Influence of contracting conditions on absenteeism

Iván Fernández-Suárez, María Botey, and Fermín Torrano*

Universidad Internacional de La Rioja (Spain)

Título: Influencia de las condiciones de contratación en el absentismo laboral.

Resumen: El propósito de esta investigación es analizar las relaciones entre las condiciones de contratación y el absentismo laboral en una muestra de 5524 trabajadores, con el fin de identificar qué segmentos (por tipo de contrato y jornada, tiempo contratado, antigüedad en la empresa y bajas por enfermedad común ocurridas en los tres últimos años) están más relacionados con la posibilidad de sufrir un proceso de incapacidad temporal en el año 2017. Se realizaron análisis descriptivos, la prueba chi-cuadrado para tablas de contingencia con dos muestras independientes y los árboles de decisión, basados en el algoritmo CHAID (*Chi-squared Automatic Interaction Detection*), para detectar las variables más importantes en la identificación de perfiles con una mayor probabilidad de sufrir una incapacidad temporal derivada de contingencias comunes. Los resultados ponen de manifiesto la existencia de diferencias entre las variables estudiadas. Se considera la modalidad de contratación un factor de riesgo importante del absentismo laboral.

Palabras clave: Absentismo laboral; Contrato; Segmentación jerárquica; CHAID.

Abstract: The aim of this paper is to analyze the relationship between hiring conditions and work absenteeism in a sample of 5524 workers in order to identify which segments (by type of contract and workday, time hired, seniority in the company and sick leaves occurred in the last three years) are more related to the possibility of suffering a temporary incapacity in 2017. Descriptive analyzes, the chi-square test for contingency tables with two independent samples, and the decision trees based on the CHAID algorithm (Chi-squared Automatic Interaction Detection) were carried out to detect the most important variables in the identification of profiles with a greater probability of suffering a temporary incapacity for common contingencies. The results show the existence of differences between the variables studied. The hiring modality is considered an important risk factor for work absenteeism.

Keywords: Work absenteeism; Contract; Hierarchical segmentation; CHAID.

Introduction

Workplace absenteeism is one of the current primary concerns at economic, social and business levels, as the absence of an employee from their post, in its different forms and causes, leads to production losses, changes the organisational environment and generates increased cost to the public Social Security system (Hassard, Teoh, Visockaite, Dewe, & Cox, 2018; López-Alonso, Ibarrondo-Dávila, & Rubio-Gámez, 2015; Mazzoni et al., 2019). In Spain, absenteeism due to temporary incapacity (work-related and non-work-related) accounts for 13.09 lost days for every active worker, one of the highest levels on an international scale (Blasco de Luna et al., 2018). When it comes to absenteeism due to common illness, the data is even more negative. According to the 7th Adecco Report on absenteeism, in 2017 86.4% of cases of temporary incapacity were related to cases of common illness, with an average duration of 17.21 days lost (see Blasco de Luna et al., 2018). This has led to an increase of 6% in comparison with 2016. The increased cost of absenteeism, added to the personal and social repercussions it can have on organisations, has meant its control is one of the primary strategic objectives of any business, region or state (Martin-Fumadó, Martí Amengual, Puig Bausili, & Arimany-Manso, 2014; Martín-Román, & Moral, 2017).

Absenteeism is regarded by several authors as a multidimensional process (Altunkaynak, 2018; Khashaba, El-Helaly, El-Gilany, Motawei, & Foda, 2018; Thurston, &

Glendon, 2018), which has been explained as much by individual factors, such as age (Fernández-García, 2013), gender (Campos-Serna, Ronda- Pérez, Artazcoz, & Benavides, 2012), the employee's nationality (Casey, Riseborough, & Krauss, 2015) or the consumption of alcohol (Luján, Santana, Alemán, & Sánchez, 2014), as by the consequences of the various social, economic and organisational conditions to which the employee is exposed in their work environment (Boada i Grau, De Diego Vallejo, Agulló Tomás, & Mañas Rodríguez, 2005). In respect of working conditions, a great number of studies highlight the relationship between absenteeism and different organisational variables, such as salary, job security and contract type (de Menezes, & do Monte, 2013). In particular, it is observed that people with a poorer state of health, and therefore more of a propensity to calling in sick, are those workers with temporary contacts and who are exposed to the worst working conditions (Virtanen et al., 2005).

Job insecurity, characterised by instability, uncertainty and lack of guarantees on working conditions, has been used as a valid indicator for the analysis of absenteeism, as it has already been identified as one of the risk factors (Jiang, & Lavaysse, 2018) and factors of work attendance (Bogaerts, De Witte, Verlinden, & Vermeulen, 2017). At the same time, differences in recruitment terms may also be determinants in absenteeism (García, Green, & Navarro Paniagua, 2018; Restrepo, & Salgado, 2013). The studies specifically place particular importance upon contract type, as higher rates of absenteeism have been evidenced in employees with casual contracts than in those with permanent contracts (Zaballa et al., 2016). Similarly, these studies indicate that the length of the working day stated in the employment contract shows a

close link with levels of absenteeism and the health of employees (Benavides, Benach, Diez- Roux, & Roman, 2000). For example, employees working temporary partial days manifest higher rates of job dissatisfaction and symptoms of musculoskeletal disorders (Benavides, & Benach, 1999).

Objetives and Hypothesis

Under the aforementioned frame of reference, the aim of this study is to analyse the variables of a contractual nature which have an influence on workplace absenteeism. The intention is thus to provide a description of which recruitment factors (type of contract and work day, contract term, longevity and illnesses suffered in recent years and their duration) may influence the propensity to call in sick in the workplace environment.

