Original article

Scand J Work Environ Health 2004;30(5):379-389

doi:10.5271/sjweh.826

Exposure to job stress factors in a national survey in France

by Cohidon C, Niedhammer I, Wild P, Guéguen A, Bonenfant S, Chouanière D

Affiliation: Institut de veille sanitaire, Département santé travail, 12, rue du Val d'Osne, 94415 Saint Maurice Cedex, France. c.cohidon@invs.sante.fr

Refers to the following texts of the Journal: 2001;27(5):0 1998;24(3):0

Key terms: exposure; factor, psychosocial; france; job stress; job-exposure matrix; matrix, job-exposure; segmentation method; survey; work

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/15529801

Original article

Scand J Work Environ Health 2004;30(5):379–389

Exposure to job-stress factors in a national survey in France

by Christine Cohidon, MD, PhD,[1] Isabelle Niedhammer, PhD,[2] Pascal Wild, PhD,[1] Alice Guéguen, PhD,[2] Sebastien Bonenfant, MSc,[2] Dominique Chouanière, MD[1]

Cohidon C, Niedhammer I, Wild P, Guéguen A, Bonenfant S, Chouanière D. Exposure to job stress factors in a national survey in France. *Scand J Work Environ Health* 2004;30(5):379–389.

Objectives Job stress is a growing epidemiologic field in France. The aim of this study was to evaluate the levels of occupational psychosocial risk factors to which the French working population is exposed. It also focused on developing a job-exposure matrix.

Methods This study used existing French national data on work conditions collected by the French Ministry of Labor in 1991, in which 20 929 workers were interviewed by questionnaire. The items of the questionnaire were retained that represented potential stressors. A principal component analysis was performed to summarize the data in terms of job-stress factors. Four independent variables (gender, age, occupation, and activity) were available, as well as scores for exposure to psychosocial risk factors. The CART (classification and regression tree) segmentation method was used to construct the job-exposure matrix.

Results Fourteen psychosocial factors were identified and interpreted on the basis of the results of the principal component analysis and deepened by experts' judgment. To take into account the well-known difference of distribution in occupations between men and women, a gender-specific matrix was developed.

Conclusions This study is the first attempt to develop a job-exposure matrix in the area of job-stress factors in France. Thus the results have allowed the assessment of exposure to 14 psychosocial factors for all of the 455 categories of the French occupational classification for men and women separately.

Key terms French national study, job-exposure matrix, psychosocial factors at work, segmentation method.

For a long time, occupational epidemiology has been limited in scope in the study of physicochemical risks. In 1979, Karasek (1) began to take an interest in the effects of job stress and to identify the role of certain psychosocial factors at work in the onset of health disorders (2), including psychological demands, decision latitude, and social support, the last mentioned subsequently being introduced by Johnson et al (3). Despite its limits, which have often been stressed (4–7), and the appearance of other models, like the effort–reward imbalance model by Siegrist and his co-workers (8–10), the model of Karasek remains a reference in the study of psychosocial factors at work. Schnall et al (6) and Kristensen (7) have proposed alternative models that are beginning to be used that include factors like physical workload, chemical or physical risk factors, job insecurity, and the use of new technologies.

As Hagberg et al evoked, during the X2001 Exposure Assessment in Epidemiology and Practice conference in Goteborg on June 2001, "Performing adequate exposure assessments is an important component of the effort to improve health and the quality of life of working populations [p 354]" (11). Following the example of the job-exposure matrices for physicochemical exposures, the countries in the forefront of the study of psychosocial factors at work, namely, the United States and the Scandinavian countries, have developed job-exposure matrices specific to these factors (12–14). Hence, for a given job often described by a job title or an occupational category, exposure levels to the factors considered are assigned. The advantage of such matrices lies in the possibility of matching an assessment of exposure to one or several occupational risk factors to a job title (15, 16).

The existing three matrices distinguish two domains, psychosocial factors (a person's perception of psychosocial work conditions and interpersonal relationships at the workplace) and physicochemical risk factors. The

1 Département Epidémiologie en Entreprise (Epidemiology Department), INRS (French National Research and Safety Institute), Vandoeuvre, France.
2 INSERM Unit 88 (National Research Institute on Health and Medicine Unit 88), Saint-Maurice, France.

Reprint requests to: Dr Christine Cohidon, Institut de veille sanitaire, Département santé travail, 12, rue du Val d'Osne, 94415 Saint Maurice Cedex, France. [E-mail: c.cohidon@invs.sante.fr]

factors explored by these three matrices are not identical; instead they broadly cover both domains. The methods used in their development are similar and are based on calculations of the exposure means (and standard deviation) per job title. In general, the data stem from national surveys on work conditions carried out with other aims. Using a mean exposure score for each job title entails the dual problem of the validity and precision of this score when the number of job titles is large. The national job-coding schemes usually contain several hundred items, and the surveys on work conditions usually lack sufficient numbers for each job title to yield valid exposure estimates. In this case, the only alternative is, as done by Johnson & Stewart (13) and Schwartz et al (14), to group jobs with neighboring job codes to obtain large enough numbers of subjects to achieve a reasonable precision. The intrinsic risk with this approach is that the scores are less specific for each job and therefore less valid. We resolved this trade-off by using a segmentation method, the CART (classification and regression tree) segmentation method, which was developed to optimize the prediction of the individual exposure (17) by the mean of the exposure groups obtained by this method. Furthermore, this method not only uses the a posteriori information from the data in the process of the job-exposure-matrix construction, but also the a priori expert-based hierarchical structure of the data.

In France, the evaluation of exposure to psychosocial work factors has been the subject of only one publication, which describes these factors according to a very global classification of job title (18). Studies are still very infrequent in this domain and are either focused on specific occupational sectors (19, 20) or limited to certain categories of employees (21). Periodic national surveys do, however, exist on the work conditions of the French population, and they explore the area of psychosocial work factors (18). The aims of this paper were (i) to characterize, on the basis of a large French national survey, the occupational risk factors to which the French working population is exposed in the area of job stress, (ii) to test the relevance of the CART segmentation method for this problem in comparing its results to the crude exposure estimates (mean and standard deviation for each occupational category of the French classification), and (iii) to discuss the advantages of such a method.

