8-14-2009

# Systematic evaluation of observational methods assessing biomechanical exposures at work

E-P Takala

I Pehkonen

M Forsman

G-A Hansson

SE Mathiassen

*See next page for additional authors*



## Recommended Citation

Takala, E-P., Pehkonen, I., Forsman, M., Hansson, G-A, Mathiassen, SE, Neumann, W. Patrick, Sjogaard, G., Veiersted, KB, Westgaard, R., Winkel, J. Systematic evaluation of observational methods assessing biomechanical exposures at work. 17th World Congress on Ergonomics, IEA2009; Aug 9-14, 2009; Beijing, China: International Ergonomics Association.

**Authors**
E-P Takala, I Pehkonen, M Forsman, G-A Hansson, SE Mathiassen, W. Patrick Neumann, G Sjogaard, KB Veiersted, R Westgaard, and J Winkel

# Systematic evaluation of observational methods assessing biomechanical exposures at work

Human Factors Engineering Lab, Ryerson University

www.ryerson.ca/hfe

E-P. Takala, I. Pehkonen, M. Forsman, G-A Hansson, SE Mathiassen, W. Patrick Neumann, G. Sjogaard, KB Veiersted, R. Westgaard and J. Winkel

Presented at the

17th World Congress on Ergonomics, IEA2009; Aug 9-14, 2009; Beijing, China: International Ergonomics Association.

For a more in-depth look on this subject, please see:

Takala, E. P., Pehkonen, I., Forsman, M., Hansson, G.-Å., Mathiassen, S. E., Neumann, W. P., Sjøgaard, G., Veiersted, B., Westgaard, R. and Winkel, J. (2010) Systematic evaluation of observational methods assessing biomechanical exposures at work. Scandinavian Journal of Work, Environment & Health, v36(1): 3-24

# Systematic evaluation of observational methods assessing biomechanical exposures at work

Esa-Pekka Takala [1], Irmeli Pehkonen [1], Mikael Forsman [2], Gert-Åke Hansson [3], Svend Erik Mathiassen [4], W. Patrick Neumann [5], Gisela Sjøgaard [6], Kaj Bo Veiersted [7], Rolf Westgaard [8], Jørgen Winkel [9]

[1] Finnish Institute of Occupational Health, Helsinki [2] Karolinska Institutet, Stockholm, [3] Occupational and Environmental Medicine, Lund, [4] Högskolan i Gävle, Gävle, [5] Ryerson University, Toronto, [6] University of southern Denmark, Odense, [7] National Institute of Occupational Health, Oslo, [8] Norwegian University of Science and Technology, Trondheim, [9] University of Gothenburg and National Research Centre for the Working Environment, Copenhagen

The aim of this project was to identify and systematically evaluate observational methods to assess workload on the musculoskeletal system. Searches in the electronic databases and other sources identified 29 observational methods. The methods were evaluated for the aspects related to their reliability and usability for different purposes. The results of evaluation will be found in internet with a tool that helps the user to search for most suitable method by sorting the methods according to the several items evaluated.

Numerous methods have been developed to assess physical workload (biomechanical exposures) in order to identify hazards leading to musculoskeletal disorders, to monitor the effects of ergonomic changes, and for research. No single method is suitable for all purposes. The selection and use of methods has often been based on tradition rather than on a critical evaluation, because individuals are generally not aware of methods outside of their own realm of experience.

### Aims

The aim of this project was to systematically and critically evaluate observational methods presented in the literature, and to provide recommendations for their use.

## METHODS

### Search and selection of reference literature

Literature searches were conducted in the following electronic databases: PubMed, Embase, CISDOC, ScienceDirect, and Google Scholar.

The searches started with several combinations of the following search terms related to observational methods (OR): *observation*, workload, lifting, manual material handling, risk assessment, task analysis, posture, ergonomic, occupational exposure*. These terms were combined (AND) with terms related to musculoskeletal system (OR): *musculoskeletal, back, neck, extremities*.

The results were first screened by title and abstract. About 580 potential references were identified, including original scientific reports, reviews and internet sources. Full texts of these references were collated in electronic (or scanned) format for further evaluation. Many original articles referred to methods used only for a specific study and the methods were so scantly described that it was not possible to make an evaluation. We also excluded methods that were not observational, e.g. measurements of postural angles from the video recordings.

### Development of the framework for evaluation

There is no generally accepted standard way to evaluate methods for the assessment of workload. The structure and contents of the evaluation framework was developed in an iterative manner.

Reliability was evaluated with "Concurrent validity" (How well does the method correspond with more valid method/s?) and "Predictive validity" (How well has the risk-estimation of the method been shown to be associated with MSDs or has predicted them?). In addition we evaluated the intra- and inter-observer repeatability. Guidelines for the evaluation were prepared to keep consensus among the evaluators.