This primary objective is detailed through different employment scenarios:

- Are employees on a temporary contract most likely to experience temporary incapacities arising from non-work-related occurrences?
- Does the working day have an influence on the likelihood of experiencing further temporary incapacity due to non-work-related occurrences?
- Is there a link between longevity in a post and the likelihood of experiencing temporary incapacity due to non-work-related occurrences?
- Are temporary incapacities experienced in recent years a potential predictor variable of present-day absenteeism?

Method

Participants

The sample is of an incidental and non-probabilistic nature and is made up of the full complement of staff hired in 2017, being 5524 people in a Spanish company in the Services sector, engaged in maintenance and cleaning tasks. Table 1 shows the characteristics of the sample with regard to its different recruitment methods.

Table 1. Contract characteristics of the sample.

Variable and Group	N	%
Gender		
Male	814	14.7%
Female	4710	85.3%
Type of Contract		
Temporary	2653	48.03%
Permanent	2871	51.97%
Type of Work Day		
Full	2130	38.56%
Partial	3394	61.44%
2017 Contract Days		
0 to 2 months	1392	25.20%
2 to 8 months	1331	24.09%
More than 8 months	2801	50.71%
Longevity		
Less than 1 year	1432	25.92%
1 to 3 years	1201	21.74%
4 to 15 years	1651	29.89%
Over 15 years	1240	22.45%

At the same time, data was obtained for each employee relating to their absenteeism due to common illness in 2017. There was particular focus on the number of cases of common illness which had been experienced and the number of sick days taken this year.

In the development of the study, the term common illness shall equate to the concept of sickness.

Variables

On account of the above, the variables which have been used (see Table 2) correspond, on the one hand, to information about employee recruitment in 2017, (type of contract and work day, contract term in 2017 and longevity), and on the other hand, on the number of periods of sick leave taken by each employee in the last 3 years (2015, 2016 and 2017), and their duration. The ratio of absenteeism due to illness was also calculated, with the result obtained by dividing the number of sick days by the number of contracted days in 2017.

Table 2. Study subject variables and coding.

Variable	Description	Possible Values
TContract	Type of Contract	0 = permanent contract
		1 = temporary contract
Work Day	Type of Work Day	0 = full work day
		1 = partial work day
ContractTm	Contract Term in 2017	1 = 0 to 2 months
		2 = 2 to 8 months
		3 = More than 8 months
ContractDays	Duration of 2017 Contract in Days	From 0 to 360
Longev	Longevity	1 = Less than 1 year
		2 = Between 1 and 3 years
		3 = Between 4 and 15 years
		4 = Over 15 years

Variable	Description	Possible Values
IllXX_Cases	Number of cases of illness which have occurred in 20XX (XX= 17, 16, 15)	From 0 to 10
XXIll01_Case	Dichotomous variable related to IllXX_Cases	0 = Has not experienced any case of illness in year XX 1 = Has experienced illness in year XX
IllXX_Durat	Duration of illnesses in 20XX (XX = 17, 16, 15)	From 0 to 360 days
Illness_Ratio	IllXX_ Durat/ContractedDays	From 0 to 1
Ill01_Ratio	Dichotomous variable related to Illness_Ratio	0 = if Illness_Ratio < .5 1 = if Illness_Ratio ≥ .5

Procedure

The process for obtaining the data was completed using the ERP (*Enterprise Resource Planning*) software in the organisation with which the study is concerned. The statistical control of workplace absenteeism was developed using Navision and Meta4 computer applications, into which all data relating to absences due to the employees' non-work-related and work-related illnesses was inputted. Similarly, all sociodemographic and contractual variables that have formed part of this study were included in both applications.

The introduction of the information relating to absenteeism was done by the Human Resources department of the aforementioned organisation. Likewise, the department for Occupational Risk Prevention conducted an investigation of all cases of accidents at work and work-related illness and the statistical analysis of absenteeism arising from both.

Data Analysis

Firstly, a descriptive analysis of the data (frequency and percentages) was produced by way of contingency tables. Furthermore, the Chi-squared test was used for contingency tables with two independent samples, with the aim being to determine the existence of significant differences between the groups being considered. Finally, the CHAID (Chi-squared Automatic Interaction Detection) (Kass, 1980) method was used, which, as is known, is a non-parametric technique for analysing the segmentation of variables and it is suitable for large samples and categorical dependent variables, as is the case here, which enables the grouping of elements described by a dependent variable according to the values of other independent variables (predictors) (Ortega, 2015).

The CHAID method analyses all the values of each potential predictor variable by using the Chi-squared statistical test, which reflects how similar or connected the variables are. From there, it selects the most significant predictor in order to form the main partition in the decision tree, in such a way that each node is comprised of those categories which are similar to the selected variable. The process continues successively until the tree is complete (Sanz, & Ponce de León, 2010).

In this kind of inductive model, it is essential that the way in which the tree's structure behaves is assessed in order to generalise the available data in a sample of different data. The technique most employed for this validation involves dividing the data into a training sample and a validation or test sample. If the model generated with the training data shows a predictive capability similar to that of the validation data, the rules may be generalised. The tree is assessed with the confusion matrix which shows the proportion of cases correctly categorised by the algorithm in relation to the whole of the categorised sample and the proportion of inaccuracies.

Selecting the optimal tree is related to the problem of overfitting. Usually, the greater the complexity the greater the overfit. In the method applied to this study, 'self-pruning' parameters were set which prevented the overfix and generated trees with little complexity (only three levels of division and 'child' nodes with more than 100 individuals), which provided a stable model with few differences between the training sample and the test.