Study population and methods

Work conditions

The survey. The data used stem from the 1991 national survey of work conditions in France carried out by the French Ministry in charge of Labor and the National Institute for Statistics and Economic Studies. This survey, first launched in 1978, was repeated in 1984, 1991, and 1998 in a sample of the French working population, whether in salaried employment or otherwise. Among the entire French working population identified in the census of the same year, the 1991 survey defined a sample basis of 1/900 representing 20 929 working men and women (18).

The data were collected by a questionnaire completed by a surveyor at the home of the person.

DARES (the research, development, studies and statistics executive of the French Ministry of Labor), forwarded the data of the 1991 survey for 20 929 workers to the Epidemiology Department of the French National Research and Safety Institute. The 1998 data were not yet available when this work was undertaken.

Potential stressors. The questionnaire used in 1991 comprised 50 major questions, exploring different aspects of work conditions. These questions often included some subquestions that were of different types, either open (eg, hour of leaving home) or categorical with different numbers of categories. Among these 50 questions, we retained 34 (excluding questions that were irrelevant for our purpose, eg, questions on the use of new technologies), from which 45 binary items were obtained and which were chosen to represent potential stressors (list of items in appendix 1). The overall rate of missing values was low (0.5%), the percentage of positive answers varying from 3.8% to 83.9% according to the item.

The psychosocial domain could be subdivided into the following four subdomains: (i) workhours, (ii) work organization, (iii) social and relational aspects, and (iv) psychocognitive strain.

Independent variables. Information on gender and age were available for each person. The following four age classes were defined: <30, 30–39, 40–49, and ≥50 years.

The job was described by two variables, namely, the PCS (Profession et Catégorie Sociale) code of the French classification of occupations (22) and the NAF (Nomenclature des Activités françaises) code, based on the European NACE (European Community Activities Nomenclature) coding, which characterizes the sector of activity (23). The PCS and NAF codes were available according to four hierarchical levels. The crudest codes comprise one digit for the PCS (6 codes) and one letter for the NAF (17 codes). The most-detailed codes comprise four digits for the PCS (455 codes) and one letter and three digits for the NAF (494 codes). To make the international comparisons easier, the PCS titles were defined as close as possible to those of the International Standard Classification of Occupations (ISCO-88) (24).

Methods

Definition of job-stress factors. This step involved defining the exposures using the information contained in the 45 psychosocial items retained. A principal component analysis (PCA) (25) was carried out using SAS software (26) by the FACTOR procedure, which allowed extraction of the factors or axes and their selection when their eigenvalue exceeded 1 (we also performed a scree plot to confirm our selection); a VARIMAX (orthogonal) rotation [with SAS also (26)] was then used to facilitate interpretation.

Only factor loadings of ≥0.40 were considered significant. Each factor was described according to its items and subsequently interpreted in terms of exposure. The retained exposure partially identified the three concepts of Karasek's model and allowed also the identification of new domains. The results of the PCA were then submitted to the judgment of job-stress experts to validate the relevance of the factors. In the end, the job-stress factors were defined as PCA factors validated or possibly modified a posteriori by the experts.

When a factor was composed of several items, exposure was defined by the algebraic sum of each item (coded 0 for absence, 1 for presence for this item). When an exposure amounted to only one item, exposure was defined by the presence or absence of this item. For clarity and to facilitate comparison between exposures, all the scores were converted to a 0–100 scale, 100 corresponding to the maximal score. When only one item existed, this score corresponded to the percentage of positive responses.

The results of the PCA were then submitted to the judgment of experts in order to validate the relevance of the factors. Then, the job stress factors were defined as PCA axes validated a posteriori by experts.

Following this first step, the scores for exposure to different psychosocial risk factors were available for each person.

Development of the matrix by the CART segmentation method. The CART segmentation method (27) allows individualization of groups that are homogeneous in terms of the exposure studied [homogeneous exposure groups (HEG)]. Individualization of the HEG is obtained in such a way that the exposure between group variance is maximum and the within-group variance is minimum while a minimum of robustness is kept and has been shown to be optimal for prediction purposes (27). The technique uses an algorithm that contains three steps. First, after the whole sample has been randomly divided into a base sample and a validation sample, an overfitting tree is grown, using the base sample. The base sample is split successively either by age, the PCS, or the NAF according to a splitting criterion proposed by Breiman et al (27), which helps prevent splits that lead to nodes that are too imbalanced. The quality of a tree is measured by its explanatory value, defined as [1 – relative cost of the tree], the relative cost of the tree being equal to the weight sum of the impurities of the terminal nodes divided by the impurity of the root node. It ranges from 0 to 1 and is comparable to the R2 that measures the explanatory value of multiple regression or the intergroup variance ratio divided by the total variance obtained in an analysis of variance. Second, this overfitted tree is successively pruned back so that a sequence of nested subtrees is obtained. The sequence goes from the simplest subtree with a unique terminal node containing the whole population to the last subtree that contains the maximum of terminal nodes. Finally, the validation sample is run though all the subtrees of the sequence to select the optimal tree, the quality of each subtree of the sequence being estimated using the validation sample. While the quality of the subtrees of the sequence always increases when estimated by the base sample, generally, when estimated by the test sample, it increases at the beginning of the sequence, reaches a maximum, and then decreases. This decrease corresponds to the overfitted part of the tree. Therefore, the optimal tree is the one that has the highest quality estimated by the validation sample. Thus the construction of the segmentation uses both samples, although, asymmetrically, the mean scores for each node (exposure group) are, however, computed using all the data. This method allows the assessment of exposure to be obtained for HEG, including those containing few persons. In addition, a specific method was developed to take into account the hierarchical structure of the NAF and PCS codes in the definition of homogeneous exposure groups (see appendix 2). For each job-stress factor, a "tree" was constructed using the CART software (28). Thus, for each "tree" (or each job-stress factor exposure), the persons were divided into terminal nodes (or HEG) described by one or several PCS codes, one or several NAF codes, one or several classes of age, and a mean score for the exposure studied. Thus the HEG differed according to the stress factor studied. Details of this procedure have been given by Guéguen et al (29).