### Evaluation

Two researchers read the available reports for each method and independently filled in the evaluation forms. After that they discussed and reached consensus of the documentation of information to be written in the forms as basic description and documentation, to be further evaluated for reliability and validity.

Based on this documentation and original reports the evaluation was done independently and blinded of each other by at least two evaluators. Discrepancies were to be resolved by discussion between evaluators to establish consensus.

According to the protocol a third evaluator was prepared to participate to the discussion if no consensus was found but this option was not needed.

## RESULTS

A total of 29 observational workload assessment methods were identified. In the following description the methods are classified in three groups according to the main focus of the method: manual material handling, upper limbs, or a general approach.

Development of methods is related to the environment and time as regards to the needs to develop a new method. Thus the new methods have had some additional properties with respect to the previous ones. Therefore the methods are presented in a chronological order.

Tables 1 to 4 show the evaluation on correspondence with valid reference, association with MSDs, repeatability between observers (inter-observer repeatability), and who would be the potential users of the method.

**Methods to assess manual material handling**

We identified eight methods aimed for the assessment of manual material handling (MMH) (Table 1).

The NIOSH Lifting Equation is probably the most referred method to assess MMH, based on biomechanical, physiological and psycho-physiological research. The result of the evaluation is one figure indicating risk. With such an index it is hardly possible to assess validity by comparison with other methods, because no 'gold standard' is available. On the contrary, the NIOSH model has served as a 'gold standard' in comparative studies of Arbouw (in Netherlands), ACGIH Lifting TLV (USA), and Washington state model. MAC (UK), ManTRA (Australia), and New Zealand code are widely used for the assessment of risks in MMH but we did not find formal studies on validity of these methods. The inter-observer repeatability of MAC and the Washington state model has been found to be moderate.

BackEST is a method developed for epidemiological research.

Table 1 Methods to assess mainly manual material handling

| Method | Correspondence with 'valid' reference | Association with MSDs | Repeatability between observers | Potential users |
|---|---|---|---|---|
| NIOSH Lifting Eq. | NA | X | - | O, R |
| Arbouw | Mod | - | - | O |
| ACGIH Lifting TLV | Mod | - | - | O |
| MAC | - | - | Good - Mod | O, W(?) |
| ManTRA | - | - | - | O, R(?),W(?) |
| NZ Code for MH | - | - | - | O, W(?) |
| Washington state ergonomic rule | Mod | X | Mod | O, W(?) |
| BackEST | Mod-Low | - | Mod | R |

Correspondence with valid reference / Repeatability between observers: Good, Moderate (Mod), Low
- = No information or conflicting result, NA= Not applicable
Association with MSDs: L = Prediction in longitudinal studies, X = Association in cross-sectional studies,
Users: W=Workers/ supervisors, O=Occupational safety/health practioners, R=Researchers

**Methods to assess workload on upper limbs**

Table 2 lists methods for the assessment of workload in upper limbs.

HSE (UK) upper limb assessment method, Stetson's checklist, Keyserling's Cumulative trauma checklist, Ketola's upper limb expert tool, and the Washington state method are checklist-type methods where each item or risk factor exceeding the criteria used in each method indicates consideration of actions at work place.

In RULA weights are given to the observed items and a sum score is calculated to describe the risk. In ACGIH HAL the hand activity and the force used are estimated with visuo-analogue scale (VAS) and the need of actions is defined from nomograms. Strain Index has an approach similar to the NIOSH Lifting Equation and a multiplicative combination of observed items is calculated using weights for the multipliers according to their magnitude observed. OCRA has a similar principle but uses more complicated observation scheme for the input of the index.

Most of these methods have been compared with the other methods and have shown moderate correspondence. Strain index and ACGIH HAL have predicted upper limb disorders in prospective studies. RULA, OCRA and the Washington state method have been associated with MSDs in cross-sectional studies.

