
The Application of Association Rules to Detect the Effects of Vaccinations against Covid-19 in the EU-27. Preliminary Estimates

Kateryna Berezka

West Ukrainian National University, Ternopil, Ukraine

e-mail: km.berezka@gmail.com

ORCID: 0000-0002-9632-4004

Olha Kovalchuk

West Ukrainian National University, Ternopil, Ukraine

e-mail: Olhakov@gmail.com

ORCID: 0000-0001-6490-9633

© 2023 Kateryna Berezka, Olha Kovalchuk

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License.

To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Berezka, K., and Kovalchuk, O. (2023). The Application of Association Rules to Detect the Effects of Vaccinations against COVID-19 in the EU-27. Preliminary Estimates. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, 27(1).

DOI: 10.15611/eada.2023.1.01

JEL Classification: C52, C53, O21

Abstract: In this research study, the authors obtained the preliminary evaluation of the impact detection of vaccinations against COVID-19 in the EU-27. The empirical basis of the study was the daily number of COVID-19 cases, vaccinations, hospitalisations, and deaths in the EU countries from March 2020 to March 2022. Rules of association were used to identify non-obvious associations between vaccinations against COVID-19 and cases of illness, hospitalisations, and deaths from COVID-19. The obtained results were used to cluster the EU countries by the level of vaccinations against COVID-19, cases of COVID-19, deaths from COVID, and COVID-19 hospitalisations for the EU member states. The K-means clustering method was used for cluster analysis. Hidden dependencies of the number of COVID-19 cases, the number of COVID-19 hospitalisations, and the number of COVID-19 deaths due to the number of vaccinations against COVID-19 by EU countries were revealed. It was established with a high probability that vaccination significantly affects the level of morbidity. For the first time, association rules were obtained, which are preliminary estimates of the relationship between the dynamics of vaccinations against COVID-19 and the dynamics of COVID-19 cases, COVID-19 hospitalisations, and deaths from COVID-19 in the EU. The results can be used to make beneficial decisions, for example, to regulate vaccination policies in individual EU countries, and predict the future consequences of the COVID-19 pandemic.

Keywords: COVID-19, association rules, k-means clustering, vaccinations, EU country, decision-making.

1. Introduction

For the third year in a row, the unprecedented economic and social upheavals caused by the COVID-19 pandemic pose a critical global threat. According to Johns Hopkins University, the number of global confirmed cases of COVID-19 was more than 521.4 million, and the number of global confirmed COVID-19 deaths exceeded 6.26 million (Donovan, 2022). Our World in Data announced that more than 11.65 billion doses of vaccinations against COVID-19 (*Coronavirus (COVID-19) Vaccinations*, 2022) have been administered worldwide. According to The World Health Organization, the number of COVID-19 deaths in the WHO European region in May 2022 exceeded 2 million (*COVID-19 deaths cross 2 million marks...*, 2022).

Uneven vaccination causes risks to the successful elimination of the coronavirus consequences, to economic recovery, and increases geopolitical tensions. Today, the economic problems related to COVID-19 are predicted to cause the global economy to fall by 2.3% compared to the pre-pandemic level and create new risks, such as inflation, rising prices, and debts. Only the fast vaccination progress and the use of the latest information technologies to analyse the situation can prompt a smooth return to pre-pandemic growth (*The Global Risks Report 2022*). Therefore, scientists have already conducted empirical research on the impact of the COVID-19 pandemic in various spheres of socio-economic life (Bieszk-Stolorz & Dmytrów, 2020; Bieszk-Stolorz & Dmytrów, 2021; Dmytrów et al., 2021) and SARS-CoV-2 vaccine effectiveness (*Results for Vaccines*, 2022). Although the pandemic has not yet been fully overcome, governments need to make effective management decisions to address its consequences in a coordinated manner, which can only be supported by effective socio-economic research methods (Berezka & Kovalchuk, 2018; Grzeskowiak & Stanimir, 2007).

The aim of the paper was to search for useful association rules to identify the hidden dependencies and links between the dynamics of vaccinations against COVID-19 and the effects of the pandemic (disease cases, hospitalisations, and deaths).

2. Data selection and description

The information base of the study was the following official statistical dataset: the daily number of cases of COVID-19 disease, vaccinations against COVID-19, hospitalisations due to COVID-19, and deaths from COVID-19 in the EU countries from March 2020 to March 2022 (*COVID-19 Data on the daily number...*, 2022).

