2007

Validity of Nordic-style questionnaires in the surveillance of upper-limb work-related musculoskeletal disorders

Alexis Descatha  
Institut National de la Recherche Médicale

Yves Roquelaure  
Laboratory of Ergonomics and Occupational Health, Angers, France.

Jean Francois Chastang  
Institut National de la Recherche Médicale

Bradley A. Evanoff  
Washington University School of Medicine in St. Louis

Maria Melchior  
Institut National de la Recherche Médicale

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/ohs_facpubs

Recommended Citation  
Descatha, Alexis; Roquelaure, Yves; Chastang, Jean Francois; Evanoff, Bradley A.; Melchior, Maria; Mariot, Camille; Ha, Catherine; Imbernon, Ellen; Goldberg, Marcel; and Leclerc, Annette, "Validity of Nordic-style questionnaires in the surveillance of upper-limb work-related musculoskeletal disorders". Scandinavian Journal of Work, Environment & Health, 58-65. 2007.

This Article is brought to you for free and open access by the Occupational Health and Safety at Digital Commons@Becker. It has been accepted for inclusion in OHS Faculty Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Validity of Nordic-style questionnaires in the surveillance of upper-limb work-related musculoskeletal disorders

by Alexis Descatha, MD,1,2 Yves Roquelaure, MD,3 Jean François Chastang, PhD,1Bradley Evanoff, MD,4 Maria Melchior, PhD,1,5 Camille Mariot, MSc,3 Catherine Ha, MD,6 Ellen Imbernon, MD,6 Marcel Goldberg, MD,1,6 Annette Leclerc, PhD1


Objectives The study aimed at comparing results of standardized Nordic-style questionnaires with those of clinical examinations in two surveys on upper-limb work-related musculoskeletal disorders.

Methods The “repetitive task” survey (1757 workers in 1993–1994 and 598 workers in 1996–1997) studied risk factors of the disorders among those exposed to repetitive work. The “Pays de la Loire” survey (2685 workers in 2002–2003) was part of a population-wide surveillance system. In both surveys, each worker completed a Nordic-style questionnaire and underwent a standardized clinical examination. The presence of at least one upper-limb work-related musculoskeletal disorder was compared, with an evaluation of sensitivity, specificity, and kappa values, with a clinical examination as reference. In the second survey, a global score of a numerical scale for the severity of symptoms at the time of the examination was evaluated in the same way (plus ROC curves).

Results Agreement between the questionnaire and the examination differed in the two surveys, from kappa 0.22 [95% confidence interval (95% CI) 0.19–0.23] in the “Pays de la Loire” survey to kappa 0.77 (95% CI 0.74–0.80) in the “repetitive task” survey in 1993–1994. Overall, sensitivity was excellent (82.3–100%). The specificity varied, from 51.1% in the “Pays de la Loire” survey to 82.4% for the ≥2 score based on the severity of symptoms in the survey.

Conclusions Nordic-style questionnaires exploring symptoms in the past year can be useful tools for monitoring upper-limb work-related musculoskeletal disorders, especially if they include numerical rating scales of symptom severity. Physical examination remains essential for a medical or clinical diagnosis.

Key terms epidemiologic surveillance; predictive value; sensitivity; specificity.

One of the dimensions of the prevention of upper-limb work-related musculoskeletal disorders is the implementation of surveillance systems. These systems are based on various data sources, such as population-based or company-wide systems. The development and use of these systems imply that appropriate tools for evaluating the health status of the worker are available (1, 2).

Some authors have proposed a multi-level model for the surveillance of upper-limb work-related musculoskeletal disorders and their risk factors, such as a first level using questionnaires and checklists, for a rapid assessment, and a second level including clinical examinations and an in-depth job analysis by trained health care providers (1, 3, 4). Various questionnaires developed for musculoskeletal disorders can be used to assess functional status (5–12). However, they are better adapted to a clinical context than to a surveillance context. The standardized Nordic questionnaire, published

1 Institut National de la Recherche Médicale (National Institute for Health and Medical Research), U687-IFR69, HNSM, Saint-Maurice, France.
2 Occupational Health Department, Poincare Teaching Hospital AP-HP, Garches, France.
3 Laboratory of Ergonomics and Occupational Health, Angers, France.
4 Division of General Medical Sciences, Washington University School of Medicine, St Louis, Missouri, United States.
5 Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King’s College, London, United Kingdom.
6 Department of Occupational Health, National Institute for Public Health Surveillance (InVS), St-Maurice, France.

