Safety Science 68 (2014) 108–120

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

The effect of alcohol and drug testing at the workplace on individual's occupational accident risk

Paulo H. Marques [a,f], Vasco Jesus [b], Serafín A. Olea [c], Valter Vairinhos [d,e], Celeste Jacinto [f,*]

[a] Universidade Europeia – Laureate International Universities, Lisboa, Portugal
[b] Instituto Superior de Estatística e Gestão de Informação, ISEGI – Universidade Nova de Lisboa, Lisboa, Portugal
[c] Departamento de Ciencias Biomédicas – Universidad de León, León, Spain
[d] CENTEC-IST, Universidade de Lisboa, Portugal
[e] CINAV, Centro de Investigação Naval, Lisboa, Portugal
[f] UNIDEMI, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

ARTICLE INFO

Article history:
Received 16 June 2013
Received in revised form 14 March 2014
Accepted 16 March 2014

Keywords:
Drug testing
Alcohol testing
Prevention
Occupational accidents
Data mining techniques

ABSTRACT

Programmes for testing Alcohol and Drugs (A&D) at the workplace, at random and by surprise, are believed to have a positive impact on safety and to reduce individual's accident risk. Despite this perception, there is limited scientific evidence and poor statistical support of this assumption. This study aims at testing whether there is such a cause-effect relationship between A&D testing and post-accident reduction, and how to quantify it. The methodology applied data-mining techniques together with classical statistics hypothesis testing. It covers a wide range of data concerning accidents, alcohol and drug tests, biographical and occupational records of a large railway transportation company in Portugal, for a period of 5½ years. Results give sound statistical evidence of individual's accident risk decrease after being tested, by quantifying the relations between A&D testing and post-testing accidents. Results also estimate the optimal testing frequency that balances testing costs and accident reduction. Optimum rates of tests per year per worker are in the ranges]0.5–1.0] in white-collars and professions at large, and]0.0–0.5] in operations/technical personnel. The fraction of accident victims that are prevented by the application of optimal frequencies are around 59% for workers onboard trains, 72% for those working near trains, and 85% for white-collars. Testing at the optimal frequency generates net savings of at least 15:1, in onboard personnel. In conclusion, testing for alcohol and drugs at workplace, at random and by surprise, has a statistically significant preventive effect in overall professions, but is stronger within white-collars.

1. Introduction and scope

The abuse of alcohol and illicit drugs can affect performance of workers (Lancet editorial, 2009) and thus constitute an additional hazard, which increases health and safety risks in the workplace. The application of A&D testing practices, especially at random, might constitute a promising strategy to identify and discourage such unwanted behaviours. The compulsory testing of these so-called psychoactive substances is expected to play a deterrent role on the abuse behaviour – which is beneficial in many ways to the individual, the employer and the society at large. This perception becomes even stronger in some specific activities, such as the transportation sector, including for instance, aviation (Li et al., 2005), maritime (O'Connor and O'Connor, 2006), or railways (UIC, 2008), in which the erroneous action of one worker can endanger the safety of thousands of persons.

The reason to embrace this study on the possible relationships between A&D testing programmes and the level of safety – measured through accident rates – is twofold: (1) it covers the workforce of a transportation company where the concern with occupational health and public safety is understandably very high and is a strategic goal and (2) the unique opportunity of having access to a comprehensive database that includes detailed information on A&D testing for all employees. This is a rare opportunity, since in various countries and many companies, A&D testing faces restrictions regarding data protection law and individual rights. However, in this particular case, the testing programme implemented has legal support, based on the recognition that the collective safety and health outweighs the individual rights to privacy

* Corresponding author. Address: UNIDEMI, Faculty of Science and Technology, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal. Tel.: +351 212 948 567; fax: +351 212 948 531.
E-mail addresses: paulo.marques@europeia.pt (P.H. Marques), vjesus@isegi.unl.pt (V. Jesus), sabao@unileon.es (S.A. Olea), valter.vairinhos@sapo.pt (V. Vairinhos), mcj@fct.unl.pt (C. Jacinto).

http://dx.doi.org/10.1016/j.ssci.2014.03.007
0925-7535/

(PCC, 1995; PP, 2009; PSCJ, 1998). Moreover, the company uses the programme not only to discourage abuses and unsafe behaviours, but also offers rehabilitation programmes to its employees, free of charge.

Improving health and safety in the workplace depends on many variables and circumstances (Lund and Aarø, 2004), thus it is difficult to demonstrate that A&D testing has indeed a preventive impact and that it may actually contribute to reduce accident rates. Even more difficult is to give statistical evidence capable of quantifying such association.

The main goal of this study is to provide sound evidence of the referred association and measure the differences in accident rates between tested and untested employees. Furthermore, the authors intend to establish the optimal frequency of testing, *i.e.*, the annual frequency of testing that is more effective in preventing accidents, as well as to provide a rough, but measurable, estimate on the return on investment.

The remaining sections of the paper provide a literature review on the subject and derived research hypotheses, followed by a summary of the methodology used and statistical approach applied. The final sections discuss the most relevant results and conclusions.

2. Literature review and derived research hypotheses

All psychoactive substances have, to a higher or lesser extent, a dysfunctional effect on work capability (Kauert, 2008; Schuckjt, 2009). A worker under the effect of a psychoactive substance becomes a hazard to him/herself and to others around. Even if the deviant behaviour is not readily visible or detectable, this person has a reduced ability to identify and control hazards – this incapacity suggests the need to address A&D abuse in the workplace (Baer and Hess, 2008).

The abuse of psychoactive substances is associated with many adverse consequences to health (Chipman et al., 2009; Degenhardt and Hall, 2012; Schuckjt, 2009) and consequently to safety at work, such as violence, accidents (Li and Bai, 2008), injuries (Trent, 1991) and absenteeism. An increasing number of countries and companies are coping with this risk (Strang et al., 2012) by the application of means of control, also in the scope of Occupational Safety and Health (OSH), such as screening for psychoactive substances in employees.

On the turn of the millennium, Kraus (2001) had carried out a systematic review of 740 publications dealing with the topics of testing for A&D in the workplace, of which only 6 presented some kind of quantification of the effect on accident rates, and all others were devoted to rather qualitative aspects, such as, philosophical, social, moral or legal arguments. Management issues and test protocols were also among the findings of these publications. After a deeper scrutiny of the 6 relevant studies, Kraus was unable to either accept or refute the hypothesis that A&D testing would lead to a reduction of accidents, because the studies reviewed suffered from several methodological shortfalls (*e.g.*: lack of a control group, or insufficient sample-size, or absence of inferential statistics). For the same reason, Kraus concluded that there was limited and biased evidence that random tests by surprise would have a stronger preventive affect, when compared with scheduled and announced tests.

To balance the lack of international standards concerning policy and practice of workplace A&D testing, the International Labour Organisation (ILO, 2003) made recommendations towards new research work aimed at evaluating *«the relationship between the consumption of alcohol and drugs and both the safety and productivity at work»* and it also suggested estimating the *«costs and benefits of screening for A&D»*.

A few economic studies on the prevention of A&D abuse at work were conducted in US, with abundance of studied cases and advanced statistics (Livingston, 1975; Ozminkowski et al., 2003; Rummel et al., 2004; Wickizer et al., 2004; Miller et al., 2007; Mehay and Webb, 2007). These revealed the relevance of balancing the costs of preventing A&D abuse and the correspondent financial return on investment.

The research by Ozminkowski et al. (2003), concerning 1791 manufacturing workers, concluded that the relation between frequency of drug testing and injuries' medical expenses was statistically significant, and had a U-shape. These findings led the authors to conclude that medical expenses resulting from accidents can be minimised if workers are subjected to drug testing at an average annual frequency of 1.68 times per worker.

In the study by Wickizer et al. (2004), 14,500 employees of 261 companies with programs for drug prevention were compared to 650,000 employees in 20,000 companies with no such programs. A statistically significant association has been demonstrated, between drug prevention programs and lower rates of occupational accidents, on services, construction and manufacturing sectors. There were small unspecified net savings associated with these programs, more so in the construction industry.

The research performed by Miller et al. (2007), covering employees of a large USA carrier, showed a statistically significant association between the A&D abuse prevention program and lower rates of occupational accidents. It was also found a benefit-cost ratio of 26 $US saved in accident reduction for each 1 $US spent in the preventive program.

On the other hand, the study by Mehay and Webb (2007), about a zero tolerance policy applied on drug prevention programs of US Navy, concluded that net benefits were negative for most plausible values of the key parameters – the deterrence effect, replacement cost, and productivity losses due to drug use.

The literature tends to agree that the number of A&D abusers detected over time decreases with continued application of tests (Taggart, 1989; French et al., 2004; Miller et al., 2007; Wenzek and Ricordel, 2008). The deterrent effect of workplace testing is usually attributed to the inhibition resulting from individual's perception of being held liable in case their state of abuse is analytically confirmed by the tests.

However, the question of the deterrent effect being or not sustainable over time, has been subjected to controversy. A review made by Cashman et al. (2009) looked for relationships between A&D testing and the supposedly reduction of accident injuries, on professional drivers of motorised vehicles. The search produced some 6000 hits, but only 19 publications received further attention. Among these, only 2 complied with the criteria established for the study, in terms of data and quality. Both covered time series of test trials in the US. The conclusions highlighted some evidence of short-time effect, but were unable to demonstrate sustainability of prevention on the long-term and the respective authors argued for the need of more research.