The empirical research was performed in Statistica software. Preliminary graphical analysis suggests that COVID-19 cases in most EU countries peaked in January 2022 (Figure 1). The countries with the highest average monthly COVID-19 cases in the study period were France, Germany, Italy, and Spain (Figure 2).

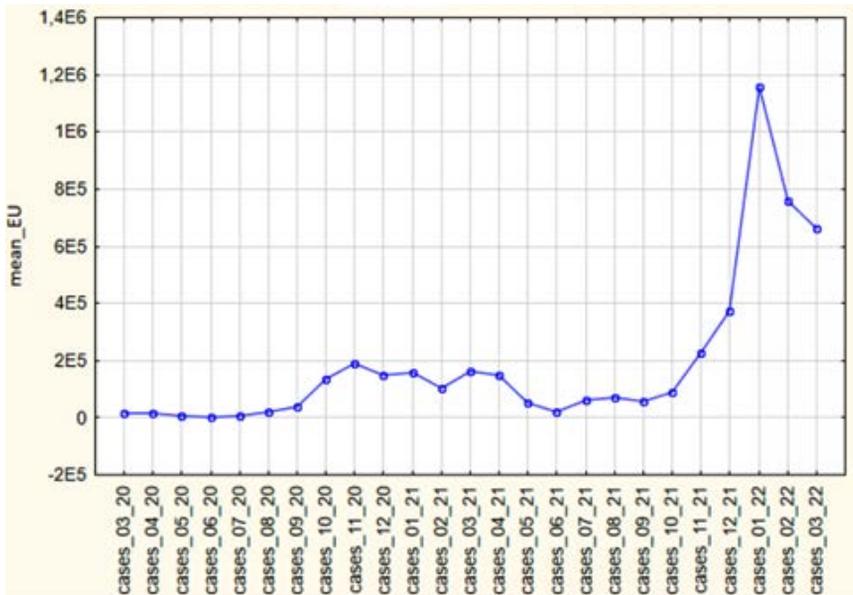


Fig. 1. Average monthly cases of COVID-19 in the EU-27 from 03.2020 to 03.2022

Source: own elaboration using Statistica software.

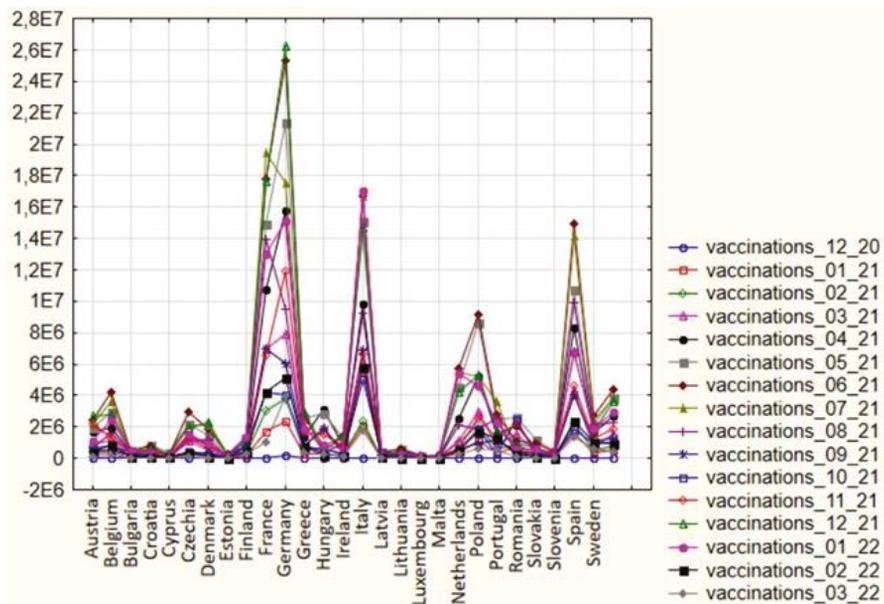


Fig. 2. Average monthly cases of COVID-19 in the EU countries from 03.2020 to 03.2022

Source: own elaboration using Statistica software.