Reprint requests to: Alexis Descatha, INSERM U687-IFR69, HNSM, 14 rue du Val d’Osne, 94415 St-Maurice Cedex, France. [E-mail: alexis.descatha@rpc.aphp.fr]
Specific questions then concentrate on each anatomical region. The anatomical areas referred to in the original version were the lower back, the shoulders, and the neck. However, similar questionnaires can be developed with a focus on upper-limb anatomical areas, as was done in several studies in France. In the “repetitive task” survey, a rather detailed questionnaire was used, with 121 questions about the upper-limbs (20–23). In the “Pays de la Loire” survey, there were fewer questions (40 items), and the severity of symptoms at the time of the examination was also assessed using a numerical, rather than a dichotomous, measure (24).

**Design**

The objectives of the two surveys differed, the “repetitive task” survey having a descriptive and etiological objective and the “Pays de la Loire” survey concentrating on the surveillance of upper-limb work-related musculoskeletal disorders at a population-wide level. The study design and the population have been described earlier (20–24). In the presentation that follows, we describe the diagnostic criteria more precisely. Appendix 1 summarizes major points concerning the questionnaires and the clinical examination.

**The “repetitive task” survey**

**Study population.** In 1993–1994, 1757 workers completed a self-administered questionnaire about their work conditions and upper-limb work-related musculoskeletal disorders. Their mean age was 38 years, and the percentage of women was 76%. All of the workers underwent a standardized clinical examination by the occupational health physician responsible for medical surveillance in their company.

The workers were selected according to occupational criteria. They were required to be exposed to repetitive work in one of the following five activity sectors: (i) assembly-line manufacture of small electrical appliances, motor vehicle accessories, or ski accessories, packaging excluded; (ii) clothing or shoe industry, packaging excluded; (iii) food industry, packaging excluded (mainly the meat industry); (iv) packaging (primarily in the food industry); (v) supermarkets (cashiers). A final control group was made up of workers from the same industries who were not exposed to repetitive work. All of the groups included both women and men, except for the supermarket cashiers, who were exclusively women.

Eighteen of the thirty-nine occupational health physicians who participated in the 1993–1994 evaluations were able to repeat the study 3 years later. The 700 workers whom they had examined in 18 different firms in 1993–1994 were thus the target population of the longitudinal study. In all, 598 (85.4%) workers...
Validity of Nordic style questionnaires for health surveillance


Clinical examination. Each worker went through a standardized clinical examination, performed by the occupational physician immediately after the worker had completed the questionnaire. A list of criteria for the diagnoses was prepared for the clinical examination. These guidelines covered 33 diagnoses in 1993–1994 and 35 in 1996–1997 (the slight changes between the two lists were limited to shoulder tendonitis). One or two regional meetings with the occupational physicians took place before the baseline survey. A presentation of the guideline and training for the standardized physical examination was included in these meetings. The presentation of the guideline was included again in the regional meetings organized before the second survey.

The following three classifications of upper-limb work-related musculoskeletal disorders were possible from the clinical examination: (i) diagnosis "proved" during the medical examination, (ii) diagnosis "proved" before the medical examination (eg, previous diagnosis by a specialist for a problem present in the last 6 months), and (iii) suspected diagnosis (not all of the criteria were met in the medical examination or the diagnosis was based on the description of symptoms in the last 6 months but no longer present at the time of the examination). Our definition of upper-limb work-related musculoskeletal disorders determined by the clinical examination included "proved", suspected, and prior diagnoses.