Within the transportation sector, in particular, there is a large international agreement for the need to control and manage the risks of working under the effect of alcohol and/or drugs. Consequently, the OECD International Transport Forum (OECD, 2010) is considering a number of legal measures to this purpose. Likewise, the Health and Safety Group of the International Union of Railways (UIC, 2008) has deemed important to implement A&D testing programmes for railway transportation workers in all activities that interact with traffic safety. Such strategy had previously been recommended by the US National Institute of Drug Abuse (Gust and Walsh, 1989; Hanson, 1993; Zwerling, 1993).

In the USA, drug screening programmes are mandatory by Presidential order (The President, 1986) for all federal workplaces. According to Miller et al. (2007, p. 565), *«the program was*

strengthened by federally mandated random drug and alcohol testing (implemented, respectively, in 1990 and 1994)».

On the other hand, very few publications report application of tests in Europe, even within sectors such as the chemical industry, where examples come from DEGUSSA (Breitstadt, 2008), EVONIK INDUSTRIES (Schiffhauer and Breitstadt, 2008) and ROCHE (Seiffert, 2008), or from the railway transportation sector, such as the French SNCF (Wenzek and Ricordel, 2008), the British NR – Network Rail (2008) or the Portuguese CP (Marques, 2009; Marques et al., 2011).

Recent studies carried out in the transportation sector, namely within road safety (Brady et al., 2009; Snowden et al., 2007) and within railways safety (Spicer and Miller, 2005; Miller et al., 2007), suggest a sustained effect of A&D testing programmes on accident reduction. Snowden et al. (2007) examined the impact of random alcohol testing, implemented in the USA since 1994, on the likelihood of large truck drivers being involved in fatal accidents under the effect of alcohol. According to these authors, controlling for the general declining trend in alcohol-involved drivers in fatal crashes showed that random alcohol testing was correlated with a 14.5% reduction in alcohol-related accidents among large truck drivers. In the same field, a study by Brady et al. (2009) provided evidence that implementation of the mandatory alcohol testing programs was associated with a reduction of 23% in alcohol involvement in fatal motor carrier crashes. Spicer and Miller (2005) studied the impact, on a large USA carrier, of a workplace peer-focused substance abuse prevention, and early intervention program, and came to the conclusion that it reduced workplace injuries. Their evidence also suggested that random drug and alcohol testing might have further reduced this risk of injury. In a later publication (Miller et al., 2007), the same authors confirmed that a peer-based program and drug testing was associated with an approximate one-third reduction in injury rate, while alcohol testing reduced the injury rate by one sixth.

It is worth mentioning that even in the organisations already applying A&D tests, many of them do it solely within the scope of Occupational Medicine exams, *i.e.*, meaning that the worker knows in advance when the test sample will be collected. This allows the abusers to learn, in time, how to deceive the system, either by practicing abstinence before the exam, or by delivering biological samples which are not their own. This shows the pertinence of studying cases of random tests, applied directly in the workplace by surprise – to enable a more reliable and realistic evaluation of the foreseen impacts.

Notwithstanding the relevance of previous studies and the great interest of their findings, it is noted that they mostly searched for associations between A&D prevention programs and occupational accidents, by aggregate results of the organisations studied, regardless of the temporal order concerning individual's subjection to preventive measures and the respective effect on occupational accidents. Some cautiousness is due, when discussing preventive effect of tests on accidents, without being able to distinguish between tests prior and post-accidents for each studied individual. Given that any means of risk control applied to someone can only influence events later on, some methodological limitations became evident. Therefore, there is limited empirical evidence of testing for A&D at the workplace being sufficiently deterrent of its abuse, as well as to what extent it may have an effect on individual's occupational accident risk.

Another issue which has not been properly answered yet deals with the testing frequency. In the rare publications reporting frequency of testing, it still remains to be confirmed the existence of a specific frequency that could be more preventive.

To fill in these gaps, this study raised two hypotheses for research:

H1. The frequency of alcohol and drug testing is negatively associated with the incidence rate of accidents occurred after the tests.

H2. Assuming that H1 is not rejected, there is an optimal frequency of tests and post-accidents that represents the most efficient frequency, beyond which, increasing the number of annual tests will result in marginal variation of accidents.

3. Methodological approach

This research work used data from October 2003 to March 2009 inclusive, covering a continuous period of 5½ years. There was a first pilot study (Marques, 2009) covering 2003–2007 data – essentially qualitative and general in scope – to explore the potential effects of several OSH interventions implemented since 2003 by the company in question. This first study suggested that the A&D testing by surprise had a positive impact in the reduction of accidents – however, such conclusion was only apparent, based on qualitative analysis and it could not be demonstrated with statistical evidence.

This fact led to the current work, involving more researchers and extending both the data coverage and the overall methodology applied. In this new study, the objective was much more specific (*c.f.* Introduction), in which the quantitative focus superimposed the qualitative analysis and it followed both descriptive and also explanatory statistical models (Sampieri et al., 2006).

From the beginning, the company's accident rates were known – whether being or not directly related to alcohol or drug intake. The historical records have hardly ever established a clear link between the accident and previous A&D intake. On the other hand, in the rare cases of substance abuse detected by random tests (roughly 1 out of 1000 tests), employees had to face disciplinary consequences and received free medical counselling and treatment. Thus, in these circumstances the authors realised that this study design could not rely solely on the very few known cases of A&D abuse and the accidents they have caused – such scarce data would not assure a statistically meaningful study. Instead, the authors decided to contrast the odds of occupational accident risk between workers with different test rates prior to accidents (both exclusively work-related).

In fact, preventive testing was already in place when the analysis began. No experimental design was in place, neither any preconception. This was not an experiment, but an observational study during a period when complete data were available, namely, accidents, alcohol and drug tests, biographical and occupational records.

3.1. Population and variables

The original dataset covered all 5407 employees of the company, including detailed records of the 31,123 A&D tests and 1589 work accidents, occurred during the 5½ years above mentioned. Employees were randomly selected by computer for A&D testing and the application of tests was unannounced. Tests application had an overall frequency of around four hundred per month. A high and unpredictable variability of individual testing came from computer arbitrary decision of "when", "where" and "who" was going to be targeted for testing, combined with the presence or absence of the target employees due to shift work and moving workplaces onboard rolling stock. Thus, despite everybody being aware of the possibility of being compulsorily tested while working, there were a number of employees who, by chance, were never subjected to any testing. In contrast, other individuals were tested several times, with different testing

frequencies and different combinations – either only alcohol, or A&D tests simultaneously. Confidentiality and anonymity of the employees were assured according to national regulations regarding A&D testing at work.

For each individual employee, the database included more than 30 variables, recording personal and occupational information, as well as safety and health data. Examples of variables are, for instance: employee ID; Date of admission in company; Sex; Date of birth; Marital status; Underage dependents (yes/no); Academic qualifications; Place of residence (city council); Company Business Unit; Shift work rotation (yes/no); Occupational category; Date of A&D test; Type of test (alcohol/drug); Weekday of test; Time of test; Counter-proof of test (yes/no); Medical fitness for work; Date of accident; Type of accident; Days lost from accident.

The term "accident" refers solely to occupational accidents and covered all cases reported during the studied period, regardless of their seriousness.

To ensure harmonised exposure and reduce biases from uncontrolled variables, the study included only employees who had been working with the company over the entire period of study – *i.e.*, all those who had either joined or left during such period were excluded from the dataset. Consequently, within the scope of this study, the entire workforce became represented by a sub-population of 3801 employees, always present during these 5½ years and a total of 29,916 records of accidents, tests, or the absence of either one or another. It should be emphasised that this restricting criterion enabled to improve homogeneity of the studied groups. With this choice, every employee included in the study panel was likely to be under the influence of the same unknown (and uncontrolled) variables. Thus, the 30 variables accounted for characterise the main difference among these employees.

Since the main objective was to verify if A&D random testing actually has an individual preventive effect, by reducing post-testing accidents (and quantifying such an association), the authors tried, as far as possible, to treat and compare data belonging to homogeneous groups of employees, *i.e.*, occupational groups performing similar tasks and exposed to the same pattern of occupational risks. This means that, within each occupational group, the experimental stimulus of being (or not) tested for A&D constituted a relevant difference. The portion untested before any accident, which emerged by chance, became the control group within each occupational risks group of this observational study.

The company had around sixty different professions exposed to a diversity of occupational risks – however, the pilot study (Marques, 2009) showed that it was possible and credible to classify them into three main categories of risk (with identical generic risks), each of which sub-divided into sub-categories (for more specific risks). Table 1 categorises these new variables (Occupational risk group and Occupational risk sub-group).

