Description of health patterns of fully vaccinated older adults hospitalized due to COVID-19 in Brazil through association rules

Descrição dos padrões de saúde de idosos totalmente vacinados e hospitalizados por COVID-19 no Brasil por meio de regras de associação

Descripción de los patrones de salud de los adultos mayores hospitalizados por COVID-19 completamente vacunados en Brasil a través de las reglas de la asociación

Thaís Barbosa de OIiveira
ORCID: https://orcid.org/0000-0002-8461-4186
State University of Maringá, Brazil
E-mail: pg403498@uem.br
Lorrany Santos Rodrigues
ORCID: https://orcid.org/0000-0002-6560-1726
University of Brasilia, Brazil
E-mail: nutricaolsr@gmail.com
Wendy Rayanne Fernandes dos Santos
ORCID: https://orcid.org/0000-0002-2204-2100
Ministry of Health, Brazil
E-mail: wendyRayanne@gmail.com
Maurício Yukio Hirata
ORCID: https://orcid.org/0000-0002-6826-9407
Oswaldo Cruz Foundation, Brazil
E-mail: mauricio.yukio22@gmail.com
Caroliny Victoria dos Santos Silva
ORCID: https://orcid.org/0000-0001-9714-3398
University of Brasilia, Brazil
E-mail: silva.caroliny@aluno.unb.br
Josmar Mazucheli
ORCID: https://orcid.org/0000-0001-6740-0445
State University of Maringá, Brazil
E-mail: jmazucheli@uem.br

Abstract
The coronavirus disease 2019 (COVID-19) is a global public health problem. Since the beginning of the pandemic, notified in March 2020, Brazil has shown high lethality from the disease in older adults. From 2012 to 2018, the country showed an increase of 20% in the older adults' population. Despite the completeness of vaccine protocols against COVID-19 in the country, there is evidence that this age group, associated with the presence of comorbidities, can be a predictor of the occurrence of hospitalization and severe symptoms due to COVID-19. In this direction, this paper aimed to identify patterns and relationships between symptoms, comorbidities, gender, Intensive Care Unit (ICU) admission, and survival status of older adults, fully vaccinated against COVID-19, hospitalized in Brazil. For this purpose, we perform association rules mining on the OpenDataSUS database. For the group of patients with comorbidity, associations with conditions of oxygen saturation (SpO_2) <95%, dyspnea and death were predominant; The female sex was associated with survival and the presence of comorbidities, while the male sex with death and admission to the ICU; for patients admitted to the ICU and who died, associations with SpO_2<95%, dyspnea, presence of comorbidities and use of ventilatory support were found. The association rule mining procedure has been shown to be useful in surveying the hospitalization profile of these patients.
Keywords: COVID-19; Symptoms; Chronic disease; Aged; Hospitalization; Data mining.

Resumo
A doença do coronavírus 2019 (COVID-19) constitui-se como um problema de saúde pública global. Desde o início da pandemia, notificada em março de 2020, o Brasil apresenta alta letalidade da doença em idosos. De 2012 a 2018, o país apresentou um aumento de 20% na população de idosos. Apesar da completude dos protocolos vacinais contra a COVID-19 no país, há evidências de que essa faixa etária, associada à presença de comorbidades, pode ser um preditor da ocorrência de internação e apresentação de sintomas graves da doença. Nessa direção, este trabalho teve como objetivo identificar padrões e relações entre sintomas, comorbidades, gênero, internação em Unidade de Terapia

Intensiva (UTI) e estado de sobrevida de idosos, totalmente vacinados contra a COVID-19, hospitalizados no Brasil. Para tanto, utilizou-se do método de mineração de regras de associação no banco de dados do OpenDataSUS. Para o grupo de pacientes com comorbidade, predominaram as associações de condições de saturação de oxigênio (SpO$_2$) <95%, dispneia e óbito. O sexo feminino associou-se à sobrevida e presença de comorbidades, enquanto o sexo masculino ao óbito e internação na UTI. Para os pacientes internados na UTI e que foram a óbito, encontrou-se associações com SpO$_2$<95%, dispneia, presença de comorbidades e uso de suporte ventilatório. O procedimento de mineração de regras de associação mostrou-se útil no levantamento do perfil de hospitalização desses pacientes.
Palavras-chave: COVID-19; Sintomas; Doença crônica; Idosos; Hospitalização; Mineração de dados.

Resumen
La enfermedad por coronavirus 2019 (COVID-19) es un problema de salud pública mundial. Desde el inicio de la pandemia, notificada en marzo de 2020, Brasil ha mostrado alta letalidad por la enfermedad en adultos mayores. Del 2012 al 2018, el país mostró un incremento del 20% en la población de adultos mayores. A pesar de la exhaustividad de los protocolos de vacunación contra la COVID-19 en el país, existe evidencia de que este grupo etario, asociado a la presencia de comorbilidades, puede ser predictor de la ocurrencia de hospitalización y síntomas graves por la COVID-19. En esa dirección, este artículo tuvo como objetivo identifica los patrones y las relaciones entre los síntomas, las comorbilidades, el género, la admisión en la Unidad de Cuidados Intensivos (UCI) y el estado de supervivencia de los adultos mayores, completamente vacunados contra COVID-19, hospitalizados en Brasil. Para ello, realizamos minería de reglas de asociación en la base de datos OpenDataSUS. Para el grupo de pacientes con comorbilidad predominaron las asociaciones con condiciones de SpO$_2$<95%, disnea y muerte; El sexo femenino se asoció con la sobrevida y la presencia de comorbilidades, mientras que el sexo masculino con la muerte e ingreso a la UCI; para los pacientes ingresados en UTI y que fallecieron se encontraron asociaciones con SpO$_2$<95%, disnea, presencia de comorbilidades y uso de soporte ventilatorio. El procedimiento de minería de reglas de asociación se ha mostrado útil para relevar el perfil de hospitalización de estos pacientes.
Palabras clave: COVID-19; Síntomas; Enfermedad crónica; Anciano; Hospitalización; Minería de datos.

1. Introduction

The coronavirus disease 2019 (COVID-19) is a major global public health emergency. It is a severe acute respiratory syndrome caused by the new coronavirus (SARS-CoV-2) first notified in China in December 2019 and recognized as causing a pandemic by the World Health Organization (WHO) on March 11, 2020 (Cash & Patel, 2020). Since its discovery until February 22, 2022, the disease had about 426 million confirmed cases and 5 million deaths worldwide (WHO, 2021). More than 190 countries were affected and those with the highest number of deaths were the United States of America (USA), Brazil, India, Mexico, and Russia. From this group, Brazil, India, and Russia make up the BRICS, a political and economic grouping of developing countries formed by these countries, in addition to China and South Africa. Zhu et al. (2021) pointed out that the number of new daily cases of COVID-19 was aggravated by the situation of social inequality and health vulnerabilities in the BRICS, Brazil being the country with the highest number of deaths.

On February 25, 2020, Brazil registered the first confirmed case of COVID-19 (Candido et al., 2020). Since then, the disease has affected more than 28 million people up to February 22, 2022, in the country, with a higher incidence among adults and higher mortality among the older adults (aged 60 years or over) (Brazil, 2021). From 2012 to 2018, there was a 20% increase in the older adult's population in Brazil (FGV, 2021). Health conditions such as the presence of chronic diseases and hospitalization characteristics and symptoms can influence the occurrence of deaths from COVID-19 in these age groups (Rocha et al., 2021). These conditions can be classified as health characteristics or circumstances that present in a persistent way, requiring fragmented or continuous active responses from health systems (Mendes, 2018).

