Computers in Biology and Medicine 131 (2021) 104249

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Discovering symptom patterns of COVID-19 patients using association rule mining

Meera Tandan, PhD [a,*], Yogesh Acharya, MD [b], Suresh Pokharel, PhD [c], Mohan Timilsina, PhD [d]

[a] Cecil G Sheps Center for Health Service Research, University of North Carolina, Chapel Hill, USA
[b] Western Vascular Institute, Galway University Hospital, Galway, Ireland
[c] The University of Queensland, St Lucia, Queensland, Australia
[d] Data Science Institute, Insight Centre for Data Analytics, National University of Ireland Galway, Ireland

ARTICLE INFO

Keywords:
COVID-19
Symptoms
Association rule mining
Chronic disease

ABSTRACT

Background: The COVID-19 pandemic is a significant public health crisis that is hitting hard on people's health, well-being, and freedom of movement, and affecting the global economy. Scientists worldwide are competing to develop therapeutics and vaccines; currently, three drugs and two vaccine candidates have been given emergency authorization use. However, there are still questions of efficacy with regard to specific subgroups of patients and the vaccine's scalability to the general public. Under such circumstances, understanding COVID-19 symptoms is vital in initial triage; it is crucial to distinguish the severity of cases for effective management and treatment. This study aimed to discover symptom patterns and overall symptom rules, including rules disaggregated by age, sex, chronic condition, and mortality status, among COVID-19 patients.

Methods: This study was a retrospective analysis of COVID-19 patient data made available online by the Wolfram Data Repository through May 27, 2020. We applied a widely used rule-based machine learning technique called association rule mining to identify frequent symptoms and define patterns in the rules discovered.

Result: In total, 1,560 patients with COVID-19 were included in the study, with a median age of 52 years. The most frequently occurring symptom was fever (67%), followed by cough (37%), malaise/body soreness (11%), pneumonia (11%), and sore throat (8%). Myocardial infarction, heart failure, and renal disease were present in less than 1% of patients. The top ten significant symptom rules (out of 71 generated) showed cough, septic shock, and respiratory distress syndrome as frequent consequents. If a patient had a breathing problem and sputum production, then, there was higher confidence of that patient having a cough; if cardiac disease, renal disease, or pneumonia was present, then there was a higher confidence of septic shock or respiratory distress syndrome. Symptom rules differed between younger and older patients and between male and female patients. Patients who had chronic conditions or died of COVID-19 had more severe symptom rules than those patients who did not have chronic conditions or survived of COVID-19. Concerning chronic condition rules among 147 patients, if a patient had diabetes, prerenal azotemia, and coronary bypass surgery, there was a certainty of hypertension.

Conclusion: The most frequently reported symptoms in patients with COVID-19 were fever, cough, pneumonia, and sore throat; while 1% had severe symptoms, such as septic shock, respiratory distress syndrome, and respiratory failure. Symptom rules differed by age and sex. Patients with chronic disease and patients who died of COVID-19 had severe symptom rules more specifically, cardiovascular-related symptoms accompanied by pneumonia, fever, and cough as consequents.

1. Introduction

The corona-virus disease 2019 (COVID-19) pandemic is a public health crisis [1]. As of December 20, 2020, a total of 76,710,234 cases and 1,693,700 deaths had been recorded globally [2]. Some European nations have managed to "flatten" the curve; other countries, like the USA, Brazil, India, and Russia are still struggling [2]. In the meantime, the UK, Germany, Spain, Poland, and Japan are experiencing a second

* Corresponding author.
E-mail addresses: meerat@email.unc.edu (M. Tandan), dryogeshach@gmail.com (Y. Acharya), s.pokharel@uq.edu.au (S. Pokharel), mohan.timilsina@insight-centre.org (M. Timilsina).

https://doi.org/10.1016/j.compbiomed.2021.104249
Received 21 December 2020; Received in revised form 25 January 2021; Accepted 25 January 2021
Available online 1 February 2021

wave. Scientists all over the world are working hard and competing to develop potential therapeutics and vaccines. The RECOVERY trial's preliminary results showed benefits of steroids in hospitalized patients with COVID-19 under respiratory support [3]. According to the World Health Organization's SOLIDARITY trial, initial drugs proposed early in the pandemic such as hydroxychloroquine, and lopinavir/ritonavir, did not reduce mortality in hospitalized patients; therefore, they were dropped from the trial [4]. At the same time, remdesivir showed promising efficacy [5], even though these trials were methodologically inferior and were designed before scientists understood the disease progression well [6]. As of December 2020, three therapeutics have been approved for the treatment of COVID-19 [7]. These include remdesivir in the USA, Japan, and Australia; dexamethasone in the UK and Japan; and favilavir in China, Italy, and Russia. However, there is still a limited understanding of these drugs' efficacy in patients with perceived contraindications, including uncontrolled diabetes, delirium, underlying malignancy, immunosuppression, or conditions in which steroids might have harmful effects [8,9].

Similarly, much progress has been made in the development of vaccines. Scientists are testing 63 vaccines in human in clinical trials, and 18 have reached the final stages [10]. Two of the vaccines, BNT162b2 by Pfizer and BioNTech and mRNA-1273 by Moderna, demonstrated 95% efficacy, therefore, they were provided emergency authorization for use in the USA, Canada, and many other countries [10]. According to data collected by Bloomberg, as of December 19, 2020, the first doses of COVID-19 shots had been given to more than 1.6 million people in four countries (the USA, the UK, China, and Russia) [11]. However, the safety profile of these vaccines in some specific subgroups, such as elderly and people with chronic comorbidities, is still unanswered. Furthermore, it is not clear whether the companies producing these vaccines would supply them as demanded, or when the world population would be vaccinated and protected against COVID-19. According to a report from the Center for Infectious Disease Research and Policy, at least 60%–70% of the human population must be immune for the COVID-19 pandemic to end [12]. Hence, they argue that the pandemic will remain for at least another 18–24 months, with hot spots popping up periodically in diverse geographic areas, while assuming at least some level of ongoing mitigating measures [12]. From this, we can easily speculate the importance of implementing appropriate public health measures, such as screening people with compatible symptoms and determining candidates for testing, quarantine, and hospital care. These measures are critical in the containment and symptomatic management of COVID-19. A wide variety of symptoms, ranging from those of a mild common cold to severe systemic complications, have been reported for COVID-19 [13]. Understanding these symptom patterns helps clinicians and healthcare workers in their clinical decision-making to provide effective supportive and therapeutic care.