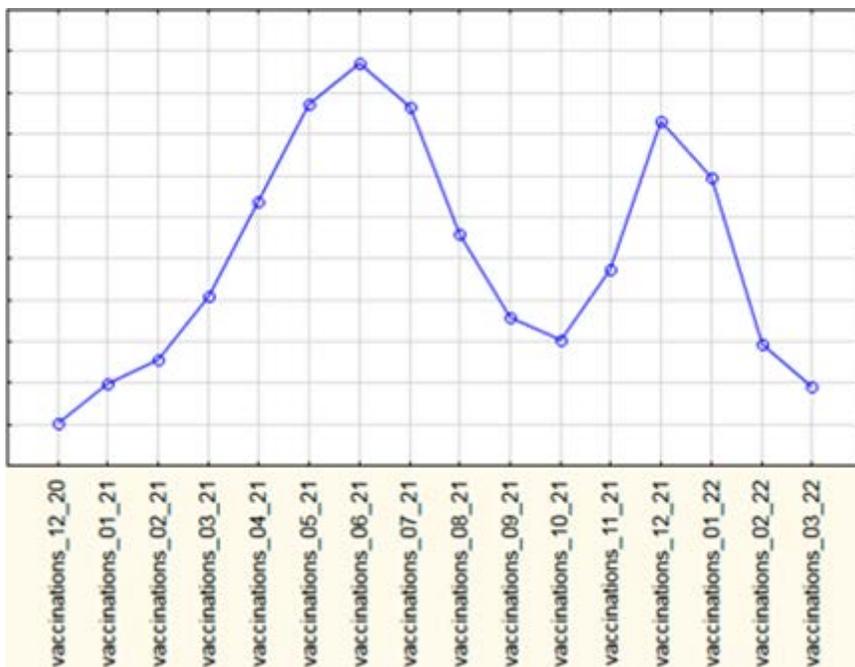


Fig. 3. Average monthly vaccinations against COVID-19 in the EU-27 from 03.2020 to 03.2022

Source: own elaboration using Statistica software.

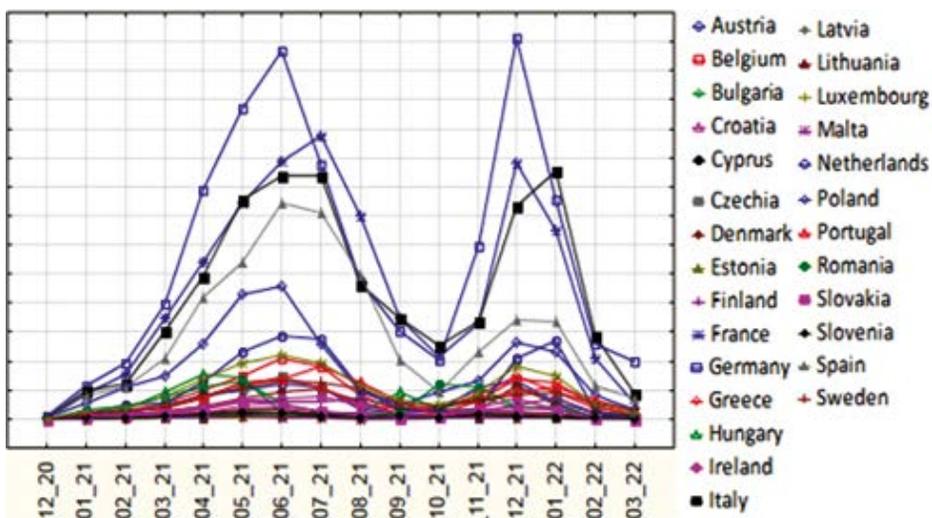


Fig. 4. Average monthly vaccinations against COVID-19 in the EU countries from 03.2020 to 03.2022

Source: own elaboration using Statistica software.

The maximum average monthly vaccinations against COVID-19 in the EU-27 were conducted from 05.2021 to 07.2022 and in 12.2021 (Figure 3). The highest number of vaccinations was carried out in Germany (December 2021), France (July 2021), Italy (January 2021), Spain (June 2021) (Figure 4).

3. Research methodology

Association rules are the data mining methods designed to identify interesting research-relevant relations between variables in large datasets. The method involves the search for strong rules (abstract associations) identified in the dataset under study using certain measures of interest (Srikant & Agraval, 1996). As a result of the application of association rules, as well as other data mining methods, previously unknown non-trivial information was obtained. The obtained knowledge describes new relations between properties and can be used to predict the values of some features based on others. This knowledge can be applied to new data with some degree of certainty. The usefulness is that this knowledge brings some effectiveness in its application. Knowledge should be presented in a form understandable to non-mathematicians.