Faced with the challenge of reducing the risk of death with the listed conditions, the elaboration of a vaccination protocol has become an ethical imperative. Therefore, the world started with the vaccination schemes against COVID-19 in December 2020, but in Brazil this only occurred from January 2021, with emergency approval by the Brazilian Health Regulatory Agency (ANVISA), with the following vaccines: CoronaVac, Pfizer, Janssen and AstraZeneca (Bee et al., 2022). Since then, government records up to December 9, 2021, indicate that Brazil has fully vaccinated (with the two-dose protocol) about 151 million citizens, a total of 75% of its population (Vaccination Brazil Platform, 2022). However, according to Moreno-Perez et al. (2022), countries

should be aware of the clinical characteristics and predictors of poor outcomes in the elderly population, despite the completeness of the vaccination scheme, to verify effective means of protection and management of health services.

Advanced data mining techniques, such as Association Rules Mining (ARM), have been used in the recognition of patterns of these conditions in populations during the COVID-19 pandemic (Shawkat et al., 2021). Data mining can be described as an intelligent method that allows the identification of useful and understandable patterns and relationships of items in a database, which has ARM as one of its oldest fields in recognizing these patterns (Williams, 2011). Association rules procedures are popularly known as Market Basket Analysis, which estimates the probability of products or groups of products being purchased simultaneously. This technique is useful in the analysis of health conditions, as it can identify symptoms, morbidities and other conditions that present themselves simultaneously in a group of individuals (Tandan et al., 2021). Therefore, considering the change in the Brazilian age structure in recent years and the emergence of health research during the COVID-19 pandemic, this study aimed to discover patterns and relationships between symptoms, comorbidities, gender, and deaths of older adults, vaccinated against COVID-19, hospitalized in Brazil. This paper is organized into methodology (section 2), the description of the results (section 3), discussion (section 4) and conclusions (section 5).

2. Methodology

2.1 Data extraction and population

This is a descriptive study of data from of older adults, vaccinated against COVID-19, hospitalized in Brazil. The data was extracted on November 23, 2021, from the OpenDataSUS on-line platform (2022), including the period from February 14h, 2021 to November 11th, 2021. The choice of the 2021-year database was due to the registration of patients' vaccination data, because vaccination against COVID-19 started in January of that year in Brazil. This database is fed by health professionals through an individual notification form of patients with flu syndrome, with a survey of demographic, symptomatic, vaccination status and morbidity profile information.

Flu syndromes are characterized by symptoms that include fever, headache, chills, sore throat, and cough. Among the potential causative agents are the viruses that cause Influenza, in addition to Adenoviruses and Coronaviruses (SARS-CoV-2). For confirmation of cases, molecular diagnosis (RT-PCR) or immunological (antibody or antigen) screening is performed. This study included patients with confirmed RT-PCR for SARS-CoV-2 infection, hospitalized, aged 60 years or over, residing in Brazil (see Figure 1).

Figure 1 - Flowchart of OpenDataSUS data used in the study.
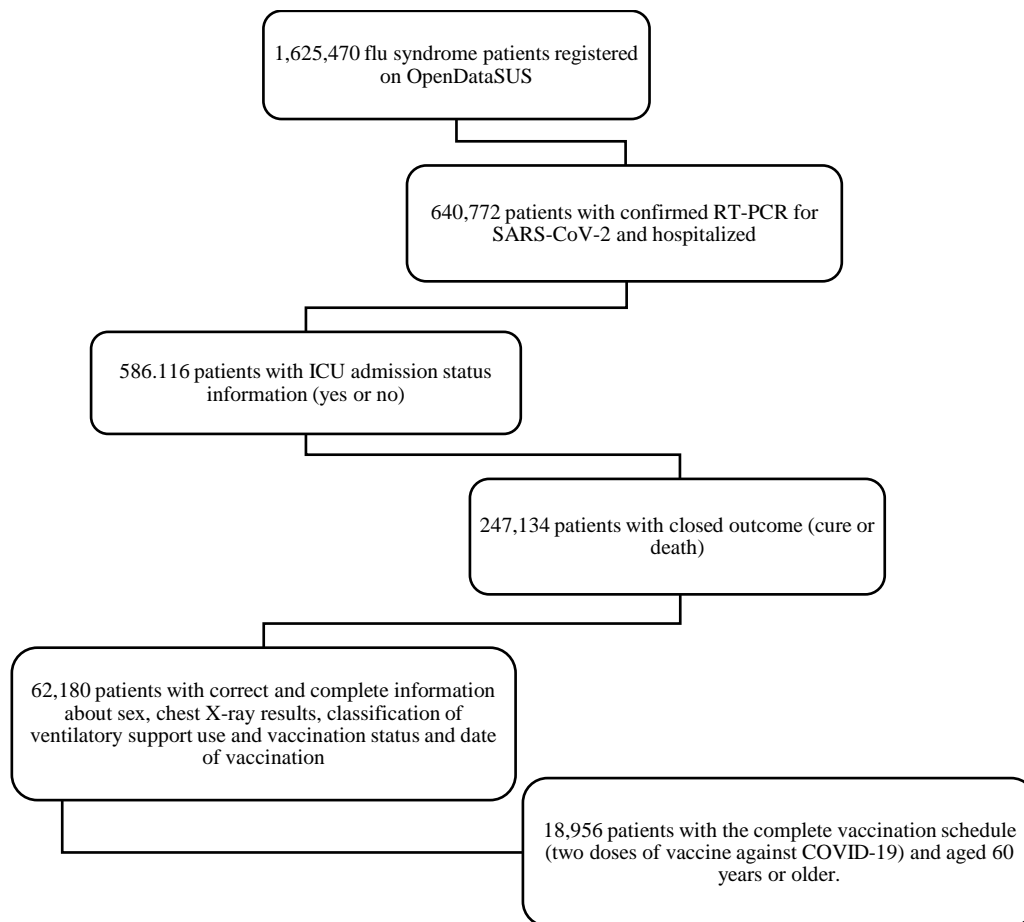
Source: Authors.

In this research, data from 62,180 hospitalized patients with complete vaccination status were used. Subsequently, a cohort of 18,956 patients aged 60 years old or over was carried out for the analysis of association rules. Originally, there were 162 variables in the database, and those containing these characteristics were excluded: (a) administrative data on hospitalization; (c) 90% to 100% of missing; (d) information regarding flu syndromes other than COVID-19; and (e) information that did not characterize symptoms, age group, gender, comorbidities and vaccination and survival status. The final data set consisted of 20 variables containing information on gender, survival, death, symptoms, chest X-ray results, comorbidities, use of ventilatory support, ICU admission information, regions of residence and vaccination status. The symptoms recorded were fever, cough, sore throat, anosmia, ageusia, dyspnea, respiratory decompensation, peripheral capillary oxygen saturation (SpO_2), diarrhea, vomiting, fatigue and abdominal pain. Comorbidities include heart disease, neuropathy, lung disease, hematopathy, asthma, diabetes, immunosuppression, nephropathy, and obesity. Brazil has 27 federative units and 5,570 municipalities. Its federative units are divided into regions, such as, North, Northeast, Central-West, Southeast and South. The regions of origin of the patients' residence were also included.