Given the purpose of the work, a few other secondary variables (new variables) were also established from the original data, namely:

- "Age", defined as numeric, giving the age of each person at the beginning of the study, in October 2003.
- "Years of work in the company", defined as numeric, giving a measure of tenure, in October 2003.
- "Occupational risk group", defined as categorical, having non-ordered values 1, 2, 3, and identifying different groups with similar occupational general risk patterns.
- "Occupational risk sub-group", defined as categorical, having non-ordered values 1a, 1b, 1c, 1d, 2a, 2b, 3a, 3b, 3c, and identifying different sub-groups with similar occupational specific risk patterns.
- "Subjection to tests before any accidents", hereafter simply called "subjection to tests". This variable was defined as categorical, having only two values for each worker (tested or untested) according to being or not subjected to testing prior to any accident. It was determined by the existence or absence, since the beginning of the study, of a test record prior to an accident event. Either being or not a victim of accident before the end of the studied period, this variable expressed the worker's previous subjection to testing.
- "Annual test frequency before any accidents", hereafter simply called "annual test frequency", defined as numeric, giving the average annual frequency of testing, for each worker, before any accident event. It was determined by the number of test records prior to accidents, divided by the period since the beginning of the study until the first accident event or, in its absence, until the end of the studied period. This process transformed the previous variable "subjection to tests before any accidents" into a variable insensitive to time. Whatever this period might have been, all workers were compared on the same time basis of test frequency, *i.e.*, tests per year (annualised variable).
- "Victim of accident after ***n*** tests", defined as categorical, having only two values for each worker (Yes or No), according to being or not a victim of accident during the period after subjection to any number of tests. It was determined by the existence or not of a first accident record since the last test until the final day of the studied period. Either tested or untested, this variable expressed the posterior occurrence or absence of accident.

The control group emerged naturally from the randomness of testing. In the last three variables, this reference group was characterised by: (a) "subjection to tests" = "untested", (b) "annual test frequency" = "0" and (c) "***n*** tests" = "0". In short, individuals of the control group were never subjected to the testing stimulus before suffering any accident.

Table 1
Frequencies of employees ever-present since 01/10/2003 through 31/03/2009, by group and sub-group of occupational categories with similar risk patterns.

Occupational risk group	Occupational risk sub-group (specific risks)	Absolute frequency (number of employees)	Relative frequency (%)	Valid percentage	Cumulative percentage
1 – Work onboard trains (moving workplaces)	1*a* – Driver	1104	29.0	29.0	29.0
	1*b* – Driver's assistant	162	4.3	4.3	33.3
	1*c* – Ticket Inspector	584	15.4	15.4	48.7
	1*d* – Chief Driver	50	1.3	1.3	50.0
2 – Work near or around trains (courtyard areas)	2*a* – Manoeuvres	155	4.1	4.1	54.1
	2*b* – Maintenance	163	4.3	4.3	58.4
3 – Work away from trains (white collars)	3*a* – Railway Station	605	15.9	15.9	74.3
	3*b* – Office work	533	14.0	14.0	88.3
	3*c* – Others, without similar risks	445	11.7	11.7	100.0
	Total	3801	100.0	100.0	

Finally, the dependent variable "Victim of accident after ***n*** tests" was also transformed into an annualised variable, as follows:

- "Incidence of accident victims after ***n*** tests", defined as numerical, having any number of workers who experienced accidents, among an homogeneous group of 1000 workers per year, with the same test frequency class. This variable was determined, for each frequency class, by the total number of accident victims posterior to testing, among 1000 workers, divided by the total period. This transformation made the variable insensitive to time. Whatever the individual period for accident occurrence might have been, all workers were compared on the same basis of annual incidence rate, *i.e.,* per thousand victims of accident among workers.
- Calculation formula was the following:

$$\text{Incidence} = (\text{accident victims} \times 1000)/(\text{total workers} \times 5.5\ \text{years}) \quad (1)$$

The creation of the above secondary variables was necessary to search for significant differences of accident occurrence between tested and untested employees. It was taken into account the subjection of each person to tests exclusively before suffering any accident. In brief, these variables allowed measuring the actual impact of tests on subsequent individual accident rates.

3.2. Statistical tools and metrics

Data analysis was made using Delphi, SPSS and Excel. Afterwards, data mining techniques were applied to identify associations between variables.

To study the complex structure of interactions between all these variables, the authors selected the output variable "Victim of accident after ***n*** tests" – consequently, all others were considered potential explanatory variables with covariance between them. For this task, the methodology applied was the CHAID (Chi-square Automatic Interaction Detector) algorithm incorporated in SPSS. As the name suggests, this is an automatic detector of interactions, based on the Chi-Square test of independence (Hand et al., 2001; Rokach and Maimom, 2008). This statistical tool produces a tree-type output, in which the most significant explanatory variable appears on the top of the tree, immediately under the output variable. Likewise, the remaining branches follow a descending hierarchy – displaying the other variables from the most significant association to the least, *i.e.*, by decreasing order of Chi-square association. The Chi-square value of each pair and the corresponding *p*-value are also plotted in the tree diagram. Each tree branch expresses the combined events of different variables leading to the output variable results.

Additionally, for a better characterisation of the strength of these associations, the authors used the coefficient Cramer *V*. This coefficient varies from 0 to 1 and it is not sensible to scale factors. Cramer *V* measures the association strength in terms of the independent variable explanatory value on the behaviour of the output variable (Acock and Stavig, 1979; Healey, 2010; Le Roy, 2012). Cramer *V* is used for nominal (categorical) variables and, despite its range values [0–1], the interpretation of strength differs from other metrics applicable to numeric variables. It can be interpreted with the following guidelines:

- If Cramer $V = 0.0$, there is no association – thus, the independent variable allows no prediction of the output variable.
- On the opposite side, if Cramer $V = 1.0$, there is a perfect association – the two variables are equal, *i.e.*, they measure the same concept and one can perfectly predict the output variable by knowing the independent variable.
- If $0.0 <$ Cramer $V < 0.10$, although the variables are associated, the strength is very weak – it has no acceptable usefulness.
- If $0.10 \leqslant$ Cramer $V \leqslant 0.19$, although the variables are associated, the strength is weak – it already has some acceptable usefulness (*i.e.*, effect).
- If $0.20 \leqslant$ Cramer $V \leqslant 0.29$, the association strength is moderate – and it has an acceptable usefulness (*i.e.*, effect).
- If $0.30 \leqslant$ Cramer $V < 1.0$, the association strength is strong – and it has an even more acceptable usefulness (*i.e.*, effect).

Furthermore, to reinforce the robustness of the findings, this study has also established the different probabilities of having an accident in case of not being tested against the case of being tested. This was done through the calculation of "1/odds ratio", which is also a metric of the association strength (Panik, 2005).

Finally, for the purpose of gross estimate of the proportion of accident victims prevented by subjection to a certain frequency of testing, a Prevented Fraction was computed as follows:

$$PF = [(AV_{F=0} - AV_{F=i})/AV_{F=0}] \times 100 \quad (2)$$

where $AV_{\mathbf{F=0}}$ is the percentage of accident victims among workers with zero tests per year (not-tested), and $AV_{\mathbf{F=i}}$ is the percentage of accident victims among workers tested at a certain frequency of *i* tests per year.

3.3. Materials and financial resources

The relevant items under this subtitle were as follows.

Materials: breathalysers (reusable breath alcohol-metres with disposable mouth adapters); non-instrumented drug test devices (disposable kits for detection of illicit drugs in samples of urine or oral fluid).
Financial: average 6.50 € per test of alcohol and 45.00 € for drugs; these values covered all expenses, including application and safe disposal of used materials. Considering the proportion of only alcohol and combined A&D tests, the average value was computed 13.08 € per application.

Finally, a simplified cost-benefit analysis, allowed to estimate the savings (in Euros) with overtime work (i.e., saved costs), by each Euro spent with tests. The net savings value, resulting from the reduction of overtime work for replacement of injured employees in accidents occurred after ***n*** tests, gave an indication of the return on investment with the tests. This economic appraisal is part of a methodology thoroughly explained by the American Industrial Hygienists Association (AIHA, 2008). This approach, called *«Industrial Hygiene Value Strategy»*, highlights that any economic evaluation of OSH interventions needs to compare the costs and the consequences of an action, as this is required to gain valid information on efficiency. It combines both quantitative and qualitative elements. The calculation process of the quantitative component is relatively complex and ample, but the global reasoning of it can be summarised in a simple way, by means of three equations to derive the monetary value of the intervention:

$$\text{Pre-intervention costs } (A) - \text{Post-intervention costs } (C) = \text{Reduced costs } (D) \quad (3)$$

$$\text{Reduced costs } (D) - \text{Intervention costs } (B) = \text{Net Savings } (E) \quad (4)$$

$$\text{Net Savings } (E) + \text{New Revenue} + \text{Other benefits } (O) = \text{Value} \quad (5)$$

In the current study, reduced costs (D) were computed as the reduction of overtime costs, whilst intervention costs (B) were the above mentioned costs of testing programme.

4. Results and discussion

4.1. Statistical results

As mentioned previously in Section 3.2, this study tested statistical associations between the output variable "Victim of accident after ***n*** tests" and all 30 other potentially explanatory variables. The study compared accident proportions between (as far as possible) homogeneous groups of employees, *i.e.*, occupational groups performing similar tasks and exposed to the same pattern of occupational risks during the same period. This assured that, within each occupational group, the stimulus of being or not being tested accounted for a relevant difference.