Given the different PCS distribution between the women and men, a gender-specific matrix was constructed for each job-stress factor. The method was also applied without taking the NAF code into account. The development of these matrices allowed the results of the CART method to be compared with the crude mean exposure scores (and standard deviation) for each PCS code. These matrices also allow levels of exposure to be assigned, within the context of epidemiologic studies, to existing data including the PCS variable, but without information on the activity sector (NAF code).

Results

Population

Of the 20 929 working people, 20 868 persons, 11 643 men and 9 225 women, were included in the matrix development study (persons for whom the NAF code had not been recorded were excluded). Table 1 describes the age and job characteristics of this population.

An examination of these criteria showed that the sample was representative of the French working population.

Definition of job-stress factors

The principal component analysis carried out on the items of the psychosocial domain highlighted 12 factors that explained 20% of the variance (table 2). Of these 12 factors, 11 comprised several items (ranging from 2 to 7). Factor 12 comprised only one item, "isolated work". Factor 10 was defined by the items "management function", interruption of work through "under productivity", and "low work social support". As these three items were related to different domains in the literature, we preferred to consider these three items separately. Thus 14 psychosocial factors were defined.

Table 3 presents, as an illustration, the mean scores for each of the job-stress factors defined, for the men and women, without distinction of age or job characteristics.

Table 1. Distribution of the population by age, occupation, and activities.

	Men		Women	
	N	%	N	%
Age				
<30 years	2589	22.2	2290	24.8
30–39 years	3586	30.8	2773	30.1
40–49 years	3241	27.8	2411	26.1
≥50 years	2227	19.1	1751	19.0
Occupational categories (one-digit codes, 6 codes)				
1. Farmers	624	5.4	393	4.3
2. Craftsmen and shopkeepers	1082	9.3	587	6.4
3. Managers and professionals	1583	13.6	667	7.2
4. Associate professionals and technicians	2398	20.6	1884	20.4
5. Clerks	1291	11.1	4445	48.2
6. Skilled and nonskilled workers	4665	40.1	1249	13.5
Activities (one-letter codes, 17 codes)				
A. Agriculture, hunting, forestry	811	7.0	458	5.0
B. Fishing, aquaculture	15	0.1	4	0.0
C. Mining industries	76	0.6	8	0.1
D. Manufacturing industries	2954	25.3	1373	14.9
E. Production and distribution of electricity, gas and water	149	1.3	43	0.5
F. Construction	1450	12.4	163	1.8
G. Commerce, car and household-goods repair	1572	13.5	1389	15.1
H. Hotels and restaurants	319	2.7	328	3.6
I. Transport and communications	879	7.5	325	3.5
J. Financial activities	301	2.6	350	3.8
K. Real estate, leasing and services to enterprises	815	7.0	751	8.1
L. Public administration	925	7.9	875	9.5
M. Education	492	4.2	982	10.6
N. Health and social work	515	4.4	1461	15.5
O. Collective, social and personal services	344	2.9	424	4.6
P. Domestic services	17	0.1	279	3.0
Q. Activities outside France	9	0.1	12	0.1
Total	11643	100	9225	100

Development of the job-exposure matrix

The full set of results will be the subject of a paper that will be published later. Here, to illustrate the CART method, we present, as an example, the "tree" concerning the exposure to "absence of latitude" for the men (figure 1). The first division of persons was based on a one-digit PCS code. Two groups with the greatest possible exposure contrasts were individualized, the persons belonging to the first four one-digit PCS codes [ie, farmers, craftsmen-shopkeepers, managers, professionals, and the associate professionals or technicians (intermediate node 2)] and persons belonging to the last two one-digit PCS codes [ie, clerks and skilled and nonskilled workers (intermediate node 7)]. Intermediate node 7 was split with regard to sector. Thus, for instance, the intermediate node 13 corresponded to blue-collar workers and clerks from the sector-grouping 3, comprising mining industries, manufacturing industries, construction, and transport as detailed in the footnote of figure 1.

This principle of segmentation was identical for the construction of the other nodes. The divisions in the left-hand part of the tree were first performed by the CART procedure on the PCS codes then on the activity codes. In the right-hand part of the tree, grouping the skilled and nonskilled workers and clerks, the segmentation was first performed on the activity by CART. The subsequent divisions were then performed on the PCS codes. Only one segmentation was carried out on the age variable (intermediate node 11). On the whole, 15 terminal nodes or homogeneous exposure groups were identified. Farmers, craftsmen-shopkeepers, and professionals (node 1) were recognized as having the lowest exposure to the lack of decision authority. Managers, with low exposure, were found in two HEG differing with respect to activity (nodes 2 and 3). People belonging to the category of associate professionals or technicians were divided into three HEG according to the type of activity, administrative and commercial staff versus technicians and foremen (nodes 4, 5, and 6). Five HEG grouped certain types of skilled and nonskilled workers and clerks (nodes 7, 10, 12, 13, and 14). Unskilled workers of the industrial sectors were the most exposed to this constraint (node 15). The use of the CART method

Table 2. Principal component analysis of the psychosocial items (N=20 868).

