Identification of gender differences in the factors influencing shoulders, neck and upper limb MSD by means of multivariate adaptive regression splines (MARS)

N. Busto Serrano[1], A. Suárez Sánchez[2,*], F. Sánchez Lasheras[3], F.J. Iglesias-Rodríguez[2], G. Fidalgo Valverde[2]

[1]*Labor and Social Security Inspectorate. Ministry of Labor, Migration and Social Security, Spain*
[2]*Department of Business Management, University of Oviedo, 33004 Oviedo, Spain*
[3]*Department of Mathematics, University of Oviedo, 33007 Oviedo, Spain*

Abstract

In the present research, models based on multivariate adaptive regression splines (MARS) are proposed to study the influence of gender in the factors affecting the development of shoulders, neck and upper limb MSD. Two different MARS models, corresponding to men and women, are constructed to identify variables with the strongest effect on the target MSD. Both models are capable to predict successfully the occurrence of the studied disorders. Men seem to be more vulnerable to physical risk factors and some other working conditions, whereas women appear to be more affected by psychosocial risk factors and activities carried out outside their working hours. According to the results, gender needs to be considered to ensure the success and effectiveness of ergonomic interventions on the whole working population.

Keywords: Musculoskeletal Disorders (MSD); gender; shoulder; neck; upper limb; multivariate adaptive regression splines (MARS).

1. Introduction

Preventing musculoskeletal disorders (MSD) has turned into one of the major Health and Safety management development areas and many ergonomic interventions have been carried out in this field (Denis et al., 2008; Yazdani et al., 2015). A lot of barriers use to be found during the implementation of these intervention programs; most of them are the same that appear in the development of general Health and Safety practices (Yazdani et al., 2018). Several hurdles are related with the difficulties to identify all factors that are involved in the development of MSD, especially when the assessor should focus on external elements that have no direct relation with working conditions. In this sense, there are several social issues such as gender, race and socioeconomics (Rasmussen et al., 2012; Mousaid et al., 2016; Barnes and Bendixsen, 2017), that affect each person individually and may have an impact on their response to hazardous conditions, low ergonomics or a high workload. The current study is focused on one of them: gender.

Gender differences have been identified in diverse occupational health areas (Messing et al., 2003; Waller et al., 2015; Artime Ríos et al., 2018). In particular, numerous studies have shown differences between women and men in MSD prevalence (Silverstein et al., 2009; Paarup et al., 2011; Widanarko et al., 2011; Laberge et al., 2012) and how job sexual division affects ergonomics (Caroly et al., 2013). Research has demonstrated a higher prevalence of shoulders, neck and upper limb MSD (the

* Corresponding author. E-mail address: suarezana@uniovi.es

target of this study) in women than in men (Messing et al., 2009). Now the key is to find out how gender affects the behavior at work and the body response to several negative work-related stimuli, and this has motivated the chosen standpoint for our analysis.

In the last years, the advances in numerical simulation (de Cos et al., 2008) and machine learning (de Andrés et al., 2011; Álvarez-Antón et al., 2013; Sánchez Lasheras et al., 2015; García Iglesias et al., 2018; Riesgo García et al., 2018; Sánchez Lasheras et al., 2018) have made possible the achievement of developments that were out of reach before. In the field of occupational risks, Suárez Sánchez et al. (2011) used a model based on support vector machines to predict work-related accidents according to working conditions, while Krzemień (2019a) implemented artificial neural network models to prevent the risk of fire in underground coal gasification processes. The implementation of machine learning techniques to ergonomics is not new either (Asensio-Cuesta et al., 2010; Suárez Sánchez et al., 2016). In the case of the present research, models based on multivariate adaptive regression splines are proposed (Guzmán et al., 2010; De Andrés et al., 2011) to study the influence of gender in the factors affecting the development of shoulders, neck and upper limb MSD.

2. Material and methods

2.1. Data set

The European Working Conditions Survey (EWCS) is conducted by the European Foundation for the Improvement of Living and Working Conditions (Eurofound, 2018). The EWCS provides an overview of working conditions of employees and the self-employed in Europe. It was first launched in 1990 and is generally conducted every five years. In each EU country, the questionnaire is administered face-to-face to a random representative sample of 'persons in employment' (i.e. employees and the self-employed).

The Eurofound datasets are stored by the UK Data Service (UKDS, 2019) and promoted online via their website. Upon request, the data are available free of charge to those who intend to use them for non-commercial purposes.

This research employs the microdata from the sixth wave of the survey (Sixth European Working Conditions Survey: 2015), conducted in 2015, when almost 44,000 workers were interviewed in 35 countries. To guarantee its validity, an expert questionnaire development group was set up to discuss the questionnaire used. The group was composed of experts and representatives of the European Commission and of international organizations (Eurofound, 2016).

The original database codified more than 740 variables (more than 370 for the sixth wave) that could be classified in the following topics:

- physical and psychosocial risk factors
- working time: duration, organization, predictability and flexibility; work–life balance
- place of work
- speed of work, pace determinants
- employee participation, human resource policies and work organization (such as task rotation); employee representation

- skills use, cognitive dimensions of work, decision-making authority, and learning in work
- employment conditions: job security and insecurity
- social relations at work: support, trust, cooperation, discrimination, violence
- gender issues: segregation, household composition, unpaid' work, extent of women in supervisory positions
- well-being and health, earnings and financial security.

In the group of items concerning health, the interviewed worker was asked to mention any health problem (from a list of ten) that they had suffered in the previous 12 months. Such problems included "Muscular pains in shoulders, neck and/or upper limbs (arms, elbows, wrists, hands, etc.)", codified as variable y15_Q78d. This was the target variable used to develop a model to predict the prevalence of such MSDs. This target variable identifies any worker who, in the previous year, suffered from any of the above-mentioned musculoskeletal symptoms.

According to the aims of the study, 350 subjects were excluded either because their gender, age or employment regime (part/full time) was not specified, or because they were under age. Thus, the final sample consisted of 43,500 subjects from the 28 EU Member States plus Albania, the former Yugoslav Republic of Macedonia, Montenegro, Serbia, Turkey, Norway and Switzerland.

Before using the machine learning technique, expert criterion was implemented to reduce the dimension of the data. As a result, 51 variables from the different topics were pre-selected as the most relevant items to try to explain the occurrence of the target disorder. In line with the approach of this study, it was essential to re-categorize the selected variables to be able to consider gender in the statistical results analysis. Thus, the following five groups were defined to classify them:

- Individual factors and health status.
- Physical risk factors.
- Psychosocial risk factors.
- Other work-related factors.
- Activities outside working hours.

Individual factors and health status includes all variables that, apparently, seem to be outside the scope of work management, so they exclusively depend on the person's biological and physiological circumstances. In the group, there are items such as age or prevalence of illness.

The physical risk factors group comprehends all variables that represent the exposure to a risk factor related with painful positions, repetitive movements, load carrying and lifting, etc.

Along the same lines, the psychosocial risk factors group fits the traditional definition of this kind of risk considered in health and safety management. For example, variables related with stress, conflicts, working rhythm and time pressure, job contents, etc. were included in this category.

Other work-related factors includes variables that have any influence on working conditions, as those traditionally connected to safety at work, polluting chemicals or physical agents, kind of activity developed, etc. Unlike the first category, in this case, each manager could act on these variables to improve working conditions by turning them healthier.

Activities outside working hours is a compendium of all those circumstances connected with free time, family care and housework. This is the main category to guarantee that gender has been considered in the data analysis, since it records the possible influence of workload and responsibilities outside the paid-working on MSD prevalence and, obviously, the differences between women and men in this sense.

Most of the variables were designed as Lickert scales, some of them were binary and a few were continuous or categorical. To describe the sample, the statistical significance of the differences between men and women was tested by means of a T-Test of difference in the case of continuous variables and a Pearson Chi-Square test in the rest.

2.2. Multivariate adaptive regression splines method (MARS)

Multivariate adaptive regression splines (MARS) is a multivariate nonparametric classification and regression technique introduced by Friedman in 1991. Its aim is to predict the values of a continuous dependent variable (in this case, y15_Q78d) from a set of independent explanatory variables.

MARS defines the functional relationship between the dependent and the independent variables by means of a group of coefficients and piecewise-defined polynomials, also called splines, of degree q (basis functions) that are entirely “driven” from the regression data (García Nieto et al., 2011; Suárez Sánchez et al., 2011; Ordóñez Galán et al., 2011; Krzemień A., 2019b). The MARS regression model is constructed by fitting basis functions to different intervals of the independent variables (Alonso Fernández et al., 2013). Generally, splines have pieces smoothly connected together to describe the behavior of the dependent variable. The degree q of the splines (2 in this work) is usually selected by achieving a compromise between performance and complexity of the model. The reason why degree 2 has been selected will be explained in the results section.