In data analysis, it is quite often necessary to determine sets of frequently occurring objects from a large set of objects. Before giving a generalised description of this problem, the authors introduced some notations and provided some definitions.

Let I be the set of objects included in the studied sets:

$$I = \{i_1, i_2, \dots, i_j, \dots, i_n\}, \quad (1)$$

where i_j is an object included in the studied sets, n is the total number of objects.

Let D be a set of transactions, each of which is a set of objects from set I :

$$D = \{T_1, T_2, \dots, T_j, \dots, T_n\}, \quad (2)$$

where T_j is a transaction that is a subset of set I ($T \subseteq I$), m is the total number of transactions available in the study.

The subset of transactions, which includes object i_j , is denoted by D_{i_j} ($D_{i_j} \subseteq D$).

An arbitrary set of objects (item set) is denoted by F ($F \subseteq I$), and the set of transactions, including set F , as D_F ($D_F \subseteq D$).

A set of objects consisting of k elements is called a k -element object set.

The support of a set of objects is the ratio of the number of transactions, including set F , to the total number of transactions:

$$Supp(F) = \frac{n(D_F)}{n(D)}, \quad (3)$$

where $n(D_F)$ is the number of transactions that include set F , $n(D)$ is the total number of transactions available in the study, $n(D) = m$.

An association rule is implication $F \rightarrow G$, where $F \subset I$, $G \subset I$, and $F \cap G = \emptyset$.

Association rules are characterised by a confidence indicator, which indicates how often the rule is true. The confidence of rule $F \rightarrow G$ is the ratio of the number of transactions that include both set F and set G to the number of transactions that include set F :

$$Conf(F \rightarrow G) = \frac{Supp(F \cup G)}{Supp(F)}. \quad (4)$$

When conducting a study, the analyst determines the thresholds – the minimum value of support for the analysed sets $Supp_{\min}$ and the minimum value of confidence $Conf_{\min}$.

Set F is called frequent if its support value is greater than the minimum specified support value:

$$Supp(F) > Supp_{\min}. \quad (5)$$

Rule $F \rightarrow G$ is an association rule if it has the support greater than $Supp_{\min}$ and the confidence greater than $Conf_{\min}$.

After the above definitions, the authors presented an algorithm for generating association rules, which consists of two steps:

- search for frequent item sets;
- formation of association rules from the sets of step 1.

Generating association rules is not an easy task. The first step is especially algorithmically difficult because it requires searching through all possible combinations of objects. As the number of objects in set I grows, the number of possible sets of objects grows exponentially.

The search for association rules is greatly facilitated by data mining.

The K -mean clustering method was used for cluster analysis. This method clusters the input dataset in quantified number (k) of groups (Anderberg, 1973).

4. Empirical results and discussion

The study used the association rules method to identify the relation between the number of vaccinations against COVID-19 (taking into account the cumulative effect) and the number of new cases of COVID-19, hospitalisations due to COVID-19, and deaths caused by COVID-19 in the EU countries from 03.2020 to 03.2022.

The analysis revealed the following association rules:

1. If the number of vaccinations against COVID-19 in one of the EU countries in the current month is higher than the average monthly number of vaccinations against COVID-19 in this country from 03.2020 to 03.2022, then with a probability of 0.9%

the number of COVID-19 cases in this country in the same month will be lower than the average monthly number of COVID-19 cases in the same country for the same period; and vice versa (Table 1).

The formal representation of the defined association rules with minimum support = 0.9, minimum confidence = 0.9, and minimum correlation = 0.9 is the following:

<IF Monthly number of vaccinations against COVID-19 >
 Limit vaccinations THEN Monthly number of cases of the disease COVID-19
 is lower than Limit cases>

Limit vaccinations – a variable in which the calculated function of the average monthly number of vaccinations against COVID-19 for the EU Member States from 03.2020 to 03.2022 is stored;

Limit cases – a variable in which the calculated function of the average monthly number of cases of COVID-19 for the EU member states from 03.2020 to 03.2022 is stored.

Table 1. Summary of association rules (vaccinations and cases)

Summary of association rules Min. support = 90.0%, Min. confidence = 90.0%, Min. correlation = 90.0% Max. size of body = 10, Max. size of head = 10						
	Body	==>	Head	Support (%)	Confidence (%)	Correlation (%)
1	0	==>	1	100.0000	100.0000	100.0000
2	1	==>	0	100.0000	100.0000	100.0000

Source: own elaboration using Statistica software.