Factors	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	Communality
Absence of latitude													
Fixed workhours	0.56	·	·	·	·	·	·	·	·	·	·	·	0.45
Control of hours	0.55	·	·	·	·	·	·	·	·	·	·	·	0.39
Being told how to do the job	0.60	·	·	·	·	·	·	·	·	·	·	·	0.43
Inability to vary deadlines	0.38	·	·	·	·	·	·	·	·	·	·	·	0.38
Work rhythm imposed by supervision	0.42	·	·	·	·	·	·	·	·	·	·	·	0.31
Application of instructions to carry out the work	0.64	·	·	·	·	·	·	·	·	·	·	·	0.47
Application of instructions in the case of an incident	0.61	·	·	·	·	·	·	·	·	·	·	·	0.40
Atypical hours													
Period of work in excess of 10 hours	·	0.57	·	·	·	·	·	·	·	·	·	·	0.43
Leaving home before 7 o'clock in the morning	·	0.38	·	·	·	·	·	·	·	·	·	·	0.36
No break for 48 consecutive hours	·	0.77	·	·	·	·	·	·	·	·	·	·	0.63
Working more than one Sunday in two	·	0.72	·	·	·	·	·	·	·	·	·	·	0.57
Working more than one Saturday in two	·	0.80	·	·	·	·	·	·	·	·	·	·	0.70
Repetitive work													
Assembly line worker	·	·	0.76	·	·	·	·	·	·	·	·	·	0.61
Work rhythm imposed by the movement of a work piece	·	·	0.80	·	·	·	·	·	·	·	·	·	0.67
Work rhythm imposed by the rate of a machine	·	·	0.74	·	·	·	·	·	·	·	·	·	0.61
Repetitive work	·	·	0.31	·	·	·	·	·	·	·	·	·	0.36
Pressure related to error consequences													
Serious consequences if error occurs	·	·	·	0.75	·	·	·	·	·	·	·	·	0.62
Financial consequences of an error	·	·	·	0.72	·	·	·	·	·	·	·	·	0.62
Dangerous consequences of an error	·	·	·	0.58	·	·	·	·	·	·	·	·	0.52
Consequences of an error on job stability	·	·	·	0.60	·	·	·	·	·	·	·	·	0.49
Contact with the public													
Working in direct contact with the public	·	·	·	·	0.73	·	·	·	·	·	·	·	0.63
Work rhythm imposed by an immediate response	·	·	·	·	0.76	·	·	·	·	·	·	·	0.63
Work rhythm imposed by a nonimmediate response	·	·	·	·	0.56	·	·	·	·	·	·	·	0.49
Experiencing situations of tension with the public	·	·	·	·	0.53	·	·	·	·	·	·	·	0.40
Absence of flexibility													
Work rhythm imposed by technical constraints	·	·	·	·	·	0.38	·	·	·	·	·	·	0.32
Work rhythm imposed by dependency on colleagues	·	·	·	·	·	0.45	·	·	·	·	·	·	0.38
Work rhythm imposed by an hourly production rate	·	·	·	·	·	0.67	·	·	·	·	·	·	0.49
Work rhythm imposed by a daily production rate	·	·	·	·	·	0.71	·	·	·	·	·	·	0.55
Inability to achieve objectives													
Insufficient time to achieve objectives	·	·	·	·	·	·	0.73	·	·	·	·	·	0.57
Insufficient information to achieve the objectives	·	·	·	·	·	·	0.66	·	·	·	·	·	0.49
Insufficient numbers to achieve the objectives	·	·	·	·	·	·	0.64	·	·	·	·	·	0.43
Concentration requirement													
Perceived inability to leave work	·	·	·	·	·	·	·	0.55	·	·	·	·	0.47
Reading small-size numbers or letters	·	·	·	·	·	·	·	0.64	·	·	·	·	0.52
Examining very small objects	·	·	·	·	·	·	·	0.69	·	·	·	·	0.50
Paying attention to brief audible or visual signals	·	·	·	·	·	·	·	0.57	·	·	·	·	0.48
Incompatibility between workhours and social rhythms													
Variable number of workdays in the week	·	·	·	·	·	·	·	·	0.58	·	·	·	0.39
Alternating hours of work	·	·	·	·	·	·	·	·	0.61	·	·	·	0.48
Night work	·	·	·	·	·	·	·	·	0.57	·	·	·	0.48
Having less than 30 minutes for lunchbreak	·	·	·	·	·	·	·	·	0.53	·	·	·	0.44
F.10 (as described on page 382)													
Low work social support	·	·	·	·	·	·	·	·	·	0.48	·	·	0.41
Management function	·	·	·	·	·	·	·	·	·	-0.37	·	·	0.50
Under-productivity	·	·	·	·	·	·	·	·	·	0.64	·	·	0.47
Polyvalence													
Frequent requirement to abandon one task for another	·	·	·	·	·	·	·	·	·	·	-0.49	·	0.48
Change of task according to requirements	·	·	·	·	·	·	·	·	·	·	-0.68	·	0.53
Isolated work	·	·	·	·	·	·	·	·	·	·	·	0.73	0.60
Percentage of inertia, before rotation	9.00	7.00	6.00	4.00	4.00	3.00	3.00	3.00	3.00	2.00	2.00	·	·
Eigenvalue, before rotation	4.13	3.35	2.67	1.82	1.65	1.46	1.35	1.30	1.23	1.14	1.05	·	·

in the case in which the PCS code and the NAF code characterize the job and where this is characterized by the PCS code alone, did not change the variance explained for this factor (55% in both cases). The explained variance was, however, usually lower for the other psychosocial factors (with the exception of atypical hours). Similar findings concerning low between-job variation were mentioned by Karasek & Theorell (2) in the case of job demand and social support–social relations issues. The NAF code did not increase the explained variance for any factor with the possible exception of "low work social support", for which the explained variance increased from 4% with the PCS alone to 7%, including the NAF and "incompatibility between workhours and social rhythms" (14% versus 18%) among the men.

The exposure scores of the HEG obtained by the CART method were compared with the crude means. For example, part of this comparison is presented in table 4. The exposure levels obtained by both methods agreed. However, for a few rare job titles for which the numbers of persons were very low (directors of secondary teaching establishments, researchers, and pharmacists), differences were observed between the two methods.

Table 3. Exposure to job-stress factors among French men and women in 1991. (SE = standard error)

Psychosocial factors	Number of items	Men (N=11 643)		Women (N=9 225)	
		Mean	SE [a]	Mean	SE [a]
Absence of latitude [b]	7	34.2	27.3	37.5	25.7
Atypical hours [b]	5	22.7	26.1	19.9	24.8
Repetitive work [b]	4	9.7	18.5	11.3	19.1
Contact with the public [b]	4	41.7	33.3	45.4	32.3
Inability to achieve objectives [b]	3	18.8	28.1	19.7	28.7
Pressure related to error consequences [b]	4	52.4	33.5	35.9	32.1
Concentration requirement [b]	4	20.0	25.7	16.7	23.4
Absence of flexibility [b]	4	22.6	25.7	16.8	23.3
Incompatibility between workhours and social rhythms [b]	4	16.6	23.8	14.4	20.3
Polyvalence [b]	2	39.1	37.0	36.9	36.1
Management function [c]	1	31.1	0.5	14.1	0.4
Isolated work [c]	1	5.4	0.2	5.2	0.2
Under-productivity [c]	1	14.2	0.3	15.2	0.4
Low work social support [c]	1	12.7	0.3	17.2	0.4

[a] For single item scores the standard error is $\sqrt{pq / n}$.
[b] Mean exposure score out of 100, exposure composed of several items.
[c] Percentage of people exposed, exposure composed of only one item.