The analysis produced CHAID classification trees for the whole sub-population (N = 3801 employees) and for each one of the three occupational risk groups of Table 1. To illustrate this kind of analysis, Fig. 1 shows the classification tree for the whole sub-population, in which the testing for A&D is simply expressed by "tested" and "untested", concerning the variable "Subjection to tests". In other words, this tree distinguishes between individuals tested and untested before suffering any accidents. Of all variables included, the latter revealed to be the most explanatory with relation to the output variable "Victim of accident after ***n*** tests" – stronger associations with p-values < 10^{-3}.

Fig. 1 gives an overall vision of the entire sub-population. The top branch (most significant association) shows that 47.0% of the untested employees (n = 727) had accidents, as opposed to a smaller proportion of 19.4% among the tested employees (n = 3074), who had accidents after testing. Such difference between 19.4% and 47.0% is statistically significant – and this fact was further corroborated by tests of hypotheses (mean comparisons and analysis of

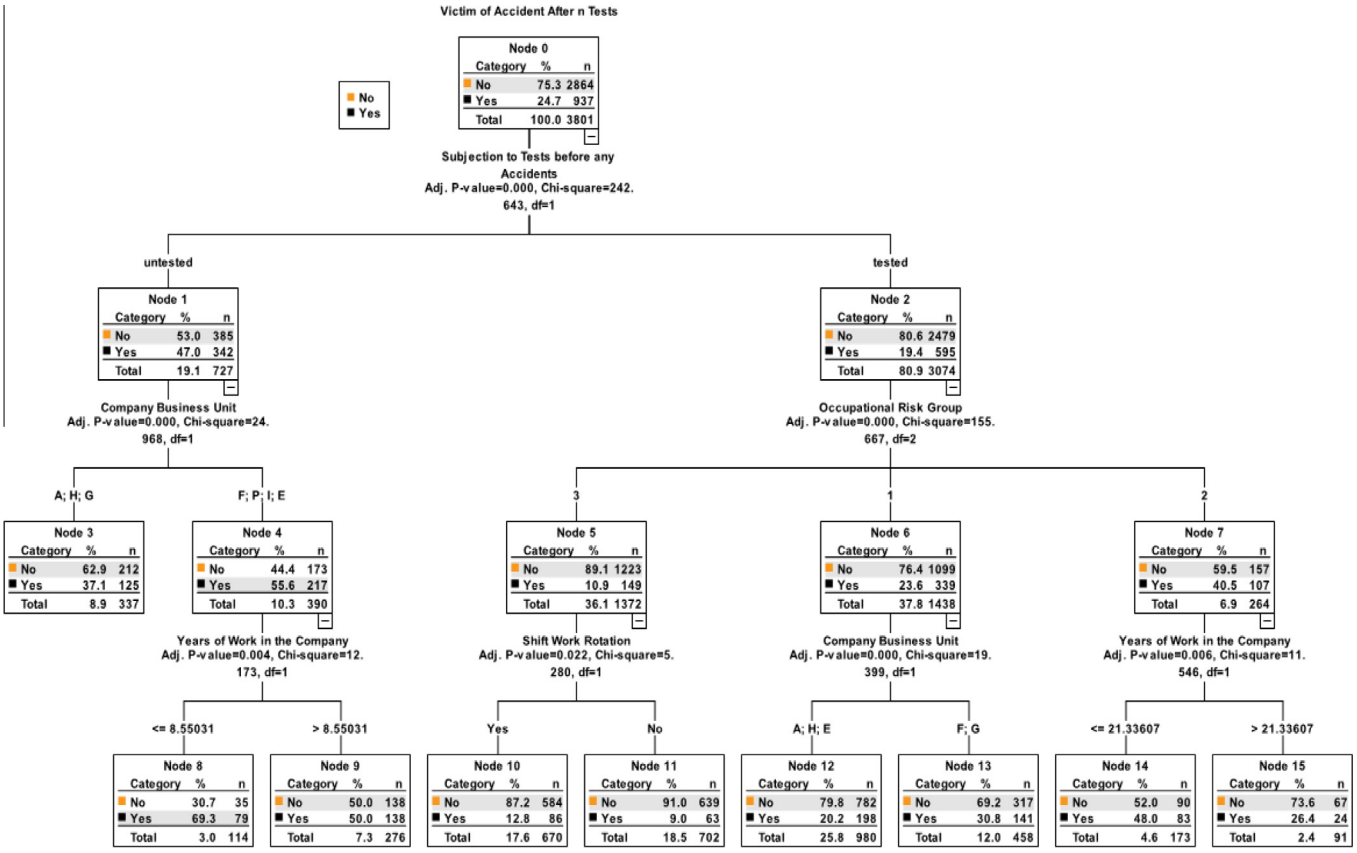

Fig. 1. Associations between the output variable "Victim of accident after ***n*** tests" and explanatory variables: "Subjection to tests before any accidents", "Company business unit", "Occupational risk group", "Years of work in the company" and "Shift work rotation", for the whole sub-population (N = 3801).

Table 2
Statistical evidence of the significant association between "Victim of accident after ***n*** tests" and explanatory variable "Subjection to tests before any accidents".

Occupational group	Association test "Victim of accident after n tests" versus "Subjection to tests before any accidents"		Strength of association "Victim of accident after n tests" versus "Subjection to tests before any accidents"	
	Pearson Chi-square	p-value	Cramer V	p-value
Whole sub-population N = 3801	242.64	<10^{-3}	0.253	<10^{-3}
Group 1 (on board trains) N_1 = 1900	76.96	<10^{-3}	0.201	<10^{-3}
Group 2[a] (near or around trains) N_2 = 318	6.39	0.011	0.142	0.011
Group 3 (away from trains) N_3 = 1583	196.82	<10^{-3}	0.353	<10^{-3}

[a] In this occupational risk group, the variable "Subjection to tests before any accidents" was not the most explanatory; in this particular case, the most significant association was with "Occupational risk sub-group" (Chi-square = 17.03; p-value < 10^{-3}).

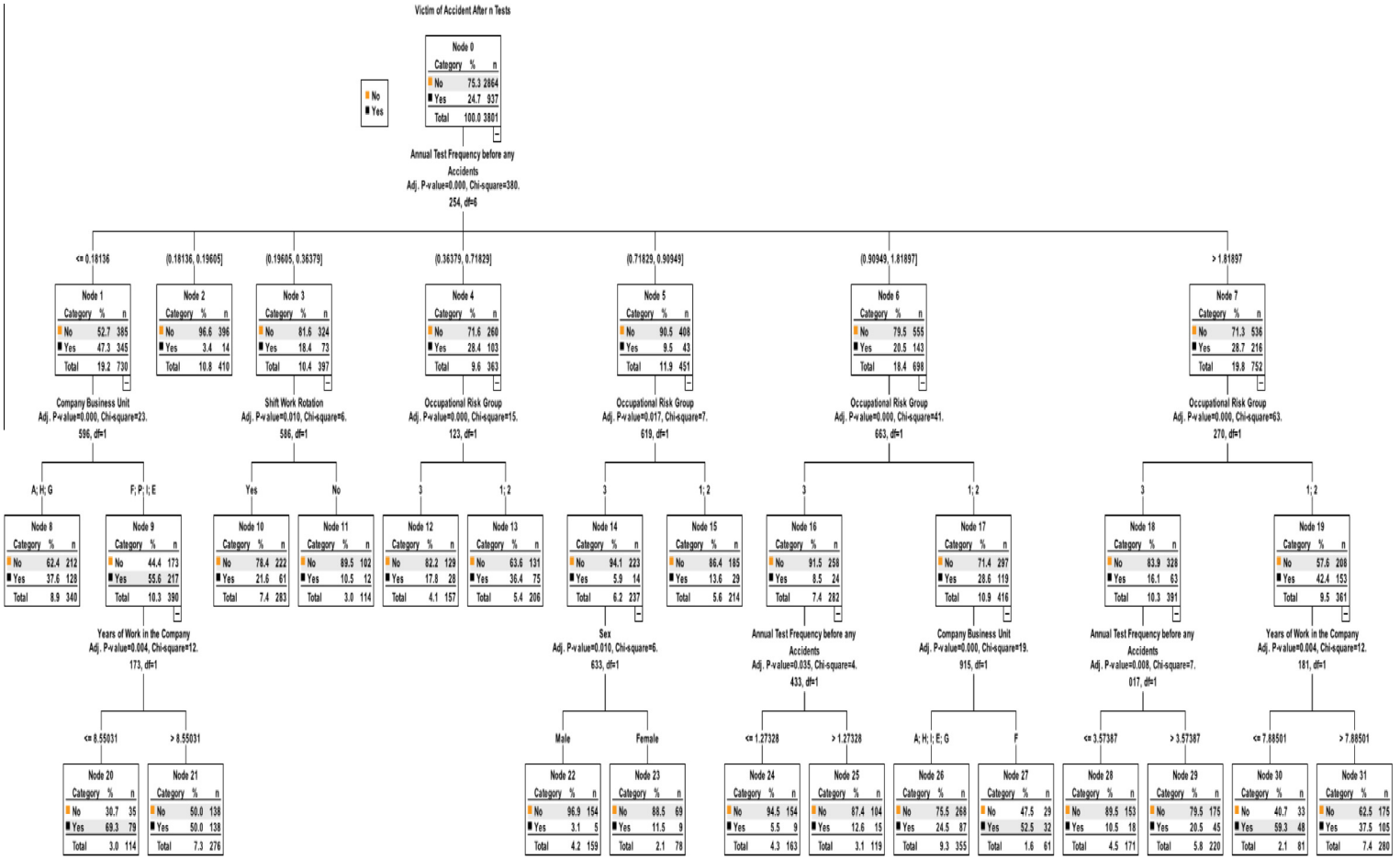

Fig. 2a. Associations between the output variable "Victim of accident after ***n*** tests" and explanatory variables: "Annual test frequency before any accidents", "Company business unit", "Shift work rotation", "Occupational risk group", "Years of work in the company" and "Sex", for the whole sub-population ($N = 3801$).

variance), as well as by application of Mann-Whitney and Kolmogorov-Smirnov tests, all of which with a p-value < 10^{-3}.