2. If the number of vaccinations against COVID-19 in one of the EU countries in the current month is higher than the average monthly number of COVID-19 vaccinations in this country from 03.2020 to 03.2022, then with a probability of 0.9% the monthly number of deaths from COVID-19 in this country in the same month will be lower than the average monthly number of deaths caused by COVID-19 (Table 2).

The formal representation of the defined association rules with minimum support = 0.9, minimum confidence = 0.9, and minimum correlation = 0.9 is the following:

< IF Monthly number of vaccinations against COVID-19 > Limit vaccinations
 THEN Monthly number of deaths from COVID-19 is lower than Limit death >

Limit death – a variable in which the calculated function of the average monthly number of deaths caused by COVID-19 for the EU member states from 03.2020 to 03.2022 is stored.

Table 2. Summary of association rules (vaccinations and deaths)

Summary of association rules Min. support = 90.0%, Min. confidence = 90.0%, Min. correlation = 90.0% Max. size of body = 10, Max. size of head = 10						
	Body	==>	Head	Support (%)	Confidence (%)	Correlation (%)
1	0	==>	1	100.0000	100.0000	100.0000
2	1	==>	0	100.0000	100.0000	100.0000

Source: own elaboration using Statistica software.

3. If the number of vaccinations against COVID-19 in the current month in one of the EU countries is higher than the average number of monthly vaccinations against COVID-19 in this country from 03.2020 to 03.2022, then with a probability of 0.5%, the average monthly number of hospitalisations due to COVID-19 in this country in the same month will be lower than the average monthly number of COVID-19 hospitalisations in the same country for the same period, and vice versa (Table 3).

The formal representation of the defined association rules with minimum support = 0.5, minimum confidence = 0.5, and minimum correlation = 0.5 is the following:

< IF Monthly number of vaccinations against COVID-19 > Limit vaccinations
THEN Monthly hospitalisations due to COVID-19
is lower than Limit hospitalizations >

Limit hospitalisations – a variable in which the calculated function of the average monthly number of COVID-19 hospitalisations for the EU member states from 03.2020 to 03.2022 is stored.

Table 3. Summary of association rules (vaccinations and hospitalisations)

Summary of association rules Min. support = 50.0%, Min. confidence = 50.0%, Min. correlation = 50.0% Max. size of body = 10, Max. size of head = 10						
	Body	==>	Head	Support (%)	Confidence (%)	Correlation (%)
1	0	==>	1	66.66667	66.6667	81.64966
2	1	==>	0	66.66667	100.0000	81.64966

Source: own elaboration using Statistica software.

Support shows what percentage of transactions support the received rule. In this case, sufficient support has been achieved so that the identified associative rule can be extended with high reliability to all EU countries.

Table 4. Limit values of the monthly number of vaccinations against COVID-19, cases of the disease COVID-19, deaths caused by COVID-19, and hospitalisations due to COVID-19 for the EU member states

Country	Limit vaccinations, % of the population	Limit cases, % of the population	Limit death, % of the population	Limit hospitalisations, % of the population
Austria	0.7	1.9	0.008	0.46
Belgium	0.8	1.5	0.012	0.57
Bulgaria	0.2	0.6	0.019	1.37
Croatia	0.4	1.0	0.015	0.69
Cyprus	0.7	2.1	0.004	0.05
Czechia	0.6	1.5	0.015	0.75
Denmark	0.8	2.2	0.004	0.19
Estonia	0.5	1.6	0.007	0.50
Finland	0.7	0.7	0.003	0.10
France	0.7	1.6	0.009	0.81
Germany	0.7	1.0	0.006	0.11
Greece	0.6	1.1	0.010	0.12
Hungary	0.6	0.7	0.018	0.86
Ireland	0.8	1.4	0.006	0.29
Italy	0.7	1.0	0.011	0.67
Latvia	0.4	1.4	0.009	0.54
Lithuania	0.4	1.1	0.010	0.40
Luxembourg	0.9	1.9	0.009	0.46
Malta	1.1	0.8	0.006	0.24
The Netherlands	0.7	1.9	0.005	0.20
Poland	0.5	0.6	0.012	0.73
Portugal	0.7	1.4	0.009	0.37
Romania	0.3	0.5	0.012	0.54
Slovakia	0.5	1.8	0.014	0.69
Slovenia	0.5	2.0	0.013	0.68
Spain	0.7	1.0	0.009	0.43
Sweden	0.7	1.1	0.008	0.34

Source: own elaboration using Statistica software.