With an unprecedented rise in global COVID-19 cases, many studies have emerged defining associated clinical disease characteristics, comorbidities, and epidemiological determinants [14,15]. However, modeling studies regarding COVID-19 that address associations between various disease determinants are scarce. Modern computing ability has made structured data extraction and mining possible, providing us with the ability to perform multiple data-related activities, such as sequential data classification, clustering, summarizing, and similarity analysis, which can be utilized to establish an association between different clinical parameters to predict likely outcomes [16]. The outbreak of COVID-19 is a significant challenge for clinicians and public health professionals. In this study, we have used data mining techniques to extract patterns of COVID-19 symptoms. These symptom pattern mining methods can act as complementary techniques to help us better understand the disease pattern in clinical settings.

Artificial Intelligence (AI) has huge potential in medicine. Companies like Alibaba developed AI solutions to help China fight against COVID-19 and predict the peak, size, and duration of the outbreak, and it was claimed to have high implementation accuracy in real-world tests in various regions of China [17]. Different types of respiratory disease can be resolved using machine learning based CT image analysis, which can effectively diagnose patients with COVID-19 [18]. It is believed that the development of COVID-19 vaccines may also be accelerated by analyzing genome sequences and molecular docking, and deploying various machine learning, and artificial intelligence techniques [19,20].

Motivation: Although some computational research related to COVID-19 has been done, most approaches have focussed on complex methods (e.g., deep neural networks) for predictions [21,22]. Simple yet explainable techniques are undervalued. The simple association rules will find every pattern in a given data set, which is useful for clinical data analysis. It further offers clinicians the option to quickly and automatically conduct well-informed diagnoses, extract invaluable information, and develop essential knowledge bases. This study discover symptom patterns in COVID-19 patients and explores symptom patterns disaggregated by age, sex, chronic condition, and mortality.

Contribution: The main contribution of this work is summarized below:

- We address the problem of automatically identifying new and useful symptom patterns in COVID-19 data using Apriori rule-based data mining algorithm.
- We demonstrated the statistically significant rules in different subgroups of patients, namely age, sex, chronic condition, and mortality.
- To the best of our knowledge, this is the first study to apply simple yet powerful rule mining algorithms to mine the frequent symptoms for COVID-19 patients. We believe these rules aid clinicians in decision making.

The rest of this article is organized as follows. In the next section 2, we provided a brief literature review on various related works. We described our methodology and data sets in section 3. In section 4, we demonstrated the experimental results that validate our approach's effectiveness. Finally, we discussed the study findings and limitations in section 5 and derived conclusions and potential future works of the study in section 6.

2. Related work

In recent days, machine learning techniques have been widely used in biomedical studies for prediction and knowledge discovery [23,24]. There are several applications of machine learning in biomedicine, such as genomic analysis [25–27], disease-gene analysis [28–31], mortality prediction [32], personalized medicine [33,34], drug discovery [35–38], prediction of adverse drug events [39–42], patient similarity [43,44], and explainable artificial intelligence methods in medicine [45–48]. One area for machine learning approaches in medicine is the association rule mining (ARM).

ARM was first proposed by R. Agrawal [49–51]. Initially, it was applied for sales data, where the task was to identify all the rules that would predict an item's occurrence (or items) based on the occurrence of other items from a given "set of transactions". The primitive idea of ARM is a brute-force approach. In this approach, all the possible rules are listed first, and those rules that do not satisfy the given condition[1] are pruned. However, this approach is computationally prohibitive due to the huge number of possible combinations. To reduce the number of candidates, R. Agrawal [51] proposed a method called Apriori. The Apriori method has two major shortcomings. First, it generate large number of candidate item sets while generating frequent item sets in a more bigger data set. Second it needs multiple scans of the database, which leads to higher computational costs. To overcome these limitations, Han et al. [52] proposed the Frequent Pattern Growth (FP growth)

[1] two thresholds values, minimum support, and minimum confidence, are used.

method. The FP growth method constructs a tree, representing the data set that maintains the association between the item sets. The FP growth has its own disadvantages. Constructing an FP tree is difficult compared to Apriori, and if the database is too large, the algorithm may not fit into shared memory. Apriori and FP growth both use a horizontal data format. Similarly, Zaki et al. [53] proposed the equivalence class clustering and bottom-up lattice transversal algorithm (Eclat) for ARM, which transforms the horizontal data format into a vertical one. The advantage of Eclat over Apriori is that it takes less database scanning. However, this approach's main drawback is that it consumes enormous memory when there are many transactions in the data set. A qualitative comparison between these state of the-art methods is shown in Table 1.

ARM is an active research field in the data mining community [54–56]. Recently, different incremental algorithms have been proposed for mining association rules to extract discovered patterns [57,58]. In the past, ARM was used to solve various problems in healthcare. Usually, there are many hidden relationships between the attributes (symptoms and diseases). Discovering these relationships help researcher to better understand a disease and its biomarkers. Some studies [59–61] have identified risk factors of heart diseases. Vladimir et al. [62] identified early childhood caries using ARM. Borah and Nath [63] came up with the concept of dynamic rare association rule mining for mining different risk factors of cardiovascular disease, hepatitis, and breast cancer. Sharma et al. [64] applied ARM for mitigating the increasing obesity problem, which is primarily caused by lack of physical exercise. Noguchi et al. [65,66] used ARM to find adverse events caused by drug-drug interactions. Ramasamy and Nirmala [67] applied ARM with an additional keyword-based clustering technique to predict disease. Kamalesh et al. [68] predicted the risk of diabetes mellitus using ARM. Pokharel et al. [69] used sequential pattern mining with gap constraint to find similarities between patients, including mortality prediction and identification of sepsis patients.

In the context of symptom mining using ARM, the study by Nahar et al. [59] demonstrated factors contributing to heart disease for male and female cohorts. Similarly, Borah et al. [63] identified symptoms and risk factors for three adverse diseases (cardiovascular disease, hepatitis, and breast cancer) using ARM. Lau et al. [70] developed constraint-based ARM across subgroups to help clinicians find useful patterns in patients with dyspepsia. Yeleswarapu et al. [71] applied ARM to extract drug symptom pairs for concept/relation extraction. This work further supports that ARM is a powerful method of capturing patient symptoms to discover new pattern in medical databases.