In general, any MARS model makes use of the following model:

$$\hat{f}(x) = \sum_{i=1}^{k} c_i \cdot B_i(x)$$

Where $B_i(x)$ represents the basis functions of the model and c_i are constant coefficients. Basis functions are defined as follows:

$$B_i(x) = \begin{cases} x & if\ \ x \geq 0 \\ 0 & otherwise \end{cases}$$

In any MARS model, for each data set of n individuals and m explanatory variables, $n \times m$ basis functions are defined. In order to prune and obtain the definitive MARS model, a two steps process is performed. Firstly, a progressive selection of basis functions leads to a very complex and overfitted model. Such model, although is able to fit the data, has poor predictive ability for new objects. To improve prediction, redundant basis functions are removed one at a time using a regression procedure. To determine which basis functions will be included in the model, the generalized cross validation (GCV) methodology is employed. In this methodology, the root mean squared residual error is divided by a penalty parameter, which depends on the complexity of the model. The GCV equation is as follows:

$$GCV(M) = \frac{\frac{1}{n}\sum_{i=1}^{k}(y_i - \widehat{f_M}(x_i))^2}{(1 - \frac{M + 1 + d \cdot M}{n})^2}$$

Where M is the number of basis functions, and the parameter d is a penalty for each base function included in the model. For our research, this parameter is 2.

In this case, MARS is implemented as a classification model, trying to predict the value of the target variable, defined as binary, and thus which subjects are expected (or not) to suffer the studied MSD. Please note that MARS models are also employed for regression problems (García Nieto et al., 2012). The usual procedure is to construct the model using a randomly selected part of the data set (in this case, 80%) and then to test the classification capabilities of the model obtained by applying it to the rest of the data set (20% of the individuals of the original sample). The outcome of the test for each individual can be either positive (expected to suffer the MSD) or negative (not expected to suffer the MSD), while the actual health status of the person may be the same (true positives, *tp*, and true negatives, *tn*) or different (false positives, *fp*, and false negatives, *fn*). The results can be presented using a confusion matrix, which records the number of correctly (*tp*, *tn*) and incorrectly (*fp*, *fn*) recognized subjects of the test sample. To evaluate the performance of the binary classification model, sensitivity and specificity are usually employed. Sensitivity (Equation 1) measures the proportion of actual positives that are correctly identified, while specificity (Equation 2) measures the proportion of actual negatives that are correctly identified (Álvarez Menéndez et al., 2010):

$$Sensitivity = \frac{tp}{tp + fn} \quad (1)$$

$$\text{Specificity} = \frac{tn}{fp + tn} \quad (2)$$

Once the MARS model is constructed, the relevance of the explanatory variables can be expressed as their contribution to the goodness of fit of the model (Suárez Sánchez et al., 2011; Suárez Sánchez et al., 2016). To determine variable importance scores, MARS calculates how much the goodness of fit is reduced when eliminating each variable. The importance of the variables is presented as a relative measure (RSS criterion): the most important variable is the one that, when eliminated, decreases the model fit the most, so it receives the highest score (100%). The scores for the rest of the explanatory variables are calculated as the ratio of the reduction in the goodness of fit of those variables to that of the most important one.

In the case of the present research, all the calculus were performed with the statistical software R (R Core Team, 2017). MARS models were calculated with the help of the `earth` package (Milborrow, 2011)

To fulfil the aims of this study the sample was split into two (women and men), and two models (MARS_1 and MARS_2) were constructed using 80% of each sample for training and 20% for testing. The purpose of this procedure was to find out whether the most relevant variables that explained the behavior of the target MSD were the same or different in each group.

3. Results

3.1. Characteristics of the sample and prevalence of the MSD

As stated above, the final sample consisted of 21,578 women and 21,922 men aged between 18 and 89 years. Figures 1 and 2 describe their distribution in terms of level of education and range of sectors. The differences in distribution between men and women were statistically significant in both variables (Pearson Chi-Square test, p-value < 0.001).

As shown in Figure 1, the most frequent level of education, both in men and women, is upper secondary education, being more frequent among men. At the same time, lower educational levels are more frequent among men, while higher educational levels (especially university studies) are more frequent among women.

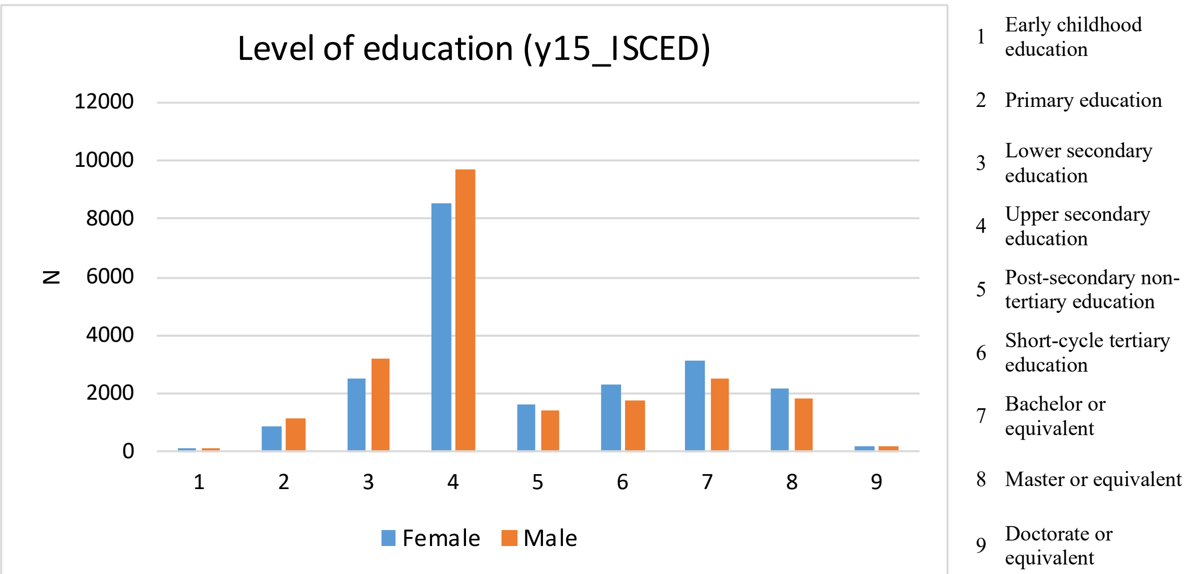
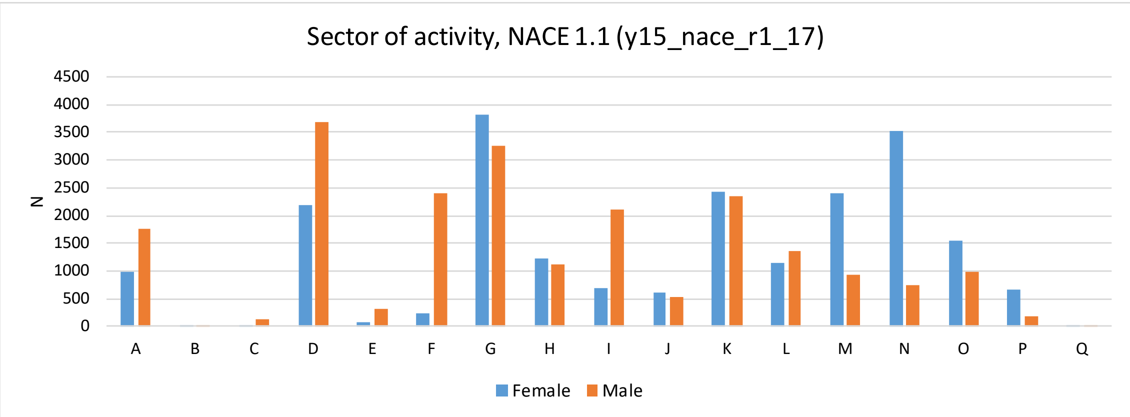


Figure 1. Distribution of the sample according to their level of education.