1. If in Austria the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.9% of the population, the number of deaths caused by COVID-19 in the same month was

lower than 0.008% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.46% of the population.

2. If in Belgium the number of vaccinations against COVID-19 in the current month was higher than 0.8% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.5% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.012% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.57% of the population.

3. If in Bulgaria the number of vaccinations against COVID-19 in the current month was higher than 0.2% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 0.6% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.019% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 1.37% of the population.

4. If in Croatia the number of vaccinations against COVID-19 in the current month was higher than 0.4% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.0% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.015% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.69% of the population.

5. If in Cyprus the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 2.1% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.004% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.05% of the population.

6. If in the Czech Republic the number of vaccinations against COVID-19 in the current month was higher than 0.6% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.5% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.015% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.75% of the population.

7. If in Denmark the number of vaccinations against COVID-19 in the current month was higher than 0.8% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 2.2% of the population, the number of deaths caused by COVID-19 in the same month was

lower than 0.004% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.19% of the population.

8. If in Estonia the number of vaccinations against COVID-19 in the current month was higher than 0.5% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.6% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.007% of the population; with a probability of 50%, the number of hospitalisations due to COVID-19 in the same month was lower than 0.50% of the population.

9. If in Finland the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 0.7% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.003% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.1% of the population.

10. If in France the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.6% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.009% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.81% of the population.

11. If in Germany the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.0% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.006% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.11% of the population.

12. If in Greece the number of vaccinations against COVID-19 in the current month was higher than 0.6% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.1% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.010% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.12% of the population.

13. If in Hungary the number of vaccinations against COVID-19 in the current month was higher than 0.6% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 0.7% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.018% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.86% of the population.

14. If in Ireland the number of vaccinations against COVID-19 in the current month was higher than 0.8% of the population, then with a probability of 90%

the number of COVID-19 cases in the same month was lower than 1.4% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.006% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.29% of the population.

15. If in Italy the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.0% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.011% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.67% of the population.

16. If in Latvia the number of vaccinations against COVID-19 in the current month was higher than 0.4% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.4% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.009% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.54% of the population.

17. If in Lithuania the number of vaccinations against COVID-19 in the current month was higher than 0.4% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.1% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.010% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.40% of the population.

18. If in Luxembourg the number of vaccinations against COVID-19 in the current month was higher than 0.9% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.9% of the population, the number of death caused by COVID-19 in the same month was lower than 0.009% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.46% of the population.

19. If in Malta the number of vaccinations against COVID-19 in the current month was higher than 1.1% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 0.8% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.006% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.24% of the population.

20. If in the Netherlands the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.9% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.005% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.20% of the population.

21. If in Poland the number of vaccinations against COVID-19 in the current month was higher than 0.5% of the population, then with a probability of 90%

the number of COVID-19 cases in the same month was lower than 0.6% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.012% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the month was lower than 0.73% of the population.

22. If in Portugal the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.4% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.009% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.37% of the population.

23. If in Romania the number of vaccinations against COVID-19 in the current month was higher than 0.3% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 0.5% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.012% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.54% of the population.

24. If in Slovakia the number of vaccinations against COVID-19 in the current month was higher than 0.5% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.8% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.014% of the population; with a probability of 50% the number of COVID-19 hospitalisations in the same month was lower than 0.69% of the population.

25. If in Slovenia the number of vaccinations against COVID-19 in the current month was higher than 0.5% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 2.0% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.013% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.68% of the population.

26. If in Spain the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.0% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.009% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.43% of the population.

27. If in Sweden the number of vaccinations against COVID-19 in the current month was higher than 0.7% of the population, then with a probability of 90% the number of COVID-19 cases in the same month was lower than 1.1% of the population, the number of deaths caused by COVID-19 in the same month was lower than 0.008% of the population; with a probability of 50%, the number of COVID-19 hospitalisations in the same month was lower than 0.34% of the population (Table 4).

The association rules helped to determine the hidden dependencies and connections between the studied variables at the stage of intelligence analysis.