In this study, these variables make up what are called health conditions. The data was converted to a "transaction" format and analyzed using the Apriori Algorithm, available in package "arules" in R (Hahsler et al., 2022).

2.2 Missing data

The OpenDataSUS database has variables that contain missing. In the case of comorbidities, symptoms, entry of radiological exams and use of ventilatory support with absent values, we imputed it as absent clinical characteristic for the individuals, as in Baqui et al. (2020a) and Baqui et al. (2020b).

2.3 Association Rule Mining (ARM)

ARM is a descriptive method of data mining that allows the identification of patterns through transactions in the frequency of co-occurrences and relationships between items, characteristics or events present in a data set (Jean-Marc, 2001). Formally, a transaction (T) is composed a set of items, in the rule form: $X \rightarrow Y$, where X and Y have no elements in common, that is, they are disjoint $X \cap Y = \emptyset$. Set X is called antecedent (left-hand side) and set Y as consequent (right-hand side). This method also allows to measure the quality of an association rule through metrics, the most common being: (1) support, (2) confidence and (3) lift (Shin, et al., 2018).

The support represents the simultaneous frequency of X and Y in relation to the total transactions (a transaction is a database row), in the form as; $support(X \rightarrow Y) = \frac{frequency\ of\ X\ and\ Y}{total\ of\ T}$. Confidence determines the predictability of the rule, measuring the frequency of simultaneous occurrence of X and Y among transactions that contain X. This metric can also be read as a conditional probability, defined as; $Pr(Y|X)$ equal $confidence(X \rightarrow Y) = \frac{frequency\ of\ X\ and\ Y}{frequency\ of\ X}$. The lift measures the dependence between items, defined as; $lift(X \rightarrow Y) = \frac{confidence(X \rightarrow Y)}{support(X) \times support(Y)}$. If $lift(X \rightarrow Y) = 1$, X and Y are independent and uncorrelated; if $lift(X \rightarrow Y) > 1$, they are positively dependent and correlated; and if $lift(X \rightarrow Y) < 1$, they are negatively dependent and correlated. These values also provide the interpretation of the times the consequent occurs among the set of antecedents.

In the health field, it is important to recognize the patterns in clinical manifestations of a disease, as this can guide the planning and organization of healthcare. Therefore, an example rule is:

$X\{dyspnea\} \rightarrow Y\{death\}$ results in [support=10%, confidence= 90%, lift=1,8]

This rule indicates that 10% of patients reported both dyspnea and death, and the probability of someone dying given they reported dyspnea is equal to 90%. The odds of death is 80% greater when dyspnea occurs.

In this study, rule extractions were performed using the Apriori Algorithm, which allows the identification of the most frequent item sets with establishment of minimum support, making the data more manageable, allowing the discovery of rules and elimination of its redundancies. The property of the Apriori Algorithm can be defined as; $\forall X, Y: (X \subseteq Y) \rightarrow Support(X) \geq Support(Y)$. So, if an item set is frequent, its subsets are too. On the other hand, if an item set is not frequent, all sets that contain it are not either. Therefore, an item set never exceeds the support of sets that contain it. To exemplify the operation performed by the Apriori Algorithm, consider the scheme shown in Figure 2.

Figure 2 – Example of an operation performed by the Apriori Algorithm and rules generation with a minimum support equal to 50%.
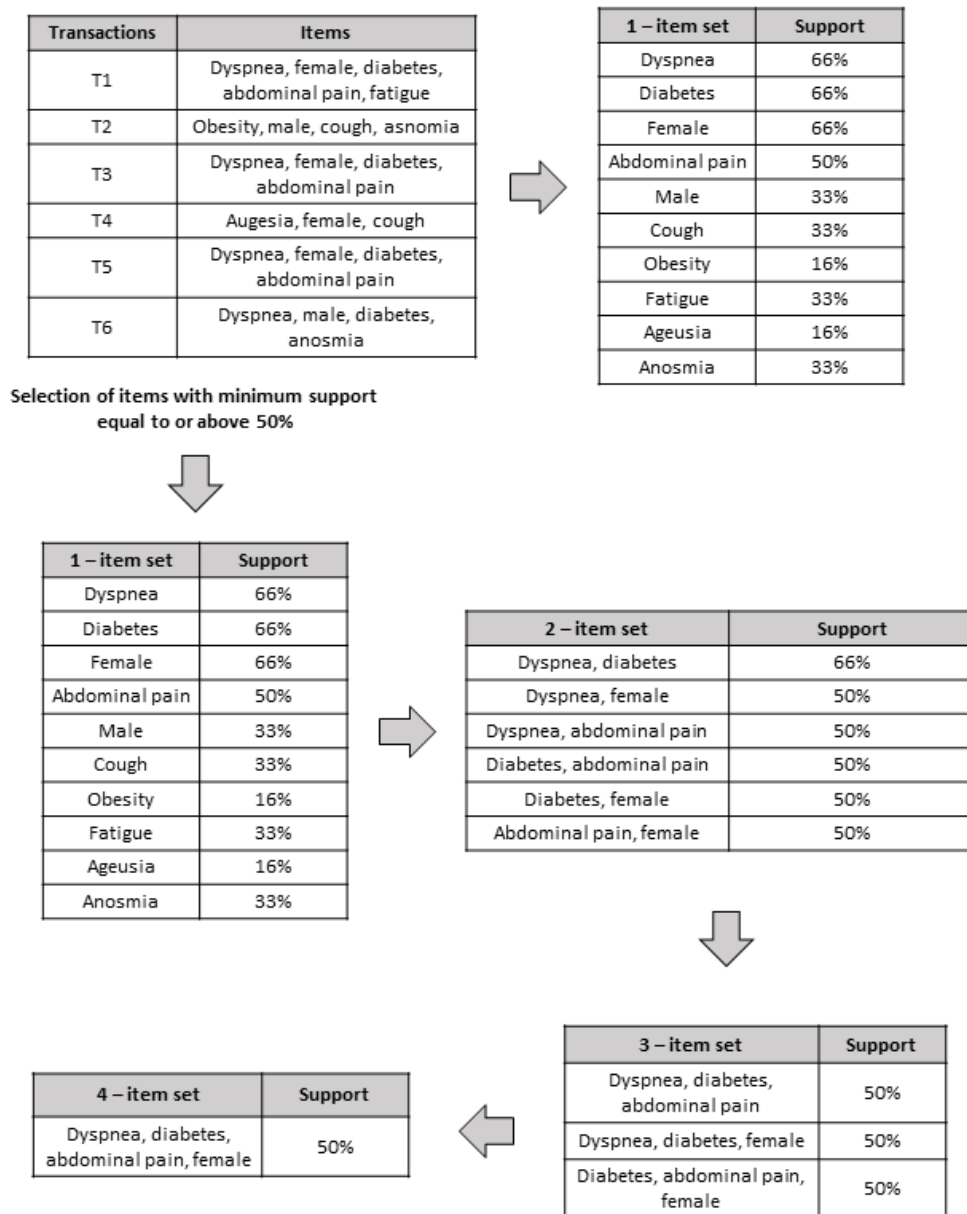
Source: Authors.

Here, a minimum support above 5% was defined. Only the top 10 rules were reported with the highest support and confidence scores.