Regarding the sectors of activity, Figure 2 presents the distribution of the sample according to the NACE (Rev.1), which is the statistical classification of economic activities in the European Community. The most frequent sector among men is manufacturing (17.1%), followed by wholesale and retail trade and repair of motor vehicles and motorcycles (15.1%), construction (11.1%) and real estate activities (10.9%). Women tend to work more frequently in wholesale and retail trade and repair of motor vehicles and motorcycles (17.7%), health and social work (16.3%), real estate activities (11.3%) and education (11.1%). While the figures in some sectors are very similar for both genders, there are also very meaningful differences in the sample studied: agriculture, manufacturing, construction and transport are clearly masculinized sectors, and education, health and social work, other service activities and activities of household are feminized ones.

It is worth commenting that section G of NACE groups together wholesale and retail trade with repair of motor vehicles and motorcycles, while they are sectors with important differences, especially for the purpose of this work. If we break down the data in specific subsections, only 8.7% of women (compared to 28.2% of men) in section G are classified in the sale, maintenance and repair of motor vehicles subsection, and most of them are in the retail (78.1%) and wholesale (13.2%) trade sectors. Moreover, considering the occupation of those individuals, according to ISCO-08 (International Standard Classification of Occupations of the International Labour Organization), most of the women in this section were service and sales workers (65%) or clerical support workers (9%), and actually only 0.2% were classified as mechanics. In contrast, 14.5% of men in section G were classified as mechanics.



A	Agriculture, hunting and forestry	**G**	Wholesale and retail trade; repair of motor vehicles and motorcycles	**M**	Education
B	Fishing	**H**	Hotels and restaurants	**N**	Health and social work
C	Mining and quarrying	**I**	Transport, storage and communication	**O**	Other service activities
D	Manufacturing	**J**	Financial intermediation	**P**	Activities of households
E	Electricity, gas, and water supply	**K**	Real estate activities	**Q**	Activities of extraterritorial organizations and bodies
F	Construction	**L**	Public administration and defense; compulsory social security		

Figure 2. Distribution of the sample according to the sector of activity, as classified in NACE Rev. 1.1.

Table 1 summarizes the rest of the variables initially considered, together with their description and the most relevant statistics. P-values recorded in Table 1 correspond to the results of statistical analysis. It is worth noting that there are very few variables where differences between male and female individuals are not statistically significant.

Very important gender differences can be found in the variables that represent an exposure to specific risks, such as vibrations, aerosols, noise, low and high temperatures, etc. The ratio of reporting of such exposure is significantly higher among men than women, with the exception of the contact with infectious materials. This is consistent with the information summarized in Figure 2: men tend to work more frequently in manufacturing and construction sectors, and women in health and social work. In the case of physical and psychosocial conditions, the situation is the opposite: the reporting ratio of exposure to factors such as dealing directly with people, working with VDU, lifting or moving people, handling conflicts or emotionally disturbing situations, etc. is significantly higher among women, with the important exception of carrying or moving heavy loads. Again, the explanation to this can be found in the types of sectors men and women are more commonly employed in.

Table 1. Description of the variables considered for the models.

				Female		Male		
Variable	**Description**	**Category***	**Type**	**Mean**	**StDev**	**Mean**	**StDev**	**p-value**
y15_Q2a	gender	1	Binary	2,00	0,00	1,00	0,00	
y15_Q2b	age	1	Continuous	43,39	12,39	43,58	12,91	0,120
y15_Q2c	employment situation	4	Categorical					< 0,001
y15_Q2d	part time/full time	4	Lickert (0-2)	1,63	0,59	1,79	0,51	< 0,001
y15_Q7	employee/self employed	4	Categorical	1,13	0,37	1,21	0,43	< 0,001
y15_Q11	kind of contract	4	Categorical	2,03	1,66	2,33	1,80	< 0,001
y15_Q29a	exposed to vibrations	4	Lickert (1-7)	1,46	1,22	2,32	1,89	< 0,001
y15_Q29b	exposed to noise	4	Lickert (1-7)	1,84	1,49	2,48	1,84	< 0,001
y15_Q29c	exposed to high temperatures	4	Lickert (1-7)	1,76	1,39	2,31	1,69	< 0,001
y15_Q29d	exposed to low temperatures	4	Lickert (1-7)	1,59	1,19	2,19	1,56	< 0,001
y15_Q29e	exposed to breathing in aerosols	4	Lickert (1-7)	1,30	0,96	2,01	1,69	< 0,001
y15_Q29f	exposed to breathing in vapours	4	Lickert (1-7)	1,32	0,95	1,63	1,28	< 0,001
y15_Q29g	exposed to contact with chemical products	4	Lickert (1-7)	1,61	1,38	1,76	1,40	< 0,001
y15_Q29i	exposed to contact with infectious materials	4	Lickert (1-7)	1,60	1,43	1,52	1,22	< 0,001
y15_Q30a	tiring or painful positions	2	Lickert (1-7)	2,86	1,99	2,94	1,96	< 0,001
y15_Q30b	lifting or moving people	2	Lickert (1-7)	1,57	1,42	1,28	0,95	< 0,001
y15_Q30c	carrying or moving heavy loads	2	Lickert (1-7)	1,97	1,51	2,63	1,87	< 0,001
y15_Q30d	sitting	2	Lickert (1-7)	3,60	2,22	3,41	2,11	< 0,001
y15_Q30e	repetitive hand or arm movements	2	Lickert (1-7)	3,87	2,30	3,76	2,23	< 0,001
y15_Q30f	dealing directly with people	3	Lickert (1-7)	4,46	2,49	3,78	2,37	< 0,001
y15_Q30g	handling conflicts	3	Lickert (1-7)	2,72	1,88	2,45	1,76	< 0,001
y15_Q30h	emotionally disturbing situatios	3	Lickert (1-7)	2,44	1,58	2,18	1,48	< 0,001
y15_Q30i	working with VDU	2	Lickert (1-7)	3,58	2,45	3,24	2,34	< 0,001
y15_Q39e	shift work	4	Binary	1,78	0,41	1,80	0,40	< 0,001
y15_Q44	work–life balance	5	Lickert (1-4)	1,88	0,74	1,97	0,77	< 0,001
y15_Q45b	too tired after work to do household jobs	5	Lickert (1-5)	2,76	1,12	2,64	1,16	< 0,001
y15_Q45c	work interference with family (time)	5	Lickert (1-5)	2,11	1,12	2,14	1,15	0,003
y15_Q45d	family interference with work (concentration)	5	Lickert (1-5)	1,77	0,91	1,73	0,91	< 0,001
y15_Q45e	family interference with work (time)	5	Lickert (1-5)	1,63	0,87	1,64	0,89	0,049
y15_Q48a	short repetitive tasks of less than 1 minute	2	Binary	0,26	0,44	0,24	0,43	< 0,001
y15_Q48b	short repetitive tasks of less than 10 minutes	2	Binary	0,41	0,49	0,38	0,49	< 0,001
y15_Q49a	working at very high speed	3	Lickert (1-7)	3,52	2,04	3,56	2,01	< 0,001
y15_Q53d	monotonous tasks	3	Binary	0,48	0,50	0,47	0,50	0,416
y15_Q54c	able to choose or change speed or rate of work	3	Binary	0,72	0,45	0,73	0,44	0,350
y15_Q61f	can take a break when needed	3	Lickert (1-5)	2,85	1,46	2,57	1,39	< 0,001
y15_Q61g	enough time to get the job done	3	Lickert (1-5)	2,01	0,99	2,02	0,98	0,334
y15_Q61m	stress	3	Lickert (1-5)	2,87	1,15	2,85	1,17	0,078
y15_Q73	thinks health or safety is at risk because of work	4	Binary	0,21	0,40	0,29	0,45	< 0,001
y15_Q74	work affects health	4	Lickert (0-2)	1,12	0,59	1,17	0,61	< 0,001

y15_Q75	general health status	1	Lickert (1-5)	2,02	0,78	1,97	0,76	< 0,001
y15_Q76	health problem lasting for more than 6 months	1	Binary	0,20	0,40	0,16	0,37	< 0,001
y15_Q77	daily activities limited because of health problems	1	Lickert (0-2)	0,13	0,39	0,11	0,36	< 0,001
y15_Q82	days lost due to sick leave in the last 12 months	1	Continuous	5,65	17,81	4,40	14,84	< 0,001
y15_Q83a	days lost due to accident in the last 12 months	1	Continuous	0,38	4,81	0,59	6,70	< 0,001
y15_Q95c	outside work: caring for children/grandchildren	5	Lickert (1-5)	3,06	1,78	2,59	1,67	< 0,001
y15_Q95d	outside work: cooking and housework	5	Lickert (1-5)	4,56	0,88	3,09	1,49	< 0,001
y15_Q95e	outside work: caring for elderly/disabled relatives	5	Lickert (1-5)	1,69	1,19	1,55	1,04	< 0,001
y15_Q95f	outside work: taking a training or education course	5	Lickert (1-5)	4,48	0,81	4,53	0,76	< 0,001
y15_Q95g	outside work: sporting, cultural or leisure activity	5	Lickert (1-5)	3,51	1,25	3,37	1,24	< 0,001
y15_Q78d	**Muscular pains in shoulders/neck/upper limbs. TARGET VARIABLE**		**Binary**	**0,47**	**0,50**	**0,41**	**0,49**	**< 0,001**

*Categories: 1 – Individual factors; 2 – Physical factors; 3 – Psychosocial factors; 4 – Other work-related factors; 5 – Outside work

There are other meaningful gender differences. The highest difference between men and women (1.5 points in a scale 1-5) is in the answer to the question whether they are involved in cooking and housework out of work. There is also an important difference in the question concerning caring for children/grandchildren and, to a lesser extent, caring for elderly/disabled relatives.
The target variable (the last one in Table 1) shows an important statistically significant difference: 47% of women reported suffering one or more of the MSDs studied, whereas only 41% of men did.