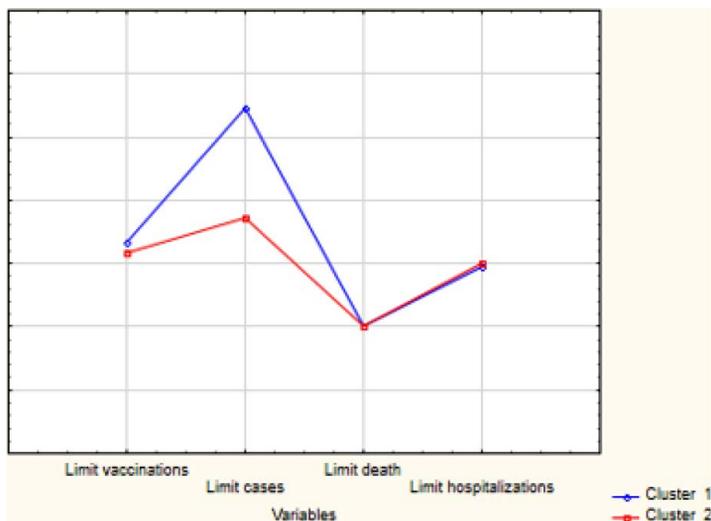


Fig. 5. Plot of means for each cluster

Source: own elaboration using Statistica software.

The obtained results were used to cluster EU countries according to the level of vaccinations against COVID-19, cases of COVID-19, deaths caused by COVID-19, and COVID-19 hospitalisations for the EU member states. To divide the EU-27 countries into groups according to the thresholds of the average values of COVID-19 cases, vaccinations against COVID-19, hospitalisations due to COVID-19, and death caused by COVID-19 by the EU country from 03.2020 to 03.2022, the cluster analysis was conducted. As a result of the application of the k-mean clustering method, two clusters were determined. The first cluster included economically developed industrial countries, such as Austria, Belgium, Cyprus, Czechia, Denmark, Estonia, France, Ireland, Latvia, Luxembourg, the Netherlands, Portugal, Slovakia, and Slovenia. The second cluster consisted of Bulgaria, Croatia, Finland, Germany, Greece, Hungary, Italy, Lithuania, Malta, Poland, Romania, Spain, and Sweden.

Table 5. Cluster means

Variable	Cluster No. 1	Cluster No. 2
Limit vaccinations	0.664286	0.584615
Limit cases	1.728571	0.861538
Limit death	0.008857	0.010692
Limit hospitalisations	0.468571	0.507692

Source: own elaboration using Statistica software.

The main criterion in the distribution of countries by selected clusters were limit cases (Figure 5). If the difference between the average threshold values of other studied indicators for the countries of the first and second groups was not more than 0.1, then for the countries of the first cluster the threshold value of COVID-19 cases was more than twice the threshold value of this indicator for the countries of the second cluster (Table 5). Thus, it can

be argued that in the countries of the first cluster, increasing the level of vaccinations against COVID-19 significantly reduced the risks of COVID-19 cases.

5. Conclusions

On the basis of empirical data on the first two years of the pandemic, a preliminary analysis of the impact detection of vaccinations in the EU-27 on COVID-19 health consequences was conducted. Non-obvious useful association rules were identified between the dynamics of vaccinations against COVID-19 and the separate consequences of a pandemic (COVID-19 cases, hospitalisations, and deaths). It was found that the dynamics of vaccination significantly affected the number of cases.

When detecting hidden dependencies of the number of COVID-19 cases and the number of deaths caused by COVID-19 from the number of vaccinations against COVID-19, 90% probability was used, and this indicates the high usefulness of the obtained rule. It contains reliable information that was not previously known but has a logical explanation. Such rules can be used to make beneficial decisions, e.g. to regulate vaccination policies in individual EU countries. The use of association rules makes it possible to find possible hidden dependencies and connections at the stage of intelligence analysis. The results obtained are only preliminary estimates and require further in-depth research, but they can already provide important information for vaccination policy in the EU-27 and can be used to build regression and forecasting models. The k-mean clustering results show that increasing the level of vaccinations against COVID-19 significantly reduces the risks of COVID-19 cases. The COVID-19 pandemic has become a global challenge and has far-reaching devastating consequences for the world, but effective management decisions, including vaccination, are needed to support economic recovery and protect the health and well-being of citizens.