This technique was chosen due to its good results both in normal populations (Peñalva-Vélez, & López-Goñi, 2014) and clinical populations (López-Goñi, Fernández- Montalvo, & Arteaga, 2012).

All the analyses were carried out using the IBM®-SPSS Statistics® package.

Results

First of all, the interrelations between the number of cases of illness which occurred in the last three years (2015, 2016 and 2017) and the variables of contract type, work day type and contract term (see Figure 1) were examined. To analyse these types of comparative data, the size of each group was divided by the total number of individuals. The number of contracted months was also taken into consideration in order to eliminate the effect of exposure time, where possible, in such a way so as the data is relative to one working month.
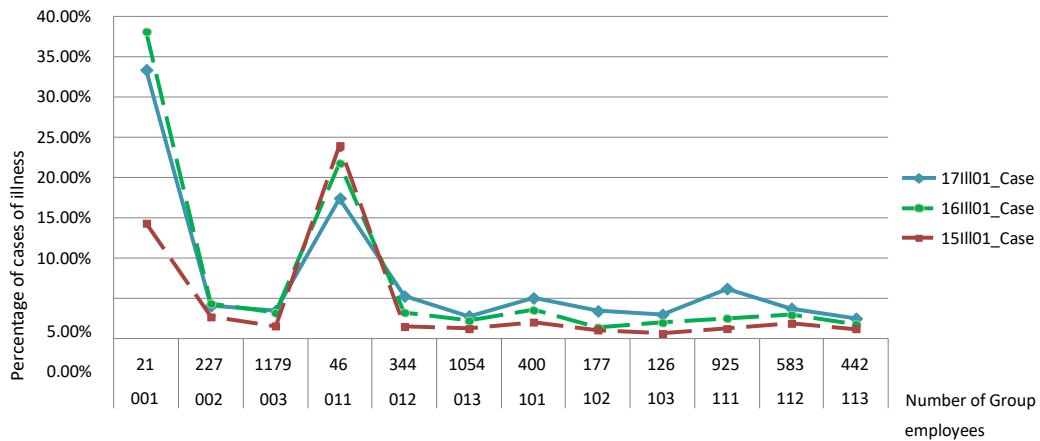
Iván Fernández-Suárez et al.

Figure 1. Percentage of cases of illness in each group by month worked.
Note: Group = ABC, where A is the Type of contract (0 = Permanent, 1 = Temporary), B is the Type of work day (0 = Full, 1 = Partial) and C is Months worked in the year (1 < = 2 months, 2 < = 8 months, 3 > 8 months).

As can be seen, the number of cases of illness in 2017 holds the same distribution between groups as in the two previous years. The existing difference stands out between the high values of the first half of the chart (permanent contract) compared with the low values in the second half (temporary contract). It is also noted that the groups on contracts of a duration shorter than two months represent a greater proportion of people with a case of illness than those of a longer duration.

If these results are compared with the average ratio of illness in each group, clear variations are observed with regard to the groups on a contract shorter than two months. This means, as can be observed in Figure 2, that the groups with short term contracts represent a greater proportion of people with a case of illness in 2017, but with a ratio of sick days lower than the groups on a longer recruitment term.
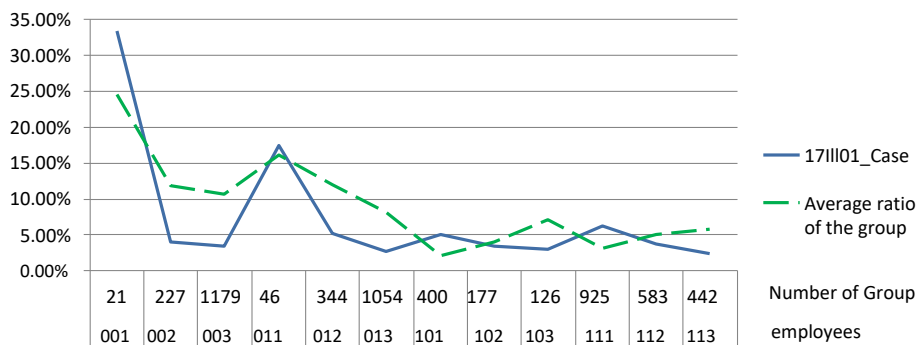

Figure 2. Percentage of cases and ratio by illness in 2017.

Secondly, four Chi-squared tests were carried out on the Group variable in Figures 1 and 2 in comparison with each of the considered variables (17Ill01_Case, 16Ill01_Case, 15Ill01_Case and Ill01_Ratio). In all cases significant differences were obtained between the different groups being considered (Table 3).

Table 3. Chi-squared tests of the Group variable in comparison with the variables 17Ill01_Case, 16Ill01_Case, 15Ill01_Case and Ill01_Ratio.

		17Ill01_Case		16Ill01_Case		15Ill01_Case		Ill01_Ratio	
	gl	Value	*p*	Value	*p*	Value	*p*	Value	*p.*
Pearson's Chi-squared test	11	403.832a	.000	462.085b	.000	199.587c	.000	94.297d	.000
Reason of verisimilitudes	11	454.306	.000	528.448	.000	237.862	.000	89.507	.000
Number of valid cases		5524		5524		5524		5524	

a.1 field (4.2%) has an anticipated frequency lower than 5. The minimum anticipated frequency is 4.90.
b.1 field (4.2%) has an anticipated frequency lower than 5. The minimum anticipated frequency is 3.92.
c.2 fields (8.3%) have an anticipated frequency lower than 5. The minimum anticipated frequency is 2.09.
d.2 fields (8.3%) have an anticipated frequency lower than 5. The minimum anticipated frequency is 1.23.