2.4 Ethical aspects

This study was carried out with publicly accessible data, with the guarantee of anonymity in the database available on public platforms in Brazil and did not require approval by an ethics committee. The legal ethical precepts established by Resolution No. 466, of December 12, 2012, of the National Health Council of Brazil were respected.

3. Results

The study population consisted of 18,956 patients, of which 52% (n= 9,931) were male, 76% (n= 14,328) had comorbidities, 86% (n= 16,290) used respiratory support, 37% (n= 7,098) were admitted to the ICU and 48% (n= 9,047) died. SpO_2 below 95% was the most common symptom, followed by dyspnea, cough and respiratory decompensation (see, Figure 3a). The most predominant comorbidities were heart disease and diabetes, while the least frequent were liver disease and Down syndrome (see, Figure 3b). The population of this study had hospitalization records among the 27 Brazilian states, with a predominance in the Southeast Region of Brazil (See, Figure 3c). It is noteworthy that 35% (n= 6,572) of the patients reported having symptoms that were not listed in the OpenDataSUS record.

Figure 3 - Relative frequency of symptoms, comorbidities, and regions of residence in older adults vaccinated against COVID-19 hospitalized in Brazil (N= 18,956).
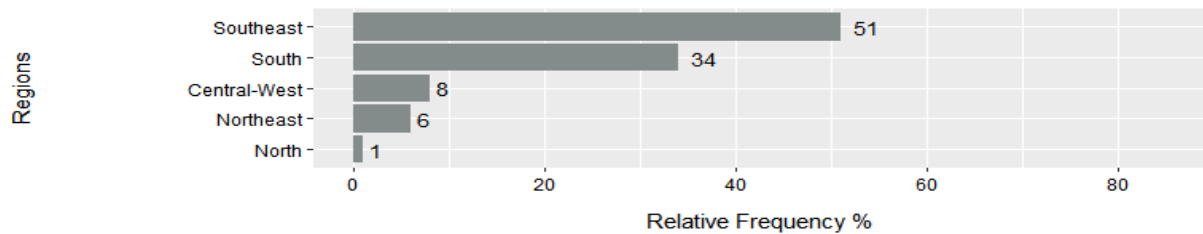
a) Symptoms
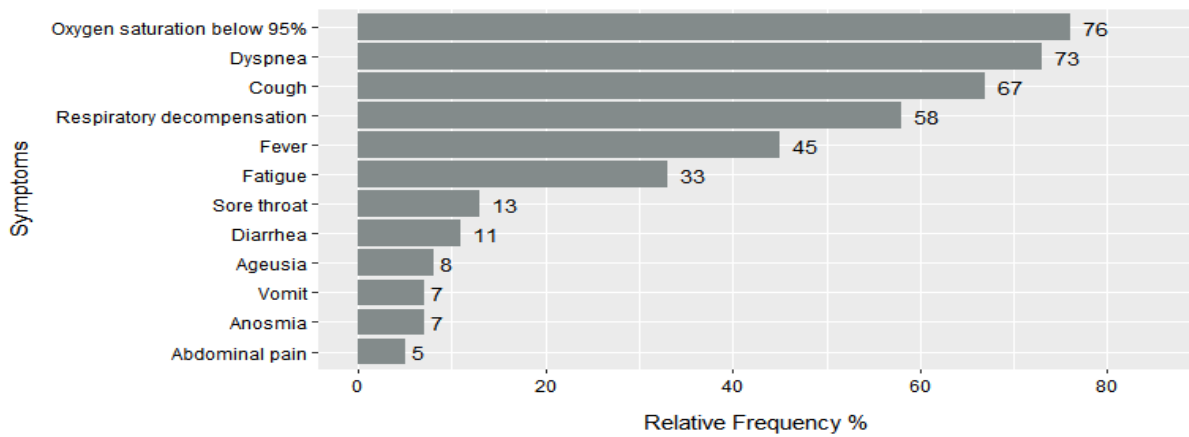
b) Comorbidity
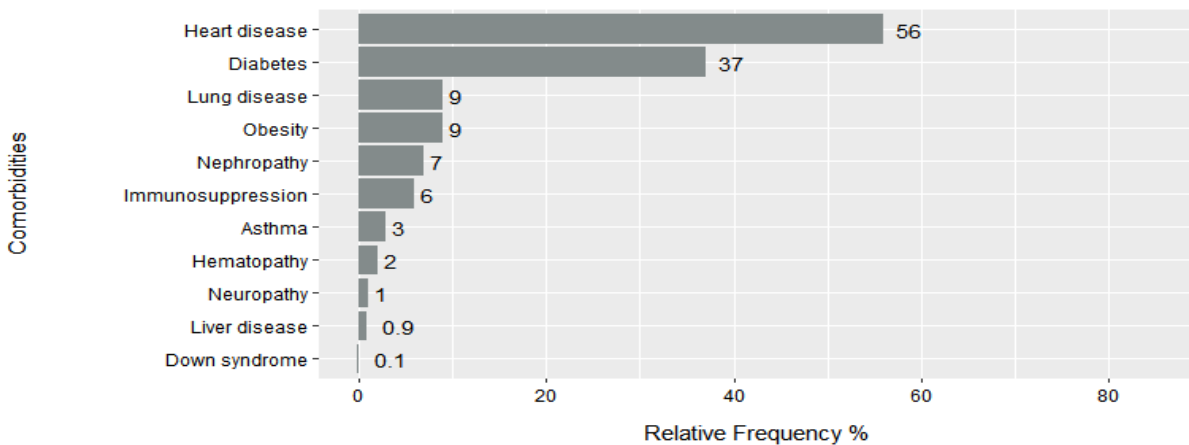
c) Regions of residence

Source: Authors.

For patients with comorbidities (n= 14,328), a total of 3,168 rules were found. Among the top 10, we observed that all conditions have a confidence equal to or above 80% (see, Table 1). That is, if a patient reported these conditions, the estimated probability that he had a comorbidity is 80%. The conditions of $SpO_2<95\%$ and death were the most predominant and, when listed individually or concomitantly with other conditions, such as respiratory decompensation, female gender, dyspnea, and ICU admission, they have a lift equal to 1.07. That is, the odds of an individual having comorbidity is 7% greater when these conditions are present.

Referring to patients without comorbidities (n= 4,628), the Algorithm discovered 1,326 rules. Among the predominant characteristics, there is the absence of radiological records, $SpO_2<95\%$, non-admission to the ICU and dyspnea. Although $SpO_2 < 95\%$ and dyspnea appear as common symptoms in this group, they occur less frequently when compared to the comorbid group. It is also observed that the reports of $SpO_2<95\%$ and dyspnea, individually or concomitantly, as well as cough are negatively correlated with the consequent, because your lift is less than 1. In other words, when the number of patients who reported the presence of this symptoms increases, the number of patients who reported not having comorbidities decreases. It is noteworthy that the use of non-invasive ventilatory support is independent of the consequent and the condition with the highest lift among those listed is survival.

Table 1 – Top 10 significant condition rules dis-aggregated by comorbidity status (N= 18,956).