3.2. MARS models

As stated above, two MARS models were constructed to describe the behavior of the target variable. Both models identified only 21 variables that are relevant to predict the prevalence of the studied MSD. Those variables are recorded in Table 2, together with a measurement of their importance according to MARS_1 (women) and MARS_2 (men). Part of them coincide in both models, although usually with different importance, but several of them only appear either in MARS_1 or in MARS_2.

Table 2. Variables selected by MARS_1 (women) and MARS_2 (men) models and measurement of their importance (RSS criterion).

MARS_1		MARS_2	
Variable	**Importance**	**Variable**	**Importance**
y15_Q30a	100.0	y15_Q30a	100.0
y15_Q75	80.4	y15_Q75	80.2
y15_Q82	80.4	y15_Q74	61.5
y15_Q74	66.2	y15_Q45b	51.6
y15_Q76	55.3	y15_Q30e	46.7
y15_Q30b	48.6	y15_Q76	42.1
y15_Q30e	48.6	y15_Q82	38.1
y15_Q2b	43.4	y15_Q29b	35.0
y15_Q29b	43.4	y15_Q29d	35.0
y15_Q45b	38.1	y15_Q2b	31.8
y15_Q61g	33.0	y15_Q30c	30.0
y15_Q61f	29.6	y15_Q30d	26.3

y15_Q95g	25.8	y15_Q39e	24.7
y15_Q53d	24.5	y15_Q45c	23.3
y15_Q73	23.3	y15_Q2d	21.9
y15_nace_r1_17	23.3	y15_Q95d	20.4
y73_Q49a	21.8	y15_Q73	19.0
y15_Q95c	20.7	y15_Q30g	17.5
y15_Q30g	17.2	y15_Q30h	16.2
y15_Q95d	14.9	y51_Q61m	13.4
y15_Q77	10.7	y15_Q30b	11.9

For each model, Figure 3 presents the weight of each of the five categories previously defined (as the percentage of variables of the model that can be classified in that category), as well as the variables that belong to each category.

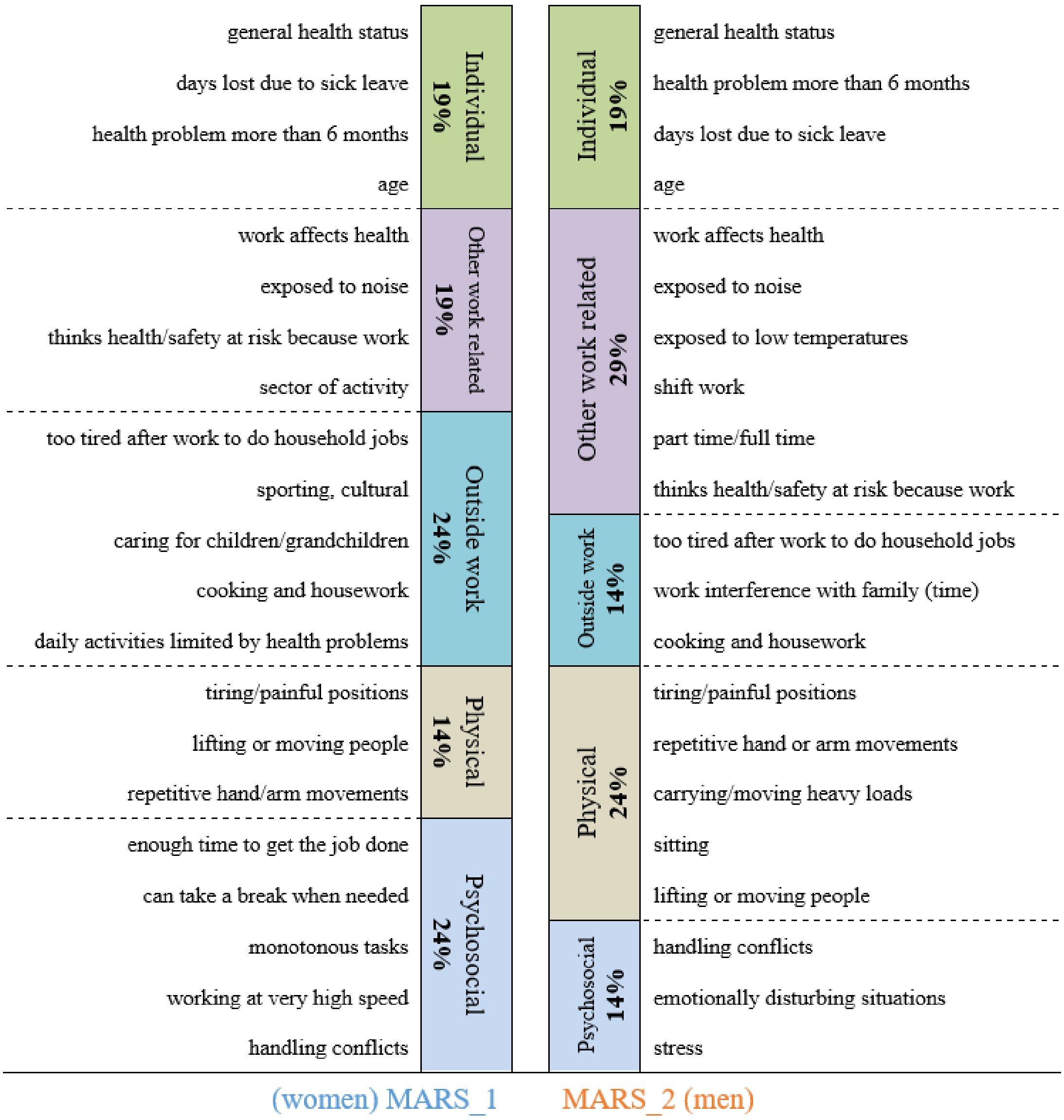


Figure 3. Weight of the different categories of factors affecting shoulders, neck and upper limb MSD.

Figures 4 and 5 present some of the splines that are part of the MARS models. These Figures not only show which variables are relevant but also explain how they influence the prevalence of the target MSD. For example, the first spline of Figure 3 shows the contribution of variable y15_Q2b (age) to the prediction of the value of the target variable (from 0 to 1) made by MARS_1: this value increases with the age of the worker, and this increase is especially important in the range of age between 30 and 50 years.