Discussion

This study allowed the assessment of exposure to job-stress factors for all the categories (PCS codes) of the French classification of occupations (22).

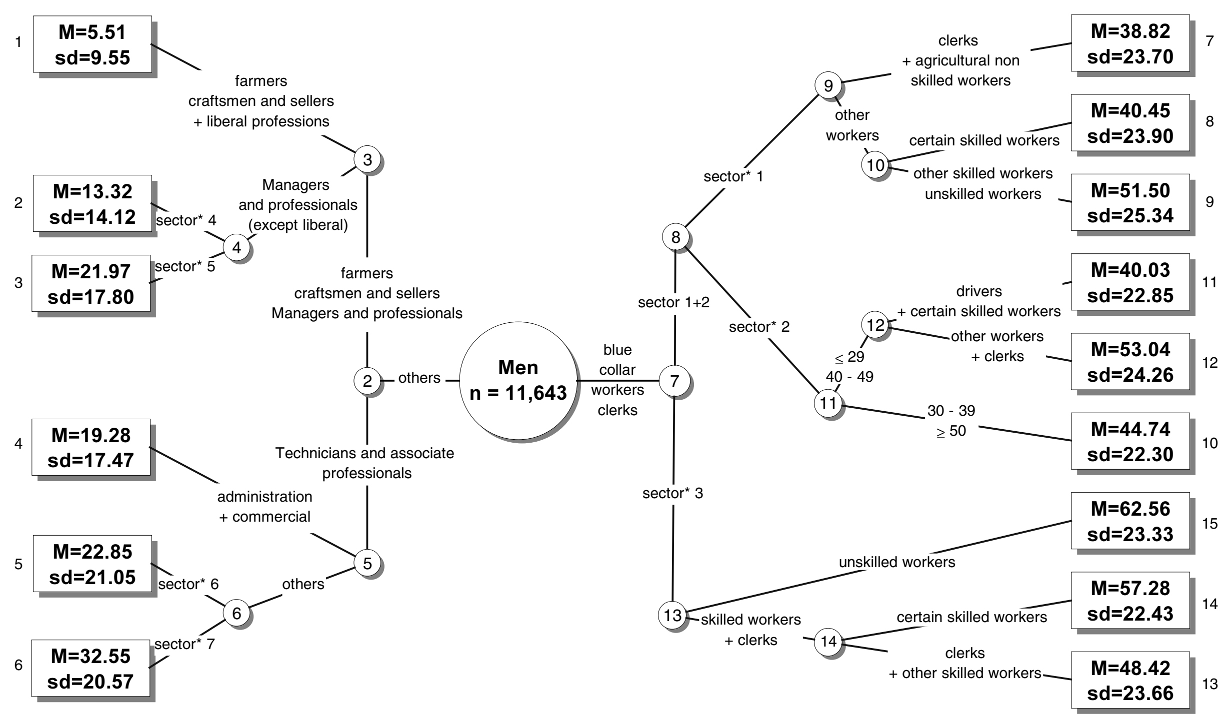

Figure 1. Segmentation tree obtained with the CART (classification and regression tree) method for the exposure to "absence of latitude" among men. * Sectors of activity (number of sector grouping): agriculture (1–6–4), fishing (1–6–4), fining industries (3–6–4), manufacturing industries (3–7–4), electricity, gas, water (2–6–5), construction (3–7–4), commerce (1–6–4), hotels (2,7–5), transport (3–7–5), financial activities (1–7–5), services to enterprises (2–6–4), public administration (2–7–5), education (1–7–5), health and social work (2–6–4), collective services (2–6–4), domestic services (1–7–4), activities outside France (1–6–5). (M = mean, sd =standard deviation)

The identification of exposure from items of the survey was one main step in this study. The validity of these results was based on the size of the study population, 20 929 persons representing a sizeable sample, and by repeating the PCA separately for the men and women with similar results (number of axes and items groups). These results were also confirmed in a confirmatory factor analysis based on tetrachoric correlations acknowledging the binary nature of the data. It was, however, necessary to compare the results stemming from the data analysis with a priori judgments based on knowledge of job stress. The PCA produced valid results in the view of the experts. Nevertheless, for four items, it was necessary to extend the criteria for factor loadings from 0.40 to 0.30.

The PCA identified the two broad areas described in the framework of Karasek's demand–control model. The "job demand" dimension is, however, partly explored by two exposures, "impossibility to achieve the objective" and "concentration requirement", while the dimension "skill discretion" was not found due to a lack of items in the original survey. Furthermore, social support, introduced secondarily by Johnson et al (3), could only be partly approximated via the "low work social support" items. Other exposures described social relations at work, such as "isolated work" and "contact with the public", but they were rather far from the dimension proposed by Johnson et al (3). Very few exposures from Siegrist's reward–effort model were explored. In fact, the dimension "job control" is probably the best dimension studied. Nevertheless, according to the literature, this dimension should play a major role as a stressor (30). On the other hand, the scope of stressors is broader than those of the models of Karasek and Siegrist, with others such as "polyvalence" or "inability to achieve objectives", both of which generate dissatisfaction and job stress.

Generalization of the results of a job-exposure matrix is often difficult when the sample does not cover all jobs. This limitation was probably less pronounced in our study because of the large sample size and its representativeness of the French working population, which we checked in terms of age and occupational categories.

