Advances in Science and Technology Research Journal

Advances in Science and Technology Research Journal, 18(7), 329–346
https://doi.org/10.12913/22998624/193217
ISSN 2299-8624, License CC-BY 4.0

Received: 2024.07.22
Accepted: 2024.09.25
Published: 2024.10.06

Profiles of Wood Processing Occupational Accident Casualties by the Size of Enterprises

Marzena Nowakowska[1], Michał Pajęcki[1*]

[1] Faculty of Management and Computer Modelling, Kielce University of Technology, al. Tysiąclecia Państwa Polskiego 7, 25-314 Kielce, Poland

* Corresponding author's e-mail: m.pajecki@tu.kielce.pl

ABSTRACT

The work covers the issues related to the diagnosis of selected occupational safety hazards in Polish wood processing enterprises. The main objective is to build models in order to identify occupational accident profiles in these enterprises on the basis of individual records characterizing the casualties, provided by Statistics Poland. The modelling task employed the latent class analysis (LCA) data mining technique. In order to enhance the process of building the LCA model and to support the procedure of selecting input variables relevant to the model, an iterative algorithm was elaborated by the authors. The impact of an enterprise size on occupational accident consequences was statistically confirmed. Following this result, LCA models were developed independently for smaller (micro and small), and for larger (medium and large) enterprises. Latent classes, presenting occupational accident profiles, were visualized in the form of heat maps. Similarities and differences between the occupational accident profiles identified for the two types of enterprises were indicated. It has been shown that employees of smaller enterprises are at greater risk of suffering more serious injury from accidents at work than employees of larger enterprises. However, in both cases, the most critical latent classes concern occupational accidents related to operating machinery; they affect workers with a low level of job seniority, and result in injuries (often traumatic amputations) involving upper limbs in particular.

Keywords: wood industry, occupational accidents, accident profiles, latent class analysis, model selection, variable assessment.

INTRODUCTION

Each type of a production activity carries various risks, in particular those connected with the workplace. Occupational accidents pose a serious social and economic problem. Research in the field of occupational safety allows for a better understanding of the nature of such accidents, occupational risk assessment, hazards identification and reduction [1]. The problem attracts the interest of a multitude of scientists from various fields, leading to the development of methods and tools supporting the analysis of accidents at work [2]. In this study, the research interest is focused on occupational accidents recorded in manufacturing industries related to wood processing.

Background

In the work, a process of literature review was carried out in two stages. First, a bibliometric analysis of bibliographic data was performed in order to find main research trends concerning occupational accidents in wood industry. Then reference was made to selected publications – the ones relating to the topic under consideration: accidents recorded in wood processing industry. The bibliographic data were retrieved and downloaded from two biggest world on-line repositories, Web of Science (WoS) and Scopus, in September, 2023. The following criteria for the bibliographic records acquisition joined by a logical conjunction were defined:

- search string imposed for article title, keywords, abstract and additional keywords

(Keywords Plus in WoS [3] and Index Keywords in Scopus [4]): ("*wood*" OR "lumber*" OR "timber*" OR "furniture*" OR "saw*mill*") AND ("manufactur*" OR "processing" OR "industr*") AND ("occupational accident*" OR "accident at work*" OR "accidents at work*" OR "industrial accident*");
- time period: "since 2000";
- document type: "article" OR "review" OR "proceedings paper".

In order to unify the resulting data set and to create a corpus of bibliographic records merged from both databases, two different data structures (WoS and Scopus) were unified. Pre-processing was done to eliminate duplicate documents as well as to perform other data cleaning procedures. The resulting data set included 127 articles (records). They were published in 73 journals; among them 56 ones contained only one article. These articles formed the basis for the knowledge domain mapping, providing a general overview of the scientific output in the field [5].

The knowledge domain mapping was processed with the use of the VOSviewer computer program [6], a tool for analysis and visualization of bibliographic data; a lexical map was created as regards occurrences and co-occurrences of keywords in scientific publications. Own elaborated database management system tools were also applied. Very important is a thesaurus developed to reduce redundant terms and to standardize them [7]. Table 1 contains an example of the thesaurus concept. The *Keyword* and *Occurrences* columns are the results from the VOSviewer program. The third column, called *Synonym*, was specially prepared to cover the respective group of the *Keyword* phrases with the same or similar meaning, and it is the authors' proposition. Except for the standard approach to the synonym creation, in order to combine phrases that differ in detail but refer to the same general concept, a term with the " +" suffix (space and plus sign) was created in the *Synonym* field allowing aggregation of the phrases. According to the VOSviewer requirements, the thesaurus file consists of the *Keyword* and *Synonym* columns and does not contain rows with the same content in both the columns. After applying the thesaurus, the number of unique key phrases ready for the knowledge map creation was reduced from 1438 to 1107 (reduction by almost 30%). Among them, there were 220 synonyms to which more than one keyword was matched.

Using VOSviewer, the lexical map of key phrases (keywords) for the 127 publications was created. With the default settings of the program (minimum number of keyword occurrences = 5), four clusters were isolated on the map, among which 100 key phrases were distributed. Table 2 shows five most often occurring key phrases within a cluster – the results from the VOSviewer program. The second column represents the number of phrases in a given cluster. For each key phrase, measures of its importance (weights) are given: occurrences, links, and total link strength [8]. The key phrases are ordered by *Occurrences* in descending order. The table also provides a brief description of each cluster.

From the point of view of this work, the *Cl*-1 cluster is the most important, as it concentrates on *occupational accidents* in *wood industry* in general. It contains 42 items. Occupational accident is the most important key phrase not only in the cluster but also in the whole map – see Figure 1. It is the result of aggregation of 16 phrases. The maximum values of all the weights (occurrences, links, total links strength) are achieved for this particular phrase. The *wood industry* phrase, attributed to the cluster, is the result of aggregation of 7 phrases. The phrase weights are as follows: *Occurrences* = 11, *Links* = 63, *Total link strength* = 153.

Table 1. Concept of synonym phrases for the thesaurus of occupational accidents bibliometric investigation

Keyword	Occurrences	Synonym
Mechanical wood-processing industry	1	Wood industry
Plywood industry	1	Wood industry
Wood industries	1	Wood industry
Wood industry	3	Wood industry
Wood processing industry	2	Wood industry
Wooden furniture industry	4	Wood industry
Woodworking industry	1	Wood industry

Note: Grey background indicates a row excluded from the thesaurus file, as required by the VOSviewer.

Table 2. The most important key phrases within bibliographic clusters concerning wood industry occupational accidents

Cluster identifier	Cluster size	Most important key phrases			
		Phrases	Occurrences	Links	Total links strength
Cl-1 Red	42	Occupational accidents	103	99	1226
		Wood	31	95	434
		Occupational risks	30	87	369
		Occupational health and safety	26	86	329
		Accident prevention	26	84	338
	The Cl-1 cluster focuses on shaping a safe working environment, for example, through risk management or occupational accident prevention.				
Cl-2 Green	25	Humans	82	99	1108
		Male	54	99	828
		Adult	54	99	837
		Female	39	91	593
		Industry	37	92	499
	The core of the Cl-2 cluster is the human factor, accompanied by issues related to health and protection of workers in the workplace.				
Cl-3 Blue	22	Middle aged	32	89	529
		Young people	18	81	308
		Adolescent	17	72	279
		Incidence	15	75	219
		Statistics	15	68	200
	As in the green cluster, the human factor plays an important role in the Cl-3 cluster. However, this occurs mostly in conjunction with age and shows that serious injuries can occur during woodworking by young and middle-aged.				
Cl-4 Yellow	11	Work-related injuries	27	86	409
		Controlled study	21	91	333
		Workers' compensation	15	64	220
		Wounds and injuries	10	71	170
		Wood manufacturing	8	48	100
	The Cl-4 cluster refers generally to the cost aspects of injuries to workers in the course of their work-related activities.				

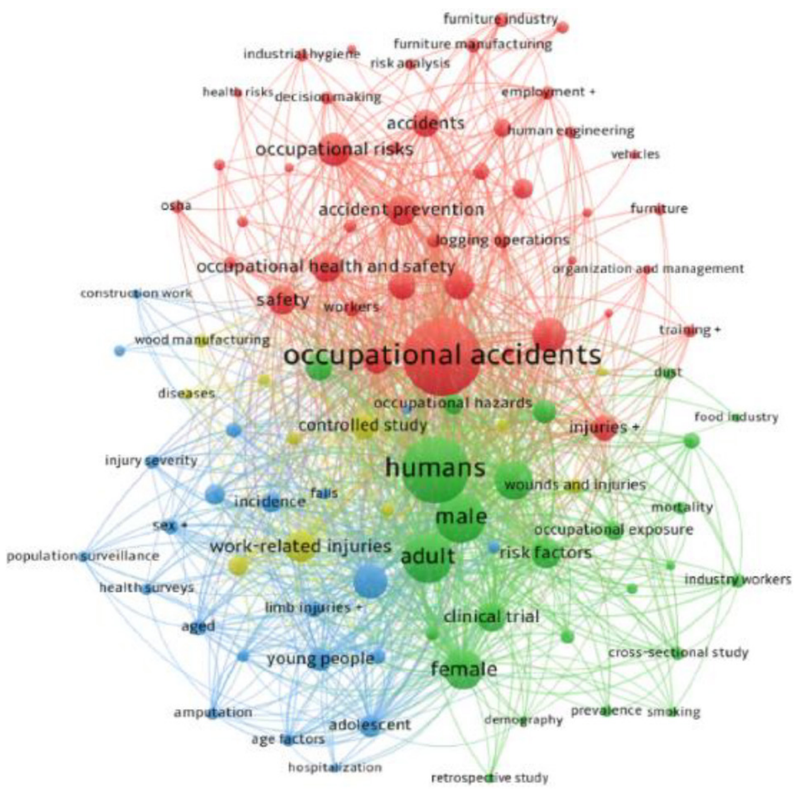

Figure 1. The map of keywords and keywords' co-occurrence in research publications on wood industry occupational accidents

Figure 2 presents the phrases with which wood industry has the strongest co-occurrence.