Experimental network of epidemiologic surveillance of upper-limb work-related musculoskeletal disorders in the Pays de la Loire region ("Pays de la Loire" survey)

The data for this investigation were collected as part of a surveillance study of work-related upper-limb musculoskeletal disorders launched by the National Institute for Health Surveillance in France and set up in the Pays de la Loire region (Loire Valley district, West-Central France, population 3 220 000) (24).

Study population. All of the occupational physicians who practice in the Pays de la Loire region (N=460) were solicited to participate in the survey. A total of 80 agreed to participate. Each participating physician followed a standardized random selection procedure for the inclusion of a sample of workers (30 for those working full-time, 15 for those working part-time). The demographics and occupational distribution of the sample was similar to those of the salaried workforce in the Pays de la Loire region and were characteristic of France. In 2 years (2002 and 2003), 2685 workers were included, of whom 42% were women. The mean age of the workers was 38 years.

Questionnaire. Participating workers were invited to complete a 3-page Nordic-style questionnaire on upper-limb work-related musculoskeletal disorders in the previous 12 months. Neck pain was not considered in the study. Each worker marked a numerical scale (0 to 10) assessing the intensity of pain in each anatomical region at the time of the examination. A global score (GS) was calculated for the numerical scale rating by summing the region scores (minimum 0, maximum 40, neck region not considered). The Nordic-style questionnaire was followed by 12 pages with questions on work exposures, as well as general and medical items.

Clinical examination. The occupational physicians were trained by the study investigators to perform a standardized physical examination, based on an international protocol for the evaluation of work-related upper-limb musculoskeletal disorders (SALTSA) (25). The training was similar to that of the “repetitive task” survey. The physicians began the examination by asking the worker about upper-limb pain in the preceding 12 months. The participants with pain in the shoulder, arm, elbow, wrist, hand, or fingers underwent a standardized localized clinical examination. The physical examination allowed the detection of the following six disorders: rotator cuff syndrome, lateral epicondylitis, ulnar nerve entrapment at the elbow (cubital tunnel syndrome), extensor or flexor tendonitis or tenosynovitis, de Quervain’s disease, and carpal tunnel syndrome. The physical examination was considered to be positive if any of the six principal upper-limb disorders was present.

Analyses

The questionnaire was considered to be positive if the worker indicated at least one symptom in any region for the recall period. The clinical examination was considered to be positive in the “Pays de la Loire”...
survey if any of the specific clinical examination tests was positive. In the “repetitive task” survey, the presence of upper-limb work-related musculoskeletal disorders according to the clinical examination was defined to include cases proved in the examination, proved before the examination, or suspected. In order to increase comparability, we performed additional analyses restricted to the six disorders in the “Pays de la Loire” survey (“restricted analysis” in appendix 1).

The validity of the questionnaire was studied by calculating Cohen’s kappa, sensitivity, specificity, and predictive values with the clinical examination as the reference.

In the “Pays de la Loire” study, we conducted additional analyses using a more restrictive definition of a “positive questionnaire”, namely, presence of symptoms at the time of the examination (intensity >0 according to the GS of the numerical rating scale). Since several thresholds for the GS of the numerical rating scale were available for the definition of a “positive questionnaire”, a receiver operating characteristic (ROC) curve was drawn to choose the optimal (best sensitivity and specificity) threshold for the GS (26).

SAS v8.2 (27) and SPSS v11.01 (28) were used for the analyses.

Results

The validity of the questionnaires, with the clinical examination as the reference, is presented in table 1. The proportion of positive questionnaires was higher than the proportion of positive clinical examinations in both surveys. In the “repetitive task” survey, the proportion of positive clinical examinations was high (47% to 71%). The kappa coefficients were fair to good in the “repetitive task” survey. In the “Pays de la Loire” survey, the value of the kappa coefficients differed according to the definition of a “positive questionnaire”. If the definition was extensive (at least one symptom in the 12 months), the agreement between the Nordic questionnaire and the clinical examination was low. If the questionnaire case definition took into account the presence of symptoms at the time of the examination (GS >0), the agreement with the clinical examination was higher.