The same CHAID analyses and other statistical tests were repeated for each of the three occupational risk groups, this time revealing the idiosyncrasies of each one. The main findings, considering solely the dependency between the response variable "Victim of accident after ***n*** tests" and "Subjection to tests" are summarised in Table 2, which also includes Cramer's V results.

In short, the type of CHAID diagrams exemplified in Fig. 1 and the statistics in Table 2 show that, in all the occupational risk groups studied, one can reject the null hypothesis. Therefore, it has been demonstrated, with a significance level of 1%, that the two variables "Victim of accident after ***n*** tests" and "Subjection to tests" are statistically associated – with a small difference in group 2, for which the significance level was 5%, essentially due the small size of this group (N = 318).

Additionally, the Cramer V values allowed to conclude that the strengths of such associations are considered moderate in the whole sub-population (0.253) and group 1 (0.201), weak in group 2 (0.142), and strong in group 3 (0.353) (Healey, 2010; Le Roy, 2012; Murteira, 1990).

Finally, the calculation of "1/odds ratio" (Panik, 2005) allowed estimating the higher probability of accident for untested employees. The results indicate how much more probable is having an accident if untested compared to tested, as being:

- 3.7 times more, in the sub-population;
- 2.6 times more, in group 1;
- 2.1 times more, in group 2;
- 7.8 times more, in group 3.

Noteworthy saying that "1/odds ratio" also quantifies the strength of the association, when comparing different groups of the same population, as is the case here.

Once the above relationship had been established and quantified (in Fig. 1, Table 2 and "1/odds ratio"), this suggested that random testing at workplace by surprise has a statistically significant effect on reducing rates of accidents at work. However, the tested workers were, overall, less time exposed to occupational risks than the 5½ years period for the untested ones, with a correspondent expected smaller portion of victims of accidents. This was precisely the reason why the authors have not interpreted the relationship between "Victim of accident after ***n*** tests" and explanatory variable "Subjection to tests" in a conclusive way. Instead, the decision was to process data further in order to render the testing issue insensitive to time of exposure to occupational risks, *i.e.*, by using the transformed variable "Annual test frequency", as shown in Figs. 2a–d.

When the initial input variable "Subjection to tests" was replaced with the time insensitive variable "Annual test frequency", the same CHAID algorithm showed (Fig. 2a) that this last variable was the first one next to the top of the tree. Again, the testing issue, either expressed only as "tested" and "untested", or expressed in annual frequency, was systematically the most explanatory.

Overall, the patterns found in the CHAID diagrams and the other statistical tests were quite similar. The analyses of the three occupational risk groups corroborated the relevant findings for the whole sub-population, with an exception in group 2 (working near or around the trains), for which the best explanatory variable (most significant in CHAID diagram) was "Occupational risk subgroup" (Chi-square = 17.035; p-value < 10^{-3}). However, a deeper scrutiny to group 2 allowed realising that, after all, being subjected to tests was also associated with the response variable, although

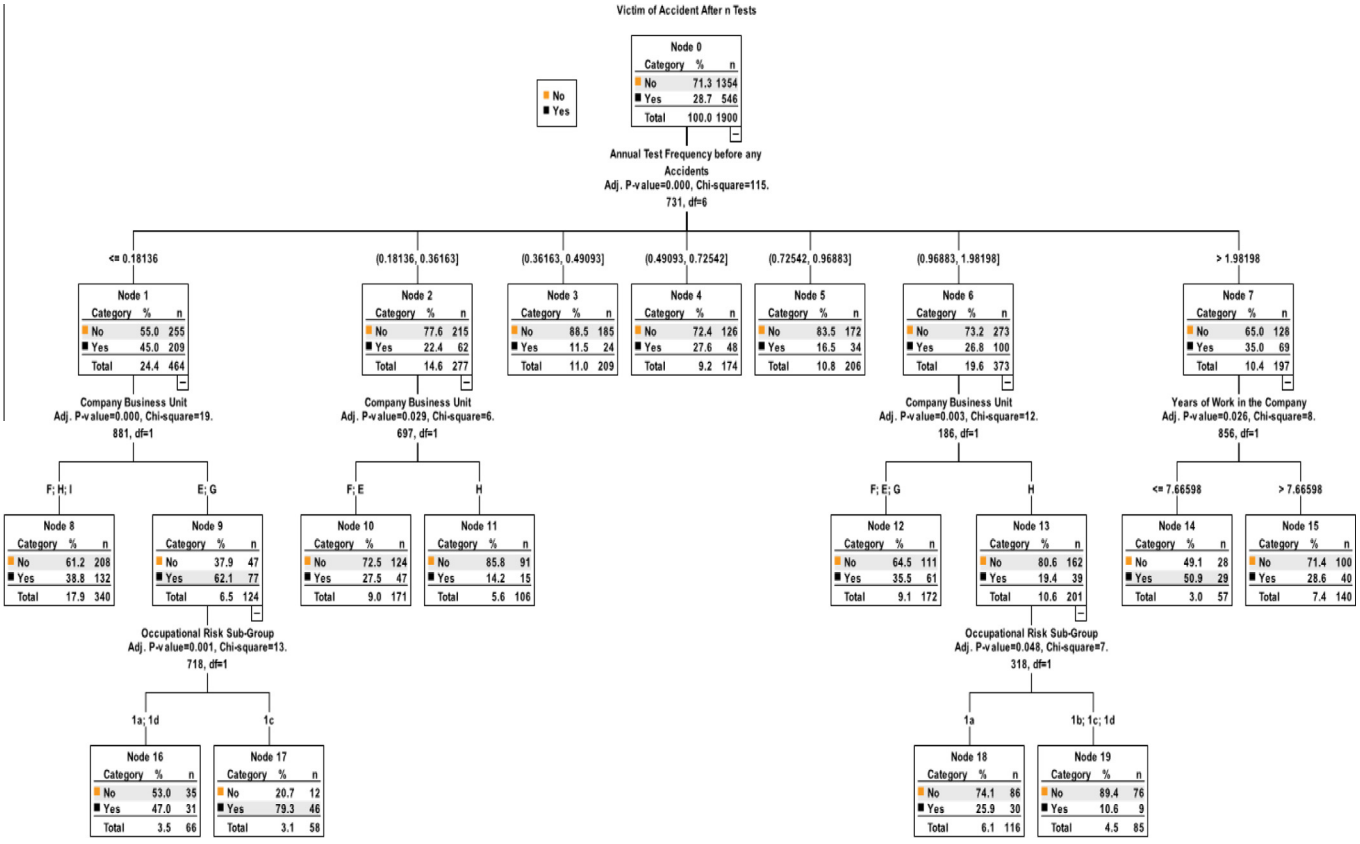

Fig. 2b. Associations between the output variable "Victim of accident after ***n*** tests" and explanatory variables: "Annual test frequency before any accidents", "Company business unit", "Years of work in the company" and "Occupational risk sub-group", in group 1 (N = 1900).