In the second stage of the literature review, publications related to the research interest of the work were identified and focused directly on them. Table 3 presents the articles selected as those that discuss occupational safety hazards in wood processing enterprises.

A large part of the articles in the table is devoted to the analysis of data obtained through surveys of enterprise workers, often in connection with the safety climate at work. There are also articles on selected cases, for example: specific types of injuries related to hearing, sight, elderly employees. Very few ones use individual data records on accidents at work obtained from institutions (including those of national level). Among the articles presented (but also others reviewed), the authors did not find the approach proposed in this paper.

Subject of the research

The main objective of the work is to build models to identify occupational accident profiles in wood manufacturing enterprises based on individual data of injured employees, taking into account the size of an enterprise. To the best of the authors' knowledge, such aspect of the analyses has not yet been considered in the literature. An investigation at a national level was undertaken. The industry sector called *Manufacture of wood and cork products, excluding furniture; manufacture of articles of straw and plaiting materials* relating to Section C *Manufacturing* according to Polish Classification of Activities was chosen in this work. It covers the manufacture of products such as veneer, wooden packaging, floor coverings, plywood, lumber products and other carpentry and joinery products [21]. Individual data records

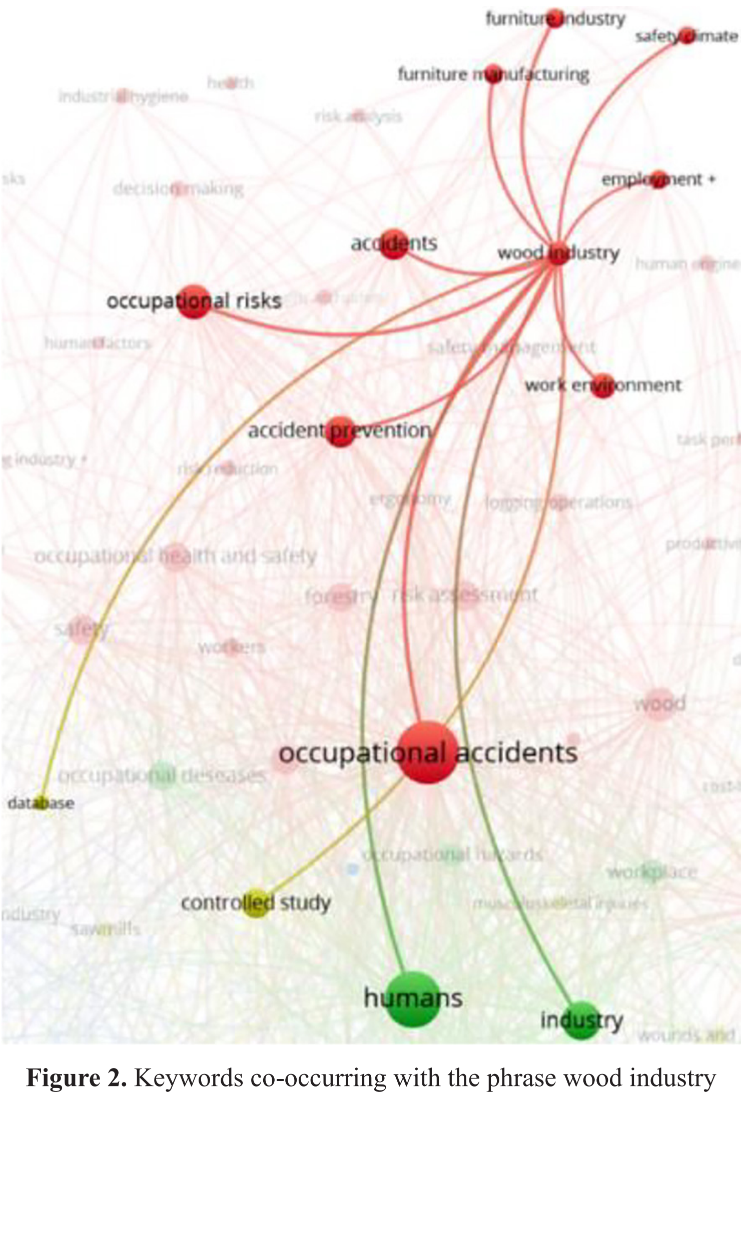

Figure 2. Keywords co-occurring with the phrase wood industry

Table 3. Selected articles on OHS in wood processing industry

Article	Main findings, achievements
The safety climate and its relationship to safety practices, safety of the work environment and occupational accidents in eight wood-processing companies [9]	Factor analysis was used to determine the structure of the safety climate in Finnish wood-processing plants on the basis of data collected through a questionnaire among workers. The authors indicated that company safety precautions are correlated with company safety practices and with the safety level of the work environment. They concluded that the better was the safety climate of the company, the lower was the accident rate.
An evaluation of hospital discharge records as a tool for serious work-related injury surveillance [10]	Hospital discharge records were extracted for a cohort of sawmills workers in British Columbia, Canada. Serious work-related injuries were identified and their most frequent causes indicated – falls, machinery related, overexertion, struck against, cutting or piercing, and struck by falling objects. It was pointed out that hospital discharge data represent an alternative and independent source of information on serious occupational accident injuries.
Work environment risk factors for injuries in wood processing [11]	The authors conducted a case-control study in Maine in wood processing enterprises in order to identify preventable injury risk factors. The data were collected through interviews. Multivariate analyses showed that variables associated with injury risk were: absence of a lockout/tagout program; high physical workload; lack of training; low level of seniority; machine-paced work or inability to take a break; male gender.
Self-organizing map and clustering algorithms for the analysis of occupational accident databases [12]	A two-level approach based on the joint use of the Kohonen Self-Organizing Map and the k-means clustering algorithm was used in order to discover the most common sequences of events leading to occupational accidents. The main purpose of the work was to analyze the effectiveness of the proposed clustering method. The dataset on non-fatal accidents at work in the Italian wood processing industry was a case study subject. The loss of control and the incorrect movements during the work with manual tools was indicated as the most critical sequence.
Job safety analysis and hazard identification for work accident prevention in para rubber wood sawmills in southern Thailand [13]	The authors conducted a cross-sectional study (with a walk-through survey). Potential OHS hazards associated with the main production processes at rubberwood sawmills were identified: high risks of exposure to wood dust and noise when sawing lumber into sheets; hand and foot injuries when struck by lumber; exposure to chemicals and fungicides; injuries due to poor ergonomics or repetitive work. Recommendations as regards the reduction of workplace hazards were formulated.
A novel tool for evaluating occupational health and safety performance in small and medium-sized enterprises: The case of the Quebec forestry/pulp and paper industry [14]	Small and medium-sized enterprises were indicated as those characterized by a higher rate of accidents at work and worse OHS results than large companies. Taking the above into account, the authors presented proprietary computer software for assessing health and safety performance. The software become a part of the standard OHS procedures used by prevention experts (so far in the forestry/pulp and paper industry in Quebec).
Large occupational accidents data analysis with a coupled unsupervised algorithm: The S.O.M. K-means method. An application to the wood industry [15]	The article focuses more on the methodology than on the problem of characterizing accidental profiles. Data on accidents at work in the wood industry were selected as a case study to illustrate the proposed cluster analysis methodology. The resulting clusters of occupational accident casualties were not discussed; selected examples were partially described. The two most critical clusters, according to the risk assessment, were related to "manual activity with hand tools" and to "free movements/manual transport" in the working area. According to the authors, their proposed method makes it possible to distinguish groups of occupational accidents, characterized by different dynamics, and also to associate a different quantification of the frequency and severity of occupational accidents with each group.
Perception of occupational risk factors in sawmills in the El Salto region of Durango, Mexico [16]	An exploratory questionnaire was used to determine the use of personal protective equipment (PPE) and the perception of safety among sawmill workers while wood processing in a selected region of Mexico. Respondent employees typically do not use full PPE during the workday, only gloves. They consider noise and vibration, but not sawdust and dust, to be the most serious occupational risk factors affecting their health. They also rate workplace safety as fair to good.
Evaluating the work environment in Turkish furniture industry from the point of occupational health and safety [17]	The work environments of SMEs in the furniture sector were studied. From the frequency analysis of measured parameters, the authors concluded that insufficient lighting, poor air quality, too noisy and dusty work environment affect the health of workers and can make them lose their attention, which can lead to injuries and fatal occupational accidents.
Nonfatal occupational injuries among workers in microscale and small-scale woodworking enterprise in Addis Ababa, Ethiopia [18]	A pretested, structured questionnaire was used to collect information from workers, and an observation checklist was used to collect the work environment data. Applying the logistic regression, the following factors were identified as significantly associated with occupational injury in wood processing plants in Ethiopia: khat chewing behavior, job dissatisfaction, work-related stress, job category, unguarded machines, and inadequate work space.
Latent class analysis for identification of occupational accident casualty profiles in the selected Polish manufacturing sector [19]	The authors identified the profiles of occupational accident casualties in Polish wood processing plants. Latent class analysis and raw data from the Statistics Poland registers were used for the study. Several patterns were identified and described on the basis of the set of variables characterizing occupational accidents casualties. Among them, the most serious one referred to incidents with casualty disability or death connected with upper limb injuries while operating machinery.
The age factor in the analysis of occupational risks in the wood industry [20]	The role of a worker age in OHS management in the wood industry was investigated on the basis of in-depth interviews carried out among occupational safety technicians and experts from Galicia (Spain). Although hard work commonly affects workers over 55, the number of accidents and absences in this age group because of occupational accidents was not higher than in other groups. Intermediate age groups, with 5 to 10 years of experience, had the highest rate of suffering from accidents at work, mainly due to overconfidence. Most experts found that preventive measures at work were not taken according to age, and even less for a given age range.

containing information on single observation objects (casualties), provided by the national data owner (Statistics Poland), were examined. Accidents can occur according to a variety of scenarios, but the identification of certain patterns can have significant preventive value.