Previous COVID-19-related studies have focused on predicting numbers of cases [72–74], and classifying COVID-19 patients from real-world x-ray data sets using complex deep neural network methods [75,76]. However, these methods focus on examining symptom patterns of COVID-19. Thus, in the current paper, we focus on simple pattern mining techniques known as ARM to provide a descriptive approach for extraction of symptom rules. No previous studies have focused on analyzing COVID-19 using ARM. In the present study, we aimed to discover the hidden relationships between symptom patterns of COVID-19 patients using ARM, which can aid in clinical decision making for the management of patients with COVID-19.

3. Methods

3.1. Study design and population

This study is a retrospective study of data from COVID-19 patients. Data were extracted on June 2, 2020 from the online platform of the Wolfram Data Repository (2020) [77]. The last date of data included was May 27, 2020.

3.2. Data management and analysis

Following the extraction of anonymized COVID-19 patients data from the online platform [77], we exported and cleaned data in R[2] version 3.4, a data management and analysis software. The primary purpose of the study was symptom mining; therefore, we created a data set for patients with symptom information and excluded all missing values. The data set had 24 variables; however, we extracted only age, sex, symptoms, chronic conditions, and mortality information. We categorized related symptoms with similar meanings for consistency (see Supplementary Table 1). For a composite, yet clinically recognized symptom pattern, the data assigns the specific disease associated with these symptoms within the symptom variable. For example, 'pneumonia' has been used within symptoms variables for patients with congregations of recognized clinical symptoms consistent with chest infection. A cleaned "symptom data" was then converted to a "transaction" format and analyzed using the Apriori algorithm, available as "arules"[3] package in R. The data management is presented in Fig. 1.

3.3. Association rule mining (ARM)

ARM discovers the pattern of frequent items or events in the data set, including the association between items or events. The pattern exposes the combination of the items or events that occur at the same time. In medicine, it is helpful to know how one disease is associated with others, for instance, diabetes and hypertension. In the context of medicine, an association rule between symptom (or disease) is expressed in the form $X \rightarrow Y$, where X and Y are a disjoint set of symptom (or disease), i.e., $X \cap Y = \varphi$. In other words, X is called the antecedent of the rule, and Y is called the consequent. Also, known as "*if→then*", "if" represents antecedent, and "then" represents consequent. Generally, the effectiveness of discovered rules is measured in terms of i) Support, ii) Confidence and iii) Lift. Formally, support can be defined as; $Support(X \rightarrow Y) = \frac{Patients\ having\ both\ X\ and\ Y}{Total\ number\ of\ patients}$. So, support determines the frequency (i.e., generality) of a rule to a given data set. Confidence can be defined as; $Confidence(X \rightarrow Y) = \frac{Patients\ having\ both\ X\ and\ Y}{Patients\ having\ X}$. Here, confidence determines how frequently symptom (or disease) Y appear in those who have X. Lift can be defined as; $Lift(X \rightarrow Y) = \frac{(Patients\ having\ both\ X\ and\ Y)/(Patients\ having\ X)}{(Fraction\ of\ patients\ having\ Y)}$, where fraction of patient having Y is number of patient having Y divided by total number of patients. Lift suggests how often symptom Y appear when symptom X appear while controlling the likely occurrence of symptoms Y. Value of lift determines the correlation between X and Y; independent ($= 1$), positive related (> 1), negative related (< 1). The drawback of the "Confidence" measures is that it might misrepresent the importance of an association. For example, in an association $X \Rightarrow Y$, Confidence score only accounts for how important item X is, but not Y. If Y is also essential in general, there would be a higher chance that a pattern containing X would contain Y, thus inflating the confidence measure. The metric lift solves this problem by measuring the strength of association between X and Y.

Fig. 2 shows data in transaction format including how ARM calculated these measures and generated rules. In this figure, we have eight patients, and the application of the rule mining algorithm gave us three rules where the antecedent (X) = [stroke] and consequent (Y) = [hypertension] in the rule1, with support score of 0.5, confidence 0.80 and lift 1.28. Support 0.5 tells that out of eight patients, four patients have "stroke and hypertension". Confidence 0.80 means 80% of the patients with stroke had hypertension. Similarly, lift 1.28 tells that "stroke" and "hypertension" are positively co-related.

In our study, we considered each patient as a single transaction. We first applied ARM to symptom data and discovered symptom rules. We then filtered the redundant rules and identified significant rules with the application of the "fisher exact" test for pattern discovery [78]. The

[2] https://www.r-project.org/.
[3] https://cran.r-project.org/web/packages/arules/.

Table 1
Qualitative comparison between different association rule mining algorithms.*tid: transaction ID.

Particular	Apriori	FP Growth	Eclat
Data format	Horizontal	Horizontal	Vertical
Search strategy	Breath first search	Depth first search	Breath first search
Memory	Candidates are saved in memory	A compact version of database is saved	Transaction sets are saved in memory
Scan database	Multiple times	Only two times	Only one time
Pros	It is very simple and straightforward to implement	– Only needs two times database scan – Use advantages of tree structure	– Only needs one time database scan – Fast support for counting
Cons	Computationally expensive as: – Need to generate large number of candidate sets – Need multiple times database scan	– FP tree may not fit in main memory – FP tree construction is cumbersome	– Intermediate *tid-lists may become too large to fit in memory – It needs more time for intersection when tid list is large