Antecedents	Consequents	Support	Confidence	Lift
{SpO$_2$<95%, died}	{With comorbidity}	31%	80%	1.06
{Dyspnea, SpO$_2$<95%, died}	{With comorbidity}	26%	80%	1.06
{SpO$_2$<95%, admitted to ICU}	{With comorbidity}	24%	80%	1.07
{Respiratory decompensation, died}	{With comorbidity}	24%	80%	1.06
{Dyspnea, admitted to ICU}	{With comorbidity}	24%	81%	1.06
{Female, dyspnea, SpO$_2$<95%}	{With comorbidity}	24%	81%	1.07
{Respiratory decompensation, SpO$_2$<95%, died}	{With comorbidity}	23%	80%	1.07
{SpO$_2$<95%, fatigue}	{With comorbidity}	22%	80%	1.06
{Respiratory decompensation, dyspnea, died}	{With comorbidity}	22%	80%	1.06
{Admitted to ICU, died}	{With comorbidity}	22%	81%	1.07
{No radiological records}	{Without comorbidity}	19%	25%	1.02
{SpO$_2$<95%}	{Without comorbidity}	17%	22%	0.92
{No admitted to ICU}	{Without comorbidity}	17%	27%	1.10
{Dyspnea}	{Without comorbidity}	17%	23%	0.94
{Cough}	{Without comorbidity}	16%	24%	0.98
{Use of non-invasive ventilatory support}	{Without comorbidity}	15%	24%	1.00
{Survived}	{Without comorbidity}	14%	28%	1.13
{Male}	{Without comorbidity}	14%	26%	1.07
{Dyspnea, SpO$_2$<95%}	{Without comorbidity}	14%	22%	0.90
{No admitted to ICU, no radiological records}	{Without comorbidity}	13%	27%	1.11

Source: Authors.

For the female group (n= 9,025), 1,202 rules were obtained. The most predominant conditions reported were survival, no ICU admission and use of non-invasive ventilatory support with confidence equal to or greater than 50% (see, Table 2). These conditions are also listed concomitantly with other critical clinical features, such as SpO$_2$ < 95% and presence of comorbidity, with lift greater than 1. For males (n= 9,931) a total of 1,267 rules were identified. Unlike the female group, in the top 10, for the male group, the conditions of death, fever and ICU admission stood out, along with critical clinical features such as dyspnea and SpO$_2$ < 95%. When death and ICU admission are reported together, the lift equals 1.09.

Table 2 – Top 10 significant condition rules dis-aggregated by gender (N= 18,956).

Antecedents	Consequents	Support	Confidence	Lift
{Survived}	{Female}	26%	50%	1.05
{No admitted to ICU, use of non-invasive ventilatory support}	{Female}	24%	50%	1.06
{Use of non-invasive ventilatory support, with comorbidity}	{Female}	24%	51%	1.07
{No admitted to ICU, with comorbidity}	{Female}	23%	51%	1.07
{SpO$_2$<95%, no admitted to ICU}	{Female}	23%	50%	1.06
{No admitted to ICU, survived}	{Female}	21%	50%	1.06
{Use of non-invasive ventilatory support, survived}	{Female}	20%	51%	1.07
{SpO$_2$<95%, use of non-invasive ventilatory support, with comorbidity}	{Female}	20%	51%	1.08
{Survived, with comorbidity}	{Female}	19%	52%	1.08
{SpO$_2$<95%, no admitted to ICU, use of non-invasive ventilatory support}	{Female}	19%	51%	1.07
{Death}	{Male}	26%	55%	1.05
{Fever}	{Male}	25%	55%	1.05
{Admitted to ICU}	{Male}	21%	56%	1.07
{Dyspnea, death}	{Male}	21%	55%	1.07
{Fever, no radiological records}	{Male}	19%	57%	1.06
{Dyspnea, SpO$_2$<95%, death}	{Male}	18%	55%	1.05
{SpO$_2$<95%, admitted to ICU}	{Male}	17%	55%	1.06
{Fever, death}	{Male}	17%	55%	1.05
{Dyspnea, Admitted to ICU}	{Male}	17%	56%	1.06
{Admitted to ICU, death}	{Male}	16%	27%	1.09

Source: Authors.

When patients were disaggregated by ICU admission status, 1,277 rules were discovered for the status of admitted (n= 7,098) and 1,803 for the status of not admitted (n= 11,858). Among the rules found for the status of admitted, the predominance of the use of invasive ventilatory support, dyspnea, death, SpO$_2$<95% and presence of comorbidity stands out (see, Table 3). Clinical conditions that were reported concurrently with the use of invasive ventilatory support had a lift greater than 2. For the non-admission status, we see clinical conditions of use of non-invasive ventilatory support, and the absence of radiological records predominate. Unlike the group of patients who were admitted to the ICU, for which death is predominant in the rules, survival is reported by those not admitted with a confidence of 80%. Furthermore, conditions that usually make up a critical clinical picture of the disease, such as dyspnea and the presence of comorbidity, when reported individually had a lift below 1, indicating that they are negatively correlated with the consequent non-admission to the ICU.

Table 3 – Top 10 significant condition rules dis-aggregated by to ICU admission (N= 18,956).

Antecedents	Consequents	Support	Confidence	Lift
{Use of invasive ventilatory support}	{Admitted to ICU}	20%	89%	2.37
{Dyspnea, SpO$_2$<95%, died}	{Admitted to ICU}	20%	60%	1.60
{SpO$_2$<95%, died, with comorbidity}	{Admitted to ICU}	19%	60%	1.60
{Dyspnea, died, with comorbidity}	{Admitted to ICU}	18%	61%	1.62
{SpO$_2$<95%, use of invasive ventilatory support}	{Admitted to ICU}	17%	89%	2.37
{Respiratory decompensation, SpO$_2$<95%, died}	{Admitted to ICU}	17%	62%	1.60
{Dyspnea, use of invasive ventilatory support}	{Admitted to ICU}	17%	89%	2.38
{Dyspnea, respiratory decompensation, died}	{Admitted to ICU}	17%	60%	1.62
{Use of invasive ventilatory support, with comorbidity}	{Admitted to ICU}	16%	89%	2.39
{Dyspnea, SpO$_2$<95%, death, with comorbidity}	{Admitted to ICU}	16%	61%	1.64
{Use of non-invasive ventilatory support}	{No admitted to ICU}	48%	76%	1.22
{No radiological records}	{No admitted to ICU}	48%	63%	1.01
{SpO$_2$<95%}	{No admitted to ICU}	46%	60%	0.96
{With comorbidity}	{No admitted to ICU}	46%	60%	0.97
{Cough}	{No admitted to ICU}	43%	64%	1.02
{Survived}	{No admitted to ICU}	42%	80%	1.28
{SpO$_2$<95%, use of non-invasive ventilatory support}	{No admitted to ICU}	38%	76%	1.21
{Use of non-invasive ventilatory support, no radiological records}	{No admitted to ICU}	36%	76%	1.22
{Use of non-invasive ventilatory support, with comorbidity}	{No admitted to ICU}	36%	75%	1.20
{Dyspnea, use of non-invasive ventilatory support}	{No admitted to ICU}	35%	75%	1.20

Source: Authors.

A total of 607 rules were discovered for surviving patients (n= 9,909) and, for patients who died (n= 9,047), a total of 288. For the first group, non-admission to the ICU was predominant, followed using non-invasive ventilatory support and the absence of radiological records with confidences above 60% (see, Table 4). The cough symptom, considered mild for the disease, reported together with not being admitted to the ICU presented the highest lift when compared to the other rules listed. Although critical conditions of the disease are also found for this group, such as SpO$_2$<95% and dyspnea, they are reported together with the most predominant characteristics, such as non-admission to the ICU and non-use of invasive ventilatory support.