The work provides additional insight into accidents at work in the wood processing industry and develops research methodology. The most important elements for scientific contribution are:
- the validity of distinguishing two types of enterprises by their size and, at the same time, the way how to classify enterprises into these groups, based on pairwise tests,
- the algorithm for the process of selecting observed variables and building latent class analysis models according to this process,
- synthetic diagnosis of occupational safety threats for two types of manufacturing enterprises in the form of heat maps,
- indication of similarities and differences between the occupational accident profiles identified for the two types of enterprises.

METHODOLOGY APPROACH

A qualitative nature is the characteristic feature of the data investigated in the research. Therefore, the latent class analysis (LCA) method was chosen to identify patterns of occupational accident casualties. LCA is a statistical tool used to build typologies based on the associations among a set of observed nominal variables [22, 23]. In the LCA model, a certain abstract qualitative variable, called a construct or a latent variable LV, is not directly observed but it reveals (manifests) its presence and intensity through other qualitative variables X_j, $j = 1, \dots, J$, whose values can be determined. These variables are indicators of the construct. The purpose of the method is to identify disjoint homogeneous subsets (clusters) in the data set on the basis of the indicators. The clusters are called latent classes and represent the values of the LV latent variable. The following form of the LCA model can be defined. It determines the probability that in the z observation the $r(z)$ vector, representing the combination of values of the indicators $X_1,\dots, X_J$, has the value equal to the q vector [23, 24]:

$$
\begin{aligned}
P(r(z) = q) &= \\
&= \sum_{c=1}^{C} \gamma_c \cdot \prod_{j=1}^{J} P(r_j(z) = q_j \mid z \in K_c) = \\
&= \sum_{c=1}^{C} \gamma_c \cdot \prod_{j=1}^{J} \rho_{q_j|c}
\end{aligned} \tag{1}
$$

where: K_c – c-th latent class, $c = 1, \dots, C$; C – the number of latent classes; γ_c – the probability of the c-th latent class, which is the probability that an observation belongs to the K_c latent class: $\gamma_c = P(K_c) = P(z \in K_c)$, $c = 1,\dots, C$, $\left(\sum_{c=1}^{C} \gamma_c = 1\right)$; $\rho_{q_j|c}$; – the conditional probability, that j-th observed variable has the value of q_j in the K_c latent class: $\rho_{q_j|c} = \mathrm{P}(q_j|c) = \mathrm{P}(r_j(z) = q_j|z \in K_c)$; q_j – the value of j-th observed variable, $q_j \in R_j$; R_j – the set of values of j-th observed variable, $j = 1,\dots, J$.

On the basis of real data (which create a sample), the γ_c and $\rho_{q_j|c}$ parameters are estimated delivering the LCA model.

The researcher does not know in advance how many latent classes are to be obtained from the data set. Usually, several LCA models are generated for a different number of classes and the best one is selected among them. This study applies the concept of selecting such best (final) model for the discussed issue by specifying:
- which observed variables play an important role in building the model,
- how many classes make up the model.

Observed variables may play more or less relevant role in the construction of the LCA model and perhaps some of them may be excluded, which will allow to limit the number of estimated parameters. In the task of selecting observed variables for the model, the discriminating ability index AR was used. It is calculated for a given LCA model with C classes as follows [19]:

$$AR(X_j|C) = \frac{1}{|R_j|} \sum_{q_j \in R_j} Range(q_j|\{c\}) \tag{2}$$

where:

$$Range(q_j|\{c\}) = \max_{c=1,\dots,C} \{\rho_{q_j|c}\} - \min_{c=1,\dots,C} \{\rho_{q_j|c}\};$$

$\rho_{q_j|c}$ and R_j are specified in Equation 1.

The AR measure allows the assessment of the importance of any X_j variable in the process of building the LCA model. It determines the average range of conditional probabilities of an observed variable in the LCA model. The measure construction and its interpretation are self-evident as they are based on probability theory. *AR* can be interpreted as an assessment of the distinguishability of latent classes in relation to the variable for which it is calculated. The *AR* value belongs to the interval [0, 1]; the maximum *AR* value occurs when the conditional probability of the variable category in one of the classes has the value of 0, and in another class the value of 1. This means that in the case of perfect distinguishability of at least two classes, *AR* is equal to 1. The nature of the measure indicates that the closer to one its value is, the stronger the discriminating ability of the observed variable becomes. In the study, the *AR* measure was used in the variable selection procedure for the LCA model and also as an indicator of importance (weight) of observed variables in defining profiles of occupational accident casualties.

An iterative backward selection type algorithm is proposed to select relevant observed variables. The diagram of the algorithm is presented in Figure 3. In the procedure, a set of temporary latent class models, with the number of classes varying in the range $C = C_{min},\ldots, C_{max}$ is used. From the series of $AR(Xj|C)$, the average value for each X_j is calculated as follows:

$$\overline{AR}(X_j) = \frac{1}{C_{max} - C_{min} + 1} \sum_{C=C_{min}}^{C_{max}} AR(X_j \mid C) \quad (3)$$

In the subsequent steps of the algorithm, the observed variable having the smallest averaged value $\overline{AR}$ of the discriminating ability index is

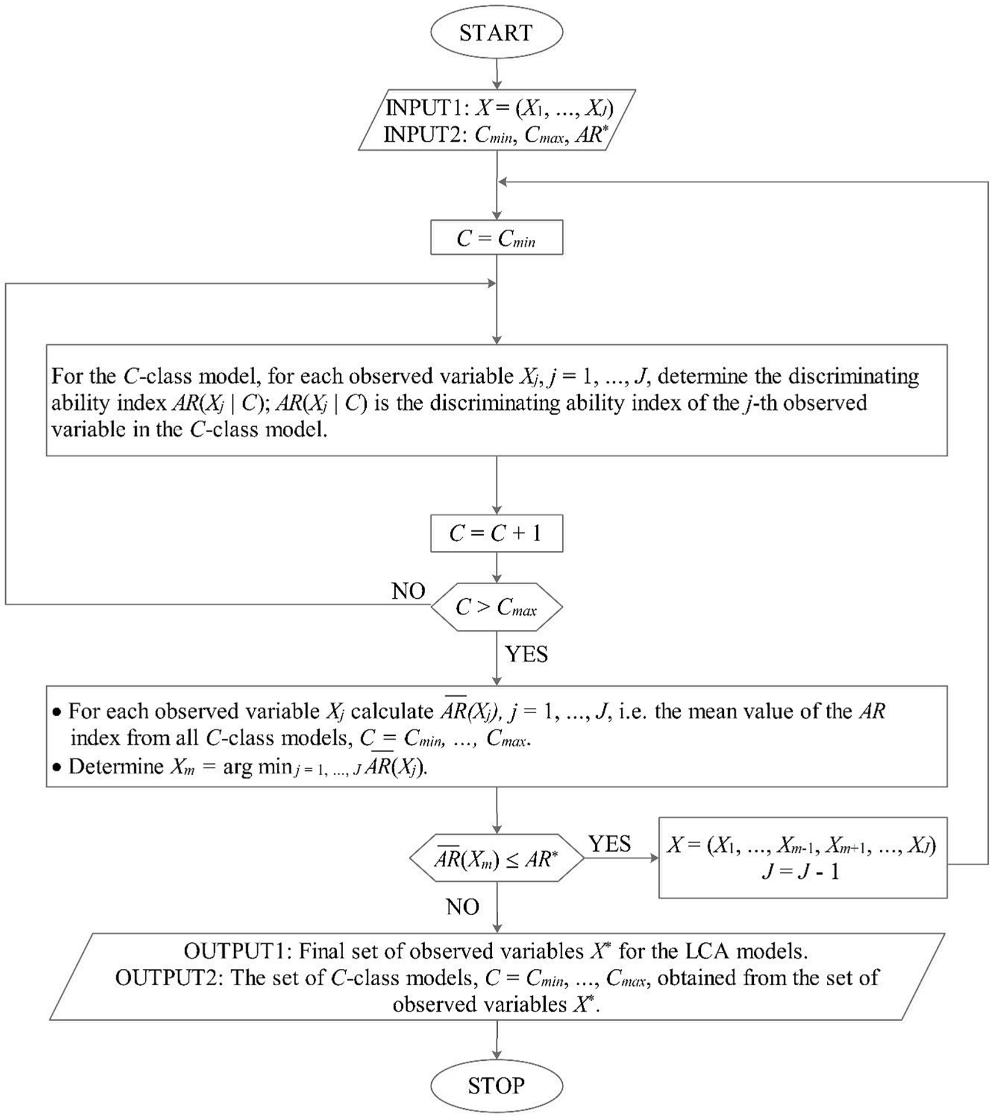

Figure 3. Algorithm for selection of relevant observed variables and generation of LCA models

excluded, as long as this value is less than a certain assumed cut-off level AR^*. The values of the parameters C_{min}, C_{max}, and AR^* are taken a priori by the researcher.