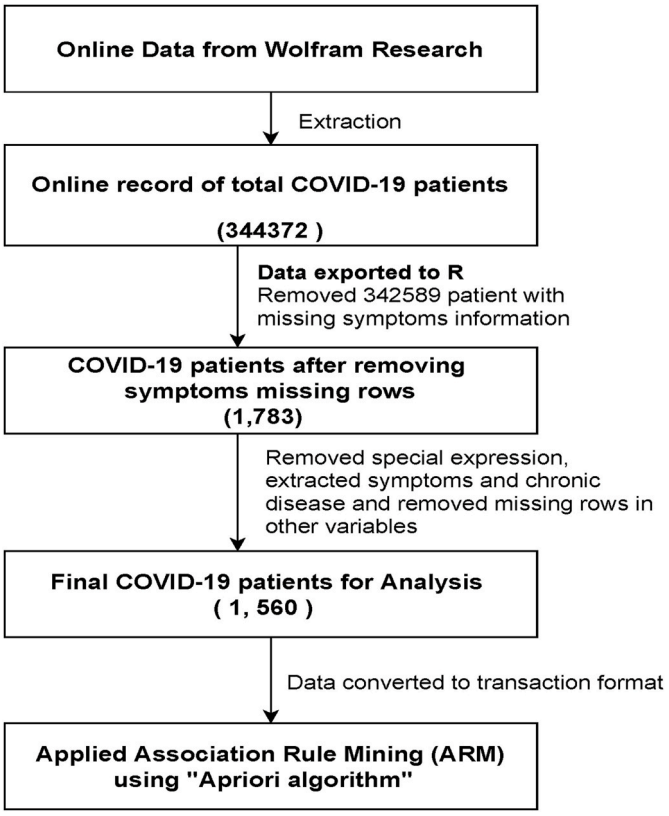


Fig. 1. Data extraction and management.

redundant rule is defined as follows: A rule $X \Rightarrow Y$ is redundant if $\exists X' \subset X$ $conf(X' \Rightarrow Y) \geq conf(X \Rightarrow Y)$, where conf is the confidence score [79]. The same approach has been used by R packages[4] for filtering the redundant rules. The following formulation can be interpreted as:

- Rule 1: $X \Rightarrow Y$ with confidence $c1$.
- Rule 2: $X' \Rightarrow Y$ with confidence $c2$ where X' is a subset of X.

Rule 1 is considered to be redundant if Rule 2 has a higher confidence than Rule 1 i.e $c1 > c2$ (where X' is a subset of X). In other words, if there exists a rule where a subset of the left hand side (LHS) can provide the right hand side (RHS) with more confidence then prior rule is said to be redundant. For the statistical test, we employed Fisher's exact test with correction for multiple comparisons to test null hypothesis that the LHS and the RHS of the rules are independent. Following this method we crafted all rules presented in Tables 2–6.

[4] https://www.rdocumentation.org/packages/arules/versions/1.6-6/topics/is.redundant.

Likewise, we added the variable sex in symptom data and followed a similar approach to discover symptom rules between male and female patients. Simultaneously, we added data on age categories (<20 years/ 20–45 years/45–65 years/>65 years), chronic disease (yes/no), and death (survived/died) independently and we discovered symptom rules between categories.

The ARM algorithm containing symptom transactions aims to construct frequent item sets, having at least a user-specified threshold. Thus, we followed the same approach as Nahar et al. [59] by setting a "confidence" threshold of 0.9, or 90%. This was because the "confidence" metric is used to rank the rules [50,80]. We set up a threshold value of minimum support above 0.001 and "lift" greater than 1 for positively correlated rules. Herein, we report only the top 10 rules with the highest support scores.

To capture rare or infrequent items, we chose low support and high confidence measures. We borrowed this idea from the study by McCormick et al. [81] for mining medical symptoms. When a symptom that rarely occurs is strongly linked with another rare symptom, it is essential to not exclude the rules characterizing these symptoms. Such rules provide valuable insight to clinicians for a novel disease like COVID-19. In other domains, such as business, the threshold with low support and high confidence will produce few rules which may not be interesting for customer analytics. We agree that constraining low support and high confidence gathers very few rules, but the results can be of great interest to clinicians, as they could explain lesser known phenomena [82]. It is often true in medical diagnosis where many symptom combinations will only manifest in a small number of patient cases. Hence, such an approach for mining the corresponding patterns and rules will support a more focused analysis of symptom discovery.

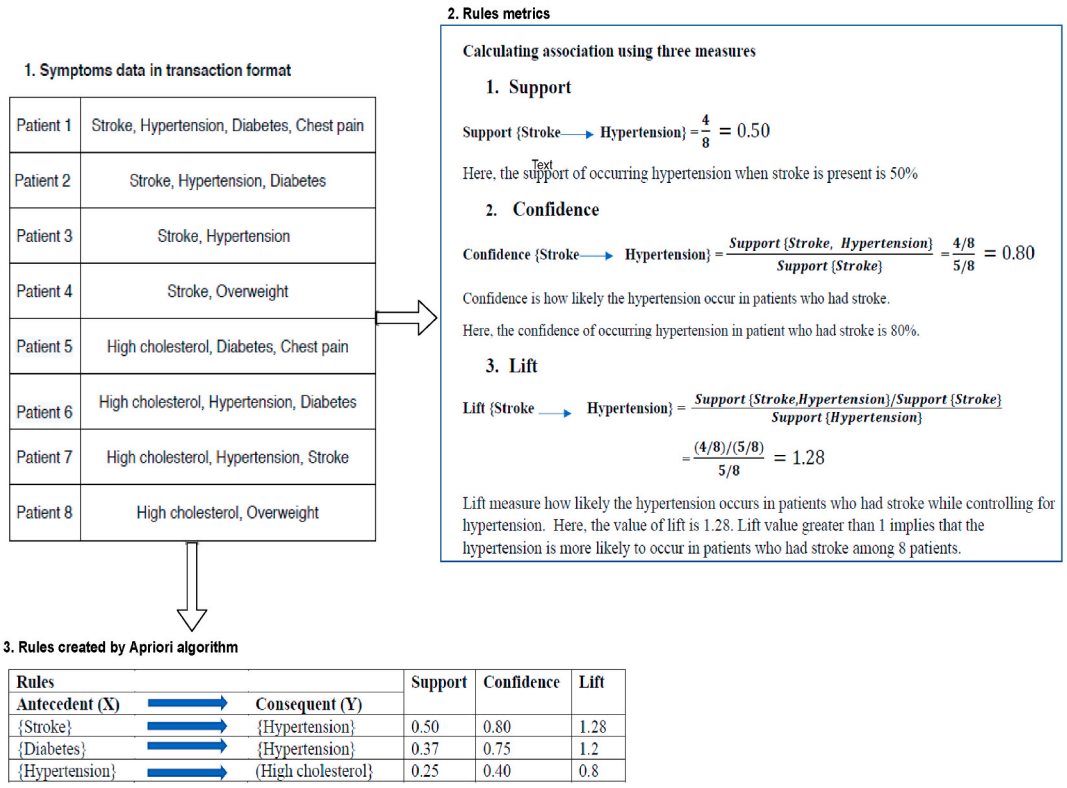
4. Results

Information was extracted for a total of 3,44,372 patients, of which 1783 had reported symptoms. We then analyzed data from 1560 patients after removing those with "missing" and "not available" values (Fig. 1). The median age was 52 years (SD ± 31.5 years; IQR 66 years), 57% of the patients were male. Of the total, 9% (147) had chronic conditions, and 8% (125) died due to COVID-19.