The negative predictive value was good in both surveys (64.5% to 100%), with a very good sensitivity (82.3% to 100%). The positive predictive value was high in the “repetitive task” survey and lower in the “Pays de la Loire” survey. The specificity was also lower in the “Pays de la Loire” survey, especially if the definition for a “positive questionnaire” was an extensive definition.

If the list of symptoms in the “repetitive task” survey was restricted to those in the “Pays de la Loire” survey, the sensitivity and specificity were similar in the two studies, but the predictive values differed (table 1). The ROC curve based on the GS score of the numerical rating scale in the “Pays de la Loire” survey showed that specificity was at 82.4% with sensitivity at 82.3% for a more restrictive case definition, with a score of at least 2 (figure 1). The area under curve was calculated at 0.85 [95% confidence interval (95% CI) 0.82–0.87].

### Table 1. Comparisons of Nordic-style questionnaire in the “repetitive task” and “Pays de la Loire” surveys, with a clinical examination as the reference. (GS = global score)

<table>
<thead>
<tr>
<th></th>
<th>Positive questionnaire</th>
<th>Positive clinical examination</th>
<th>True negative cases</th>
<th>False negative cases</th>
<th>False positive cases</th>
<th>True positive cases</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>True</td>
<td>Negative</td>
<td>True</td>
</tr>
<tr>
<td><strong>95% CI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>value</td>
<td>value</td>
<td>value</td>
</tr>
<tr>
<td><strong>Repetitive task</strong> survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete analysis a</td>
<td>1198</td>
<td>68.2</td>
<td>1155</td>
<td>65.7</td>
<td>490</td>
<td>69</td>
<td>112</td>
</tr>
<tr>
<td>1993–1994 (N=1757)</td>
<td>387</td>
<td>64.7</td>
<td>423</td>
<td>70.7</td>
<td>136</td>
<td>75</td>
<td>39</td>
</tr>
<tr>
<td>Restricted analysis b</td>
<td>1198</td>
<td>68.2</td>
<td>818</td>
<td>46.6</td>
<td>531</td>
<td>28</td>
<td>408</td>
</tr>
<tr>
<td>1993–1994 (N=1757)</td>
<td>387</td>
<td>64.7</td>
<td>306</td>
<td>51.2</td>
<td>170</td>
<td>41</td>
<td>122</td>
</tr>
<tr>
<td><strong>Pays de la Loire</strong> survey (N=2685)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Nordic questionnaire c</td>
<td>1490</td>
<td>55.5</td>
<td>345</td>
<td>12.8</td>
<td>1195</td>
<td>0</td>
<td>1145</td>
</tr>
<tr>
<td>Nordic questionnaire, GS &gt;0 d</td>
<td>731</td>
<td>27.2</td>
<td>345</td>
<td>12.8</td>
<td>1897</td>
<td>57</td>
<td>443</td>
</tr>
<tr>
<td>Nordic questionnaire, GS ≥2 d</td>
<td>696</td>
<td>25.9</td>
<td>345</td>
<td>12.8</td>
<td>1928</td>
<td>61</td>
<td>412</td>
</tr>
</tbody>
</table>

a Including all of the disorders in the clinical examination.
b Analyses restricted to six disorders.
c Including questions about last week or last year or both.
d Nordic style questionnaire with the GS of a numerical scale based on the severity of symptoms at the time of the clinical examination.
Validity of Nordic style questionnaires for health surveillance

Discussion

This study illustrates the use of Nordic-style questionnaires in two surveys with different designs and populations (appendix 1). In the “repetitive task” survey the population was highly exposed to repetitiveness, with a high prevalence of upper-limb work-related musculoskeletal disorders (40% to 70% in the clinical examination depending on the criteria). In the “Pays de la Loire” survey, the population was representative of the working population, with a lower prevalence of upper-limb work-related musculoskeletal disorders (13% in the clinical examination). Despite the differences between the two surveys, the sensitivity and specificity were similar if the analyses were restricted to the six diagnoses of the “Pays de la Loire” survey. The positive predictive values differed, however, as expected, since a positive predictive value depends on the prevalence of the outcome.