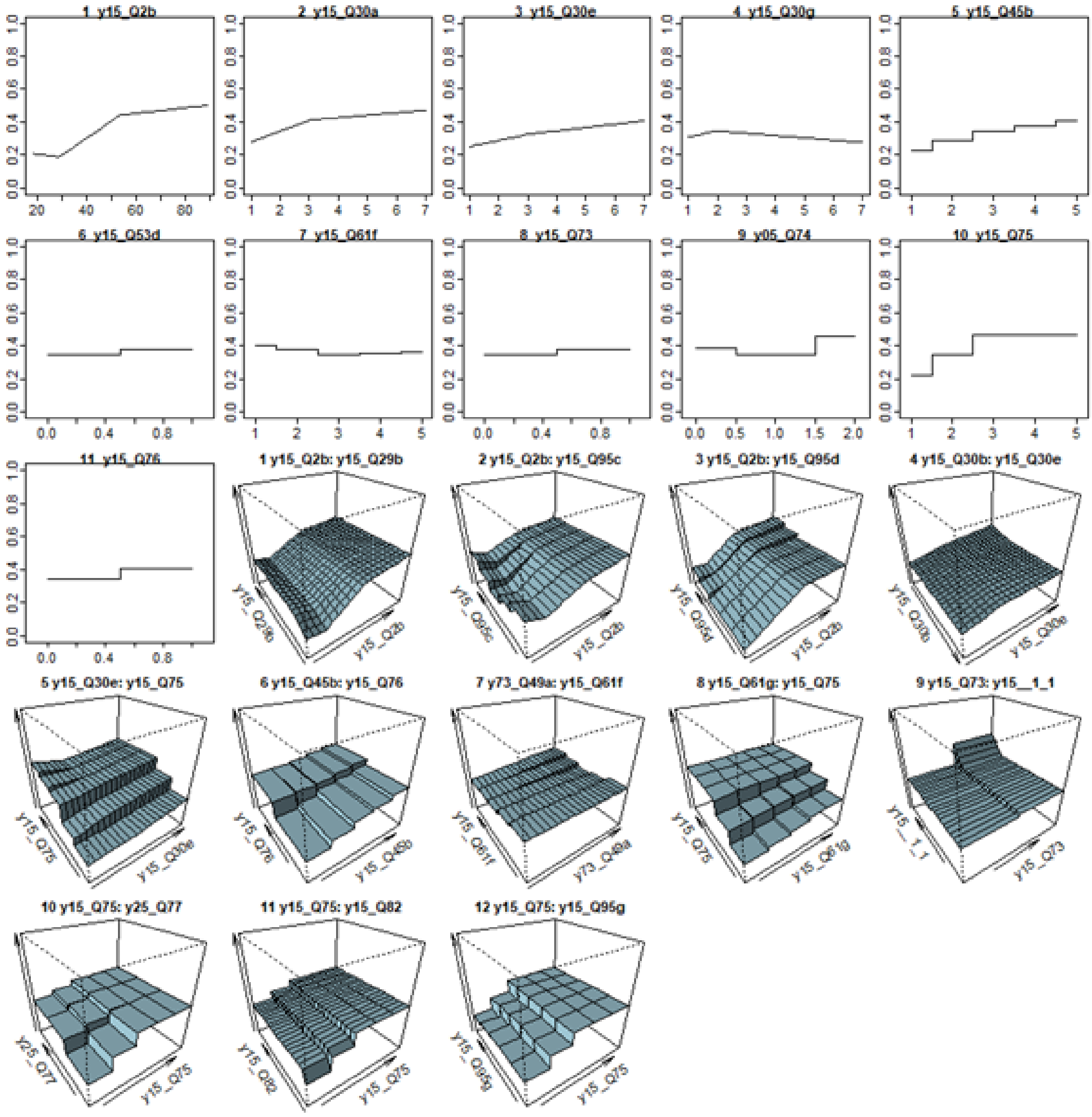


Figure 4. Relevant variables according to MARS_1 (women) model (0 no reported MSD; 1 reported MSD).

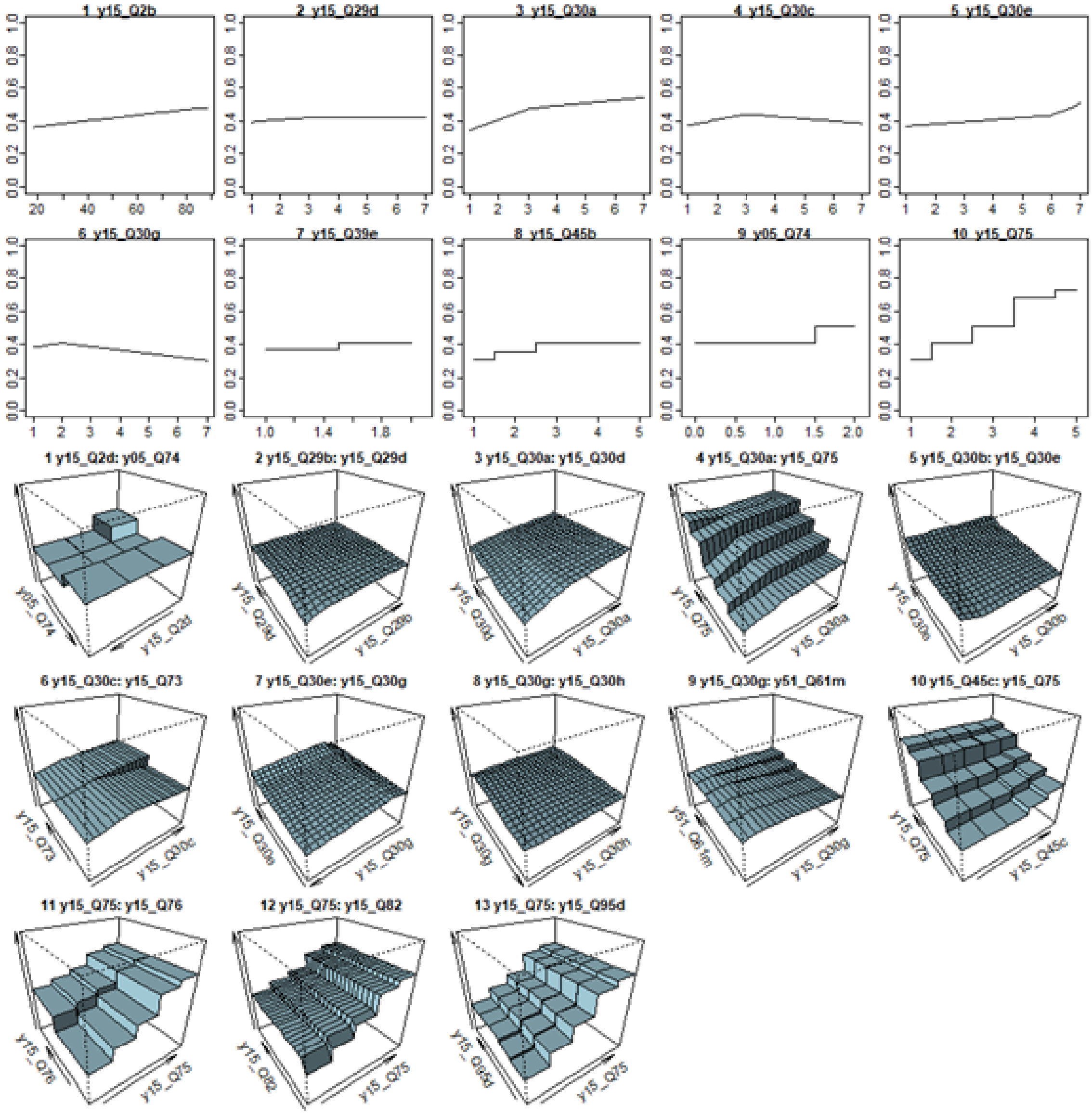


Figure 5. Relevant variables according to MARS_2 (men) model (0 no reported MSD; 1 reported MSD).

Tables 3 and 4 show the confusion matrix obtained when applying each MARS model to the test samples (20% of the original ones). The sensitivity and specificity of both models are very similar, as shown in Table 5. Increasing the grade of the models from 2 to 3 produced much more complex models without significant improvement in their performance.

Table 3. The confusion matrix for MARS_1 (women).

	Recognized	
Class	**As negative**	**As positive**
Negative	1,673 (tn)	705 (fp)
Positive	590 (fn)	1,366 (tp)

Table 4. The confusion matrix for MARS_2 (men).

	Recognized	
Class	**As negative**	**As positive**
Negative	2,069 (tn)	852 (fp)
Positive	478 (fn)	962 (tp)

Table 5. Measurement of the performance of both models.

	MARS_1	MARS_2
Sensitivity	69,84%	66,81%
Specificity	70,35%	70,83%

4. Discussion

As shown in the Results, both MARS models are shaped by 21 variables. It is easy to notice from Figure 3 that the weight of each category to explain the prevalence of shoulders, neck and upper limb MSD is different for each gender. Thus, it is inevitable to discuss these results attending to the behavior of each category.

4.1. Individual factors and health status.

This category seems to have the same weight in both models, while it is true that variables included in it appear earlier in MARS_1 (women) model, so their importance is higher in women.

The first of these variables, which takes the second place in importance in both models, is "General health status" (y15_Q75). In the case of women, it is followed by "Days lost due to sick leave in the last 12 months" (y15_Q82, in third position), and the "Existence of health problems lasting for more than 6 months" (y15_Q76, in fifth position). In the case of men, the order of appearance of these two variables is inverted and they occupy the seventh (y15_Q82) and sixth (y15_Q76) position respectively. Finally, "Age" (y15_Q2b) is included again in both models, in the eighth position for women and in the tenth position for men.