Fever (67%) was the most common symptom, followed by cough (37%), malaise/body soreness (11%), pneumonia (11%), and sore throat (8%). Headache, sputum production, nausea, diarrhoea, respiratory distress syndrome, and septic shock were each reported in 1–5% patients. Symptoms such as myocardial infarction, heart failure, and renal disease were reported in less < 1% (Fig. 3a). The frequency of chronic hypertension was 5%, diabetes 4%, and kidney and coronary heart disease 1% (Fig. 3b).

4.1. Symptom rules

We discovered 71 significant rules for the data that included symptom-only information and excluded other variables in the data set. The top 10 symptom rules by highest support scores are presented in Table 2. Among the top 10 rules, cough was the most common



Rules			Support	Confidence	Lift
Antecedent (X)	→	Consequent (Y)			
{Stroke}	→	{Hypertension}	0.50	0.80	1.28
{Diabetes}	→	{Hypertension}	0.37	0.75	1.2
{Hypertension}	→	(High cholesterol}	0.25	0.40	0.8

Fig. 2. Example of data in transaction format, measures and rules.

consequent (4), followed by septic shock (2), respiratory distress syndrome (2), and pneumonia and nausea (1 each). If a patient had a breathing problem and sputum production, there was a 100% confidence that he or she had a cough. Similarly, patients with respiratory failure and septic shock had pneumonia as a consequent.

For a demonstration purpose, the rules are visualized in Fig. 4. Let's take the example of rule number ten represented by **R10** in a green node. There are three symptom nodes "diarrhoea", "anorexia" and "nausea" represented by pink nodes. These three nodes forms a rule where the antecedent is "diarrhoea", "anorexia" and the consequent is "nausea". Both the nodes in the antecedent have outgoing links, which are of pink color pointed toward the **R10** nodes. Similarly there is a outgoing link in **R10**, which is in green color pointing toward the consequent "nausea" node.

In patients with chronic conditions only (n = 147), the algorithm discovered two significant rules. If a patient had diabetes, prerenal azotemia, and coronary bypass surgery (antecedent), then this patient had a higher confidence of presenting hypertension (consequent) (Table 2).

When patients were disaggregated by age, 12 significant rules were discovered for those <20 years of age, 20 for 20–45 years, 8 for 45–65 years, and 16 for >=65 years. The topmost rule for patients <20 years of age was [conjunctivitis, rhinorrhea]; for 20–45 years, [dry mouth, sore throat]; for 45–65 years, [nausea, weakness]; and for >=65 years, [anorexia, fever] (Table 3). In patients >45 years, heart related-symptoms(e.g., heart failure, cardiac arrhythmia, myocardial infarction), and respiratory problems (e.g.,pneumonia, sore throat) comprised most of the rules.

ARM generated 33 rules for males and 36 for females (Table 4). A difference in symptom rules was observed between the sexes. The top rules in males were [malaise/body soreness, weakness], [cough, diarrhoea] and [fever, malaise/body soreness, pneumonia] while; those in females were [cough, rhinorrhea, sore throat], [pneumonia, rhinorrhea], and [fever, sore throat, weakness].

Eleven rules were generated for patients with chronic conditions and 49 rules were generated for patients without chronic conditions. The symptoms were mild for those without chronic conditions- [headache, malaise/body soreness], [sore throat, weakness], and [headache, rhinorhea]. The symptom were more serious for those with chronic conditions- [cardiac arrhythmia, septic shock], [respiratory failure, septic shock], and [hypertension, renal disease] (Table 5).

Similarly, 76 symptom rules were identified for patients who survived COVID-19, and 7 were identified for those who died Table 6. The symptom rules among patients who died were more severe and complicated than in those who survived. The most common rules discovered in patients who died of COVID-19 were [cardiac arrhythmia, septic shock], [cardiac arrhythmia, respiratory distress syndrome], and [myocardial infarction, respiratory failure].

4.2. Run-time comparison of the rule mining algorithms

We performed a run time comparison of the state of the art rule-based algorithms in our whole COVID-19 symptom data sets. We provided the same parameters for support and confidence for Apriori, FP growth, and Eclat for a fair evaluation. Fig. 5 shows the computational time for the rule extraction task. From the experiment, we observed that the simple Apriori algorithm completed more quickly than the other algorithms. However, the difference was marginal. As the data set was very small, this might be one reason we did not see a high computational time advantage of the FP growth method, which is claimed to be scalable for large transaction data sets.

5. Discussion

We discovered symptom rules for COVID-19 patients using ARM techniques. This commonly used data mining application determines the patterns of items or events [83,84]. We performed this analysis to determine whether differences occurred in symptom rules of COVID-19 according to age group, sex, presence of chronic disease, and mortality status. Our study reported a relatively higher proportion of fever, cough,

Table 2
Top 10 significant rules for symptom and top 2 for chronic condition.