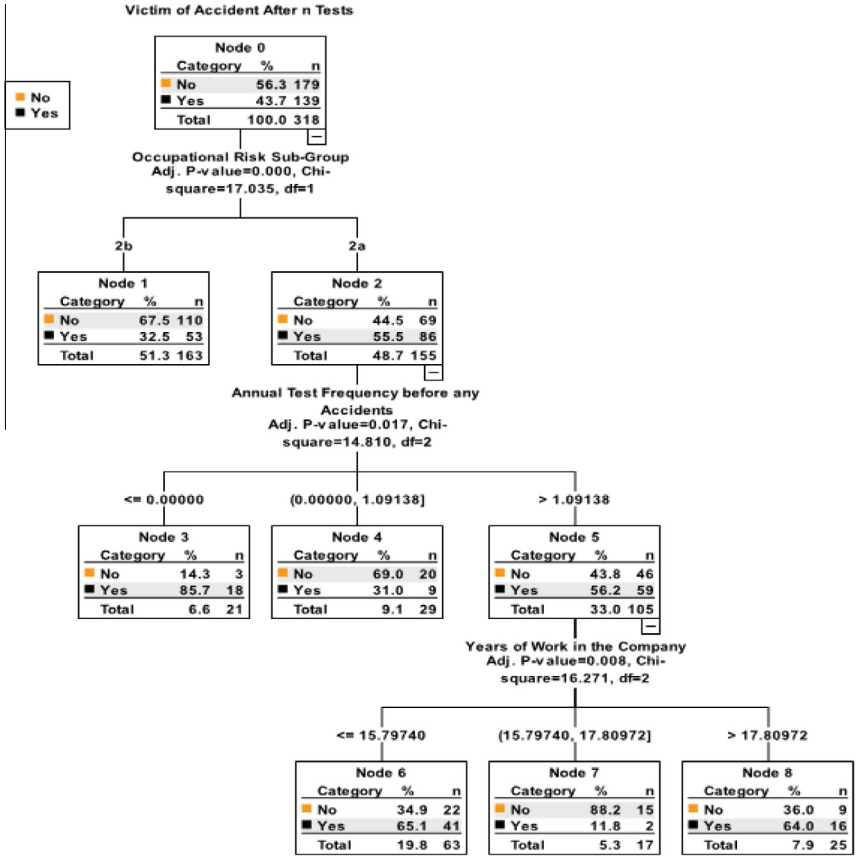

Fig. 2c. Associations between the output variable "Victim of accident after ***n*** tests" and explanatory variables: "Occupational risk sub-group", "Annual test frequency before any accidents", and "Years of work in the company", in group 2 ($N = 318$).

with a lesser strength of association. Furthermore, some classes of "Annual testing frequency" were of low dimension (below 30 employees), due to the smaller size of group 2 itself ($N = 318$).

Thus, it was verified the existence of associations between occurrence of accident victims and previous testing frequency – at 1% significance level, with a small difference in group 2, for which the significance level was 5%. Additionally, the strength of these associations was measured by Cramer V coefficient – whose values allowed concluding that they are moderate in the whole sub-population (0.270), as well as in groups 1 (0.226) and 2 (0.247), and strong in group 3 (0.364) (Healey, 2010; Le Roy, 2012; Murteira, 1990).

Once the expected negative association between accident occurrence and prior tests was confirmed, this study focused on the annual test frequency and accident rates – both annualised variables. The objective was to identify how far an organisation should go in terms of testing effort. The interest was to find out the optimal frequency, above which there is no benefit in increasing testing, *i.e.*, the frequency of tests at which the accident rates are minimised. In Figs. 2a–d, one can see that, among the statistically different classes of testing frequency, one of them reveals itself as the one with less accident occurrence after A&D testing. The authors have interpreted this pair as an optimal frequency.

To clarify this result of primal importance, the set of Figs. 3a–d depict the sub-population (3a), group 1 (3b), group 2 (3c) and group 3 (3d). These graphical representations are more intuitive than CHAID diagrams, showing the same reality by annual test frequency intervals, but using frequency classes easier for realistic implementation. In these figures, $F = 0.0$ stands for zero tests per year per worker, equivalent to not-tested – this was the class used for control.

For the sake of good statistical practice, the authors tried to create frequency classes with approximately equal number of cases, hence the classes]0–0.5],]0.5–1],]1–2], *i.e.*, classes of similar magnitude. Frequency 0.5 is a mathematical average – it means that workers in this class had been tested once in two consecutive years. Thus, there was a residual tail on the right-hand side of this scale that could not be further sub-divided, for the simple reason that very few workers had been subjected to frequencies much higher than 2 tests per year.

From these figures, one can easily see a pattern in the behaviour of proportion of accident victims ("Incidence of accident victims after ***n*** tests") in relation to prior annual frequency of subjection to tests. Similarly to the quantified statistics given before, the weakest preventive impact of testing seems to occur in group 2 (working near or around trains), whereas the strongest impact (*i.e.,* sustained effect) occurs in group 3, which is essentially composed of white-collar workers. The latter becomes a useful finding, since these professionals exist in most organisations. Hence, the results in this case might be useful for many other sectors of activity and are not restricted to transportation.

An optimal frequency was already expected by common sense and the authors also expected some kind of marginal variation above this point. In contrast, these results give evidence that after the strongest impact attained at the optimal frequency, the accident rate increases, although it never gets as high as in the control group (untested), *i.e.*, the testing preventive efficiency weakens. This fact became statistically evident. Although the explanation was not searched for, it is possible that this phenomenon is, as least partly, explained by a habituation effect, but such hypothesis still needs further study.

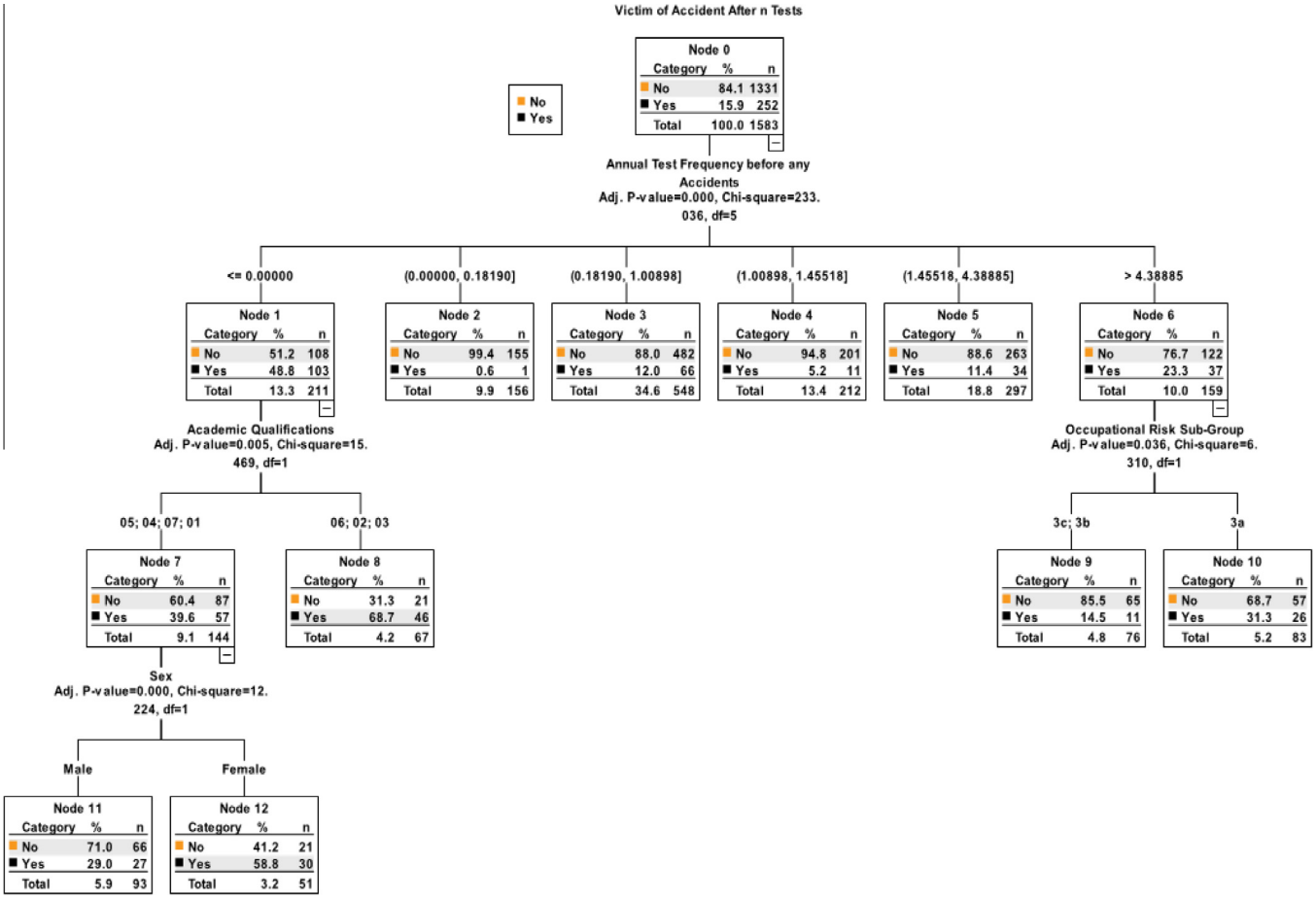

Fig. 2d. Associations between the output variable "Victim of accident after ***n*** tests" and explanatory variables: "Annual test frequency before any accidents", "Academic qualifications", "Occupational risk sub-group" and "Sex", in group 3 ($N = 1583$).

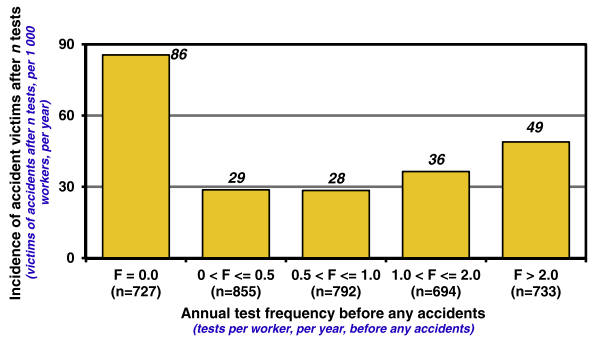

Fig. 3a. Incidence of accident victims, by subjection to tests before any accidents, in sub-population ($N = 3801$).