For the second group, we can observe that ICU admission stands out in the top 10, as well as SpO$_2$<95%, dyspnea, use of invasive ventilatory support and the presence of comorbidity. It is noteworthy that the conditions of admission to the ICU and use of invasive ventilatory support have a confidence of 90% and the highest lift among the other conditions reported. In other words, among the individuals who died, the chances of presenting these conditions are equal to 90%. Regarding lift, the odds of death is 88% greater when these conditions are reported together.

Table 4 – Top 10 significant condition rules dis-aggregated by survivor status (N= 18,956).

Antecedents	Consequents	Support	Confidence	Lift
{No admitted to ICU}	{Survived}	42%	67%	1.28
{Use of non-invasive ventilatory support}	{Survived}	39%	62%	1.19
{No admitted to ICU, no radiological records}	{Survived}	33%	69%	1.32
{No admitted to ICU, use of non-invasive ventilatory support}	{Survived}	32%	67%	1.28
{Use of non-invasive ventilatory support, no radiological records}	{Survived}	31%	64%	1.22
{SpO_2<95%, use of non-invasive ventilatory support}	{Survived}	30%	61%	1.17
{Cough, no admitted to ICU}	{Survived}	30%	70%	1.34
{No admitted to ICU, with comorbidity}	{Survived}	30%	66%	1.26
{SpO_2<95%, no admitted to ICU}	{Survived}	30%	65%	1.25
{Dyspnea, use of non-invasive ventilatory support}	{Survived}	28%	60%	1.16
{Admitted to ICU}	{Died}	27%	73%	1.52
{SpO_2<95%, admitted to ICU}	{Died}	23%	75%	1.57
{Dyspnea, admitted to ICU}	{Died}	22%	75%	1.57
{Admitted to ICU, with comorbidity}	{Died}	22%	74%	1.55
{Use of invasive ventilatory support}	{Died}	20%	87%	1.86
{Dyspnea, SpO_2<95%, admitted to ICU}	{Died}	20%	76%	1.59
{Admitted to ICU, no radiological records}	{Died}	20%	71%	1.49
{SpO_2<95%, admitted to ICU, with comorbidity}	{Died}	19%	76%	1.59
{Admitted to ICU, use of non-invasive ventilatory support}	{Died}	18%	90%	1.88
{Respiratory decompensation, admitted to ICU}	{Died}	18%	76%	1.59

Source: Authors.

It is noteworthy that no significant association rules were found according to the patients' regions of residence and results of radiological examinations.

4. Discussion

Using the ARM technique, this study identified rules of association for symptoms, presence of comorbidities, genders, ICU admission and survivor status of older adults vaccinated against COVID-19 and hospitalized in Brazil. Regarding symptoms, SpO_2<95%, dyspnea and cough were the most frequent. The first two symptoms were reported simultaneously with ICU admission and death in the group of patients with comorbidities. The findings of the most frequent symptoms in this research are similar to those found in other scientific literature (Alimohamadi et5 al., 2020; Fu et al., 2020), and corroborates that of Deng et al. (2020), who collected clinical data from patients in two hospitals in Wuhan, verifying that these symptoms, when reported together with the presence of comorbidity, were also frequent in patients with advanced age, resulting in cardiac and pulmonary complications. Consistent with this, Jacobs et al. (2020), in a prospective cohort carried out in New Jersey with hospitalized patients, identified that the predominance of dyspnea occurs among patients aged 65 to 75 years, with the lowest classification

of quality of life, physical and mental health. These studies were conducted before patients were vaccinated, but recent studies such as the one by Alsaffar et al. (2022), carried out with hospitalized patients who received a single dose of vaccines against COVID-19, still showed the older adults had higher mortality rates than the younger group.

According to Andryukov and Besednova (2021), even if the elderly has their vaccination protocol against COVID-19 complete, it is necessary for there to be policies that monitor the effectiveness of the vaccine in this target audience, since the presence of comorbidity and aging cycles can lead to severe disease. Also, according to the same authors, clinical trials with this population are scarce, due to their vulnerability and, in addition, health services do not always aim to measure the effects of vaccines over time. In our study, we also observed that the association rules for patients without comorbidities had lower support and confidence than those for patients with comorbidities, indicating that there is a predominance of symptom patterns and other hospitalization characteristics among patients with comorbidities.

Furthermore, in this paper, $SpO_2<95\%$ and dyspnea were symptoms that appear simultaneously with female gender in the group of patients with comorbidities. According to Atkins et al. (2020), hospitalized women with COVID-19 may have a higher prevalence of chronic kidney disease and asthma than men, which, added to the deteriorating clinical status, can lead to death. The effect of the pandemic can be further enhanced by considering women vulnerable to poverty and with diminished access to health services (Connor et al., 2020). Some authors such as Mi et al. (2020), Jin et al. (2020), Capuano et al. (2020), Lipsky and Hung (2020) identified that men tend to have higher odds of death than women. Raimondi et al. (2021), in an observational study of COVID-19 patients hospitalized in Bergamo, Italy, pointed out that women may have lower mortality rates than men, but once severe symptoms appear, their probability of dying is like that of men. In this study, the survival characteristic stood out in the top 10 rules for the female group and death for the male group.

$SpO_2<95\%$, dyspnea, use of invasive ventilatory support, and the presence of comorbidity are predominant among the association rules for patients admitted to the ICU and those who died, with confidence above 60%. The clinical conditions of symptoms are related to an increased risk of severity and mortality (Elghazaly et al., 2022). When admission to the ICU is reported concomitantly with use of invasive ventilatory support and death, the confidence was equal to 90%. The use of ventilatory support in hospitalized patients presents risks of developing ventilator-associated pneumonia, recurrent in ICU, with risk factors such as the use of endotracheal tubes, nasogastric contaminated respiratory equipment; permanence of decubitus in the supine position; and preference for nasal intubation (Koenig & Truwit, 2006).

Chang et al. (2021), in a systematic review and meta-analysis study of COVID-19 and its clinical manifestations in ICU patients, found that the use of ventilatory support, acute kidney injury and acute respiratory distress syndrome (ARDS) are risk factors for mortality. $SpO_2<95\%$ can be a predictor of in-hospital mortality in patients with COVID-19 and requires healthcare services to focus on reversing hypoxemia to reduce chances of respiratory decompensation (Mejía et al., 2020). Despite vaccination being referred to as the main mean of preventing mortality from COVID-19, the presence of comorbidity remains a co-determining factor of this outcome in older adults (Lv et al., 2021; PrabhuDas et al., 2021). Given this, it is necessary for governments to invest not only in the production of vaccines, but also in mechanisms to combat chronic diseases.

In Brazil, health is a right established by the Constitution of the Federative Republic of 1988 and regulated by Laws Nº8.080/90 and Nº8.142/90 that regulate the Unified Health System (SUS) – public, universal, and comprehensive. In 2006, the National Primary Care Policy (PNAB) was established, updated in 2012 and 2017, which presents guidelines and principles that are consistent with the universal right to health established by the SUS, and which are applied to guarantee access to basic health services, defined as Primary Health Care (PHC).