Symptoms (N = 1560)					
Rules	Antecedents	Consequents	Support	Confidence	Lift
Rule 1	{Breathing problem, Sputum}	{Cough}	0.004	1.0	2.7
Rule 2	{Respiratory failure, Septic shock}	{Pneumonia}	0.002	1.0	8.7
Rule 3	{Sputum, Weakness}	{Cough}	0.002	1.0	2.7
Rule 4	{Conjuctivites, Rhinorhea}	{Cough}	0.001	1.0	2.7
Rule 5	{Cardiac arrythmia,Renal disease}	{SepticShock}	0.001	1.0	70.9
Rule 6	{Cardiac arrythmia, Renal disease}	{Respiratory distress syndrome}	0.001	1.0	37.1
Rule 7	{Fever, Heart failure}	{Cough}	0.001	1.0	2.7
Rule 8	{Pneumonia, Renal disease}	{SepticShock}	0.001	1.0	70.9
Rule 9	{Pneumonia, Renal disease}	{Respiratory Distress Syndrome}	0.001	1.0	37.1
Rule 10	{Anorexia, Diarrhoea}	{Nausea}	0.001	1.0	48.8
Chronic Conditions (N = 147)					
Rule 1	{Diabetes, Prerenal azotemia}	{Hypertension}	0.001	1.0	19.3
Rule 2	{Cornary bypass surgery, Diabetes}	{Hypertension}	0.001	1.0	19.3

Note: Green circles (nodes) represent top 10 rules by support, R1 = rule1 and so on. Higher the support value, the larger the green nodes. Pink circles (nodes) represent symptom. Symptom with pink arrows towards rules (e.g,sputum - > R1) are antecedents. Symptom with green arrow outwards rules (e.g, cough < - R1) are consequents.

pneumonia, malaise/body soreness, and sore throat in COVID-19 patients, as indicated in a recent systematic review including 148 studies [85]. Overall, our study demonstrated a significant difference in the presenting symptoms between younger and older adults, males and females, patients with and without chronic conditions, and those who survived and died due to COVID-19.

Most of the top ten symptom rules consisted of mild symptoms such as breathing problems, sputum production, weakness, conjunctivitis, rhinorrhea, anorexia, and diarrhoea. Except for the last two symptoms, the consequent was cough. It has been argued that cough alone is not a reliable symptom in COVID-19 patients [86]. However, our study demonstrated cough as a top consequent for patients presenting with fever and heart failure, indicating that cough should not be neglected as a mild symptom. Severe symptoms equally appeared in the top ten rules: the most important rule was cardiac arrhythmia and renal disease; patients were more likely to develop septic shock if these two symptoms occurred. The other rule was pneumonia and renal disease, for which respiratory distress syndrome was an associated consequent. Rules including severe symptoms were much stronger than those including mild ones, and these symptoms were linked to chronic conditions and the survival of patients, indicating the critical importance of symptom identification and management of COVID-19 cases. As per our understanding, this is the first study conducted to date to define symptom rules in COVID-19 patients using machine learning (ARM) techniques, which limits us from comparing our findings with those from other studies. However, some exploratory and review studies have reported similar symptoms among COVID-19 patients, as detailed in our study [87–89].

Table 3
Top 10 significant symptom rules dis-aggregated by age (N = 1560).

<20 years					
Rules	Antecedents	Consequents	Support	Confidence	Lift
Rule 1	{Conjuctivites, Rhinorhea}	{<20 years}	0.001	1.0	10.3
Rule 2	{Dry mouth, Headache}	{<20 years}	0.001	1.0	10.3
Rule 3	{Headache, Sputum}	{<20 years}	0.001	1.0	10.3
Rule 4	{Cough, Rhinorhea, Weakness}	{<20 years}	0.001	1.0	10.3
Rule 5	{Malaise/body soreness, Sore throat, Weakness}	{<20 years}	0.001	1.0	10.3
Rule 6	{Cough, Sore throat, Weakness}	{<20 years}	0.001	1.0	10.3
Rule 7	{Fever, Headache, Rhinorhea, Weakness}	{<20 years}	0.001	1.0	10.3
Rule 8	{Cough, Headache, Rhinorhea, Sore throat}	{<20 years}	0.001	1.0	10.3
Rule 9	{Fever, Headache, Rhinorhea, Sore throat}	{<20 years}	0.001	1.0	10.3
Rule 10	{Fever, Headache, Sore throat, Weakness}	{<20 years}	0.001	1.0	10.3
20–45 years					
Rule 1	{Dry mouth, Sore throat}	{20–45 years}	0.004	1.0	3.3
Rule 2	{Diarrhoea, Fever, Rhinorhea}	{20–45 years}	0.003	1.0	3.3
Rule 3	{Dry mouth, Fever}	{20–45 years}	0.003	1.0	3.3
Rule 4	{Diarrhoea, Sore throat}	{20–45 years}	0.003	1.0	3.3
Rule 5	{Fever, Pneumonia, Sore throat}	{20–45 years}	0.003	1.0	3.3
Rule 6	{Cough, Malaise/body soreness, Sputum}	{20–45 years}	0.002	1.0	3.3
Rule 7	{Hypertension, Renal disease }	{20–45 years}	0.001	1.0	3.3
Rule 8	{Nausea, Non respiratory symptoms}	{20–45 years}	0.001	1.0	3.3
Rule 9	{Dry mouth, Weakness}	{20–45 years}	0.001	1.0	3.3
Rule 10	{Diarrhoea, Headache}	{20–45 years}	0.001	1.0	3.3
45–65 years					
Rule 1	{Nausea, Weakness}	{45–65 years}	0.003	1.0	5.1
Rule 2	{Cough, Heart failure}	{45–65 years}	0.001	1.0	5.1
Rule 3	{Fever, Heart failure}	{45–65 years}	0.001	1.0	5.1
Rule 4	{Cough, Myocardial infraction}	{45–65 years}	0.001	1.0	5.1
Rule 5	{Fever, Myocardial infraction}	{45–65 years}	0.001	1.0	5.1
Rule 6	{Dry mouth, Pneumonia}	{45–65 years}	0.001	1.0	5.1
Rule 7	{Cough, Sore throat, Sputum}	{45–65 years}	0.001	1.0	5.1
Rule 8	{Breathing problem, Cough, Weakness}	{45–65 years}	0.001	1.0	5.1
>65 years					
Rule 1	{Anorexia, Fever}	{>65 years}	0.004	1.0	2.5
Rule 2	{Diarrhoea, Fever, Nausea}	{>65 years}	0.004	1.0	2.5
Rule 3	{Cardiac arrythmia, Pneumonia}	{>65 years}	0.003	1.0	2.5
Rule 4	{Anorexia, Cough}	{>65 years}	0.003	1.0	2.5
		{>65 years}	0.003	1.0	2.5

(*continued on next page*)

Table 3 (*continued*)

<20 years					
Rules	Antecedents	Consequents	Support	Confidence	Lift
Rule 5	{Breathing problem, Rhinorhea}				
Rule 6	{Breathing problem, Sore throat}	{>65 years}	0.003	1.0	2.5
Rule 7	{Anorexia, Diarrhoea}	{>65 years}	0.001	1.0	2.5
Rule 8	{Anorexia, Breathing problem}	{>65 years}	0.001	1.0	2.5
Rule 9	{Anorexia, Malaise/ body soreness}	{>65 years}	0.001	1.0	2.5
Rule 10	{Breathing problem, Nausea}	{>65 years}	0.001	1.0	2.5

Table 4
Top 10 significant symptom rules dis-aggregated by sex.