The statistical methods most often used to construct a job-exposure matrix in fields other than that of job stress are based on expert judgment and constructed a priori. The alternative to this method is the a posteriori use of information stemming from individual measurements, either by means of apparatus or questionnaire (31). In our study, as in the three existing psychosocial job-exposure matrices, we mainly used the second method, using data already collected. The standard method of developing a matrix is based on calculating an average exposure for each job title. Besides this method, we used a second technique, the CART method, the main advantage being its ability to characterize HEG by using several variables. In this study, the job was described not only with the most detailed occupational category possible (455 codes) but also by activity (496 codes) and age (4 classes). The principle of developing matrices conceals the variability of exposure within job titles. The problem is maybe even more important in the field of job stress since the evaluation of exposure is often subjective. Reducing this drawback involves improving the evaluation of exposure by taking into account, for example, the duration of exposure, its pattern, or, as in the case of our study, gender, age, and a detailed description of the job. This improved evaluation is possible with the CART method but more difficult with a standard method. In France, the combination of all NAF and PCS codes provides more than 300 000

Table 4. Comparison between the CART (classification and regression tree) method and the crude means according to the Profession et Catégorie Sociale (PCS) code, for the factor "absence of latitude" (extract, details only for job categories 1, 2, 3). (SE = standard error)

PCS code	CART				N
	Mean	SE	Mean	SE	
1. Farmers	5.79	0.39	·	·	624
10. Farmers	5.79	0.39	·	·	624
11. Small farm	5.94	0.82	5.51	0.22	142
12. Medium size farm	6.03	0.73	5.51	0.22	206
13. Big farm	5.54	0.53	5.51	0.22	276
2. Craftsmen and shopkeepers	5.29	0.27	·	·	1082
21.Craftsmen	5.71	0.38	5.51	0.22	634
22.Shopkeepers	4.76	0.41	5.51	0.22	360
23. General managers (>10 workers)	4.22	0.96	5.51	0.22	88
3. Managers and professionals	15.44	0.41	·	·	1583
31. Liberal professionals	5.84	0.92	5.51	0.22	181
32. Managers					501
33. Administrative managers	17.77	1.29	15.19	0.43	176
34. Teaching, life science and health professionals	·	·	·	·	236
341					
3411. Higher education teaching professionals	33.33	1.59	31.34	1.52	114
3414. Teaching department managers	16.19	2.74	31.34	1.52	15
3415. University teaching professionals	13.71	3.35	15.19	0.43	25
342					
3421. Researchers	6.21	1.97	15.19	0.43	23
343					
3431. Hospital medical doctors	15.92	3.09	15.19	0.43	35
3432. Nonprivate medical doctors	22.08	4.46	15.19	0.43	11
3433. Psychologists	14.29	0.00	15.19	0.43	7
3434. Hospital medical students	.				–
3435. Hospital pharmacists	30.95	8.59	15.19	0.43	6
35. Creative or performing artists	16.37	1.73	15.19	0.43	89
36. Enterprise managers					901
37. Directors and chief executives	13.88	0.65	15.19	0.43	460
38. Production and operations department managers	15.41	0.73	15.19	0.43	441
4. Associate professionals and technicians	28.12	0.42	·	·	2398
5. Clerks	45.69	0.65	·	·	1291
6. Skilled and nonskilled workers	51.06	0.36	·	·	4665

possibilities. Most are improbable combinations, but more than 10 000 are plausible. The sample size is insufficient to provide exposure estimations for each combination, and thus grouping in HEG is necessary. The method adopted in this study was optimized in this case since, to group job titles, it takes into account the information from the data and the hierarchical structure of the classifications defined a priori. However, the indicator "variance explained" has shown that including the NAF code generally does not improve the variance explained. It means that the PCS code, in the four-digit level, is precise enough to explore a large part of the sector activity. If other information less correlated to the PCS, like the size of the company or the work unit, were available, the CART method would allow it to be incorporated and would perhaps yield a better predictor. Furthermore, it does not only rely on the a posteriori use of information stemming from individual measurements, but we also used a priori information, as defined earlier by experts who set up the hierarchical structure of both the PCS job code and the NAF code for activity sector. Thus the segmentation not only uses the study data but also the relations between the categories [ie, that the finer groupings (in 4 digits for the PCS) belong to broader categories (2 digit)]. The inclusion of this hierarchical structure is an interesting aspect of CART. This feature can, for instance, be observed in table 4, in which the job code 3414 is grouped with other job codes belonging to the same broad group rather than with another group with a closer mean score.

Examination of the tree constructed for the "absence of latitude" shows the distribution of certain a priori job categories with similar exposure in distinct HEG. For example, nonskilled craftsmen were distributed into four different HEG (on account of different activity) and their exposure score varied from 44 to 63 with two HEG with very similar exposure levels (53 and 51.5). This is the result, first, of the information contained in the data of the survey used and, second, of the very principle of constructing the matrix, based on the systematic splitting into two groups by taking into account the hierarchical codes of the data. The tree-construction technique, which involves using half the population as a base and the other half as validation, could also explain these situations; when numbers are low, it may be that the exposure of the base sample differs somewhat from that of the validation sample.

Overall the exposure scores of the HEG agree well with crude means based solely on job title. However, as already mentioned, for some job titles with very few persons, this may not be true. It is not clear whether this is an advantage or a drawback of the method. The optimality criterion on which CART is based may not always be relevant. Anyway, when few data points exist, the validity of a mean score is always questionable.

Depending on the nature of the exposure and on its distribution in the population, exposure may only be weakly contrasted between jobs on certain occasions. In this case, the segmentation technique is not efficient. Thus a variance of less than 10% explained by the construction of the tree should lead to the rejection of the study of this exposure. An exception can be made in a case in which a job title with a small number of people is particularly exposed when compared with the remainder of the population; then the division should be retained.

Comparison of the results observed in France with those of the three other matrices in this domain has proved difficult due to a lack of homogeneity in the choice of exposures considered, their definition, and the method used. Moreover, only the Finnish publication provides results stemming from the job-exposure matrix (12), but in a general way and not per job title. Matrices on job-stress factors are not easy to transpose from one country to another. Indeed, the perception of the work environment and the behavioral responses to a stressful situation are socially coded, and the response of each individual falls within the scope of a cultural register. Moreover, it would be even more interesting to have available matrices, similar for the exposure definitions and the method of development allowing comparisons between countries.