In the LCA method, as in other cluster analysis methods, it is important to determine the number of clusters (latent classes). A universal and unambiguous way to decide about this is not known in the literature [25, 26] and it is still under discussion [27–30]. In the process of selecting the number of latent classes, the Bayesian information criterion (BIC) is the most commonly used and trusted fit index for model comparison [22, 30–32]. Other information measures are also applied, such as: adjusted Bayesian information criterion (ABIC), Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), sample adjusted Bayesian information criterion (SABIC) [23, 33–34]. Also popular are: measures based on entropy, the lo-mendel-rubin (LMR) test, or the bootstrap likelihood ratio test (BLRT) [22–26, 30]. In any case, the nature of the issue under study and the subjective approach of the researcher imply the choice of the measures used.

In this work, the following quality measures were considered for selection of the final LCA model out of several candidates, with the number of classes varying from C_{min} to C_{max}: BIC, CAIC, ABIC, entropy-based measure E, and the discriminating ability index *AR*.

PREPARATION OF DATA

Individual data records on accidents at work for 2008–2017 in Poland registered in productions enterprises, obtained from Statistics Poland (GUS), were the subject of the research. The study focuses on the industry sector called *Manufacture of wood and cork products, excluding furniture; manufacture of articles of straw and plaiting materials* and it refers to accidents that happened in a direct connection with the production process. Thus, only records that met the following criteria were selected (the denotation in italics originate from the Polish statistical accident card):

- people injured in accidents at work – *Industrial workers and craftsmen*, *Operators and assemblers of machines and devices*, and *Employees doing simple works*;
- accident location – *Industrial production sites*,
- work process – *Production, processing, storage*.

In Poland, the size of an enterprise is classified according to the number of employees (without recalculation to full-time employment) as follows:

- 0 – self-employed persons with no other employees,
- 1 – up to 9 employees,
- 2 – 10–49 employees,
- 3 – 50–249 employees,
- 4 – 250–499 employees,
- 5 – more than 499 employees.

According to the literature [35, 36], employees of smaller enterprises are at greater risk than employees of larger enterprises in terms of the severity of occupational accidents (including loss of life). In order to test the applicability of the above statement to the analyzed data and to propose a possible aggregation of enterprise size classifications, multiple comparison tests for equality of proportions of seriously injured casualties were conducted. Because tests of multiple comparisons are not independent, the Bonferroni correction (considered to be relatively conservative) was applied [37, 38]. The results of the tests in each pair of enterprise size are summarized in Table 4. The null hypotheses H0 state that proportions in both populations are the same.

The *p*-values, calculated after the Bonferroni correction, imply that in six out of fifteen cases the null hypothesis of the equality of proportions should be rejected. The test results indicate the possibility of creating two variants of the enterprise size aggregation:

- (0, 1, and 2) for the test numbers: 1, 2, and 6, respectively, and (3, 4, and 5) for the test numbers: 13, 14, and 15, respectively;
- (1 and 2) for the test number: 6, and (0, 3, 4, and 5) for the test numbers: 3, 4, 5, 13, 14, and 15, respectively.

Taking into account the logic and meaning of the values, it was proposed to classify the size of an enterprise according to the pair of aggregates indicated in the first variant and to create the variable *WK* having two values (the following notations are used later in the paper):

- wk1; micro and small enterprises employing up to 49 people;
- wk2; medium and large enterprises employing 50 people or more.

In the work, attributes characterizing the accidents casualties are marked with the symbol P*xx*, where *xx* stands for the number of an item from the

Table 4. Test for equality of proportions of persons seriously injured in accidents at work by enterprise size

Test number	Pairwise population by enterprise size	Chi-square statistic	*p*-value	*p*-value after Bonferroni correction
1	0 and 1	0.2618	0.6089	1
2	0 and 2	0.9310	0.3346	1
3	0 and 3	2.8459	0.0916	1
4	0 and 4	4.3015	0.0381	0.5712
5	0 and 5	3.7516	0.0528	0.7913
6	1 and 2	5.6738	0.0172	0.2583
7	1 and 3	41.7432	<.0001	<.0001*
8	1 and 4	51.5309	<.0001	<.0001*
9	1 and 5	44.0053	<.0001	<.0001*
10	2 and 3	30.1105	<.0001	<.0001*
11	2 and 4	38.5189	<.0001	<.0001*
12	2 and 5	30.7957	<.0001	<.0001*
13	3 and 4	5.2654	0.0218	0.3263
14	3 and 5	2.3656	0.1240	1
15	4 and 5	0.4127	0.5206	1

Note: * The null hypothesis is rejected at $\alpha = 0.05$.

statistical accident card (for example, P02 means the age of the injured person). Table 5 provides a summary of the variables considered in the LCA modeling. The information in the table is divided into two data sets: wk1 and wk2, according to the enterprise size (the *WK* variable), as explained earlier. Observations that did not provide information were removed (for example, P09 – *Injured body part = Unknown or undefined*). Data transformation was proposed, mainly the aggregation of values or variables, which helped to solve the problem of rare categories. In this process, the substantive meaning of variables was taken into account. The presence of heterogeneous categories made it possible to examine how the LCA model deals with such a set of data. The P289 variable, which speaks for *Casualty injury severity*, is the result of the aggregation of two attributes: P28 (*Accident consequence*) and P29 (*Inability to work*).

LCA MODELS BY ENTERPRISE SIZE – THE RESULTS

The algorithm for building the LCA models and selecting relevant variables in the process of the model's creation (Fig. 3) was used for both the data sets wk1 and wk2. The number of latent classes varied from $C_{min} = 2$ to $C_{max} = 15$. The following identifiers for the resulting models were used in the process: wk1-LCA2,..., wk1-LCA15 for the wk1 data set and wk2-LCA2,..., wk2-LCA15 for the wk2 data set, where the number at the end of the identifier indicates the number of latent classes in the model. The cut-off value AR^* for the observed variables was assumed to be 0.2 (this constitutes the variable selection criterion). Out of 12 indicators (Table 5), the criterion was met by 9 (P02, P05, P06, P08, P09, P21, P26,

Table 5. Characteristics of the research data

Observed variables and their descriptive values (The information is taken from the Polish statistical accident card)	Value codes	*wk*1 data set [%]	*wk*2 data set [%]
P01 – Casualty gender			
Male	1	92.75	82.23
Female	2	7.25	17.77
P02 – Casualty age			
Up to 24 years old	1	17.42	17.28
25–34 years old	2	30.21	30.23
35–44 years old	3	24.95	25.60
45–54 years old	4	18.11	18.89
Over 54 years	5	9.32	8.01

Table 5. Cont.

P05 – Casualty occupation			
Industrial workers, craftsmen, and employees doing simple works	1	78.24	63.67
Operators and assemblers of machines and devices	2	21.76	36.33
P06 – Enterprise job seniority			
Up to 5 years (inexperienced workers)	1	69.56	66.86
6–10 years (mature workers)	2	15.37	16.23
Over 10 years (veteran workers)	3	15.07	16.91
P07 – Hours from start of work to accident			
Below 4	1	46.35	47.06
4–7	2	48.98	48.20
8 and more	3	4.67	4.74
P08 – Injury type			
Wounds and superficial injuries	1	52.43	57.36
Bone fractures	2	18.69	15.25
Displacements, dislocations, sprains and strains	3	6.95	12.86
Traumatic amputations (loss of body parts)	4	13.28	4.85
Various other injuries – a heterogeneous category, that includes various injury types	5	8.66	9.68
P09 – Injured body part			
Head, neck	1	5.64	7.52
Body – a heterogeneous category, that includes: Thoracic and lumbar spine, Torso and internal organs, Whole body and its various parts, Other body part	2	3.91	5.21
Upper limbs	3	71.60	65.62
Lower limbs	4	18.85	21.65
P16 – Time of year			
Spring months	1	25.10	26.20
Summer months	2	25.59	24.68
Autumn months	3	24.31	24.97
Winter months	4	25.00	24.15
P21 – Activity performed at accident time			
Operating machinery	1	55.52	43.52
Working with tools and objects	2	25.69	30.56
Transport at the workplace	3	12.72	15.39
Presence at the accident scene without doing work	4	6.08	10.52
P26 – Material factor as injury source			
Buildings, structures, surfaces	1	5.31	7.44
Another factor – a heterogeneous category, that includes various injury sources	2	8.91	9.62
Hand tools	3	10.16	8.96
Machines and devices	4	46.99	37.23
Materials, objects, products, machine parts	5	28.63	36.74
P27 – Main accident cause			
Defect of material factor	1	21.50	14.35
Misuse of material factor	2	13.94	13.98
Inappropriate work organization	3	10.70	12.03
Safety neglect	4	53.86	59.64
P289 – Casualty injury severity; a new variable, defined on the basis of the variables: P28 (Accident consequence) and P29 (Inability to work)			
Slight – accident resulting in inability to work for 0–29 days	1	36.75	50.26
Medium – accident resulting in inability to work for 30–89 days	2	43.18	35.90
Serious – severe or fatal accident or accident causing inability to work for at least 90 days	3	20.07	13.84
Number of observations		**3916**	**9822**