Finally, decision trees were used to identify the types of profiles which manifest a higher probability of experiencing an illness in 2017. One decision tree was analysed for every variable considered (2017 cases of illness and illness ratio).

Initially, for the first tree, the sample was arbitrarily split into two units: training sample (60%) and test sample (40%). Hereafter, explanator variables were selected, type of con-

tract, work day, contract term, longevity and number of instances of illness in 2016 and 2015. Some self-pruning parameters were defined for the tree: three levels of division and child nodes with more than 100 individuals, with the aim of guaranteeing the establishment of a non-overfixed tree. The results obtained are shown in Figure 3:
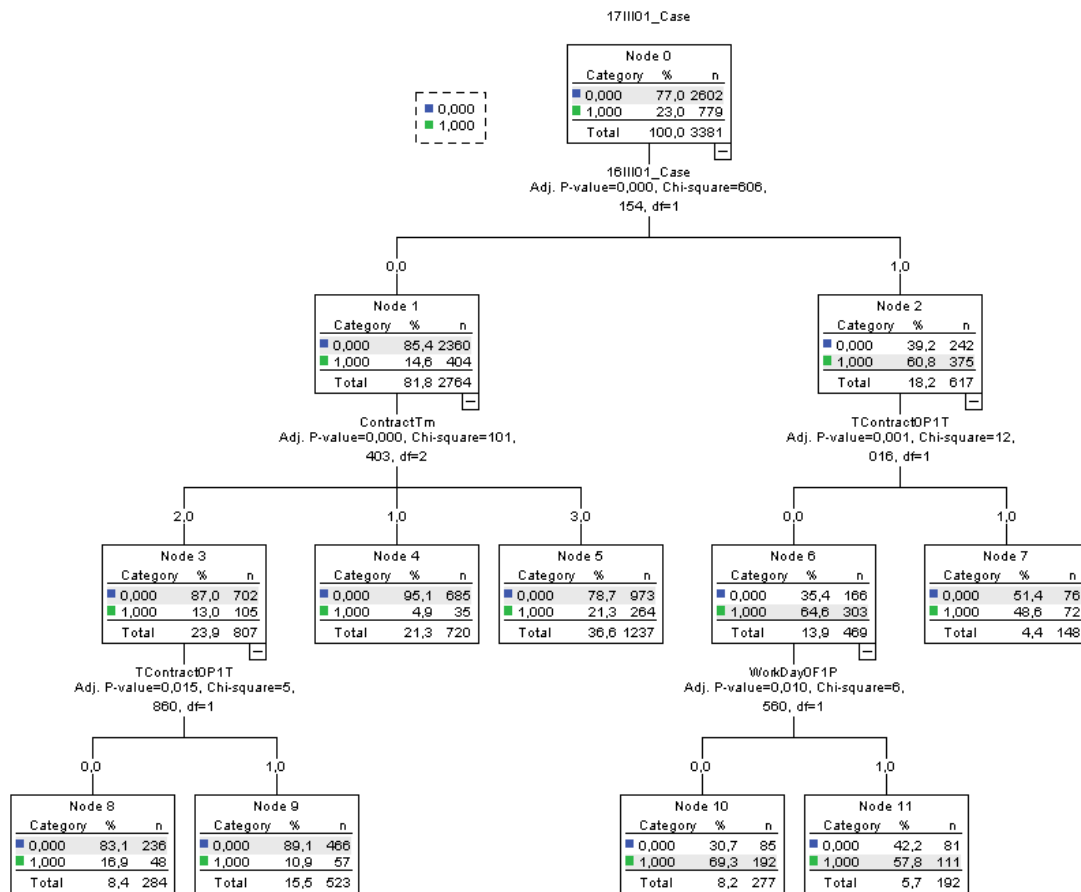

Figure 3. Decision tree for the variable of cases of illness in 2017.

The tree produced accommodates a percentage of people for each node with no absenteeism and another complementary batch of people with absenteeism. The algorithm used to determine the predictive value for all people in a node is as follows: 0 is predicted if its associated probability is greater than a fixed amount on the decision boundary and where 1 is lower.

To correctly interpret the tree, the importance of the weight of every independent variable on the tree was taken into account, which is dependent upon the depth the variable on the tree appears with and the number of times it is present. The most important variable was the existence or non-existence of cases of illness in the previous year (100%). The second, and less relevant, variable was the cases of illness

connected to the two previous years (20%). The rest of the variables manifested an influence of less than 15%.

The confusion matrix produced for a threshold value (decision *threshold* or boundary) equal to 50 produced an accuracy rate percentage of 81% in the training sample and 79.6% in the test sample, which indicates it is a stable tree (the variation between both samples was very low) with a very high accuracy rate.

The assessment of the performance of binary classifiers was conducted through the ROC (*Receiver Operating Characteristic*) curve and the PRC (*Precision-Recall*) curve (Figure 4). Whilst various indicators or measures exist for assessing a predictor for a binary classifier, the work of de Saito and Rehmsmeier (2015) was taken as a reference, as can be seen in Table 4.

Table 4. Description of predictor assessment measures for binary classifiers by Saito and Rehmsmeier (2015).