Males					
Rules	Antecedents	Consequents	Support	Confidence	Lift
Rule 1	{Malaise/body soreness, Weakness}	{Male}	0.008	0.9	1.6
Rule 2	{Cough, Diarrhoea}	{Male}	0.005	1.0	1.8
Rule 3	{Fever, Malaise/body soreness, Pneumonia}	{Male}	0.005	1.0	1.8
Rule 4	{Cough, Fever, Headache, Malaise/ body soreness}	{Male}	0.004	1.0	1.8
Rule 5	{Anorexia, Fever}	{Male}	0.004	1.0	1.8
Rule 6	{Breathing problem, Malaise/body soreness}	{Male}	0.004	1.0	1.8
Rule 7	{Fever, Headache, Malaise/body soreness, Sore throat}	{Male}	0.004	1.0	1.8
Rule 8	{Heart failure, Pneumonia}	{Male}	0.003	1.0	1.8
Rule 9	{Headache, Malaise/ body soreness, Weakness}	{Male}	0.003	1.0	1.8
Rule 10	{Anorexia, Cough}	{Male}	0.003	1.0	1.8
Females					
Rule 1	{Cough, Rhinorhea, Sore throat}	{Female}	0.006	0.9	2.1
Rule 2	{Pneumonia, Rhinorhea}	{Female}	0.004	1.0	2.4
Rule 3	{Fever, Sore throat, Weakness}	{Female}	0.004	1.0	2.4
Rule 4	{Cough, Fever, Rhinorhea, Sore throat}	{Female}	0.004	1.0	2.4
Rule 5	{Fever, Pneumonia, Weakness}	{Female}	0.003	1.0	2.4
Rule 6	{Cough, Dry mouth}	{Female}	0.003	1.0	2.4
Rule 7	{Diarrhoea, Sore throat}	{Female}	0.003	1.0	2.4
Rule 8	{Breathing problem, Rhinorhea}	{Female}	0.003	1.0	2.4
Rule 9	{Fever, Rhinorhea, Weakness}	{Female}	0.003	1.0	2.4
Rule 10	{Cough, Malaise/body soreness, Sputum}	{Female}	0.002	1.0	2.4

Patient age significantly determine the clinical feature and prognosis of the disease. In our study, the age-wise distribution of symptom patterns showed similar rules for patients below the age of 45, with the exception of hypertension and renal disease in the 20–45 year age group. It is difficult to differentiate why hypertension and renal disease make up an essential rule in these age groups; however, universally above 90% of the patients with renal disease have hypertension [90]. The other

Table 5
Top 10 significant symptom rules dis-aggregated by presence of chronic condition (N = 1560).

With chronic conditions					
Rules	Antecedents	Consequents	Support	Confidence	Lift
Rule 1	{Cardiac arrythmia, Septic shock}	{ With chronic disease }	0.002	1.0	10.6
Rule 2	{Respiratory failure, Septic shock}	{ With chronic disease }	0.002	1.0	10.6
Rule 3	{Hypertension, Renal disease}	{ With chronic disease }	0.001	1.0	10.6
Rule 4	{Cardiac arrythmia, Renal disease }	{ With chronic disease }	0.001	1.0	10.6
Rule 5	{Cough, Heart failure}	{ With chronic disease }	0.001	1.0	10.6
Rule 6	{Fever, Heart failure}	{ With chronic disease }	0.001	1.0	10.6
Rule 7	{Pneumonia, Renal disease }	{ With chronic disease }	0.001	1.0	10.6
Rule 8	{Cough, Myocardial infraction}	{ With chronic disease }	0.001	1.0	10.6
Rule 9	{Fever, Myocardial infraction}	{ With chronic disease }	0.001	1.0	10.6
Rule 10	{Headache, Sputum}	{ With chronic disease }	0.001	1.0	10.6
Without chronic conditions					
Rule 1	{Headache, Malaise/body soreness}	{Without chronic disease}	0.021	1.0	1.1
Rule 2	{Sore throat, Weakness}	{Without chronic disease}	0.009	1.0	1.1
Rule 3	{Headache, Rhinorhea}	{Without chronic disease}	0.006	1.0	1.1
Rule 4	{Malaise/body soreness, Pneumonia}	{Without chronic disease}	0.006	1.0	1.1
Rule 5	{Fever, Headache, Sore throat}	{Without chronic disease}	0.006	1.0	1.1
Rule 6	{Cough, Headache, Sore throat}	{Without chronic disease}	0.006	1.0	1.1
Rule 7	{Diarrhoea, Nausea}	{Without chronic disease}	0.005	1.0	1.1
Rule 8	{Cough, Diarrhoea}	{Without chronic disease}	0.005	1.0	1.1
Rule 9	{Diarrhoea, Rhinorhea}	{Without chronic disease}	0.004	1.0	1.1
Rule 10	{Rhinorhea, Weakness}	{Without chronic disease}	0.004	1.0	1.1

common symptoms in the rules were fever, cough, dry mouth, headache, sore throat, body soreness, sputum production, and rhinorrhea, which are consistent with the symptoms reported by Liu et al. in young and middle-aged hospitalized COVID-19 patients [91]. In patients between the ages of 45–65 years, symptom rules mostly comprised of cardiac symptoms, such as heart failure,and myocardial infarction, accompanied by fever, and cough. This could be attributed to the clinical co-morbidities in these subgroups of patients compared to younger patients. In contrast, in patients above 65 years of age, the symptom patterns were more often breathing problems followed by pneumonia and other mild symptoms (cough, fever, anorexia, diarrhoea, and nausea). Breathing difficulties and pneumonia are frequently reported clinical presentation with longer disease courses in COVID-19 patients over 60 years older [91–93]. Upon comparison with symptoms presented in the literature, overall, our study showed similar findings; most of the younger adults have ear, nose, and throat-related symptoms, while older

Table 6
Top 10 significant symptom rules dis-aggregated by patients status (survived vs. died) (N = 1560).