References

- Anderberg, M. (1973). *Cluster analysis for applications*. Academic Press.
- Berezka, K., & Kovalchuk, O. (2018). Correspondence analysis as a tool for computer modeling of sustainable development. *Econometrics. Advances in Applied Data Analysis*, 22(4), 9-23. <https://doi.org/10.15611/eada.2018.4.01>
- Bieszk-Stolorz, B., & Dmytrów, K. (2020). Assessment of the similarity of the situation in the EU labour markets and their changes in the face of the COVID-19 pandemic. *Sustainability*, 14(6), 3646.
- Bieszk-Stolorz, B., & Dmytrów, K. (2021). A survival analysis in the assessment of the influence of the SARS-CoV-2 pandemic on the probability and intensity of decline in the value of stock indices. *Eurasian Economic Review*, 11(2), 363-379.
- Coronavirus (COVID-19) Vaccinations. (2022). Our world in data. <https://ourworldindata.org/covid-vaccinations>

- COVID-19 deaths cross 2 million mark in the European region – WHO.* (2022). Reuters. <https://www.reuters.com/world/europe/covid-19-deaths-cross-2-million-mark-european-region-who-2022-05-12>
- COVID-19 data.* (2022). March. OWID. <https://github.com/owid/covid-19-data/tree/master/public/data>.
- Data on the daily number of new reported COVID-19 cases and deaths by EU/EEA country* (2022). March. ECDC. <https://www.ecdc.europa.eu>
- Dmytrów, K., Landmesser, J., & Bieszk-Stolorz, B. (2021). The connections between COVID-19 and the energy commodities prices: evidence through the Dynamic Time Warping method. *Energies*, 14(13), 4024.
- Donovan, D. (2022). *U.S. officially surpassed 1 million COVID-19 deaths*. Johns Hopkins University. Coronavirus Resource Center. <https://coronavirus.jhu.edu>
- Grzeskowiak, A., & Stanimir, A. (2007). Possibilities of applications of chosen multivariate techniques to socio-economic researches. The analysis of comparability of results. *Prace Naukowe Akademii Ekonomicznej im. Oskara Langego we Wrocławiu*, (1180).
- Results for Vaccines.* (2022). The Economist. <https://www.economist.com>
- Srikant, R., & Agrawal, R. (1996). *Mining quantitative association rules in large relational tables*. ACM SIGMOD Conference. <https://www.semanticscholar.org>
- The Global Risks Report 2022.* (2022). 17th Edition. The World Economic Forum. <https://www3.weforum.org>

Stosowanie zasad stowarzyszenia w celu wykrywania skutków szczepień przeciwko COVID-19 w UE-27. Wstępne szacunki

Streszczenie: W tym badaniu autorzy uzyskali wstępną ocenę skuteczności wykrywania szczepień przeciwko COVID-19 w UE-27. Empiryczną podstawą badania jest dzienna liczba zachorowań na COVID-19, szczepień, hospitalizacji i zgonów w krajach UE w okresie od marca 2020 do marca 2022 r. Reguły asocjacji posłużyły do zidentyfikowania nieoczywistych powiązań między szczepieniami przeciwko COVID-19 oraz przypadków zachorowań, hospitalizacji i zgonów z powodu COVID-19. Uzyskane wyniki posłużyły do grupowania krajów UE według poziomu szczepień przeciwko COVID-19, przypadków choroby, zgonów z jej powodu oraz hospitalizacji dla krajów członkowskich UE. Do analizy skupień zastosowano metodę k-średnich grupowania. Ujawniono ukryte zależności liczby zachorowań na COVID-19, liczby hospitalizacji z powodu COVID-19 oraz liczby zgonów z powodu COVID-19 w związku z liczbą szczepień przeciwko COVID-19 w krajach UE. Stwierdzono z dużym prawdopodobieństwem, że szczepienia istotnie wpływają na poziom zachorowalności. Po raz pierwszy uzyskano reguły asocjacyjne, które są wstępnymi szacunkami zależności między dynamiką szczepień przeciwko COVID-19 a dynamiką zachorowań na COVID-19, hospitalizacji z tego powodu i zgonów w krajach UE. Wyniki mogą posłużyć do podejmowania korzystnych decyzji, np. do uregulowania polityki szczepień w poszczególnych krajach UE i przewidywania przyszłych konsekwencji choroby.

Słowa kluczowe: COVID-19, zasady asocjacji, grupowanie k-średnich, szczepienia, kraj UE, podejmowanie decyzji.