The matrices (with the NAF and PCS codes or the PCS code alone) could be used by researchers, occupational physicians, practitioners, and prevention specialists. Such matrices have several advantages and drawbacks. First, use of a job-exposure matrix in epidemiology is cheaper than other means of obtaining an estimation of job-stress exposure. Second, and more importantly, a self-reported job-stress factor may be affected by personality traits, and such reporting bias may inflate the association between exposure and health in studies in which both are based on self-reporting. This is not a problem if job-stress is quantified through a job-exposure matrix. A possible drawback of a multicriteria CART classification is that all the variables on which the groupings are based must be available. This requirement is usually not a problem with PCS and age. If the classification is based on NAF or another less common variable, it could limit the standard use of the job-exposure matrix. However the main drawback is that, for many job-stress factors, the variation is within jobs rather than between jobs, even if additional information is taken into account. This problem leads to exposure misclassification, and, hence, to health effects of stress being estimated with very high imprecision. Such study designs can therefore only be exploratory.

Such an exploratory epidemiologic study was the motivation of our study. Its aim was to provide an evaluation of exposure to job-stress factors for a more wide-ranging project whose objective was to link existing French health data collected in the early 1990s (mortality and morbidity indicators) to exposure data by means of the PCS code. If the contemporary health studies were to be analyzed, the 1998 exposure data would of course be more adequate.

Acknowledgments

Our thanks go to the DARES of the French Ministry of Labor, which made the data of the Survey of Working Conditions available without any financial counterpart.

The data presented in this report were the subject of an oral communication presented at the X2001 Exposure Assessment in Epidemiology and Practice conference (Göteborg on June 2001).

References

1. Karasek RA. Job demands, job decision latitude, and mental strain: implications for job redesign. Adm Sci Q 1979; 24:285–308.
2. Karasek RA, Theorell T. Healthy work: stress, productivity, and the reconstruction of working life. New York (NY): Basic Book; 1990. p 381.
3. Johnson JV, Hall EM, Theorell T. Combined effects of job strain and social isolation on cardiovascular disease morbidity and mortality in a random sample of the Swedish male working population. Scand J Work Environ Health 1989; 15:271–9.
4. Baker B. The study of stress. Ann Rev Public Health 1985, 6:367–81.
5. Kristensen TS. Job stress and cardiovascular disease: a theoretic critical review. J Occup Health Psychol 1996;1;(3):246–60.
6. Schnall PL, Landsbergis PA, Baker D. Job strain and cardiovascular disease. Annu Rev Public Health 1994;15:381–411.
7. Kristensen TS. The demand-control-support model: methodological challenges for future research. Stress Med 1995; 11:17–26.
8. Siegrist J. Adverse health effect of high-effort / low reward conditions. J Occup Health Psychol 1996;1:27–41.
9. Niedhammer I, Siegrist J. Facteurs psychosociaux au travail et maladies cardio-vasculaires: l'apport du modèle du Déséquilibre Efforts/Récompenses [Psychosocial factors at work and cardiovascular diseases: contribution of the effort–reward imbalance model]. Rev Epidemiol Sante Publique, 1998;46:398–410.
10. Bosma H, Peter R, Siegrist J, Marmot M. Two alternative job stress models and the risk of coronary heart disease. Am J Public Health 1998;88:68–74.
11. Broadening the view of exposure assessment [consensus report]. Scand J Work Environ Health 2001;27(5):354–7.
12. Kauppinen T, Toikkanen J, Pukkala E. From cross-tabulations to multipurpose exposure information systems: a new job-exposure matrix. Am J Ind Med 1998;33:409–17.
13. Johnson JV, Stewart WF. Measuring work organization exposure over the life course with a job-exposure matrix. Scand J Work Environ Health 1993;19: 21–8.
14. Schwartz JE, Piepper CF, Karasek RA. A procedure for linking psychosocial job characteristics data to health surveys. Am J Public Health 1988;78(8):904–9.
15. Buring JE, Evans DA, Fiore M, Rosner B, Hennekens C. Occupation and risk of death from coronary heart disease. J Am Med Assoc 1987;258(6):791–2.
16. Schrijvers CTM, Van de Mheen HD, Stronks K, Mackenbach JP. Socioeconomic inequalities in health in the working population: the contribution of working conditions. Int J Epidemiol 1998;27:1011–8.
17. Breiman L. Statistical modeling: the two cultures. Stat Med 2001;16(3):199–231.
18. Dussert F, Vinck L. Conditions, organisation du travail et nouvelles technologies en 1991 [Job factors, work organisation and new technologies in 1991]. Paris: Ministère du travail, de l'emploi et de la formation professionnelle; 1993. Dossiers statistiques du travail et de l'emploi du n°90–91–92
19. Niedhammer I, Goldberg M, Leclerc A, David S, Bugel I, David S. Psychosocial work environment and subsequent depressive symptoms in the Gazel cohort. Scand J Work Environ Health 1998;24(3):197–205.
20. Niedhammer I, Bugel I, Goldberg M, Leclerc A, Guéguen A. Psychosocial factors at work and sickness absence in the GAZEL cohort: a prospective study. Occup Environ Med 1998;55:735–41.
21. Derriennic F, Laville A. Eléments de Synthèse. In : Touranchet A, Volkoff S, editors. Age, travail et santé: etudes sur les salariés âgés de 37 à 52 ans: enquête ESTEV 1990 [Age, work and health]. Paris: Les editions du Institut National de la Santé et de la Recherche Médicale [French Institute of Health and Medical Research] (INSERM); 1996. p 440.
22. Nomenclature et codes. 2ème édition. Nomenclature des professions et catégories socioprofessionnelles PCS [French classification of occupations PCS]. Paris: Institut National de la Statistique et des Études Économiques (INSEE); 1994. p 1–417.
23. Nomenclatures d'activités (NAF) et de produits (CPF). Paris: Institut National de la Statistique et des Études Économiques (INSEE); 1992. Journal Officiel de la République Française, p 1–417.
24. International Labour Office (ILO). ISCO-88 International standard classification of occupations. Geneva: ILO; 1990. p 457.
25. Lebart L, Horineau A, Warwick K. Multivariate descriptive statistical analysis. New York (NY): J Wiley; 1984.
26. SAS institute Inc. SAS/STAT user's guide, Release 6.03 edition. Cary (NC): SAS Institute Inc; 1988. p 1028.
27. Breiman L, Friedman JH, Ohshen RA, Stone CJ. Classification and regression trees. Belmont: Wadsworth; 1984.
28. Steinberg D, Colla P. CART: a supplementary module for SYSTAT. Evanston (IL): SYSTAT Inc; 1992.
29. Guéguen A, Goldberg M, Bonenfant S, Martin JC. Using a representative sample of workers for constructing the SUMEX French general population-based job-exposure matrix. Occup Environ Medicine. In press.
30. Schnall P, Belkic K, Landbergis P, Baker D. The workplace and cardiovascular disease. Occup Med 2000;15(1)1–369.
31. Bouyer J, Hémon D. Retrospective evaluation of occupational exposures in population-based case-control studies: general overview with special attention to job exposure matrices. Int J Epidemiol 1993;22 Suppl 2: S57-S64.