Table 6. Quality measures for the LCA models

Model	G^2	BIC	CAIC	ABIC	E
wk1-LCA2	14904.89	15343.35	15396.35	15174.94	0.71
wk1-LCA3	13890.79	14552.62	14632.62	14298.42	0.70
wk1-LCA4	13499.92	14385.11	14492.11	14045.12	0.74
wk1-LCA5	13145.62	14254.18	14388.18	13828.39	0.74
wk1-LCA6	12811.57	14143.49	14304.49	13631.91	0.72
wk1-LCA7	12536.30	14091.59	14279.59	13494.22	0.70
wk1-LCA8	12339.38	14118.04	14333.04	13434.86	0.70
wk1-LCA9	12216.82	14218.84	14460.84	13449.87	0.71
wk1-LCA10	12095.05	14320.44	14589.44	13465.68	0.71
wk1-LCA11	11982.94	14431.70	14727.70	13491.15	0.71
wk1-LCA12	11889.69	14561.81	14884.81	13535.46	0.72
wk1-LCA13	11807.23	14702.72	15052.72	13590.58	0.72
wk1-LCA14	11746.45	14865.31	15242.31	13667.37	0.71
wk1-LCA15	11660.59	15002.81	15406.81	13719.08	0.73
wk2-LCA2	20135.87	20567.91	20614.91	20418.55	0.71
wk2-LCA3	17455.41	18108.06	18179.06	17882.44	0.65
wk2-LCA4	16421.00	17294.27	17389.27	16992.38	0.67
wk2-LCA5	15477.94	16571.83	16690.83	16193.67	0.68
wk2-LCA6	14662.84	15977.35	16120.35	15522.92	0.69
wk2-LCA7	14172.12	15707.25	15874.25	15176.55	0.70
wk2-LCA8	13777.77	15533.51	15724.51	14926.54	0.67
wk2-LCA9	13455.09	15431.45	15646.45	14748.21	0.68
wk2-LCA10	13138.21	15335.19	15574.19	14575.68	0.70
wk2-LCA11	12898.51	15316.10	15579.10	14480.33	0.70
wk2-LCA12	12675.40	15313.61	15600.61	14401.57	0.70
wk2-LCA13	12486.63	15345.46	15656.46	14357.15	0.72
wk2-LCA14	12321.46	15400.91	15735.91	14336.33	0.72
wk2-LCA15	12192.68	15492.74	15851.74	14351.89	0.72

P27, and P289) and 8 (P02, P05, P06, P08, P09, P21, P26, and P289) variables for wk1 and wk2, respectively. Of the resulting models, the best model was selected for each enterprise group.

Table 6 and scree plots in Figure 4 present the quality measures for the wk1 models. The numbers of latent classes in the LCA model are marked on the horizontal axis (they also indicate the numbers of the models). The decrease in the values of BIC and CAIC measures is relatively gentle starting from the wk1-LCA4 model. The lowest values of these measures occur for the wk1-LCA7 model. For the ABIC measure, the mild decline starts from the wk1-LCA7 model, reaching a minimum for the wk1-LCA8 model. The entropy-based E index value is at least 0.70, indicating the sufficient class differentiability in each model [39]. Figure 5 shows the values of the AR index for the considered nine observed variables, in relation to each C-class model. Starting with the wk1-LCA6 model, six variables form a group that plays a major role in the discrimination of latent classes: P21, P26, P08, P06, P09 and P289.

Table 6 and scree plots in Figure 6 present the quality measures for the wk2 models. The decline in BIC and CAIC is relatively mild starting from the wk2-LCA7 model. For the ABIC measure, the mild decline starts from the wk2-LCA10 model, reaching a minimum for the wk1-LCA14 model. The entropy-based E index value in each model is at least 0.60, which speaks for sufficient class differentiability [39]. Figure 7 shows the values of the AR index for the considered eight observed variables, in relation to each C-class model. Starting with the wk2-LCA6 model, six variables

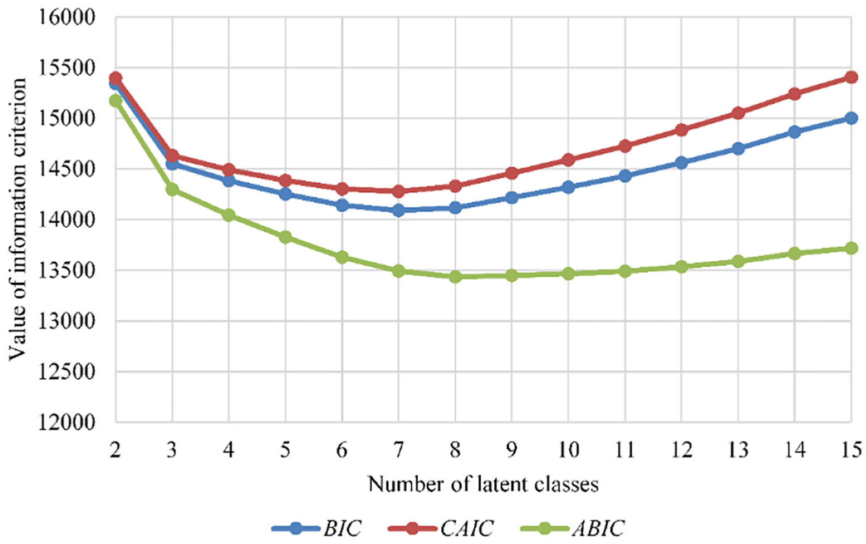

Figure 4. Scree plots of information measures for the wk1 LCA models

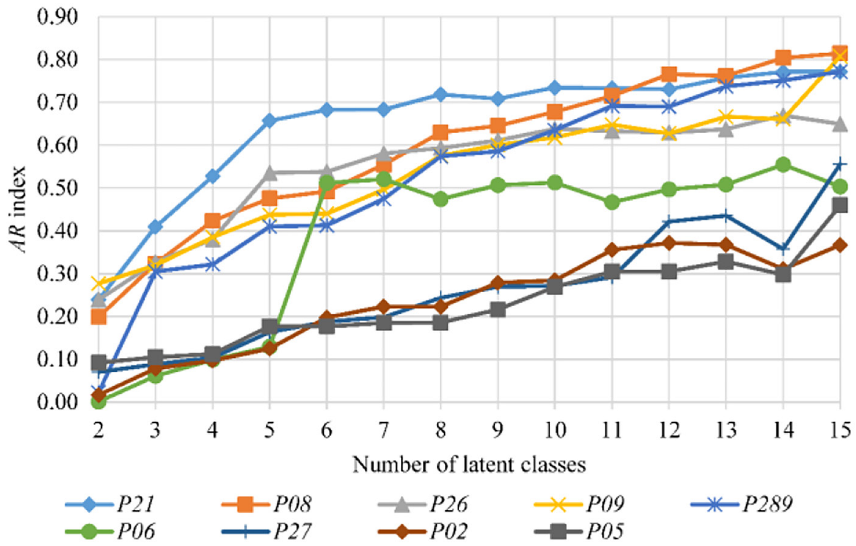

Figure 5. AR values of observed variables in the wk1 LCA models

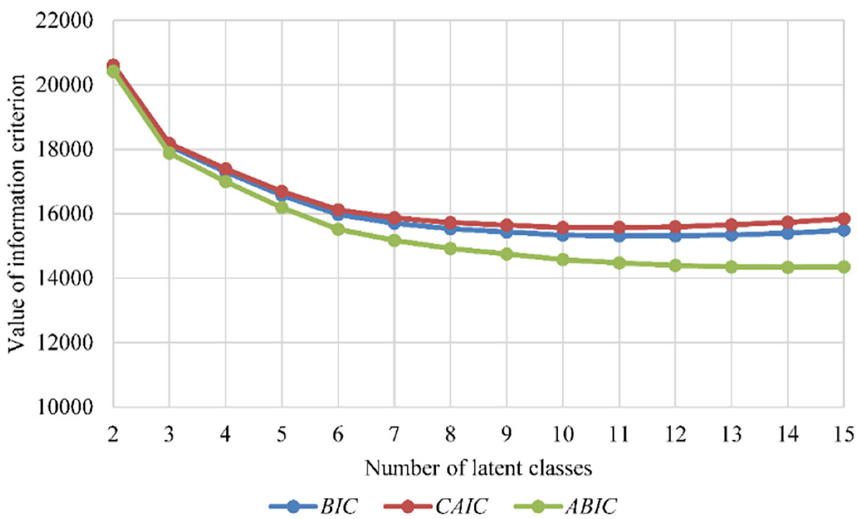

Figure 6. Scree plots of information measures for the wk2 LCA models

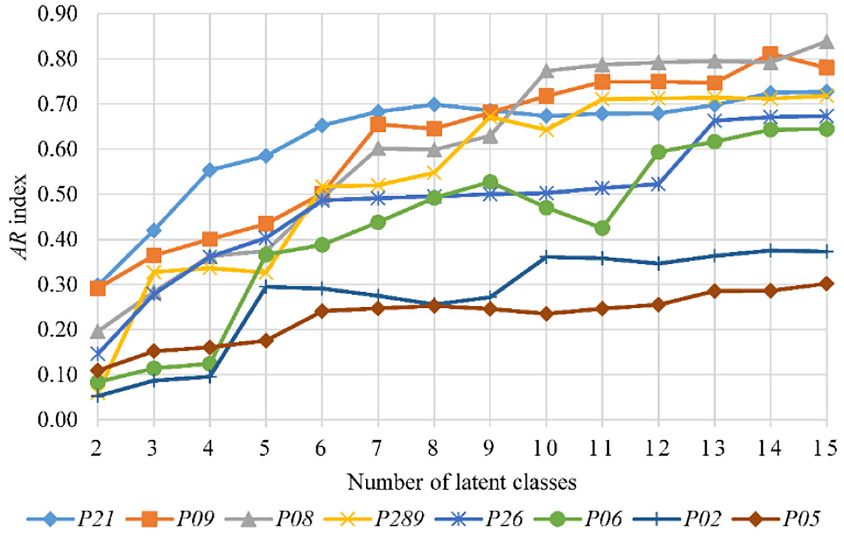

Figure 7. AR values of observed variables in the wk2 LCA models

form a group that plays largest major role in the discrimination of latent classes: P21, P09, P08, P289, P26 and P06.