The administration of the questionnaire, especially its venue and context, could have had an effect on the reported prevalence of symptoms. In a study conducted among bus drivers, the prevalence rates of upper-limb work-related musculoskeletal disorders differed between survey formats, including a general health questionnaire with a Nordic style questionnaire completed before a periodic health examination and a health questionnaire completed after the periodic health examination (29). The authors concluded that prevalence increased if the participants focused their interest on the musculoskeletal system. In some contexts, the workers might also tend to under-report symptoms if they are afraid of possible consequences with respect to their jobs. In order to avoid information bias, workers must be confident about data protection.

In this study, we considered the clinical examination as the method of reference. The clinical examination in the two surveys included information on current and past symptoms and physical findings, as recommended and usually done in regular medical activity. However, a questionnaire is more formal and describes a rather long period of the history of symptoms, whereas a standardized physical examination generally describes only the current situation.

The use of a clinical examination as the reference method can be questioned (30). In the “Pays de la Loire” survey, only major disorders were investigated, and participants suffering from disorders without a clear diagnosis were not considered cases. This procedure could have decreased the positive predictive value of the clinical examination. However, symptoms also represent an important dimension. In a follow-up study comparing 129 clinically examined cases of upper-limb work-related musculoskeletal disorders and 655 controls, Nordlund & Ekberg (31) concluded that the level of questionnaire-based self-reported musculoskeletal symptoms predicts future health problems. Feuerstein et al (32) described a tool for predicting clinical outcomes for patients with upper-limb work-related musculoskeletal disorders, and this tool takes into account pain severity. Baron et al (16) concluded that the reliability and validity of symptom data were acceptable for the purposes of workplace ergonomics programs.

The authors of the Nordic questionnaire indicate that the validity of the questionnaire, studied in a small sample, was good (13). Olhsson et al (33) reported the sensitivity and specificity of a screening questionnaire for neck and upper limb complaints in a sample of 165 women. The results for sensitivity were similar to our findings (92% for the shoulder, 66%–79% for other regions). The specificity was similar to that found in the “repetitive task” survey and in the “Pays de la Loire” survey with a GS of >0. Palmer et al (17) assessed the validity of a Nordic-style questionnaire administered to a
sample of 105 hospital outpatients. Regional pain reports proved to be sensitive in relation to specific upper-limb work-related musculoskeletal disorders, in particular for cervical spondylosis (sensitivity 90% for neck pain in the last year), adhesive capsulitis (sensitivity 100% for shoulder pain in the last year), lateral epicondylitis (sensitivity 90% for elbow pain in the last year) and Raynaud’s phenomenon (sensitivity 78% for hand or wrist pain in the last year). However, with the exception of reported finger blanching in patients with Raynaud’s phenomenon, the specificity was low (range 33%–38%). Silverstein et al (34), in a study comparing a data source for the surveillance of upper-limb work-related musculoskeletal disorders, found that the sensitivity of a symptoms questionnaire, when compared with a physical examination plus interview, was relatively high (78% to 88%), but the specificity (21%–38%) and positive predictive value (31%–50%) were low.

In a preliminary study we conducted using data collected among workers in a shoe factory, the sensitivity of the questionnaire ranged from 65% to 70%, with a high negative predictive value (18). In this study, the possibility of bias in the prevalence rate had been discussed, with a high prevalence of upper-limb work-related musculoskeletal disorders according to the clinical examination (36% to 44%) and with only 46% to 58% positive questionnaires (18). This finding may have been due to the underreporting of symptoms when workers were not confident about the protection and use of the data. Sensitivity may have been overestimated in the “Pays de la Loire” survey, in which the sensitivity was 100%; in that study a physical examination was performed only if the interview of the worker indicated the presence of symptoms at the time of the examination. If the interview by the occupational physician and the questionnaire give similar information, then sensitivity is expected to be high. However, the agreement between symptoms reported by the Nordic-style questionnaire and by the physician-led interview was poor (kappa 0.27, 95% CI 0.25–0.29), 39% of the workers having been misclassified. We found the strongest disagreement if the questionnaire was positive and the interview was negative (38.9%, N=1045). This difference could be partly explained by the period of symptoms, which was longer in the questionnaire than in the interview.