		Predicted	
Observed	0-negative	TN-True Negative	FP-False Positive
	1-positive	FN-False Negative	TP-True Positive

Measure	Formula
ACC	(TP+TN)/(TP+TN+FN+FP)
ERR	(FP+FN)/(TP+TN+FN+FP)
SN, TPR, REC	TP/(TP+FN)
SP	TN/(TN+FP)
FPR	FP/(TN+FP)
PREC, PPV	TP/(TP+FP)
MCC	(TP*TN-FP*FN)/((TP++FP)*(TP+FN)*(TN+FP)*(TN+FN))$^{1/2}$
$F_{0,5}$	1,5*PREC*REC/(0,25*PREC+REC)
F_1	2*PREC*REC/(PREC+REC)
F_2	5*PREC*REC/(4*PREC+REC)

ACC: accuracy; ERR: error rate; SN: sensitivity; TPR: true positive rate; REC: recall; SP: specify; FPR: false positive; PREC: precision; PPV: positive value; MCC: Matthews correlation coefficient; F: F score; TP: true positive; FP: false positives; FN: false negatives

The ROC curve shows the figure obtained upon comparing the accuracy (ACC = *accuracy* or proportion of successes) against the sensitivity (SN = *sensitivity*). The curve associated with the research problem produces an area of .693.
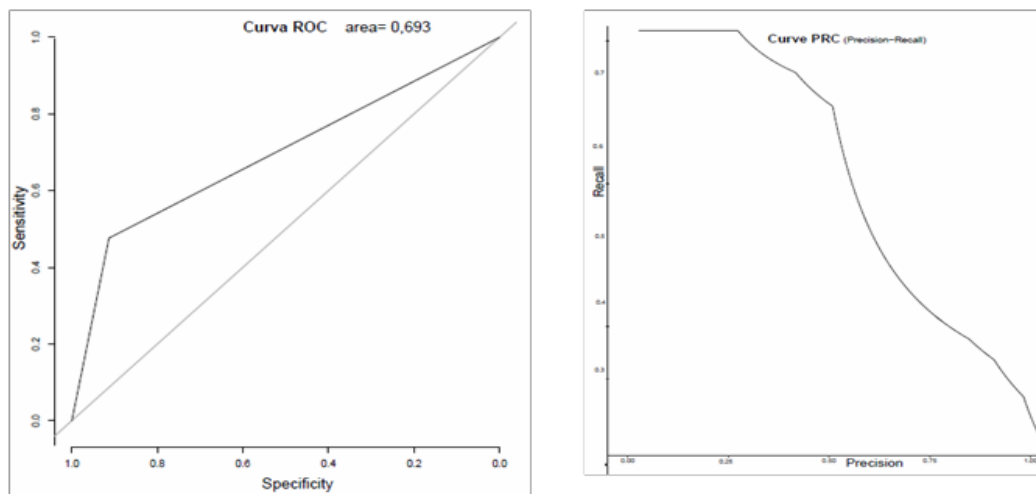
Figure 4. ROC curve and PRC curve.

The sample used in this study presents an unbalanced target variable (it contains just 23% of employees with an instance of illness in 2017). According to Saito and Rehmsmeier (2015), to assess a predictive value in unbalanced samples, it is more appropriate to use the PRC curve, which compares precision (PREC = precision) against sensitivity (SN = *sensitivity* and REC = *recall*).

The aim of this investigation is to identify the maximum number of employees with a case of illness, despite loss of accuracy. In other words, maximising sensitivity. Therefore, according to the PRC curve, it was decided that the threshold equal to .18 should be selected, which produces a confusion matrix with a sensitivity equal to .82 and accuracy of .34.

On the basis of the above, it was determined that nodes 5, 7, 10 and 11 have a predictive value of 1 (probability that cases of illness exist in 2017 > .18), therefore, the employees categorised into these nodes are the target of this study, as their prediction is that they experienced a case of illness in 2017.

The segmentation obtained from the group of workers is as follows:

- Segment 1 (node 10): workers who experienced some sort of illness in 2016, on a permanent contract and working a full day (probability .693, number of successes = 111, number of errors = 81).

- Segment 2 (node 11): workers who experienced some sort of illness in 2016, on a permanent contract and working a partial day (probability .578, number of successes = 192, number of errors = 85).

- Segment 3 (node 7): workers who experienced some sort of illness in 2016, on a temporary contract (probability .486, number of successes = 72, number of errors = 76).

- Segment 4 (node 5): workers who did not experience any illness in 2016, on a contract term greater than 8 months (probability .213, number of successes = 264, number of errors = 973).

On the other hand, from the second decision tree, identification of those segments of workers who represent a ratio of illness greater than 50% was attempted, taking into account the variables associated with the contract. Firstly, workers were selected who presented with a case of illness in 2017, providing a total of 1290 workers. The ratio of illness