On the basis of all the considered measures and, additionally, the insight into the estimated parameters, the wk1-LCA7 and wk2-LCA8 models were selected for further analysis. Their qualitative assessment is considered to be better than that of other models for respective data sets. The two obtained LCA models for the two groups of manufacturing enterprises illustrated in the form of heat map are shown in Table 7. The map presents synthetic diagnosis of occupational safety threats. The layout of the map is defined by the ordering of the analyzed observed variables according to their identifiers. The bolded columns show the empirical distributions of the analyzed variables and provide references for the other part of the map. The map cells represent the estimators of the conditional probabilities $\rho_{q_j|c}$. The estimators of the probabilities γ_c for the latent classes are given at the top of the maps.

OCCUPATIONAL ACCIDENTS BY ENTERPRISE SIZE – THE DISCUSSION

The number of latent classes varies between the two groups of enterprises. Occupational accidents profiles for micro and small enterprises (wk1) are identified by the 7-class model wk1-LCA7 estimated on the basis of 9 observed variables; its latent classes (clusters) were given the following–labels (identifiers): wk1-K1,..., and wk1-K7. In the case of medium and large enterprises (wk2), such profiles are identified by the 8-class model wk2-LCA8, for which 8 observed variables were used; compared to the wk1-LCA7 model, the P27 variable is absent. In this case, the wk2-LCA8 model latent classes were given the following labels: wk2-K1,..., and wk2-K8. The first part of each identifier (wk1 or wk2) informs about the model to which a class belong. The second part (K1 to K7 or K1 to K8) is used to indicate a particular cluster in the model.

In the two models, the relevance of the observed variables, determined by the values of the AR index, is different. Also, the importance ranking of the variables is different – see Figure 8. On the basis of the estimated conditional probabilities for the P289 variable, the latent classes obtained in the wk1-LCA7 and wk2-LCA8 models were assigned to one of three groups as follows, reflecting the degree of casualty injury:

a) mild accident category when $P(P289 = 3) < 0.1$ and $P(P289 = 1) > 0.5$; contains latent classes: wk1-K3, wk1-K7, wk2-K2. wk2-K4, wk2-K5, and wk2-K6,
b) intermediate accident category when it is neither mild nor alert accident category; contains latent classes: wk1-K4, wk1-K5, wk1-K6. wk2-K1, and wk2-K3,
c) alert accident category when $P(289 = 3) > 0.25$ and $P(P289 = 2) + P(289 = 3) > 0.5$; contains latent classes: wk1-K1, wk1-K2, wk2-K7, and wk2-K8.

The proposed classification is not the same as the values of the observed variable P289 and the above

Table 7. The heat maps of occupational accident casualty profiles for the wk1-LCA7 and wk2-LCA8 models

Estimated $\gamma_c \rightarrow$			0.066	0.219	0.187	0.118	0.150	0.164	0.096		0.048	0.150	0.104	0.139	0.180	0.124	0.138	0.118
Variable	Value	wk1	wk1-K1	wk1-K2	wk1-K3	wk1-K4	wk1-K5	wk1-K6	wk1-K7	wk2	wk2-K1	wk2-K2	wk2-K3	wk2-K4	wk2-K5	wk2-K6	wk2-K7	wk2-K8
P02	1	**0.17**	0.09	0.22	0.31	0.12	0.15	0.00	0.27	**0.17**	0.06	0.00	0.05	0.32	0.33	0.27	0.11	0.10
	2	**0.30**	0.25	0.31	0.37	0.33	0.34	0.13	0.38	**0.30**	0.23	0.13	0.26	0.43	0.39	0.37	0.24	0.30
	3	**0.25**	0.27	0.21	0.19	0.28	0.25	0.36	0.21	**0.26**	0.33	0.38	0.32	0.17	0.15	0.22	0.29	0.26
	4	**0.18**	0.25	0.17	0.10	0.18	0.18	0.32	0.09	**0.19**	0.26	0.31	0.25	0.07	0.10	0.11	0.24	0.24
	5	**0.09**	0.13	0.09	0.03	0.09	0.08	0.19	0.05	**0.08**	0.12	0.18	0.12	0.01	0.02	0.03	0.11	0.10
P05	1	**0.78**	0.65	0.81	0.79	0.69	0.77	0.84	0.83	**0.64**	0.61	0.63	0.49	0.62	0.68	0.61	0.74	0.65
	2	**0.22**	0.35	0.19	0.21	0.31	0.23	0.16	0.17	**0.36**	0.39	0.37	0.51	0.38	0.32	0.39	0.26	0.35
P06	1	**0.70**	0.61	0.87	0.92	0.67	0.76	0.14	0.80	**0.67**	0.61	0.20	0.50	0.92	0.94	0.80	0.63	0.64
	2	**0.15**	0.15	0.09	0.07	0.20	0.13	0.32	0.14	**0.16**	0.18	0.31	0.21	0.08	0.06	0.13	0.20	0.17
	3	**0.15**	0.24	0.04	0.00	0.12	0.11	0.54	0.07	**0.17**	0.21	0.49	0.29	0.01	0.00	0.07	0.17	0.19
P08	1	**0.52**	0.08	0.31	0.92	0.59	0.13	0.66	0.83	**0.57**	0.14	0.85	0.12	0.67	0.84	0.88	0.47	0.08
	2	**0.19**	0.35	0.16	0.01	0.02	0.65	0.12	0.07	**0.15**	0.05	0.03	0.23	0.03	0.07	0.03	0.15	0.67
	3	**0.07**	0.51	0.01	0.03	0.01	0.16	0.01	0.02	**0.13**	0.10	0.06	0.60	0.18	0.03	0.02	0.01	0.16
	4	**0.13**	0.00	0.44	0.01	0.00	0.03	0.17	0.03	**0.05**	0.00	0.00	0.01	0.00	0.01	0.01	0.31	0.02
	5	**0.09**	0.06	0.07	0.02	0.38	0.03	0.04	0.06	**0.10**	0.71	0.06	0.04	0.11	0.06	0.07	0.06	0.07
P09	1	**0.06**	0.01	0.01	0.02	0.33	0.01	0.04	0.04	**0.08**	0.23	0.15	0.02	0.13	0.04	0.11	0.00	0.00
	2	**0.04**	0.09	0.02	0.00	0.20	0.03	0.00	0.00	**0.05**	0.64	0.03	0.04	0.05	0.01	0.01	0.00	0.03
	3	**0.72**	0.18	0.95	0.88	0.21	0.53	0.90	0.82	**0.66**	0.03	0.66	0.16	0.44	0.91	0.81	0.98	0.66
	4	**0.19**	0.72	0.03	0.10	0.25	0.43	0.06	0.14	**0.22**	0.10	0.16	0.78	0.37	0.04	0.07	0.02	0.31
P21	1	**0.56**	0.08	0.88	0.76	0.38	0.31	0.77	0.00	**0.44**	0.35	0.49	0.10	0.23	0.81	0.01	0.85	0.32
	2	**0.26**	0.15	0.12	0.17	0.27	0.23	0.18	0.98	**0.31**	0.22	0.31	0.06	0.21	0.17	0.98	0.13	0.36
	3	**0.13**	0.21	0.00	0.07	0.21	0.44	0.05	0.01	**0.15**	0.31	0.15	0.21	0.41	0.01	0.01	0.01	0.28
	4	**0.06**	0.56	0.01	0.00	0.14	0.02	0.00	0.01	**0.11**	0.12	0.05	0.63	0.15	0.00	0.00	0.01	0.04
P26	1	**0.05**	0.50	0.00	0.01	0.07	0.05	0.00	0.01	**0.07**	0.06	0.02	0.49	0.06	0.00	0.02	0.00	0.06
	2	**0.09**	0.28	0.03	0.07	0.19	0.10	0.04	0.06	**0.10**	0.23	0.11	0.23	0.12	0.05	0.04	0.03	0.08
	3	**0.10**	0.01	0.06	0.04	0.05	0.01	0.04	0.71	**0.09**	0.03	0.06	0.00	0.02	0.01	0.51	0.01	0.07
	4	**0.47**	0.11	0.82	0.67	0.19	0.13	0.70	0.00	**0.37**	0.19	0.36	0.18	0.26	0.65	0.00	0.81	0.21
	5	**0.29**	0.09	0.09	0.21	0.50	0.72	0.22	0.21	**0.37**	0.49	0.45	0.10	0.55	0.28	0.42	0.14	0.57
P27	1	**0.22**	0.09	0.30	0.16	0.27	0.12	0.29	0.16									
	2	**0.14**	0.02	0.15	0.13	0.10	0.19	0.15	0.18									
	3	**0.11**	0.14	0.09	0.09	0.17	0.16	0.06	0.07									
	4	**0.54**	0.76	0.46	0.63	0.46	0.52	0.50	0.59									
P289	1	**0.37**	0.27	0.10	0.59	0.62	0.18	0.32	0.68	**0.50**	0.47	0.69	0.33	0.81	0.61	0.82	0.09	0.06
	2	**0.43**	0.40	0.48	0.41	0.20	0.62	0.50	0.27	**0.36**	0.30	0.26	0.47	0.16	0.36	0.16	0.51	0.67
	3	**0.20**	0.33	0.42	0.00	0.18	0.20	0.18	0.05	**0.14**	0.22	0.05	0.20	0.03	0.03	0.02	0.40	0.27

percentage threshold values depend on the researcher's decision in the context of data specificity.