In our study, the values for specificity ranged from 51% to 82%. The rather low values may have been due to the fact that the time period of the questionnaire and the physical examination differed, between the past year or the past six months for symptoms in the questionnaire and present time for the physical examination, at least in the “Pays de la Loire” survey. Some workers could have been sick and could have recovered during this period; such an occurrence would have artificially increased the number of false positive cases and would have decreased the specificity of the questionnaire. In the “repetitive task” survey, the physical examination included symptoms no longer present at the examination (or “prior diagnosis” when made by a specialist) in the “suspected diagnosis” category. The specificity was thus very high in this study (>75%).

In the “Pays de la Loire” survey, the specificity increased from 51% to 81% if the answers according to the numerical scale of pain intensity for each region at the time of the consultation were taken into account. The lower specificity in the “Pays de la Loire” survey could be explained by the limited number of diagnoses in the physical examination, and this limited number is consistent with the results obtained for the “repetitive task” survey with a restricted list of diagnoses.

In conclusion, Nordic-style questionnaires exploring symptoms in the past year can be considered useful tools for the surveillance of upper-limb work-related musculoskeletal disorders, especially if they include numerical rating scales on symptom severity. The physical examination remains essential for the establishment of a medical or clinical diagnosis. For other purposes, questionnaires remain useful tools, giving information on functional, psychological, and psychosocial dimensions of musculoskeletal disorders.

Acknowledgments

The “Pays de la Loire” survey was conducted with financial support from the National Institute for Public Health Surveillance (InVS).

We are grateful to the study group on repetitive work involved in this survey, as well as to the 39 occupational physicians who participated. We are also grateful to the 80 occupational physicians involved in the sentinel network for the “Pays de la Loire” survey. We thank Lawrence J Fine for his contribution to the preparation of the questionnaire of the “repetitive task survey” in 1992.

Reference

4. Ricci MG, De Marco F, Occhipinti E. Criteria for the health...

Received for publication: 18 May 2006
### Appendix 1

The questionnaire and the clinical examination criteria used in the “repetitive task” and “Pays de la Loire” surveys

<table>
<thead>
<tr>
<th>Questionnaire on upper-limb work-related musculoskeletal disorders</th>
<th>“Repetitive task” survey</th>
<th>“Pays de la Loire” survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place</td>
<td>The worker alone in the occupational health service</td>
<td>The worker alone in the occupational health service</td>
</tr>
<tr>
<td>Number of pages</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Place in the questionnaire</td>
<td>At the end</td>
<td>At the end</td>
</tr>
<tr>
<td>Symptoms</td>
<td>Pain and discomfort</td>
<td>Pain and discomfort</td>
</tr>
<tr>
<td>Recall period</td>
<td>Six months</td>
<td>Six months</td>
</tr>
<tr>
<td>Region</td>
<td>Neck or shoulder to fingers</td>
<td>Neck or shoulder to fingers</td>
</tr>
<tr>
<td>Drawing</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Criterion for positive case</td>
<td>At least one symptom in the last 6 months</td>
<td>At least one symptom in the last 6 months</td>
</tr>
<tr>
<td>Clinical examination</td>
<td>Standardized criteria specific for the study</td>
<td>Standardized criteria specific for the study</td>
</tr>
<tr>
<td>Criteria</td>
<td>SALTSA criteria</td>
<td></td>
</tr>
<tr>
<td>Number of clinicians</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>1993–1994</td>
<td>·</td>
<td></td>
</tr>
<tr>
<td>1996–1997</td>
<td>·</td>
<td></td>
</tr>
<tr>
<td>Number of disorders</td>
<td>·</td>
<td></td>
</tr>
<tr>
<td>1993–1994</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>1996–1997</td>
<td>·</td>
<td></td>
</tr>
<tr>
<td>Uncommon disorders</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Specific tests</td>
<td>Systematic specific tests</td>
<td>Systematic specific tests</td>
</tr>
<tr>
<td>Blinding of examiners (towards symptoms)</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

---