The PHC has the potential to solve up to 80% of a population's health problems, as well as reduce the development of chronic and infectious diseases, hospitalization, and mortality rates (Pan American Health Organization, 2021). Passos et al. (2020) list that chronic diseases prevalent in Brazilians are preventable by PHC services and the main ones are ischemic heart

diseases, ischemic and haemorrhagic stroke, chronic obstructive pulmonary disease, and diabetes. All of them are configured as risk factors for mortality from COVID-19 (Nishiga et al., 2020). Misra-Hebert et al. (2021) identified that the pandemic affected users' access to PHC services for monitoring chronic diseases, such as diabetes, as higher levels of glycated hemoglobin were associated with lower chances of health care (in person or virtual) after the beginning of the pandemic.

According to Assis et al. (2021), the regions of Brazil with PHC populational coverage above 75% have lower COVID-19 loads and better early care for symptomatic and comorbid patients, preventing the occurrence of severe cases of the disease. In view of this, it is necessary to invest in PHC actions and in the expansion of its population coverage for the prevention and control of comorbid older adult patients, combined with the application of vaccines, to reduce mortality from COVID-19 in this group.

5. Final Considerations

In the present study, it was observed that the main characteristics of hospitalization among older adults vaccinated against COVID-19 show a profile of presence of comorbidities, critical symptoms of COVID-19 ($SpO_2 < 95\%$ and dyspnea) and use of ventilatory support. The data mining methodology by association rules, in particular the Apriori Algorithm, proved to be useful in surveying the profile of these patients.

That said, there is a need to visualize the profile of patients so that the SUS can organize itself to meet the constant demographic transition in Brazil, to promote the quality of life of this age group and reduce their hospitalization and mortality rates. Among the limitations of this study, we found the use of secondary data, which have probable underreporting and inputting errors. Nor were there any information regarding the variants of the etiological agent of COVID-19, nor of the laboratories producing the vaccines. Furthermore, it is noteworthy that it is not possible to assert causality between the items found in the association rules. Thus, it is recommended that future studies investigate probable causality between the findings presented here.

References

Alimohamadi, Y., Sepandi, M., Taghdir, M., & Hosamirudsari, H. (2020). Determine the most common clinical symptoms in COVID-19 patients: a systematic review and meta-analysis. *Journal of preventive medicine and hygiene*, 61(3), 304–312. https://doi.org/10.15167/2421-4248/jpmh2020.61.3.1530.

Alsaffar, W. A., Alwesaibi, A. A., Alhaddad, M. J., Alsenan, Z. K., Alsheef, H. J., Alramadan, S. H., Aljassas, H. A., Alsaghirat, M. A., & Alzahrani, H. J. (2022). The Effectiveness of COVID-19 Vaccines in Improving the Outcomes of Hospitalized COVID-19 Patients. *Cureus*, 14(1), e21485. https://doi.org/10.7759/cureus.21485.

Andryukov, B. G., & Besednova, N. N. (2021). Older adults: panoramic view on the COVID-19 vaccination. *AIMS public health*, 8(3), 388–415. https://doi.org/10.3934/publichealth.2021030.

Assis, S. J. C., Lopes, J. M., Guedes, M., Sanchis, G., Araujo, D. N., & Roncalli, A. G. (2021). Primary health care and social isolation against COVID-19 in Northeastern Brazil: Ecological time-series study. *PloS one*, 16(5), e0250493. https://doi.org/10.1371/journal.pone.0250493.

Atkins, J. L., Masoli, J., Delgado, J., Pilling, L. C., Kuo, C. L., Kuchel, G. A., & Melzer, D. (2020). Preexisting Comorbidities Predicting COVID-19 and Mortality in the UK Biobank Community Cohort. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 75(11), 2224–2230. https://doi.org/10.1093/gerona/glaa183.

Baqui, P., Bica, I., Marra, V., Ercole, A., & van Der Schaar, M. (2020a). Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *The Lancet Global Health*, 8(8), 1018-1026.

Baqui, P., Marra, V., Alaa, A. M., Bica, I., Ercole, A., & van der Schaar, M. (2021b). Comparing COVID-19 risk factors in Brazil using machine learning: the importance of socioeconomic, demographic and structural factors. *Scientific reports*, 11(1), 15591. https://doi.org/10.1038/s41598-021-95004-8.

Bee, G. R., Pinto, D. D., da Silva, A. C. C. A., Oliveira, T., & Arrigo, J. da S. (2022). Vacinas contra COVID-19 disponíveis no Brasil / Vaccines against COVID-19 available in Brazil. *Brazilian Journal of Development*, 8(1), 6246–6263. https://doi.org/10.34117/bjdv8n1-422.

Brazil (2021). Coronavirus panel.

Candido, D. S., Claro, I. M., de Jesus, J. G., Souza, W. M., Moreira, F., Dellicour, S., Mellan, T. A., du Plessis, L., Pereira, R., Sales, F., Manuli, E. R., Thézé, J., Almeida, L., Menezes, M. T., Voloch, C. M., Fumagalli, M. J., Coletti, T. M., da Silva, C., Ramundo, M. S., Amorim, M. R., & Faria, N. R. (2020). Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science (New York, N.Y.)*, 369(6508), 1255–1260. https://doi.org/10.1126/science.abd2161.

Capuano, A., Rossi, F., & Paolisso, G. (2020). Covid-19 Kills More Men Than Women: An Overview of Possible Reasons. *Frontiers in cardiovascular medicine*, 7, 131. https://doi.org/10.3389/fcvm.2020.00131.

Cash, R., Patel, V. (2020). Has COVID-19 subverted global health? *The Lancet*, 395 (10238) 1687-1688.

Chang, R., Elhusseiny, K. M., Yeh, Y. C., & Sun, W. Z. (2021). COVID-19 ICU and mechanical ventilation patient characteristics and outcomes-A systematic review and meta-analysis. *PloS one*, 16(2), e0246318. https://doi.org/10.1371/journal.pone.0246318.

Connor, J., Madhavan, S., Mokashi, M., Amanuel, H., Johnson, N. R., Pace, L. E., & Bartz, D. (2020). Health risks and outcomes that disproportionately affect women during the Covid-19 pandemic: A review. *Social science & medicine (1982)*, 266, 113364. https://doi.org/10.1016/j.socscimed.2020.113364.

Deng, Y., Liu, W., Liu, K., Fang, Y. Y., Shang, J., Zhou, L., Wang, K., Leng, F., Wei, S., Chen, L., & Liu, H. G. (2020). Clinical characteristics of fatal and recovered cases of coronavirus disease 2019 in Wuhan, China: a retrospective study. *Chinese medical journal*, 133(11), 1261–1267. https://doi.org/10.1097/CM9.0000000000000824.

Elghazaly, S., Abu, E. S. H., Samy, S., AbdelFatah, M., Hassany, M., Khader, Y., Afifi, S., & Eid, A. (2022). The Impact of Comorbidities on COVID-19 Severity and Mortality in Egypt iproc. *Jmir Publications*, 8(1):e36576. doi: 10.2196/36576.

FGV - Fundação Getúlio Vargas (2021). Older adults Maps.