The selection of this kind of variables by the MARS models demonstrate that a good health status is a key to prevent MSD, so interventions oriented to maintain a healthier lifestyle in the workforce have a "raison d'être". On the other hand, considering gender, there is room for debate (Härenstam, 2009) about how social roles could be an obstacle to improve women's health status. This is especially relevant if the same kind of interventions are implemented for the whole of the workforce without keeping in mind external factors as those included in the category of *Activities outside working hours*, that are substantially different in men and women. Actually those *Activities outside working hours*, which are significantly more relevant in MARS_1 (women) to predict the studied MSD, have also a very strong influence on the general health status, but are essentially out of the control of the employer.

It is also important to remark that individual factors are deeply linked to ergonomics. In this sense, a good ergonomic assessment implies having in mind differences in workers including anthropometrics or functional body characteristics, where gender is a differentiating element (Côté, 2012).

4.2. Other work related factors.

This category, which comprehends all those work-related factors that are not included in physical or psychosocial risk factors, starts to show relevant differences attending to gender. These factors represent almost 29% for men and just 19% for women. This is a crucial issue when it comes to carrying out an intervention program to reduce MSD at work: if it focusses on workplace conditions, as is frequently done, its impact might be higher on men than on women.

The first variable that appears in both cases is the perception that "Work affects the health of the worker" (y15_Q74), taking the third position in importance in the case of men and the forth position in the case of women. That makes sense considering that this item could be summarizing a set of circumstances that concern working conditions, as carrying out risky tasks, or suffering poor environmental conditions (exposure to noise -both- and low temperature -men-).

In the case of women, the "Sector of activity" (y15_nace_r1_17) is one of the variables selected by MARS to explain the target MSD. That seems to be logical, now that there are plenty of different tasks with a high impact on the health status of workers, specifically talking about shoulders, neck and upper limb MSD. Such tasks are deeply related with certain jobs, as cleaning (Bell and Steele, 2012; Calvet et al., 2012), textile jobs (Karimi et al., 2016) or health care (Salerno et al., 2012; Neupane et al., 2016). Indeed, this is consistent with the content of Figure 2 and the comments previously made: health and social work and activities of household are two of the most feminized sectors in the sample.

In the case of men, other variables come into play that are strongly connected to work organization and schedules, such as "Working full or part time" (y15_Q2d) or "Shift working" (y15_Q39e). These are factors that act synergically with other individual characteristics, especially with age, that also appears in MARS_2 and accelerates the health deterioration in relation with working on shifts (Oginska et al., 1993).

As previously suggested, the main idea that these results bring is that it is easier to design and develop an intervention program based on workplace conditions with a good impact on men than on women. This is because such workplace conditions are internal factors within the control of the organization, that can be evaluated and improved, and have deep and positive impact avoiding or reducing MSD in men.

4.3. Activities outside working hours.

This is the most important category to consider the influence of gender on the prevalence of MSD and how a double workload (professional and personal) has an irreversible impact on women's health. This category is, with *Psychosocial factors*, the one that affects women the most, representing almost 24% of the variables in MARS_1 model. On the other hand, for men, it barely represents 14%.

There are some remarkable differences between women and men. First, it must be mentioned that men's items are mainly related with the obstacles to find a balance between family time and work, which does not necessarily lead to assume that men carry out the housework. An incompatibility between these two sides of life is confirmed, but there is no evidence of which of both is chosen as the imperative one and which the sacrificed one is.

In the case of women, even though the model includes one item linked with free time activities and hobbies, the other four factors point out the possible impact of housework on MSD prevalence, as other studies have emphasized before (Habib et al., 2010; Habib et al., 2012).

It is important to remark that there is one item that only appears on MARS_1: "Caring for children and grandchildren" (y15_Q95). This seems to be one of the main conclusions obtained from the current study: how caring work is usually the responsibility of women, as social roles based on gender frequently stablish nowadays, and how this affects women's general health status, particularly MSD prevalence.

Once again, these findings recommend considering gender on ergonomic programs to ensure their effectiveness, because ignoring it might prevent from reaching real and ongoing improvements on women. They also point to the idea that public policies promoting gender equality could be indispensable to improve working conditions in women. There seems to be a vague zone where the borderline between companies and administration's responsibilities is not clear.

4.4. Physical risk factors

It was expectable that physical risk factors at work would be extremely linked to the occurrence of the target MSD, so the results of the MARS models contribute to validate, in a technical way, the expert criteria applied to the data selection and filtering. Physical risk factors have an influence in both models, but there are significant differences between men and women.

This category represents nearly 24% of variables in the case of men and 14% for women. In the same direction as earlier analysis, this fixes the idea that ergonomic interventions exclusively aimed at addressing physical risk factors are unintentionally more focused on men's circumstances.

Ergonomics used to be the warhorse issue in some sectors that are traditionally masculinized, such as construction (Valero et al., 2016). Thus, it is necessary to make a distinction between physical risk factors that typically affect one or the other gender: women are more likely to suffer in works implying repetitive tasks and postural statism, whereas men carry out tasks that involve more physical efforts (Chatigny and Riel, 2014, Clays et al., 2020). In fact, the factor related with "Carrying or moving heavy loads" (y15_Q30c) appears only in MARS_2.

Some recent works propose including personal characteristics (including gender) to improve the performance of classical risk assessment methods for this type of factors, such as the NIOSH lifting equation (Barim et al., 2019).

It can be concluded that performing risk assessment on workplaces is still a good ergonomic strategy when the target workforce is composed mostly of men, provided that diversity is attended to through specific measures to avoid absence of protection for women. Nevertheless, as stated before, the ideal approach would be designing an intervention program attending each individual's characteristics, including gender.

4.5. Psychosocial factors

Gender identity, social behavior and psychosocial factors are related concepts, and because of that their assessment becomes more difficult as well as the design of effective intervention programs. For example, even how masculinity is constructed in working environments (Stergiou-Kita et al., 2016) could be an obstacle to act against psychosocial risks in many masculinized sectors.

As previous studies have pointed out (Solidaki et al., 2010; Haukka et al., 2011; Eatough et al., 2012; Suárez Sánchez et al., 2016; Busto Serrano et al., 2018), psychosocial factors have shown an impact on MSD prevalence. However, considering gender, the influence of psychosocial risks on MSD is just the opposite to that of *Physical risk factors*: the weight of this category is higher for women.

It seems that "Having to handle conflicts" (y15_Q30g) affects in a similar way both genders, since it appears almost in the same position in both models: 18th for men and 19th for women. However, on the other hand, there are many differences in the remainder factors of this category: MARS_2 (men) includes items such as "Stress" (y15_Q61m) and "Emotionally disturbing situations" (y15_Q30h), while MARS_1 (women) includes variables related with work rhythm and job contents (y15_Q61g, y15_Q61f, y15_Q53d and y15_Q49a).

It must be kept in mind that these differences can also determine the course of psychosocial intervention based on the gender of the workforce. Once again, it must be underlined that considering gender is the right strategy to deal with psychosocial risks, as it happened with physical or other work related factors.

All the differences that have been found between genders regarding psychosocial factors and their influence on MSD could have numerous causes and origins that should be deeply studied. In fact, a new line of research based on these ideas could be established.

5. Conclusions

The origin of MSD is linked to numerous factors, but it must be noted that these factors have different behavior attending to gender. Two different MARS models, corresponding to each gender, that identify variables with the strongest influence in the prevalence of shoulders, neck and upper limb MSD have been calculated, confirming this conclusion.

Men seem to be more vulnerable to physical risk factors and some working conditions such low temperature or shift work, while women appear to be more affected by psychosocial related working conditions and activities carried out outside their working hours.

Analyzing the results obtained, it must be said that any intervention to reduce the MSD prevalence at work must be done including gender considerations. In this way, men need interventions more aimed at addressing physical risk factors and some other work-related conditions, such as environmental damaging agent's reduction, and actions over organizational factors. On the other hand, talking about women, an effective ergonomic intervention to reduce MSD prevalence must keep in mind the strong influence of psychosocial risks factors, as well as the consequences of the activities carried out outside work (such as housework and caring work), and how they contribute to increase gender differences. In conclusion, employers have to consider work-family balance in their organizational practices and policies.

How the gender difference works and where its causes lay could open new lines of research that probably need new data. This could be done through the design of a new gender sensitive tool, such as a survey focused on working conditions and health status.

Regarding the results presented, it can be concluded that, up to now, companies have mainly applied programs that would be effective on men but would fall short on women. Also, it can be said that a successful intervention against MSD must be carried out both by companies and public administrations, since some preventive measures would demand legislative changes.