Figure 9 shows the distribution of the adopted classification by the enterprise size, considering the estimated probability of occurrence of the latent classes. In the group of medium and large enterprises (wk2), the probability of occurrence of accident profiles classified in the mild category is highest, close to 0.6, while in the group of micro and small enterprises (wk1) it is lowest, slightly exceeding 0.28. The probability of occurrence of accident profiles from the intermediate category is lowest in the group of medium and large enterprises, amounting to 0.15, while in micro and small enterprises it is highest, at the level of 0.43. Both groups of enterprises have similar probabilities of occurrence of accident profiles from the alert category, although in the group of micro and small enterprises it is slightly higher (0.28 for wk1 vs. 0.26 for wk2). Micro and small wood processing enterprises have a considerably higher threat of more severe injuries in occupational accidents than medium and large enterprises (71.7% for wk1 vs. 40.8% for wk2); this conclusion is in consent with the outcome of some other studies [14, 17]. On the basis of the heat maps, the accident profiles were characterized as follows.

Mild accident category

Similar patterns occur in both groups of enterprises, with wounds and superficial injuries predominating. They contain the following classes:

- wk1-K3 and wk2-K5, which can be labeled as a light severity of injury involving upper limbs during machine operation among inexperienced workers,
- wk1-K7 and wk2-K6, which can be labeled as a very light severity of injury involving upper limbs when working with tools or objects among inexperienced workers.

In micro and small enterprises, neglect of safety is the main cause of accidents. In medium and large enterprises, there are additional patterns: wk2-K2 – a very light severity of injury involving upper limbs during the production process among experienced workers, and wk2-K4 – a very light severity of injury involving limbs during various activities among inexperienced workers.

Intermediate accident category

In each of the wk1 and wk2 groups of enterprises, there is a fuzzy, heterogeneous profile of accidents at work, which is assigned to the intermediate accident category, represented by the wk1-K4 and wk2-K1 latent classes, respectively. In these two clusters, the differentiation of conditional probabilities for the observed variables is smaller than in the remaining latent classes of both models wk1-LCA7 and wk2-LCA8. The clusters are also characterized by a significant share of heterogeneous categories. Differences between classes wk1-K4 and wk2-K1 are noticeable in terms of the estimated distributions of the observed variables: *Injury type* (P08) and *Injured body part* (P09).

The remaining classes in the wk1-LCA7 model differ from the remaining class in the wk2-LCA8 model:

- wk1-K5 can be labeled as a moderate severity of limb injuries (often fractures) during the production process or its support among inexperienced workers; wk1-K6 can be labeled as a moderate severity of upper limb injuries (often lacerations and superficial injuries) when operating machines among experienced workers,
- wk2-K3 can be labeled as a moderate severity of lower limb injuries (often dislocations, sprains, sprains and strains) during work not directly related to the production processes.

Alert accident category

There are similar profiles in both groups of enterprises as regards alert accident category. They are represented by classes wk1-K2 and wk2-K7, which can be labeled as a very high severity of upper limb injuries (often amputations) while operating machinery. However, the estimated probability for inexperienced workers is distinctly greater in wk1-K2 than in wk2-K7.

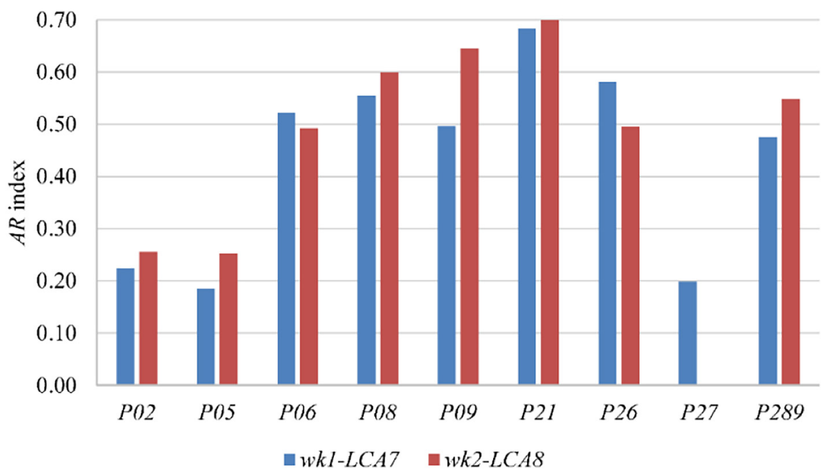

Figure 8. AR values of observed variables in the wk1-LCA7 and wk2-LCA8 models

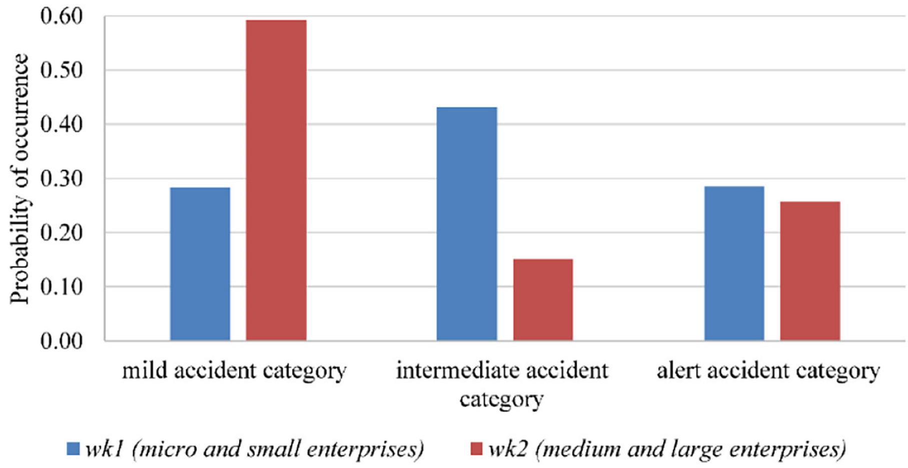

Figure 9. Distribution of occupational accident profiles classification by enterprise size

There are also different alert accident profiles in the two groups of enterprises:

- wk1-K1, which can be labeled as a high severity of lower limb injuries (usually dislocations, sprains, and tears) during work not directly related to the production process,
- wk2-K8, which can be labeled as a high severity of injuries to upper limbs (usually fractures) during the production processes or their support.

CONCLUSIONS

The work covers the issues related to the diagnosis of selected occupational safety hazards in wood processing enterprises at the country level – Poland. Data on accidents at work provided by Statistics Poland, which include individual records characterizing the accident casualties, were used.

Opinions of other researchers that the size of an enterprise has an impact on the severity of the consequences of accidents at work have been confirmed through the test for equality of proportions with the Bonferroni correction. On the basis of the test results, a division of the data set into two subsets was proposed as follows: data related to micro and small enterprises, employing up to 49 people, and data related to medium and large enterprises, employing 50 people or more.

The identification of profiles of occupational accident casualties was carried out using the LCA data mining technique. The process of building the LCA model was carried out independently for each of the two groups of enterprises, resulting in latent classes that defined accident profiles. A proprietary algorithm indicating the observed variables that, due to their high discriminatory property should be included in the LCA model, was proposed. Occupational accident profiles were illustrated graphically in the form of heat maps. Latent classes were individually labeled and compared in groups formed according to the degree of harm to the casualty. The LCA method proved to be a good tool in analyzing qualitative occupational accident data. It made it possible to extract rational patterns of accidents at work. General conclusions drawn on the basis of the research are presented below.

1. In terms of the quantity aspect of occupational accident threats, micro and small enterprises (wk1) differ from medium and large enterprises (wk2). The probability of the alert and intermediate accident category is significantly higher for the former (71.7%) than for the latter (40.8%). Thus, employees of smaller enterprises (wk1) may be at greater risk of suffering more serious injury from an accident at work than employees of larger enterprises (wk2).
2. In both groups of enterprises, the most severe work accidents are related to operating machinery. Such incidents mainly affect workers with little job seniority and injuries involve particularly upper limbs. The casualties often suffer from traumatic amputations.
3. In micro and small enterprises, high injury severity of accidents occurs during activity not directly related to production process, but due to the worker finding himself at the accident site. Displacements, dislocations, sprains, and strains of lower limbs were the most common injuries.
4. In the case of medium and large enterprises, high injury severity was characterized by accidents

occurring during the production process or its support. Injuries, mostly fractures, were sustained to limbs (more often upper ones).

Despite the introduction of legal regulations for the protection of workers' health and life in Poland, they may be insufficient for high-risk production enterprises, such as wood processing ones. In the light of the obtained results, it also appears that certain health and safety requirements may not be fully met, or the weak link may be the employee himself (not always through their own fault). Enterprises should be targeted with appropriate information and prevention programs, as well as appropriate inspection activities by competent institutions. In particular, it seems reasonable to extend increased/special attention to smaller enterprises.

Acknowledgement

The authors thank the Statistics Poland for providing individual raw data for the research. This research was supported by: Kielce University of Technology, Faculty of Management and Computer Modelling, Kielce, Poland, grant number 04.0.06.00/1.02.001 SUBB.ZKIS.24.001.