Fu, L., Wang, B., Yuan, T., Chen, X., Ao, Y., Fitzpatrick, T., Li, P., Zhou, Y., Lin, Y. F., Duan, Q., Luo, G., Fan, S., Lu, Y., Feng, A., Zhan, Y., Liang, B., Cai, W., Zhang, L., Du, X., Li, L., & Zou, H. (2020). Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: A systematic review and meta-analysis. *The Journal of infection*, 80(6), 656–665. https://doi.org/10.1016/j.jinf.2020.03.041.

Hahsler, M., Buchta, C., Grun, B., & Hornik, K. (2022). Arules: Mining Association Rules and Frequent Itemsets, Package Version 1.7-3 of rproject.org.

Jacobs, L. G., Gourna Paleoudis, E., Lesky-Di Bari, D., Nyirenda, T., Friedman, T., Gupta, A., Rasouli, L., Zetkulic, M., Balani, B., Ogedegbe, C., Bawa, H., Berrol, L., Qureshi, N., & Aschner, J. L. (2020). Persistence of symptoms and quality of life at 35 days after hospitalization for COVID-19 infection. *PloS one*, 15(12), e0243882. https://doi.org/10.1371/journal.pone.0243882.

Jean-Marc, A. (2001). Data mining for association rules and sequential patterns: sequential and parallel algorithms. *Springer Science & Business Media*.

Jin, J. M., Bai, P., He, W., Wu, F., Liu, X. F., Han, D. M., Liu, S., & Yang, J. K. (2020). Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. *Frontiers in public health*, 8, 152. https://doi.org/10.3389/fpubh.2020.00152.

Koenig, S. M., & Truwit, J. D. (2006). Ventilator-associated pneumonia: diagnosis, treatment, and prevention. *Clinical microbiology reviews*, 19(4), 637–657. https://doi.org/10.1128/CMR.00051-05.

Lipsky, M. S., & Hung, M. (2020). Men and COVID-19: A Pathophysiologic Review. *American journal of men's health*, 14(5), 1557988320954021. https://doi.org/10.1177/1557988320954021.

Lv, G., Yuan, J., Xiong, X., & Li, M. (2021). Mortality Rate and Characteristics of Deaths Following COVID-19 Vaccination. *Frontiers in medicine*, 8, 670370. https://doi.org/10.3389/fmed.2021.670370.

Mejía, F., Medina, C., Cornejo, E., Morello, E., Vásquez, S., Alave, J., Schwalb, A., & Málaga, G. (2020). Oxygen saturation as a predictor of mortality in hospitalized adult patients with COVID-19 in a public hospital in Lima, Peru. *PloS one*, 15(12), e0244171. https://doi.org/10.1371/journal.pone.0244171.

Mendes, E. V. (2018). The care of chronic conditions in primary health care. *Revista Brasileira em Promoção da Saúde*, 31 (2), 1-3.

Mi, J., Zhong, W., Huang, C., Zhang, W., Tan, L., & Ding, L. (2020). Gender, age and comorbidities as the main prognostic factors in patients with COVID-19 pneumonia. *American journal of translational research*, 12(10), 6537–6548.

Misra-Hebert, A. D., Hu, B., Pantalone, K. M., & Pfoh, E. R. (2021). Primary Care Health Care Use for Patients With Type 2 Diabetes During the COVID-19 Pandemic. *Diabetes care*, 44(9), 173–174. https://doi.org/10.2337/dc21-0853.

Moreno-Perez, O., Ribes, I., Boix, V., Martinez-García, M. Á., Otero-Rodriguez, S., Reus, S., Sánchez-Martínez, R., Ramos, J. M., Chico-Sánchez, P., Merino, E., & On behalf the COVID-19 ALC research group (2022). Hospitalized patients with breakthrough COVID-19: Clinical features and poor outcome predictors. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases,* 118, 89–94. https://doi.org/10.1016/j.ijid.2022.02.007.

Nishiga, M., Wang, D. W., Han, Y., Lewis, D. B., & Wu, J. C. (2020). COVID-19 and cardiovascular disease: from basic mechanisms to clinical perspectives. *Nature reviews. Cardiology*, 17(9), 543–558. https://doi.org/10.1038/s41569-020-0413-9.

Pan American Health Organization. (2021). What is primary health care?

Passos, V. M. A., Champs, A., Teixeira, R., Lima-Costa, M., Kirkwood, R., Veras, R., Nascimento, B. R., Nogales, A. M., Schmidt, M. I., Duncan, B. B., Cousin, E., Naghavi, M., & Souza, F. M. (2020). The burden of disease among Brazilian older adults and the challenge for health policies: results of the Global Burden of Disease Study 2017. *Population health metrics*, 18(Suppl 1), 14. https://doi.org/10.1186/s12963-020-00206-3.

OpenDataSUS Platform by Federal Government (2022).

PrabhuDas, M., Fuldner, R., Farber, D., Kuchel, G.A., Mannick, J., Nikolich-Zugich, J., Sen, R., & Turner, J. (2021). Research and resource needs for understanding host immune responses to SARS-CoV-2 and COVID-19 vaccines during aging. *Nat Aging* 1, 1073–1077. https://doi.org/10.1038/s43587-021-00156-x.

Raimondi, F., Novelli, L., Ghirardi, A., Russo, F. M., Pellegrini, D., Biza, R., Trapasso, R., Giuliani, L., Anelli, M., Amoroso, M., Allegri, C., Imeri, G., Sanfilippo, C., Comandini, S., Hila, E., Manesso, L., Gandini, L., Mandelli, P., Monti, M., Gori, M., & HPG23 Covid-19 Study Group (2021). Covid-19 and gender: lower rate but same mortality of severe disease in women-an observational study. *BMC pulmonary medicine*, 21(1), 96. https://doi.org/10.1186/s12890-021-01455-0.

Rocha, R., Atun, R., Massuda, A., Rache, B., Spinola, P., Nunes, L., Lago, M., & Castro, M. C. (2021). Effect of socioeconomic inequalities and vulnerabilities on health-system preparedness and response to COVID-19 in Brazil: a comprehensive analysis. *The Lancet. Global health*, 9(6), 782–792. https://doi.org/10.1016/S2214-109X(21)00081-4.

Shawkat, M., Mahmoud, B., & Ali, I.E. (2021). "A Novel Approach of Frequent Itemsets Mining for Coronavirus Disease (COVID-19)." *European Journal of Electrical Engineering and Computer Science*, 5(2), 5-12.

Shin, D. P., Park, Y. J., & Seo, J. (2018). Association Rules Mined from Construction Accident Data. *KSCE J Civ Eng,* 22 (4), 1027–1039 (2018). https://doi.org/10.1007/s12205-017-0537-6.

Tandan, M., Acharya, Y., Pokharel, S., & Timilsina, M. (2021). Discovering symptom patterns of COVID-19 patients using association rule mining. *Computers in biology and medicine*, 131, 104249. https://doi.org/10.1016/j.compbiomed.2021.104249.

Vaccination Brazil Platform by Federal Government (2022).

WHO - World Health Organization (2021). *Panel COVID-19 by WHO*.

Williams, G. (2011). Data mining with Rattle and R: The art of excavating data for knowledge discovery. *Springer Science & Business Media.*

Zhu, J., Yan, W., Zhun, L., & Liu, J. (2021). COVID-19 pandemic in BRICS countries and its association with socio-economic and demographic characteristics, health vulnerability, resources, and policy response. *Infect Dis Poverty*, 10 (97).