was calculated by dividing the number of sick days due to illness by the number of contracted days in 2017. Subsequently, the sample was split into two groups: a test sample (30%) and a training sample (70%). On this occasion, the target group (workers with a ratio greater than 50%) accounted for 25% of the sample, meaning that, as above, this tree is also an unbalanced sample. Explanatory variables were selected: type of contract, type of work day, contract term and longevity. The results obtained are shown in the decision tree compiled in Figure 5:
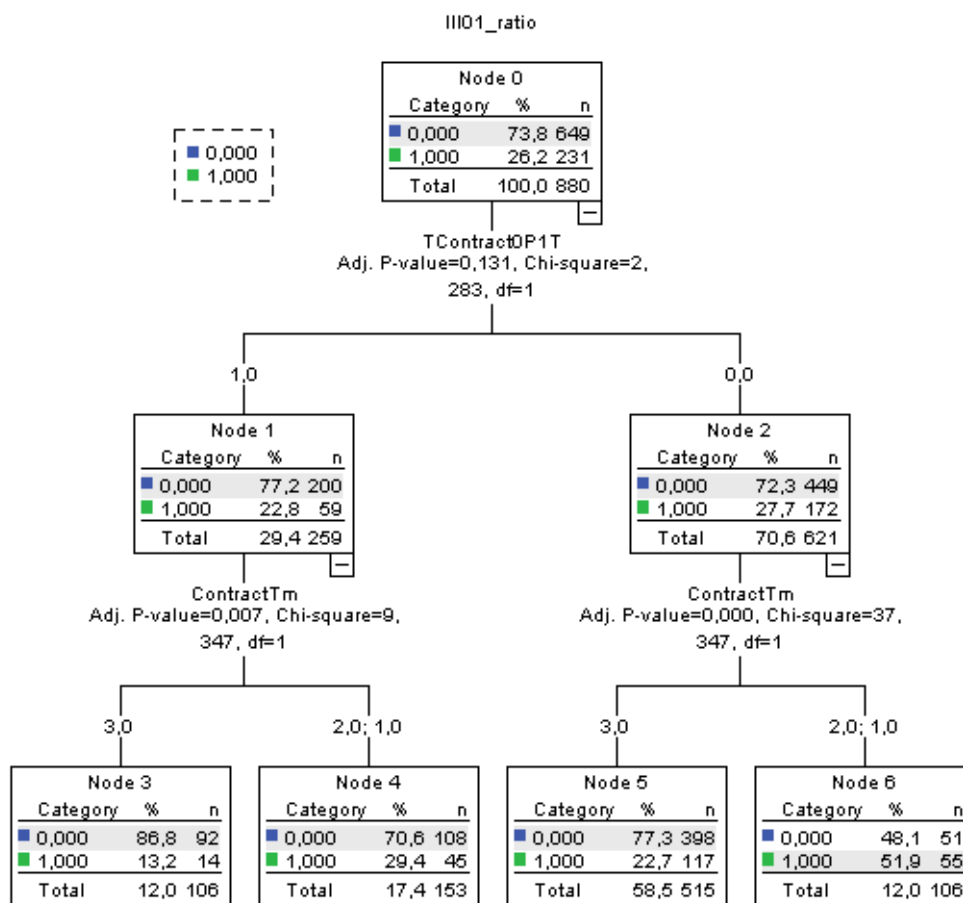

Figure 5. Decision tree for the dependent variable ill01_ratio.

To interpret this tree, the importance indicator is made available, as before, for the weight of every independent variable in the tree, with the variable of contract term being of greatest importance (with 100%), followed by type of contract and longevity (each with 50%). Finally, the type of work day is not of importance in this tree.

The confusion matrix obtained by a threshold value equal to .50 produced an accuracy rate percentage of 74.2% in the training sample and 76.6% in the test sample. This indicates that it is a stable tree (the variation between training sample and test sample is very low), with a very high success rate.

The threshold used in this matrix is .5 and provides a sensitivity equal to .23 and accuracy equal to .52. Being an unbalanced tree, the PRC curve was used and the threshold value equal to .25 was selected, which increased sensitivity equal to .43, losing a certain level of accuracy (= .38).

Finally, the population segments predicted as ill01_ratio greater than .5 were as follows:

- Segment 1 (node 6): workers on a permanent contract and a contract term greater than 8 months (probability .52, number of successes = 55, number of errors = 51).

- Segment 2 (node 4): workers on a temporary contract

and a contract term lower than 8 months (probability .29, number of successes = 45, number of errors = 108).

Discussion

The results obtained reveal the existence of relationships between the recruitment method and workplace absenteeism. In particular, with regard to the first hypothesis, if workers on a temporary contract are most likely to experience illness, the results obtained do not allow for this hypothesis to be tested, since it is evidenced that workers on this type of contract (temporary) show a lesser percentage of cases of common illness than workers on a permanent contract. These results contradict previous different studies, such as that of Virtanen et al. (2005). Nevertheless, they must consider the specific effects of the activity of cleaning and the ageing population, as well as the greater daily workload allocated to those on permanent contracts compared with those on temporary contracts, such as mediating variables of the results, which should be studied through future research in this field.

Furthermore, an assessment was carried out to see if the work day had an influence on absenteeism. In relation to this matter, segments 1 and 2 of the first CHAID tree highlight the fact that the partial work day provides less probability of suffering a case of common illness than the full work day, which can be explained by exposure time, which is one of the most relevant factors in physical health problems. Taking into account the fact that work place absenteeism in the cleaning sector is intimately related to a musculoskeletal aetiology, the result is completely understandable in line with the investigations of Estryn-Béhar and Van Der Heijden (2012), Gallis (2013), Moreno, Moncada, Llorens and Carrasquer (2010), and Vercruyssen and Van de Putte (2013).

Thirdly, the relationships between longevity and the possibility of experiencing illness were examined, evidencing the fact that longevity is one of the variables of greatest importance in the case of illness ratio. These conclusions must be taken with caution, given the peculiarity and specificity of the sample, and the relationship between the type of contract and the median age of the employees.

Finally, the findings made enable the last hypothesis to be verified, because in the analyses carried out, the variable 'cases of common illness in the three reference years' pre-sents the same behaviour. Likewise, in the first CHAID tree, in which the variables cases of common illness in 2015 and 2016 featured, both are of great importance. In our opinion, the aforementioned history of absenteeism provides clear indicators of future absenteeism. The influence on the health of past cases may be one of the justifications, even though the inclusion of individual variables, such as personality and self-reported health, is recommended for use in future studies.