To close the present contribution, it must be said that, applying the occupational ergonomics principle of "adapting working conditions to each person", gender needs to be considered to ensure the success and effectiveness of ergonomic interventions on the whole working population, addressing both physical and psychosocial risk factors.

Acknowledgements

The authors wish to acknowledge the European Foundation for the Improvement of Living and Working Conditions (Eurofund), as well as the UK Data Service, for providing us with the results of the database used in this work.
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Alonso Fernández, J.R., Díaz Muñiz, C., Garcia Nieto, P.J., de Cos Juez, F.J., Sánchez Lasheras, F., Roqueñí, M.N., 2013. Forecasting the cyanotoxins presence in fresh waters: A new model based on genetic algorithms combined with the MARS technique. Ecological engineering 53, 68-78. https://doi.org/10.1016/j.ecoleng.2012.12.015

Álvarez Antón, J.C., García Nieto, P.J., de Cos Juez F.J., Sánchez Lasheras, F., González Vega, M., Roqueñí Gutiérrez, M.N., 2013. Battery state-of-charge estimator using the SVM technique. Applied Mathematical Modelling 37 (9), 6244-6253. https://doi.org/10.1016/j.apm.2013.01.024

Álvarez Menéndez, L., de Cos Juez, F.J., Sánchez Lasheras, F., Álvarez Riesgo, J.A., 2010, Artificial neural networks applied to cancer detection in a breast screening programme. Mathematical and Computer Modelling 52 (7-8), 983-991. https://doi.org/10.1016/j.mcm.2010.03.019

Artime Ríos, E., Suárez Sánchez, A., Sánchez Lasheras, F., Seguí Crespo, M.M., 2018. Neural Comput & Applic. https://doi.org/10.1007/s00521-018-3581-3

Asensio-Cuesta, S., Diego-Mas, J.A., Alcaide-Marzal, J., 2010. Applying generalised feedforward neural networks to classifying industrial jobs in terms of risk of low back disorders. Int. J. Ind. Ergonom. 40, 629-635. http://dx.doi.org/10.1016/j.ergon.2010.04.007

Barim, M.S., Sesek, R.F., Capanoglu, M.F., Drinkaus, P., Schall, M.C., Gallagher, S., Davis, G.A., 2019. Improving the risk assessment capability of the revised NIOSH lifting equation by incorporating personal characteristics. Appl. Ergon. 74, 67-73.

Barnes, K.L., Bendixsen, C.G., 2017. "When This Breaks Down, It is Black Gold": Race and Gender in Agricultural Health and Safety. J. Agromedicines. 22(1), 56-65. https://doi.org/10.1080/1059924X.2016.1251368

Bell, A.F., Steele, J.R., 2012. Risk of musculoskeletal injury among cleaners during vacuuming. Ergonomics 55(2), 237-47. https://doi.org/10.1080/00140139.2011.592605

Busto Serrano, N.M., García Nieto, P., Suárez Sánchez, A., Sánchez Lasheras, F., Riesgo Fernández, P., 2018. A Hybrid Algorithm for the Assessment of Influence of Risk Factors in the Development of Upper Limb Musculoskeletal Disorders. In: de Cos Juez F. et al. (eds) Hybrid Artificial Intelligent Systems. HAIS 2018. Lecture Notes in Computer Science 10870, 634-646. https://doi.org/10.1007/978-3-319-92639-1_53

Calvet, B., Riel, J., Couture, V., Messing, K., 2012. Work organisation and gender among hospital cleaners in Quebec after the merger of "light" and "heavy" work classifications. Ergonomics 55(2), 160-72. https://doi.org/10.1080/00140139.2011.576776

Caroly, S., Major, M.E., Probst, I., Molinié, A.F., 2013. Le genre des troubles musculo-squelettiques. Interventions ergonomiques en France et au Canada. Travail, genre et sociétés. 29, 49-67.

Chatigny, C., Riel, J., 2014. La santé et la sécurité des élèves en centre de formation professionnelle: aproche, représentations, et genre. Pistes: Perspectives Interdisciplinaires sur le Travail et la Santé. 16(4):1-28.

Clays, E., Hallman, d., Oakman, J., Holtermannd, A., 2020. Objectively measured occupational physical activity in blue-collar workers: What is the role of job type, gender and psychosocial resources? Appl. Ergon. 82, 102948.

Côté, J.N., 2012. A critical review on physical factors and functional characteristics that may explain a sex/gender difference in work-related neck/shoulder disorders. Ergonomics 55(2), 173-82. https://doi.org/10.1080/00140139.2011.586061

De Andrés, J., Lorca, P., de Cos Juez, F.J., Sánchez-Lasheras, F., 2011. Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). Expert Systems with Applications 38 (3), 1866-1875. https://doi.org/10.1016/j.eswa.2010.07.117

De Andrés, J., Sánchez-Lasheras, F., Lorca, P., de Cos Juez, F.J., 2011. A hybrid device of self organizing maps (SOM) and multivariate adaptive regression splines (MARS) for the forecasting of firm's bankruptcy. Accounting & Management Information Systems 10, 351-374.

De Cos Juez, F.J., Sánchez Lasheras, F., García Nieto, P.J. Álvarez-Arenal, A., 2008. Non-linear numerical analysis of a double-threaded titanium alloy dental implant by FEM. Applied Mathematics and Computation. 206 (2), 952-967. https://doi.org/10.1016/j.amc.2008.10.019

Denis, D., St-Vincent, M., Imbeau, D., Jetté, C., Nastasia, I., 2208. Intervention practices in musculoskeletal disorder prevention: a critical literature review. Appl. Ergon. 39(1):1-14. https://doi.org/10.1016/j.apergo.2007.02.002

Eatough, E.M., Way, J.D., Chang, C.H., 2012. Understanding the link between psychosocial work stressors and work-related musculoskeletal complaints. Appl. Ergon. 43, 554-563. https://doi.org/10.1016/j.apergo.2011.08.009

Eurofound (European Foundation for the Improvement of Living and Working Conditions), 2018. European Working Conditions Survey Integrated Data File, 1991-2015. [data collection]. 7th Edition. UK Data Service. SN: 7363, http://doi.org/10.5255/UKDA-SN-7363-7

Eurofound, 2016. Sixth European Working Conditions Survey – Overview report, Publications Office of the European Union, Luxembourg.

Eurofund Website: Sixth European Working Conditions Survey: 2015. https://www.eurofound.europa.eu/surveys/european-working-conditions-surveys/sixth-european-working-conditions-survey-2015. Last accessed: 15/02/2019.

Friedman, J.H., 1991. Multivariate adaptive regression splines (with discussion), Ann. Stat. 19, 1–141.

García Iglesias, D., Roqueñí Gutiérrez, N., De Cos, F.J., Calvo, D., 2018. Analysis of the High-Frequency Content in Human QRS Complexes by the Continuous Wavelet Transform: An Automatized Analysis for the Prediction of Sudden Cardiac Death. Sensors 18(2), 560. https://doi.org/10.3390/s18020560

García Nieto, P.J., Martínez Torres, J., de Cos Juez, F.J., Sánchez Lasheras, F., 2012. Using multivariate adaptive regression splines and multilayer perceptron networks to evaluate paper manufactured using Eucalyptus globulus. Appl. Math. Comput. 219 (2), 755-763. https://doi.org/10.1016/j.amc.2012.07.001

Garcia Nieto P.J., Sánchez Lasheras, F., de Cos Juez, F.J., Alonso Fernández, J.R., 2011. Study of cyanotoxins presence from experimental cyanobacteria concentrations using a new data mining methodology based on multivariate adaptive regression splines in Trasona reservoir (Northern Spain) J. Hazard. Mater. 195, 414-421. https://doi.org/10.1016/j.jhazmat.2011.08.061

Guzmán, D., de Cos Juez, F.J., Sánchez Lasheras, F., Myers, R., Young L., 2010. Deformable mirror model for open-loop adaptive optics using multivariate adaptive regression splines, Opt. Express 18 (7), 6492-6505. https://doi.org/10.1364/OE.18.006492

Habib, R.R., Fathallah, F.A., Messing, K., 2010. Full-time homemakers: workers who cannot "go home and relax". Int. J. Occup. Saf. Ergon. 16(1), 113-28. http://doi.org/10.1080/10803548.2010.11076833

Habib, R.R., El Zein, K., Hoieii, S., 2012. Hard work at home: musculoskeletal pain among female homemakers. Ergonomics. 55(2): 201-11. http://doi.org/10.1080/10.1080/00140139.2011.574157

Härenstam, A., 2009. Exploring gender, work and living conditions and health suggestions for contextual and comprehensive approaches. Scand. J. Work. Environ. Health. 35(2), 127-33.

