data from different occupational groups and guidelines with risk action limits are needed for those interpretations. However, a first version of such guidelines are already published as a report in Swedish from Lund University (Hansson et al., 2016), and there are ongoing studies on large material of technical measurements of Danish blue collar workers (Jørgensen et al., 2013).

CONCLUSIONS

The first of the two described projects where observational methods were evaluated showed that those methods, especially in ratings of postures, repetitions and forces, have low reliability. Also the inter-method reliability is often low. Therefore the risk assessment is heavily dependent on the observer and on the method chosen. The new technical measurements have high validity and reliability. They may be used by practitioners to investigate the postures and movements exposure in a work group in an efficient way. And the new Iphone application may be use to quickly and quantitatively compare work techniques (in training) and work stations (in production system development). In the search for practical and reliable risk assessment methods to be used in prevention of musculoskeletal disorders, today's low-cost electronic devices should be utilized to a higher degree. As the technical methods are much more reliable, they should replace observations, when it is possible in practice, and in designing of new risk assessment tools (which should be developed in collaboration with practitioners. Such new tools may combine measurements and checklists in smartphone applications).

ACKNOWLEDGEMENTS. The projects reported, the evaluation of observational methods, and the development and evaluation of technical methods were supported by the Swedish Research Council for Health, Working Life and Welfare (FORTE project no. 1212–1202), and AFA Insurance (Dnr: 120237), respectively. I am also thankful for all work by all co-workers in those two projects, who were in the first project: Ida-Märta Rhen, Kristina Eliasson, Teresia Nyman, Peter Palm, Katarina Kjellberg, Per Lindberg, Natalja Balliu and Xuelong Fan; and in the second project: Camilla Dahlqvist, Liyun Yang, Jens Wahlström, Carl Lind, Wim Grooten and Gert-Åke Hansson.

REFERENCES

Dahlqvist, C., Hansson, G.-Å. & Forsman, M. 2016. Validity of a small low-cost triaxial accelerometer with integrated logger for uncomplicated measurements of postures and movements of head, upper back and upper arms. *Applied Ergonomics* **55**, 108–116.

Davies, M.& Fleiss, J.L. 1982. Measuring agreement for multinomial data. *Biometrics* **38**, 1047–1051.

Douwes, M. & de Kraker, H. 2009. Hand arm risk assessment method (HARM), a new practical tool.In: *Proceedings of IEA 2009*. Beijing.

Eliasson, K. & Nyman, T. 2013. Practitioners knowledge, usage and need concerning risk management. In: *Proceedings of the Nordic Ergonomic Society annual conference.* Reykjavik, Iceland.

Eliasson, K., Palm, P., Nyman, T. & Forsman, M. 2017. Inter- and intra-observer reliability of risk assessment of repetitive work without any specific method. *Applied Ergonomics* **62**, 1–8.

European agency for safety and health at work 2010. *OSH in Figures: Work-related Musculoskeletal Disorders in the EU - Facts and Figures*. Luxembourg, pp. 103–150.

Forsman, M., Eliasson, K., Rhén, I.M., Nyman, T. & Kjellberg, K. 2015. Symposium – Observational Risk Assessment Methods. In: *Proceedings of IEA 2015*, Melbourne.

Hallgren, K.A. 2012. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology* **8**, 23–34.

Hansson, G.-Å., Arvidsson, I. & Nordander, C. 2016. Riktvärden för att bedöma risken för belastningsskador, baserade på tekniska mätningar av exponeringen (In Swedish; English title: 'Recommended action limit values for musculoskeletal risk assessments, based on technical exposure measurements'). *Rapport nr 4/2016.* Arbets- och miljömedicin – Lund. Sweden.

Jørgensen, M.B., Korshøj, M., Lagersted-Olsen, J., Villumsen, M., Mortensen, O.S., Skotte, J., Søgaard, K., Madeleine, P., Thomsen, B.L. & Holtermann, A. 2013. Physical activities at work and risk of musculoskeletal pain and its consequences: protocol for a study with objective field measures among blue-collar workers. *BMC Musculoskelet Disord*. doi: 10.1186/1471-2474-14-213.

Kalkis, H., Roja, Z. & Kalkis, V. 2014. Physical load analysis in hotel cleaning work. *Agronomy Research* **12**(3), 843–850.

Kjellberg, K., Lindberg, P., Nyman, T., Palm, P., Rhen, I.-M., Eliasson, K., Carlsson, R., Balliu, N. & Forsman, M. 2015. Comparisons of six observational methods for risk assessment of repetitive work - results from a consensus assessment. In: *Proceedings, 19th Triennial Congress of the International Ergonomics Association*, Melbourne.

Kong, Y.K., Lee, S.Y., Lee, K.S. & Kim, D.M. 2017. Comparisons of ergonomic evaluation tools (ALLA, RULA, REBA and OWAS) for farm work. *Int J Occup Saf Ergon*. **17**, 1–19.

Korshøj, M., Skotte, J.H., Christiansen, C.S., Mortensen, P., Kristiansen, J., Hanisch, C., Ingebrigtsen, J. & Holtermann, A. 2014. Validity of the Acti4 software using ActiGraph GT3X accelerometer for recording of arm and upper body inclination in simulated work tasks. *Ergonomics* **57**(2), 247–53.

Skotte, J., Korshøj, M., Kristiansen, J., Hanisch, C. & Holtermann, A. 2014. Detection of physical activity types using triaxial accelerometers. *J. Phys. Activity Health* **11**(1) 76–84.

SWEA - Swedish Work Environment Authority. AFS 2012:2 Ergonomics. *www.av.se.*

Takala, E.P., Pehkonen, I., Forsman, M., Hansson, G.A., Mathiassen, S.E., Neumann, W.P., Sjogaard, G., Veiersted, K.B., Westgaard, R.H. & Winkel, J. 2010. Systematic evaluation of observational methods assessing biomechanical exposures at work. *Scand. J. Work Environ. Health* **36**, 3–24.

Trask, C., Mathiassen, S.E., Wahlström, J. & Forsman, M. 2014. Cost-efficient assessment of biomechanical exposure in occupational groups, exemplified by posture observation and inclinometry. *Scand. J. Work Environ. Health* **40**, 252–265.

Yang, L. & Forsman, M. 2015. Development and validation of a novel iOS application for measuring arm inclination. In: *Proceedings of the Nordic Ergonomic Society annual conference*. Lillehammer, Norway.

Yang, L., Grooten, W.J.A. & Forsman, M. 2017. An iPhone application for upper arm posture and movement measurements. *Applied Ergonomics*. In press.

Westgaard, R.H. & Winkel, J. 1997. Ergonomic intervention research for improved musculoskeletal health: A critical review. *International Journal of Industrial Ergonomics* **20**, 463–500.