Among the limitations of this study is the lack of control over moonlighting, which is very common in a sector affected by cases of subrogation (such is the cleaning sector) which may affect the total work days of the population assessed. On the other hand, the non-probabilistic type of sample used (for convenience) makes the results neither generalisable nor conclusive. Likewise, the age of the workers, as previous different studies indicate (e.g. Fernández-García, 2013), also appears to influence cases of absenteeism, as is consistent with the ageing of the working population which Spain and Europe are experiencing. It is therefore recommended that this variable be included in subsequent studies.

Finally, although a great number of statistical methods exist for carrying out the task of classification (e.g. methods of logistic regression, discriminant analysis, neural networks, rule induction, Bayesian networks, etc.), decision trees offer clear advantages which allow all possible solutions to be fully proposed. Furthermore, their graphical representation facilitates interpretation (Chien, Wang, & Cheng, 2007). Likewise, Fleiss et al. (1972) put forward the benefits of this technique against others in the generation of diagnoses. This decision and classification technique has been widely used in the area of cognitive psychology, psychometry and the psychology of creativity (e.g., Batchelder, & Riefer, 1999; Chun et al., 2013). Nevertheless, at the same time, decision trees present different limitations, such as the large quantity of data required (available in this study) or the fact that, if costs and profits are attributed to each of the actions, it is possible to apply the maximised expected value as a decision criterion, which requires the decision maker to support the risk of downfall. In this investigation, the decision criteria has been that of identifying the maximum number of employees with a case of illness (with moderate error), without assessing the financial benefits which may be brought, a matter which should be the subject of a study in future investigations in this field.

References

Altunkaynak, B. (2018). A statistical study of occupational accidents in the manufacturing industry in Turkey. *International Journal of Industrial Ergonomics, 66*, 101-109. doi:10.1016/j.ergon.2018.02.012

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57-86.

Benavides, F., & Benach, J. (1999). Types of employment and health: Analysis of the Second European Survey on Working Conditions. *Gaceta Sanitaria, 13*(6), 425- 430. doi:10.1016/S0213-9111(99)71402-0

Benavides, F. G., Benach, J., Diez-Roux, A. V., & Roman, C. (2000). How do types of employment relate to health indicators? Findings from the Second European Survey on Working Conditions. *Journal of Epidemiology and Community Health, 54*(7), 494-501. doi:10.1136/jech.54.7.494

Blasco de Luna, F. J., Guada, J., Sánchez-Robles, B., Blázquez Agudo, E. M, Checa Martín, J.L…, Súarez, A. (2018). *VII Informe Adecco sobre absentismo.* Madrid: Adecco. Retrieved from https://www.adeccoinstitute.es/descarga-el-vii-informe- adecco-sobre-absentismo-laboral/

Boada i Grau, J., De Diego Vallejo, R., Agulló Tomás, E.A., & Mañas Rodríguez, M.M. (2005). Absenteeism from work as consequent of organizational variables. *Psicothema, 17*(2), 212-218.

Bogaerts, Y., De Witte, H., Verlinden, H., & Vermeulen, V. (2017). Job insecurity and presenteeism: An indirect effect through burnout. *Gedrag en Organisatie, 30*(1), 3-24. Retrieved from https://www.researchgate.net/publication/317758911

Campos-Serna, J., Ronda-Pérez, E., Artázcoz, L., & Benavides, F.G. (2012). Desigualdades de género en salud laboral en España [Gender inequalities in occupational health in Spain]. *Gaceta Sanitaria, 26*(4), 343-351. doi:10.1016/j.gaceta.2011.09.025

Casey, T. W., Riseborough, K. M., & Krauss, A. D. (2015). Do you see what i see? Effects of national culture on employees' safety-related perceptions and behavior. *Accident Analysis and Prevention, 78*, 173-184. doi:10.1016/j.aap.2015.03.010

Chien, C., Wang, W., & Cheng, J. (2007). Data mining for yield enhancement in semiconductor manufacturing and an empirical study. *Expert Systems with Applications, 33*(1),192–198. doi:10.1016/j.eswa.2006.04.014

Chun Fu, L., Yu-chu, Y., Yu, H. H., & Ray, I. C. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education, 68*, 199–210. doi:10.1016/j.compedu.2013.05.009

de Menezes, T. A., & do Monte, P. A. (2013). Relationship between the type of contract work and dedication of workers in employment: Evidence for metropolitan regions. *Nova Economia, 23*(1), 51-72.

Estryn-Béhar, M., & Van Der Heijden, B. I. J. M. (2012). Effects of extended work shifts on employee fatigue, health, satisfaction, work/family balance, and patient safety. *Work, 41*(SUPPL.1), 4283-4290. doi:10.3233/WOR-2012-0724-4283

Fernández-García, R. (2013). ¿Por qué los jóvenes se accidentan tanto? [Why do young people get so injured?] *Gestión práctica de riesgos laborales: Integración y desarrollo de la gestión de la prevención, 101*, 42-47.