Haukka, E., Leino-Arjas, P., Ojajärvi, A., Takala, E.P., Viikari-Juntura, E., Riihimäki, H., 2011. Mental stress and psychosocial factors at work in relation to multi-site musculoskeletal pain: A longitudinal study of kitchen workers. Eur. J. Pain. 15, 432-38. http://doi.org/10.1016/j.ejpain.2010.09.005

Karimi, N., Moghimbeigi, A., Motamedzada, M., Roshanaei, G., 2016. Evaluation of Related Risk Factors in Number of Musculoskeletal Disorders Among Carpet Weavers in Iran. Sa. Health. Work. 7(4), 322-325. http://doi.org/10.1016/j.shaw.2016.04.004

Krzemień, A., 2019a. Dynamic fire risk prevention strategy in underground coal gasification processes by means of artificial neural networks. Archives of Mining Sciences 64 (2019), 1, 3-19. https://doi.org/10.24425/ams.2018.124991

Krzemień, A., 2019b. Fire risk prevention in underground coal gasification (UCG) within active mines: Temperature forecast by means of MARS models. Energy 170, 777-790. https://doi.org/10.1016/j.energy.2018.12.179

Laberge, M. Vézina, N., Saint-Charles, J., 2012. Safe and healthy integration into semiskilled jobs: does gender matter? Work. 41(1), 4642-9 http://doi.org/10.3233/WOR-2012-0102-4642

Messing, K., Punnett, L., Bond, M., Alexanderson, K., Pyle, J., Zahm, S., Wegman, D., Stock, S.R., de Grosbois, S., 2003. Be the fairest of them all: challenges and recommendations for the treatment of gender in occupational health research. Am J Ind Med. 43, 618-629. http://doi.org/10.1002/ajim.10225

Messing, K., Stock, S.R., Tissot, F., 2009. Should studies of risk factors for musculoskeletal disorders be stratified by gender? Lessonsfrom the 1998 Québec Health and Social Survey. Scand. J. Work Environ. Health. 35(2), 96-112.

Milborrow, S., 2011. Derived from mda:mars by T. Hastie and R. Tibshirani. earth: Multivariate Adaptive Regression Splines. R package, http://www.milbo.users.sonic.net/earth.

Mousaid, S., De Moortel, D., Malmusi, D., Vanroelen, C., 2016. New perspectives on occupational health and safety in immigrant populations: studying the intersection between immigrant background and gender. Eth. Health. 21(3), 251-67. http://doi.org/10.1080/13557858.2015.1061103

Neupane, S., Nygard, C.H., Oakman, J., 2016. Work-related deteminants of multi-site musculoskeletal pain among employees in the health care sector. Work. 54(3), 689-97. http://doi.org/10.3233/WOR-162320

Oginska, H., Porkorski, J., Oginski, A., 1993. Gender, ageing, and shiftwork intolerance. Ergonomics. 36 (1-3), 161-8. http://doi.org/10.1080/00140139308967868

Ordóñez Galán, C., Sánchez Lasheras, F., de Cos Juez, F.J., Bernardo Sánchez, A. 2017. Missing data imputation of questionnaires by means of genetic algorithms with different fitness functions. Journal of Computational and Applied Mathematics 311, 704-717. https://doi.org/10.1016/j.cam.2016.08.012

Paarup, H.M., Baelum, J., Holm, J.W., Manniche, C., Wedderkopp, N., 2011. Prevalence and consequences of musculoskeletal symptoms in symphony orchestra musicians vary by gender: a cross-sectional study. BMC Musculoskelet. Disord. 7, 12-223. http://doi.org/10.1186/1471-2474-12-223

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rasmussen, C.D., Jorgensen, M.B., Carneiro, I.G., Flyvholm, M.A., Olesen, K., Sogaard, K. Holtermann, A., 2012. Partipation of Danish and immigrant cleaners in a 1-year worksite intervention preventing physical deterioration. Ergonomics. 55(2), 256-64. http://doi.org/10.1080/00140139.2011.592651

Riesgo García, M. V., Krzemień, A., Manzanedo del Campo, M. Á., Escanciano García-Miranda, C., Sánchez Lasheras, F., 2018. Rare earth elements price forecasting by means of transgenic time series developed with ARIMA models. Resour. Policy 59, 95-102. https://doi.org/10.1016/j.resourpol.2018.06.003

Salerno, S., Livigni, L., Magrini, A., Talamanca, I.F., 2012. Gender and ergonomics: a case study on the “non.formal” work of women nurses. Ergonomics. 55(2), 140-6. http://doi.org/10.1080/00140139.2011.637134

Sánchez Lasheras, F., de Cos Juez, F. J., Suárez Sánchez, A., Krzemień, A., Riesgo Fernández, P., 2015. Forecasting the COMEX copper spot price by means of neural networks and ARIMA models. Resour. Policy 45, 37–43. https://doi.org/10.1016/j.resourpol.2015.03.004

Sánchez Lasheras, J.E., González Donquiles, C., García Nieto, P.J. et al., 2018. Neural Comput & Applic. https://doi.org/10.1007/s00521-018-3503-4

Silverstein, B., Fan, Z.J., Smith, C.K., Bao, S., Howard, N., Spielholz, P., Bonauto, D., Viikari-Juntura E., 2009. Gender adjustment or stratification in discerning up per extremity musculoskeletal disorder risk? Scand. J. Work Environ. Health. 35(2): 113-26.

Solidaki, E., Chatzi, L., Bitsios, P., Markatzi, I., Plana, E., Castro, F., Palmer, K., Coggon, D., Koqevinas, M., 2010. Work-related and psychological determination of multisite musculoskeletal pain. Scand. J. Work. Environ. Health. 36: 54-61.

Stergiou-Kita, M., Mansfield, E., Colantonio, A., Moody, J., Mantis, S., 2016. What's gender got to do with it? Examining masculinities, health and safety and return to work in male dominated skilled trades. Work. 54(3): 721-33. https://doi.org/10.3233/WOR-162322

Suárez Sánchez, A., Riesgo Fernández, P., Sánchez Lasheras, F., de Cos Juez, F.J., García Nieto, P.J., 2011. Prediction of work-related accidents according to working conditions using support vector machines. Appl. Math. Comput. 218, 3539-3552. https://doi.org/10.1016/j.amc.2011.08.100

Suárez Sánchez, A., Iglesias-Rodríguez, F.J., Riesgo Fernández, P., de Cos Juez, F.J., 2016. Applying the K-nearest neighbor technique to the classification of workers according to their risk of suffering musculoskeletal disorders. Int. J. Ind. Ergonom. 52, 2016, 92-99. https://doi.org/10.1016/j.ergon.2015.09.012

UKDS (UK Data Service) Website: https://www.ukdataservice.ac.uk/. Last accessed: 15/02/2019.

Valero, E., Sivanathan, A., Bosché, F., Abdel-Wahab, M., 2016. Musculoskeletal disorders in construction: A review and a novel system for activity tracking with body area network. Appl. Ergon. 54:120-30. https://doi.org/10.1016/j.apergo.2015.11.020

Waller, S., Bradley, M., Hosking, I., Clarkson P.J., 2015. Making the case for inclusive design. Appl. Ergon. 46, 297-303. http://dx.doi.org/10.1016/j.apergo.2013.03.012

Yazdani, A., Neumann, W.P., Imbeau, D., Bigelow, P. Pagell, M., Wells, R., 2015. Prevention of musculoskeletal disorders within management systems: A scoping review of practices, approaches, and techniques. Appl. Ergon. 51, 255-62. https://doi.org/10.1016/j.apergo.2015.05.006

Yazdani, A., Wells, R., 2018. Barriers for implementation of successful change to prevent musculoskeletal disorders and how to systematically address them. Appl. Ergon. 73, 122-140. https://doi.org/10.1016/j.apergo.2018.05.004

Widanarko, B., Legg, S., Stevenson, M., Devereux, J., Eng, A., Mannetje, A., Cheeng, S., Douwes, J., Ellison-Loschmann, L., McLean, D., Pearce, N., 2011. Prevalence of musculoskeletal symptoms in relation to gender, age, and occupational/industrial group. Int. J. Ind. Ergon. 41(5), 561-572. https://doi.org/10.1016/j.ergon.2011.06.002