Received for publication: 3 April 2002

Appendix I

Psychosocial items of the work condition survey used to develop the matrix

Workhours
- Leaving home before 7 o'clock in the morning
- Fixed hours of work
- Variable number of workdays in the week
- Alternating workhours
- Night work
- Period of work in excess of 10 hours
- No break for 48 consecutive hours
- Working more than one Sunday in two
- Working more than one Saturday in two

Work organization
- Control of hours
- Inability to vary deadlines
- Work rhythm imposed by supervision
- Application of instructions to carry out the work
- Application of instructions in the case of an incident
- Assembly line worker

(Work organization continued)
- Work rhythm imposed by the movement of a work piece
- Work rhythm imposed by the rate of a machine
- Work rhythm imposed by an immediate response
- Work rhythm imposed by a nonimmediate response
- Work rhythm imposed by technical constraints
- Work rhythm imposed by dependency on colleagues
- Work rhythm imposed by a hourly production rate
- Work rhythm imposed by a daily production rate
- Having at least a 30-minute lunchbreak
- Frequent requirement to abandon one task for another
- Change of task according to requirements
- Insufficient time to achieve objectives
- Repetitive work
- Isolated work

Social and relational aspects
- Working in direct contact with the public
- Experiencing situations of tension with the public
- Being told how to do the job
- Insufficient information to achieve the objectives
- Insufficient numbers to achieve the objectives
- Management function
- Lack of cooperation

Cognitive strain
- Serious consequences if error occurs
- Financial consequences of an error
- Dangerous consequences of an error
- Consequences of an error on job stability
- Perceived inability to leave work
- Reading small-size numbers or letters
- Examining very small objects
- Paying attention to brief audible or visual signals
- Lack of work

Appendix II

Statistical methods for dealing with hierarchical variables

Because the segmentation software does not take hierarchical variables into account, we developed a method to manage this type of variable.

For a hierarchical variable with 5 levels, such as the NAF nomenclature, the first level is indexed by i ($i = 1, \cdots, I$) and includes I categories. The second level is indexed by $j(i)$ ($j(i) = 1, \cdots, J_i$) and includes

$$\sum_{i=1}^{I} J_i$$

categories. We indexed levels 3, 4 and 5 in a similar manner.

We assumed that a split can occur at any of the 5 levels. If the split occurs at level 1, the hierarchical variable is a nominal variable with I categories and yields a set of $2^{I-1}-1$ splits, as would any nominal variable.

If the split occurs at level 2, it separates the observations that have categories depending on a single category, noted i^* of the 1st level, from all the other observations. There are $2^{J_{i^*}}-2$ splits associated with category i^*, and, in all,

$$\sum_{i=1}^{I} 2^{Ji} - 2I$$

splits at the second level. To take these splits into account in using a segmentation software such as CART, which manages nominal variables, I nominal variables must be created to explore the second level of the hierarchical variable. The i[th] variable has J_i+1 categories, noted as $\{0(i), 1(i), 2(i), \cdots, j(i), \cdots, J_i(i)\}$, and the $0(i)$ category consolidates all the second-level categories $\{j(i')$ with $i' \neq i\}$.

At the 3rd, 4th, and 5th levels, we proceed similarly, creating new nominal variables.

Set of splits from the NAF and PCS codes

In the PCS (Profession et Catégorie Sociale) nomenclature, the 2-digit codes of level 3 are subdivided into categories that are themselves divided into 4-digit codes in level 4; the codes from the same category begin with the same three digits. We created an intermediate level, 3', coded with 3 digits, thereby transforming the PCS into a 5-level hierarchical variable. For example, code 46 of level 3, "intermediate administrative and sales occupations in companies" led to 25 4-digit level 4 codes. These 25 codes were consolidated into the following six categories: (i) "general management and administration", (ii) "sales", (iii) "information, communication, performing arts", (iv) "transportation, tourism", (v) "banks, insurance", and (vi) "hotels and restaurants"; these categories made up the codes of the intermediate level 3'. Code 46 is thus divided into the following six 3-digit codes: 461, 462, 463, 464, 465 and 466.

The introduction of this additional level made it possible to reduce the computation time considerably by reducing the number of splits to be examined: when we consider the PCS in 4 levels, the number of splits from code 46 is $2^{25} - 2$ (ie, 33 554 430), while, if we consider it with 5 levels (4 initial levels and 1 intermediate level), the number of splits is much lower (710).

NAF has $2^{16} - 1$ splits for the first level, including, for example, the split that separates sections $\{A,C,I,M,N\}$ from the other sections. These splits can be explored with a 17-category nominal variable. At the 2nd level, if we consider, for example, the splits from section D, there are 14 subsections from D noted $\{DA,DB, \cdots,DN\}$ which lead to $2^{14} - 2$ splits on level 2 from section D including, for example, the split that separates subsections {DC, DD, DE, DG} from all other subsections. These splits can be explored by using a 15-category nominal variable $\{0(D),DA,DB, \cdots,DN\}$, where 0(D) represents all sections other than D.

In all, 331 nominal variables were created to explore the nomenclatures, 200 for the NAF (Nomenclature de Adtivités françaises) and 131 for the PCS.