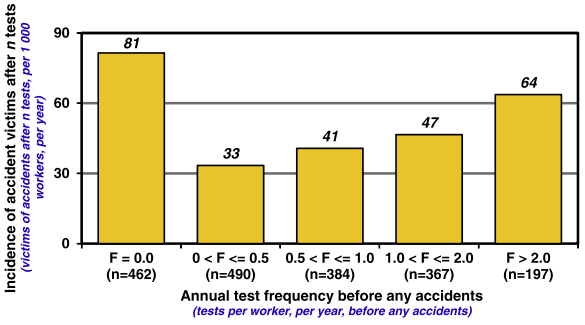

Fig. 3b. Incidence of accident victims, by subjection to tests before any accidents, in group 1 ($N = 1900$).

To assess the expected usefulness of the preventive effect of testing at these practical optimal frequencies for each occupational group, Table 3 compares proportions of victims of accidents, between untested workers and tested at optimal frequencies (all at 1% significance level).

In short, findings of the present study include odds ratios and prevented fractions pointing out a relevant proportional occurrence of fewer accidents among tested workers as compared with untested ones, and also a statistically significant association between a certain frequency of tests and a minimum of accidents. Furthermore, not only the existence but also a relevant strength of this association was established.

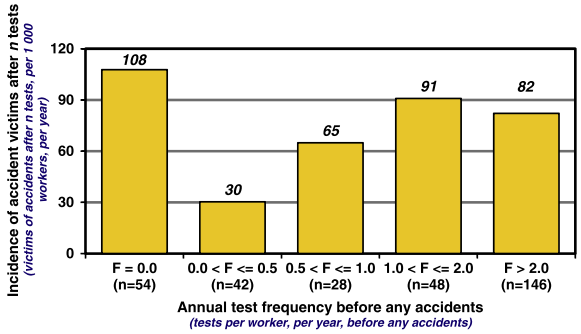

Fig. 3c. Incidence of accident victims, by subjection to tests before any accidents, in group 2 ($N = 318$).

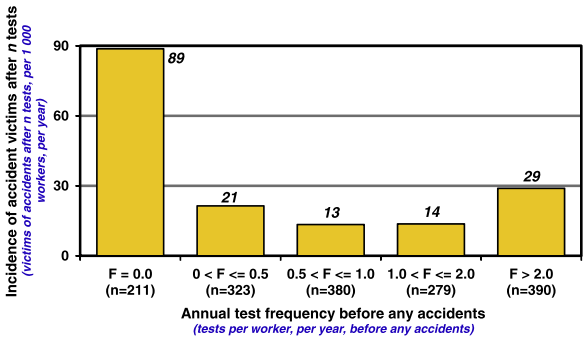

Fig. 3d. Incidence of accident victims, by subjection to tests before any accidents, in group 3 ($N = 1583$).

Table 3
Comparison between proportions of victims of accidents, between untested workers and tested at optimal frequencies.

Occupational group	Whole sub-population $N = 3801$	Group 1 (on board trains) $N_1 = 1900$	Group 2 (near or around trains) $N_2 = 318$	Group 3 (away from trains) $N_3 = 1583$
Optimal testing frequency (tests per year per worker)	$0.5 < F \leqslant 1.0$	$0 < F \leqslant 0.5$	$0 < F \leqslant 0.5$	$0.5 < F \leqslant 1.0$
Accident victims among untested (relative frequency) (%)	47	44.8	59.3	48.8
Accident victims among tested at optimal frequency (relative frequency) (%)	15.7	18.4	16.7	7.4
Prevented Fraction of accident victims (relative frequency) (%)	66.7	59.0	71.9	84.9

Roughly speaking, these main findings converge and expand the knowledge from a few other studies also involving large numbers and detailed statistical analyses (Cashman et al., 2009; Miller et al., 2007; Ozminkowski et al., 2003; Wickizer et al., 2004). For instance, the current findings also point at a U-shape curve with an optimal frequency of testing, representing a minimum cost, as found by Ozminkowski et al. (2003).

On the other hand, the present study leveraged significant differences of accidents between tested and untested employees, owing to the novel and clear distinction of individual subjection to tests exclusively before any accidents.

4.2. Economic appraisal of testing

This was a basic (but rather conservative) estimation of the economic value of the A&D testing programme. To avoid biases and controversy, it only covered group 1, because this category includes onboard professions (drivers, their assistants and ticket inspectors), who are absolutely essential to run a train. If a worker from this group suffers an accident with lost days, someone else is called in to replace him/her, which implies costs with overtime work. These extra costs are well known in the company in question and this allowed to compare the average costs with application of tests in group 1 against the money saved from the non-expenditure with overtime work, due to the reduction of accidents occurred after subjection to tests at the optimal frequency. Accordingly to the methodology explained by the American Industrial Hygienists Association (AIHA, 2008), *c.f.* Section 3.3, this estimate showed a net saving of about 15 € for each 1 € actually invested in testing. This rough estimate converges with the benefit-cost ratio of 26:1 found by Miller et al. (2007) in a more in-depth economic appraisal. Despite different methods and different cost items, both studies have established that random testing in the workplace can be cost-effective.

4.3. Limitations and contributions

Statistical limitations include a few variables with instantaneous data values that might have changed over the study period – however this is quite frequent in longitudinal studies. Also noteworthy is the very low number of cases with high testing frequencies (very few over 2 tests per year per worker) – and these were all aggregated in the last frequency interval ($F > 2$).

Methodological limitations include the existence of non-homogeneous groups with relation to a few variables, as well as the fact that there was no distinction between those tested only for alcohol or simultaneously for alcohol and drugs. The estimation of financial return used a simplified approach, accounting only for one category of saved costs (overtime work) and it did not adjust the time value of money.

On the other hand, a number of strengths and contributions are worthy mentioning, such as:

- The rarity of data (outside USA, very few companies and few countries carry out random A&D testing at workplace).
- The large amount of data (around 30,000 records) and the long period covered (5½ years).
- The quantification facet and robustness of the statistical analysis (producing the same results using different methods), thus increasing the reliability of results.
- The finding of an optimal testing frequency, enabling better management of resources.
- The broad social and cross-sector usefulness of the findings, especially considering that the strongest preventive impact was found in the white-collar professionals, who exist in every organisation, facilitating transferability of knowledge.

5. Conclusion

The study compared accident rates between (as far as possible) homogeneous groups of employees, *i.e.*, occupational groups performing similar tasks and exposed to the same pattern of occupational risks during the same period. This assured that, within each occupational group, the stimulus of being or not tested for A&D accounted for a relevant difference.

The findings consistently supported that groups tested for A&D at random and unannounced, reported substantially lower accident rates, after any number of tests, than the spontaneously untested group.

The association between testing and reduction of individual post-testing accidents was statistically confirmed and measured through different statistical techniques, delivering similar results and allowing to draw meaningful conclusions. These conclusions emerged from the contrast of accident rates (per year) after tests (per year), between homogeneous groups of workers, only differing on their test frequency. Thus, all other things being equal, the different individual frequencies of subjection to testing were likely to be responsible for different outcomes. Therefore, this research work was able to support the assumption that the application of A&D testing at the workplace, at random and by surprise has a

significant preventive effect on individual's occurrence of accidents.

It was also demonstrated that, for each occupational group, there is an optimal frequency of annual testing associated with a minimum accident rate, above which the increase in testing loses impact and becomes less efficient in terms of prevention. An interesting finding in this study was the fact that the positive impact of prevention is stronger and more effective within white-collar professionals. The impact on the operations and technical workers was also statistically significant, showing moderate to strong association strength. These findings are useful for cross-sector application, since white-collar workers are common in all activity sectors.

Furthermore, the study also established a sound evidence of a positive return on investment, through a conservative and simplified estimate.

With regard to the optimal testing frequencies observed, they led to convincing values of the fraction of accident victims that are prevented. If this information remains unknown to employees, the application of tests at these particularly low frequencies may be useful, in terms of maximising prevention while minimising costs.

Future developments of this work might focus on the reason(s) why accident rates increase above the optimal testing frequency, despite never becoming as high as in the case of the untested workers. It could also explore different specific occupational groups and establish their optimal testing frequencies.

Acknowledgement

Recognition is due to the railway company *CP-Comboios de Portugal* for rendering this study possible and for providing all the necessary data.

References

Acock, A.C., Stavig, G.R., 1979. A measure of association for nonparametric statistics. Soc. Forces 57 (4).

AIHA – American Industrial Hygiene Association, 2008. Develop and test general strategy components. In: Demonstrating the Business Value of Industrial Hygiene. American Industrial Hygiene Association, p. 56, <www.ihvalue.org>.

Baer, H.P., Hess, M., 2008. Dangerous behavior in companies – individuals at risk and high-risk behavior. In: Breitstadt, R., Kauert, G. (Eds.), The Worker – Risk Factor and Reliability. Shaker Verlag, Aachen, pp. 57–65.

Brady, J.E., Baker, S.P., DiMaggio, C., McCarthy, M.L., Rebok, G.W., Li, G., 2009. Effectiveness of mandatory alcohol testing programs in reducing alcohol involvement in fatal motor carrier crashes. Am. J. Epidemiol. 170 (6), 775–782.