REFERENCES

1. Józwik J., Pietras P. Investigation and assessment of occupational risk on the metal cutting machine tool stand. Advances in Science and Technology Research Journal 2013; 7(20): 47–54, doi: 10.5604/20804075.1073057.
2. Yilmaz F., Ozcan M.S. A risk analysis and ranking application for lifting vehicles used in construction sites with integrated AHP and Fine-Kinney approach. Advances in Science and Technology Research Journal 2019; 13(3): 152–161, doi: 10.12913/22998624/111779.
3. Clarivate. KeyWords Plus generation, creation, and changes, from http://support.clarivate.com/ScientificandAcademicResearch/s/article/KeyWords-Plus-generation-creation-and-changes?language=en_US (Accessed: 10.02.2023).
4. Elsevier. Scopus Content Coverage Guide, from https://www.elsevier.com/?a=69451 (Accessed: 15.04.2023).
5. Boyack K.B. Mapping knowledge domains: Characterizing PNAS. In: Proceedings of the National Academy of Sciences of the United States of America National Academy of Sciences, USA, 2004; 101(1): 5192–5199.
6. Van Eck N.J., Waltman L. VOSviewer Manual. Manual for VOSviewer version 1.6.15, Universiteit Leiden, Leiden, Netherlands, from https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.15.pdf (Accessed: 11.12.2022).
7. BydTermCymru. The meaning of 'term' and 'terminology standardization'. Llywodraeth Cymru Welsh Government, from https://gov.wales/bydtermcymru/how-to-use/meaning-term-and-terminology-standardization (Accessed: 12.09.2019).
8. Brzozowska-Rup K., Nowakowska M. Bibliometric Studies on Renewable Energy—Poland Compared to Other EU Countries. Energies 2022; 15(13): 4577, doi: 10.3390/en15134577.
9. Varonen U., Mattila M. The safety climate and its relationship to safety practices, safety of the work environment and occupational accidents in eight wood-processing companies. Accident Analysis & Prevention 2000; 32(6): 761–769, doi: 10.1016/s0001-4575(99)00129-3.
10. Alamgir H., Koehoorn M., Ostry A., Tompa E., Demers P. An evaluation of hospital discharge records as a tool for serious work related injury surveillance. Occupational and Environmental Medicine 2006; 63(4): 290–296, doi: 10.1136%2Foem.2005.026047.
11. Holcroft C.A., Punnett L. Work environment risk factors for injuries in wood processing. Journal of Safety Research 2009; 40(4): 247–255, doi: 10.1016/j.jsr.2009.05.001.
12. Palamara F., Piglione F., Piccinini N. Self-Organizing Map and clustering algorithms for the analysis of occupational accident databases. Safety Science 2011; 49(8–9): 1215–1230, doi: 10.1016/j.ssci.2011.04.003.
13. Thepaksorn P., Thongjerm S., Incharoen S., Siriwong W., Harada K., Koizumi A. Job safety analysis and hazard identification for work accident prevention in para rubber wood sawmills in southern Thailand. Journal of Occupational Health 2017; 59(6): 542–551, doi: 10.1539/joh.16-0204-CS.
14. Tremblay A., Badri A. A novel tool for evaluating occupational health and safety performance in small and medium-sized enterprises: The case of the Quebec forestry/pulp and paper industry. Safety Science 2018; 101: 282–294, doi: 10.1016/j.ssci.2017.09.017.
15. Comberti L., Demichela M., Baldissone G., Fois G., Luzzi R. Large occupational accidents data analysis with a coupled unsupervised algorithm: The S.O.M. K-Means Method. An Application to the Wood Industry. Safety 2018; 4(4): 51, doi: 10.3390/safety4040051.
16. Aragón-Vásquez A.Y., Silva-Lugo E.D., Nájera-Luna J.A., Hernández-Díaz J.C., Hernández

F.J., Cruz-Carrera R.D. Perception of occupational risk factors in sawmills in the El Salto region of Durango, Mexico. Revista Chapingo Serie Ciencias Forestales y del Ambiente 2019; 25(2): 253–268, doi: 10.5154/r.rchscfa.2019.01.005.

17. Karademir D., Koc K. Evaluating the work environment in turkish furniture industry from the point of occupational health and safety. Fresenius Environmental Bulletin 2020; 29(4A): 2639–2646.
18. Mulugeta H., Tefera Y., Gezu M. Nonfatal occupational injuries among workers in microscale and small-scale woodworking enterprise in Addis Ababa. Ethiopia. Journal of Environmental and Public Health 2020, 2020: 6407236, doi: 10.1155/2020/6407236.
19. Nowakowska M., Pajęcki M. Latent class analysis for identification of occupational accident casualty profiles in the selected Polish manufacturing sector. Advances in Production Engineering & Management 2021; 16(4): 485–499, doi: 10.14743/apem2021.4.415.
20. Araújo-Vila N., Toubes D.R., Fraiz-Brea J.A. The age factor in the analysis of occupational risks in the wood industry. Healthcare 2022; 10(7): 1355, doi: 10.3390/healthcare10071355.
21. Regulation of the Council of Ministers of 24 December 2007 on the Polish Classification of Activities (PKD), Journal of Laws of 2007, no. 251, item 1885, as amended.
22. Tei J.Y., Coxe S., Cham H. Statistical Power to Detect the Correct Number of Classes in Latent Profile Analysis. Structural Equation Modeling 2013; 20(4): 640–657, doi: 10.1080/10705511.2013.824781.
23. Collins L., Lanza S. Latent class and latent transaction analysis: with applications in the social, behavioral, and health sciences. A John Wiley & Sons. Inc., Hoboken, New Jersey, USA, 2010, doi: 10.1002/9780470567333.
24. Lanza S.T., Bray B.C., Collins L.M. An introduction to latent class and latent transition analysis. In: Schinka J.A., Velicer W.F., Weiner I.B. (ed.), Handbook of psychology, 2nd ed., Wiley, Hoboken, New Jersey, USA, 2013; 2: 691–716.
25. Weller B.E., Bowen N.K., Faubert S.J. Latent class analysis: a guide to best practice. Journal of Black Psychology 2020; 46(4): 287–311, doi: 10.1177/0095798420930932.
26. Masyn K.E. Latent class analysis and finite mixture modelling. In: Little T.D. (ed.), The Oxford handbook of quantitative methods: Statistical analysis, Oxford University Press, New York, NY, USA, 2013: 551–611, doi: 10.1093/oxfordhb/9780199934898.013.0025.
27. Dziak J.J., Coffman D.L., Lanza S.T., Li R., Jermiin L.S. Sensitivity and specificity of information criteria. Briefings in Bioinformatics 2020; 21(2): 553–565, doi: 10.1093/bib/bbz016.
28. Nylund-Gibson K., Choi A.Y. Ten frequently asked questions about latent class analysis. Translational Issues in Psychological Science 2018; 4(4): 440–461, doi: 10.1037/tps0000176.
29. Dziak J.J., Lanza S.T., Tan X. Effect size. statistical power and sample size requirements for the bootstrap likelihood ratio test in latent class analysis. Structural Equation Modeling 2014; 21(4): 534–552, doi: 10.1080/10705511.2014.919819.
30. Nylund K.L., Asparouhov T., Muthén B.O. Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. Structural Equation Modeling; 14(4): 535–569, doi: 10.1080/10705510701575396.
31. Killian M.O., Cimino A.N., Weller B.E., Hyun Seo C. A systematic review of latent variable mixture modeling research in social work journals. Journal of Evidence-Based Social Work 2019; 16(2): 192–210, doi: 10.1080/23761407.2019.1577783.
32. Petersen K.J., Qualter P., Humphrey N. The application of latent class analysis for investigating population child mental health: a systematic review. Frontiers in Psychology 2019; 10(1214), doi: 10.3389/fpsyg.2019.01214.
33. Lanza S.T., Dziak J.J., Huang L., Wagner A.T., Collins L.M. Proc LCA & Proc LTA Users' Guide (Version 1.3.2), PennState: The Methodology Center, The Pennsylvania State University, Philadelphia, USA, from https://www.methodology.psu.edu/ (Accessed: 20.01.2021).
34. Kim S.-Y. Determining the number of latent classes in single- and multiphase growth mixture models. Structural Equation Modeling 2014; 21(2): 263–279, doi: 10.1080/10705511.2014.882690.
35. Abdalla S., Apramian S.S., Cantley L.F., Cullen M.R. Occupation and Risk for Injuries. In: Mock C.N., Nugent R. Kobusingye O., Smith K.R. (ed.), Injury Prevention and Environmental Health, 3rd edition, Chapter 6, The International Bank for Reconstruction and Development, The World Bank, Washington (DC), USA, 2017: 97–132, doi: 10.1596/978-1-4648-0522-6_ch6.
36. Sinclair R.C., Cunningham T.R. Safety activities in small businesses. Safety Science 2014; 64: 32–38, doi: 10.1016/j.ssci.2013.11.022.
37. Haynes W. Bonferroni Correction. In: Dubitzky W., Wolkenhauer O., Cho K.H., Yokota H. (ed.), Encyclopedia of Systems Biology, Springer, New York, NY, USA, 2013 from https://doi.org/10.1007/978-1-4419-9863-7_1213 (Accessed: 16.03.2023).
38. Gelman A., Hill J., Yajima M. Why we (usually) don't have to worry about multiple comparisons. Journal of Research on Educational Effectiveness 2012; 5(2): 189–211, doi: 10.1080/19345747.2011.618213.
39. Asparouhov T., Muthén B. Auxiliary variables in mixture modeling: three-step approaches using mplus. Structural Equation Modeling 2014; 21(3): 329–341, doi: 10.1080/10705511.2014.915181.