Fleiss, J. L, Spitzer, R. L, Cohen, J., & Endicott, J. (1972). Three computer diagnosis methods compared. *Archives of General Psychiatry, 27*(5), 643–649. doi:10.1001/archpsyc.1972.01750290057011

Gallis, C. (2013). Increasing productivity and controlling of work fatigue in forest operations by using prescribed active pauses: A selective review. *Croatian Journal of Forest Engineering, 34*(1), 103-112. Retrieved from https://hrcak.srce.hr/116733

García Mainar, I., Green, C. P., & Navarro Paniagua, M. (2018). The effect of permanent employment on absenteeism: Evidence from labor reform in Spain. *ILR Review, 71*(2), 525-549. doi:10.1177/0019793917717226

Hassard, J., Teoh, K. R. H., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of Occupational Health Psychology, 23*(1), 1-17. doi:10.1037/ocp0000069

Jiang, L., & Lavaysse, L. M. (2018). Cognitive and affective job insecurity: A meta- analysis and a primary study. *Journal of Management, 44*(6), 2307-2342. doi:10.1177/0149206318773853

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. A*pplied Statistics, 29*(2), 119–127. Retrieved from https://www.jstor.org/stable/2986296?seq=1

Khashaba, E., El-Helaly, M., El-Gilany, A. H., Motawei, S. M., & Foda, S. (2018). Risk factors for non-fatal occupational injuries among construction workers: A case– control study. *Toxicology and Industrial Health, 34*(2), 83-90. doi:10.1177/0748233717733853

López-Alonso, M., Ibarrondo-Dávila, M. C., & Rubio-Gámez, C. (2015). Analysis the cost of prevention in construction sites. *Informes de la Construccion, 67*(537). doi:10.3989/ic.13.062

López-Goñi, J. J., Fernández-Montalvo, J., & Arteaga, A. (2012). Predictive validity of the EuropAsi: Clinical diagnosis or composite scoring? *Journal of Substance Abuse Treatment, 42*(4), 392-399. Retrieved from https://academica-

e.unavarra.es/bitstream/handle/2454/28358/Predictive_validity.pdf?s equence=3&i sAllowed=y

Luján, L. M. B., Santana, N. R., Alemán, S. T., & Sánchez, J. A. S. (2014). Consumption of alcohol in canary islands: Trends and the most important sociodemographic factors associated (1997-2009}. *Revista Espanola de Nutrición Comunitaria, 20*(2), 57-66. doi:10.7400/RENC.2014.01.2.5012

Martin-Fumadó, C., Martí Amengual, G., Puig Bausili, L., & Arimany-Manso, J. (2014). Temporary disability and its legal implications. *Medicina Clinica, 142*(SUPPL. 2), 37-42. doi:10.1016/S0025-7753(14)70070-3

Martín-Román, A., & Moral, A. (2017). A methodological proposal to evaluate the cost of duration moral hazard in workplace accident insurance. *European Journal of Health Economics, 18*(9), 1181-1198. doi:10.1007/s10198-017-0878-6

Mazzoni, C. F., Campos, M. V. G., Sirqueira, C. A. G., Krause, F., de Kraker, H., Könemann, R., & Douwes, M. (2019). Developing a methodology for a participatory ergonomics evaluation process: Human performance and productivity cycle. *Proceedings of the 20th Congress of the International Ergonomics Association.* IEA 2018. Advances in Intelligent Systems and Computing, vol 824. Springer, Cham. doi:10.1007/978-3-319-96071-5_190

Moreno, N., Moncada, S., Llorens, C., & Carrasquer, P. (2010). Double presence, paid work, and domestic-family work. *New solutions: A journal of environmental and occupational health policy : NS, 20*(4), 511-526. doi:10.2190/NS.20.4.h

Ortega, J. L. (2015). Diferencias y evolución del impacto académico en los perfiles de Google Scholar Citations: una aplicación de árboles de decisión. *Revista Española de Documentación Científica, 38*(4), e102. doi:10.3989/redc.2015.4.1225

Peñalva-Vélez, A., & López-Goñi, J. J. (2014). Competencias ciudadanas en el alumnado de magisterio: la competencia intercultural personal. *Revista Electrónica Interuniversitaria de Formación del Profesorado (REIFOP), 17*(2), 139-153. Retrieved from https://www.redalyc.org/pdf/2170/217031054011.pdf

Restrepo, C., & Salgado, E. (2013). Types of contracts and worker absenteeism in Colombia. *Journal of Business Research, 66*(3), 401-408. doi:10.1016/j.jbusres.2012.04.006

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE, 10*(3), e0118432. doi:10.1371/journal.pone.0118432

Sanz, E., & Ponce de León, A. (2010). Claves en la aplicación del algoritmo Chaid. Un estudio del ocio físico deportivo universitario. *Revista de Psicología d*el *Deporte, 19*(2), 319-333. Retrieved from https://www.redalyc.org/pdf/2351/235116352011.pdf

Thurston, E., & Glendon, A. I. (2018). Association of risk exposure, organizational identification, and empowerment, with safety participation, intention to quit, and absenteeism. *Safety Science, 105*, 212-221. doi:10.1016/j.ssci.2018.02.012

Vercruyssen, A., & Van de Putte, B. (2013). Work-family conflict and stress: indications of the distinctiveness of role combination stress for Belgian working mothers. *Community, Work and Family, 16*(4), 351-371. doi:10.1080/13668803.2013.776515

Virtanen, M., Kivimäki, M., Joensuu, M., Virtanen, P., Elovainio, M., & Vahtera, J. (2005). Temporary employment and health: A review. *International Journal of Epidemiology, 34*(3), 610-622. doi:10.1093/ije/dyi024

Zaballa, E., Martínez, J. M., Duran, X., Albertí, C., Gimeno Ruiz de Porras, D., & Benavides, F.G. (2016). Incidence of sickness absence by type of employment contract: One year follow-up study in Spanish salaried workers. *Archives of Public Health, 74*(1). doi:10.1186/s13690-016-0152-x. Taken from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5037859/