Breitstadt, R., 2008. Drug screening as contribution to safety. In: Breitstadt, R., Kauert, G. (Eds.), The Worker – Risk Factor and Reliability. Shaker Verlag, Aachen, pp. 100–103.

Cashman, C.M., Ruotsalainen, J.H., Greiner, B.A., Beirne, P.V., Verbeek, J.H., 2009. Alcohol and drug screening of occupational drivers for preventing injury. Cochrane Database Syst. Rev. (2). Art. No.: CD006566. http://dx.doi.org/10.1002/14651858.CD006566.pub2.

Chipman, M., Jin, Y.L., Drowsy, 2009. Drowsy drivers: the effect of light and circadian rhythm on crash occurrence. Saf. Sci. 47 (10), 1364–1370.

Degenhardt, L., Hall, W., 2012. Extent of illicit drug use and dependence, and their contribution to the global burden of disease. Lancet 379, 55–70.

French, M.T., Roebuck, M.C., Alexandre, P.K., 2004. To test or not to test: do workplace drug testing programs discourage employee drug use? Soc. Sci. Res. 33 (1), 45–63.

Gust, S.W., Walsh, J.M., 1989. Drugs in the Workplace: Research and Evaluation Data. NIDA Research Monograph, National Institute of Drug Abuse, Rockville, vol. 91.

Hand, D., Mannila, H., Smyth, P., 2001. A Systematic Overview of Data Mining Algorithms. Principles of Data Mining. The Massachusetts Institute of Technology Press, Cambridge, Massachusetts, pp. 145–151.

Hanson, M., 1993. Overview on drug and alcohol testing in the workplace. Bull. Narcotics 45 (2), 3–44.

Healey, J.F., 2010. The Essentials of Statistics – A Tool for Social Research, second ed. Wadsworth, Belmont, p. 293.

ILO, 2003. Drug testing. Alcohol and Drug Problems at Work – The Shift to Prevention. International Labour Organisation, Geneve, p. 87.

Kauert, G., 2008. The effects of illicit drugs – a comparison of the potencial for influence. In: Breitstadt, R., Kauert, G. (Eds.), The Worker – Risk Factor and Reliability. Shaker Verlag, Aachen, pp. 23–27.

Kraus, J.F., 2001. The effects of certain drug-testing programs on injury reduction in the workplace: an evidence-based review. Int. J. Occup. Environ. Health 7 (2), 103–108.

Roy, Le, M.K., 2012. Research Methods in Political Science. Cengage, Belmont, CA, p. 193, 196.

Li, Y., Bai, Y., 2008. Comparison of characteristics between fatal and injury accidents in the highway construction zones. Saf. Sci. 46 (4), 646–660.

Li, G., Baker, S.P., Lamb, M.W., Qiang, Y., Melissa, L., McCarthy, M.L., 2005. Characteristics of alcohol-related fatal general aviation crashes. Accid. Anal. Prev. 37 (1), 143–148.

Livingston, W., 1975. Betriebliche Alkoholismus-Programme in U.S. Firmen: Ein Untersuchungsbericht des Long Range Planning Service, Stanford Research Institute, London.

Lund, J., Aarø, L.E., 2004. Accident prevention – presentation of a model placing emphasis on human, structural and cultural factors. Saf. Sci. 42 (4), 271–324.

Marques, P.H., 2009. Occupational safety management in the Portuguese trains, focused on behavioural risks' control. In: Guedes Soares, C., Jacinto, C., Teixeira, A.P. & Antão, P. (Eds.), Riscos Industriais e Emergentes, Edições Salamandra, Lisboa, vol. 2, pp. 911–930 (in Portuguese, ISBN 978-972-689-233-5).

Marques, P.H., Jesus, V., Vairinhos, V., Abajo Olea, S., Jacinto, C., 2011. The control of alcohol and drugs and occupational accidents at the "Trains of Portugal": data analysis. In: Arezes et al. (Eds.), International Symposium on Occupational Safety and Hygiene – SHO'2011, Guimarães, Portugal, pp. 373–377 (ISBN: 978-972-99504-7-6).

Mehay, S., Webb, N., 2007. Workplace drug prevention programs: does zero tolerance work? Appl. Econ. 39 (21), 2743–2751.

Miller, T.R., Zaloshnja, E., Spicer, R.S., 2007. Effectiveness and benefit-cost of peer-based workplace substance abuse prevention coupled with random testing. Accid. Anal. Prev. 39 (3), 565–573.

Murteira, B., 1990. Probabilities and Statistics, 2nd ed. McGraw-Hill, Lisboa (pp. 412–413, in Portuguese).

Network Rail, 2008. Drugs and alcohol policy Testing. In: Health and Safety Handbook, NR, London, version 1.0, pp. 4–9.

O'Connor, P.J., O'Connor, N., 2006. Work-related maritime fatalities. Accid. Anal. Prev. 38, 737–741.

OECD– International Transport Forum, 2010. Summary document. In: Drugs and Driving, Detection and Deterrence. Organisation for Economic Cooperation and Development, OECD Publishing, Paris,, p. 13.

Ozminkowski, R.J., Mark, T.L., Goetzel, R.Z., Blank, D., Walsh, J.M., Cangianelli, L., 2003. Relationships between urinalysis testing for substance use, medical expenditures, and the occurrence of injuries at a large manufacturing firm. Am. J. Drug Alcohol Abuse 29 (1), 151–167.

Panik, M.J., 2005. Advanced Statistics from an Elementary Point of View. Elsevier Academic Press, London, p. 693.

PCC– Portuguese Constitutional Court, 1995. Acórdão n° 319/95. Acórdãos do Tribunal Constitucional, vol. 31 (in Portuguese).

PP– Portuguese Parliament, 2009. Lei n.° 7/2009. Diário da República – Iª Série, no 30, de 12 de fevereiro, pp. 926–1029 (in Portuguese).

President, The., 1986. Executive order 12564: drug-free federal workplace. Fed. Regul. 51, 32889–32893.

PSCJ– Portuguese Supreme Court of Justice, 1998. Acórdão do processo 97S243, de 24/06/1998.www.dgsi.pt/jstj.nsf/954f0ce6ad9dd8b980256b5f003fa814/56e61e8c22fafdd6802568fc003b7ae1?OpenDocument (in Portuguese, accessed March 2013).

Rokach, L., Maimom, O., 2008. Advanced decision trees. In: Data Mining With Decision Trees - Theory and Applications. World Scientific Publishing Co., Singapore, pp. 72–73.

Rummel, M., Rainer, L., Fuchs, R. 2004. Alkohol im Unternehmen, Verlag Hogrefe, Göttingen.

Sampieri, R.H., Collado, C.R., Lúcio, P.B., 2006. Search Models,, .. Metodologia de Pesquisa, 3rd ed. McGraw-Hill, São Paulo, Brasil (pp.152–247, in Portuguese).

Schiffhauer, N., Breitstadt, R., 2008. Keeping Tabs on Drug Abuse. Dräger Review, vol. 96. Springer, Lübeck, pp. 10–15.

Schuckjt, M.A., 2009. Alcohol-use disorders. The Lancet 373, 492–501.

Seiffert, B., 2008. A field report of a drug screening of trainees. In: Breitstadt, R., Kauert, G. (Eds.), The Worker – Risk Factor and Reliability. Shaker Verlag, Aachen, pp. 104–108.

Snowden, C., Miller, T., Waehrer, G., Spicer, R., 2007. Random alcohol testing reduced alcohol-involved fatal crashes of drivers of large trucks. J. Stud. Alcohol Drugs 68 (5), 625–633.

Spicer, R., Miller, T., 2005. Impact of a workplace peer-focused substance abuse prevention and early intervention program. Alcohol. Clin. Exp. Res. 29 (4), 609–611.

Strang, J., Babor, T., Caulkins, J., Fischer, B., Fozcroft, D., Humphreys, K., 2012. Drug policy and public good: evidence for effective interventions. Lancet 379, 71–83.

Taggart, R.W. 1989. Results of the drug testing program at Southern Pacific Railroad. National Institute of Drug Abuse Research Monograph No. 91, U.S. Government Printing Office, Washington, DC.

The Lancet Editorial, 2009. Alcohol misuse needs a global response. The Lancet 373, 433.

Trent, R.B., 1991. Emergency room evidence on the role of alcohol intoxication in injury at work in the US. Saf. Sci. 14 (3–4), 241–252.

UIC, Occupational Health and Safety Group, 2008. Developing management arrangements for the Control of Safety Risks related to the Influence of Alcohol, Drugs and/or Psychoactive Medication. International Union of Railways, Paris, pp. 1–17.

Wenzek, M., Ricordel, I., 2008. Cannabis et sécurité du travail. Photographie de l'evolution de sa detection au sein des contrôles de stupéfiants depuis 2004 à la SNCF. Annales Pharmaceutiques Françaises 66, 255–260 (in French).

Wickizer, T.M., Kopjar, B., Franklin, G., Joesch, J., 2004. Do drug-free worplace programs prevent occupational injuries? Evidence for Washington State. Health Serv. Res. 39 (1), 91–110.

Zwerling, C., 1993. Current practice and experience in drug and alcohol testing in the workplace. Bull. Narcotics 45 (2), 115–196.