Table 2. Methods to assess workload on upper limbs. For explanation of symbols see Table 1.

| Method | Correspondence with 'valid' reference | Association with MSDs | Repeatability between observers | Potential users |
|---|---|---|---|---|
| HSE UL | - | - | - | O, W(?) |
| RULA | Mod- Low | X | Good - Mod | O, R |
| Stetson's checklist | - | - | Mod | R |
| Keyserling's Cumulative trauma checklist | Mod | - | Mod - Low | O, R |
| Ketola's upper limb expert tool | Mod-Low | - | Mod | O |
| Strain index | Mod | L, X | Good - Mod | O, R |
| ACGIH HAL | Mod | L, X | Good - Mod | O, R |
| OCRA | Mod | X | - | O, R |
| Washington state ergonomic rule | - | X | Mod | O, W(?) |

Table 3. General methods to assess workload. For explanation of symbols see Table 1.

| Method | Correspondence with 'valid' reference | Association with MSDs | Repeatability between observers | Potential users |
|---|---|---|---|---|
| OWAS | Mod | X | Good | R |
| AET | - | - | - | R |
| Posture targeting | - | - | - | O, R |
| PLIBEL | Mod(?) | - | Mod | O |
| PATH | Good - Mod | - | Good - Mod | O, R |
| REBA | - | - | Mod - Low | O, R(?) |
| LUBA | - | - | - | O(?),R(?) |
| QEC | Good | X | Mod | O, W, R(?) |
| Washington state ergonomic rule | Mod | X | Mod | O, W(?) |

Table 4. General methods to assess workload, developed for computerized input. For explanation of symbols see Table 1.

| Method | Correspondence with 'valid' reference | Association with MSDs | Repeatability between observers | Potential users |
|---|---|---|---|---|
| ERGAN (Arban) | - | - | - | ? |
| HARBO | Mod | - | Good | R |
| PEO | Mod | X | Good - Mod | R |
| TRAC | Mod | X | Good - Mod | R, O |
| Postural workload evaluation system by Chung | - | - | - | R, O(?) |

### General methods to assess workload

Tables 3 and 4 list general observational methods. The methods in table 4 have been developed for computerized input of data; even though software has later been developed for several methods listed in Tables 1-3.

OWAS is the most referred method to observe and code working postures. Other methods like PATH and TRAC have adopted postures developed for OWAS. The time consuming observation scheme and the decision rules based only on frequency distribution of items limit its use outside of research.

AET was aimed for general classifying of jobs but it also has items related to biomechanical exposures. Posture targeting is aimed to code postures of body parts and has illustrative output. In LUBA the rating of postures is based on psycho-physiological experiments. Chung's method is a video based extension of LUBA.

PLIBEL is an ergonomic checklist for screening of risk factors. REBA is an extension of RULA and gives a single sum score describing risk, although not validated in scientific studies. Validity and usability testing have been essential in the development of QEC that gives separate sum scores for different body parts.

PATH and TRAC use observation of work actions in addition to the postures. PEO was developed to measure exactly the duration and frequency of different postures. HARBO was developed parallel with PEO for the observation of postures by position of hands in epidemiologic studies.

## DISCUSSION

Systematic observation methods started to appear in the scientific literature some 30 years ago. The first ones, like OWAS for the whole body or RULA for the upper limbs, have been widely used and referred to since their original publication. The sampling of targets for observation (usually work tasks) has been either systematic, so as to get a frequency distribution of postures or actions (OWAS), or has – more often – been focused merely on "problematic situations". In the nineties, videos and computer software allowed methods to be developed that include assessment of variation of load over time (e.g. PEO, TRAC, PATH). The output of methods has been either descriptive profiles of the observed items, or the observed factors have been combined to an index describing risk (e.g. NIOSH lifting equation, Strain Index, OCRA). In the recent decades, formal studies on reproducibility and validity of the methods have been requested. Differences have then been found between the results obtained by using different observational methods simultaneously to assess the same target.

This study was limited to references which were available in common electronic databases. We found no simple strategy to combine the search terms to be effective and therefore extensive searches with known names of the methods as well as with the options of "related references" were performed. Still the references were restricted mainly to peer reviewed scientific publications. It is probable that much more methodological development has been done and reported in, for example, conference proceedings or as academic dissertations.

There is no standard way to perform a systematic analysis of methods assessing workload. The assessment of the validity of observational methods is problematic because the definition of a 'gold standard' is difficult. In several reports detailed measurements from video recordings have served as a more valid reference. Video recordings have several limitations due to the visual restrictions and projection error. Postures can be recorded with accurate technical measures. The correspondence of observations and technical measures has not been reported to be high except in few studies. In the comparisons of observational methods with detailed measurements from video or technical measures, the strict cut-off limits for the categories have been used in the analysis. If the real observations are close to these limits, even a small systematic error of the observers can result to misclassification and poor correspondence. A sensitivity analysis with different cut-off limits of the technical methods should be included in this kind of comparisons.

### Internet site of the results

Detailed information of the evaluated methods with full references will be found in

www.ttl.fi/workloadexposuremethods

The site also includes a tool to sort the methods according to the items used in the evaluation so that the user can select and compare methods for his/her purposes.

*Acknowledgments.* This study was financially supported by the Nordic Council of Ministers (Project 411040 - 70107)