Survived					
Rules	Antecedents	Consequents	Support	Confidence	Lift
Rule 1	{Cough, Fever}	{Survived}	0.267	0.9	1.0
Rule 2	{Fever, Malaise/body soreness}	{Survived}	0.082	1.0	1.1
Rule 3	{Cough, Malaise/body soreness}	{Survived}	0.054	1.0	1.1
Rule 4	{Fever, Sore throat}	{Survived}	0.052	1.0	1.1
Rule 5	{Fever, Weakness}	{Survived}	0.046	0.9	1.0
Rule 6	{Breathing problem, Fever}	{Survived}	0.045	0.9	1.0
Rule 7	{Fever, Rhinorhea}	{Survived}	0.037	1.0	1.1
Rule 8	{Cough, Sore throat}	{Survived}	0.034	1.0	1.0
Rule 9	{Fever, Headache}	{Survived}	0.031	1.0	1.0
Rule 10	{Cough, Rhinorhea}	{Survived}	0.029	1.0	1.1
Died					
Rule 1	{Cardiac arrythmia, Septic shock}	{Died}	0.002	1.0	12.5
Rule 2	{Cardiac arrythmia, Respiratory distress syndrome}	{Died}	0.002	1.0	12.5
Rule 3	{Myocardial infraction, Respiratory failure}	{Died}	0.002	1.0	12.5
Rule 4	{Cardiac arrythmia, renal disease}	{Died}	0.001	1.0	12.5
Rule 5	{Pneumonia, renal disease}	{Died}	0.001	1.0	12.5
Rule 6	{Myocardial infraction, Respiratory distress syndrome}	{Died}	0.001	1.0	12.5
Rule 7	{Breathing problem, Respiratory distress syndrome}	{Died}	0.001	1.0	12.5

adults have breathing difficulties, anorexia, diarrhoea, fever, and fatigue [93].

Sex-wise distribution of COVID-19 showed that males were more susceptible to infections than females [94,95]. On a similar note, our study showed different symptom rules for males and females, which contrasts with results from the study by Liu et al. [96]; which exhibited no difference in symptomatology between the sexes. In our research, most of the rules discovered for males included malaise/body soreness, cough, anorexia, headache, and pneumonia; those for females included sore throat and rhinorrhea. Fever was equally presented in both sexes, whereas heart failure was reported only in males. This is consistent with the general notion that men have increased incidence of cardiovascular disease, including viral myocarditis, compared to females, with the exception of hypertension [97,98].

Chronic disease is a complex phenomenon and an independent risk factor of increased severity and death in critical COVID-19 patients [99]. Preexisting chronic conditions strongly correlates with the severity of disease and admission to intensive care units [96]. In our study, chronic hypertension and diabetes were relatively higher, as reported in earlier studies [99,100]. If a patient had had diabetes, coronary bypass surgery, or prerenal azotemia, the occurrence of hypertension was more likely, which implies that patients with chronic hypertension are more susceptible to severe COVID-19 or the risk of fatal disease outcomes. A recent study showed strong association of hypertension with mortality of COVID-19 patients [101]. In hospitals, patients with hypertension as an underlying health condition were 1.6–3.1 times more likely to die from COVID-19 [99,102].

Symptoms exhibited by patients with chronic conditions are critical to case management. In our study, patients with chronic diseases showed more severe symptom rules compared to patients without chronic conditions. Importantly, severe symptom rules such as [cardiac arrhythmia, septic shock, respiratory failure] and [hypertension, heart failure, and renal disease] suggest the requirement of exceptional management and treatment, as underlying disease greatly affect patients survival [99]. Furthermore, rules also included cough and fever in conjunction with heart failure and myocardial infarction, providing useful insight into the role of cough and fever in COVID-19 patients with chronic diseases that needs careful consideration.

Similarly, our study identified a difference in the symptom rules between patients who survived or died of COVID-19. Among those who died, the rules included severe symptoms (cardiac arrhythmia, septic shock, respiratory distress syndrome, myocardial infarction, renal disease, and pneumonia). These symptoms are consistent with those in a retrospective study conducted in Wuhan, China, which reported similar severe symptoms among patients who did not survive COVID-19 infections [99].

ARM is a structured method of discovering frequent patterns in a data set and forming noticeable rules among regular patterns. In medicine, applications of ARM can vary. For instance, it can be used to discover frequent disease patterns in specific geographic areas [84], understand trends in diagnose and diagnostic test requirements in emergency departments [103], diagnose hyperlipidemia [104], extract patterns of heart disease and the prediction of heart attacks [105] and select appropriate medicine for a disease based on a patient's description [106]. ARM specific to COVID-19 has not been previously applied, even though studies that use machine learning algorithms to investigate radiological findings are available in the literature [107,108].

5.1. Limitation

This study is based on retrospective data available online with limited patient level variables, restricting a robust analysis. Similarly, some of the composite symptoms are assigned within a disease, making it impossible to ascertain the individual symptoms reported within them. Furthermore, the online nature of these data did not allow us to explain how these data were collected and made available; hence, we cannot rule out information collection bias in the study. Furthermore, a large chunk of data was missing in the data set, which may have caused misrepresentation of the patient's population. Therefore, we recommend applying ARM to the primary data sets that are available from hospitals or primary care settings to produce a more reliable and accurate result.

There is a caveat for the confidence metric in the ARM technique when a negative correlation exists between the two sets, for instance $\neg X \Rightarrow Y$. In most of the cases, when examining negatively correlated rule, lower support and confidence are preferred. The positive symptoms are often obvious; however, negative symptoms are subtler and more difficult to recognize and diagnose [109]. Therefore, it is very pressing that researchers. However, in this study, we have not looked into negatively correlated rules. Furthermore, rules discovered by algorithms require clinical validation and verification. This is an important limitation of our study.

6. Conclusion and future work

The most frequent symptoms in our study included fever, cough, pneumonia, sore throat, and breathing problems. Additionally, respiratory distress syndrome, nausea, septic shock, and respiratory failure represented one to five percent of symptoms among COVID-19 patients. ARM techniques identified significantly different symptom rules for COVID-19 between younger and older patients, male and female patients, patients with and without chronic conditions, and those who

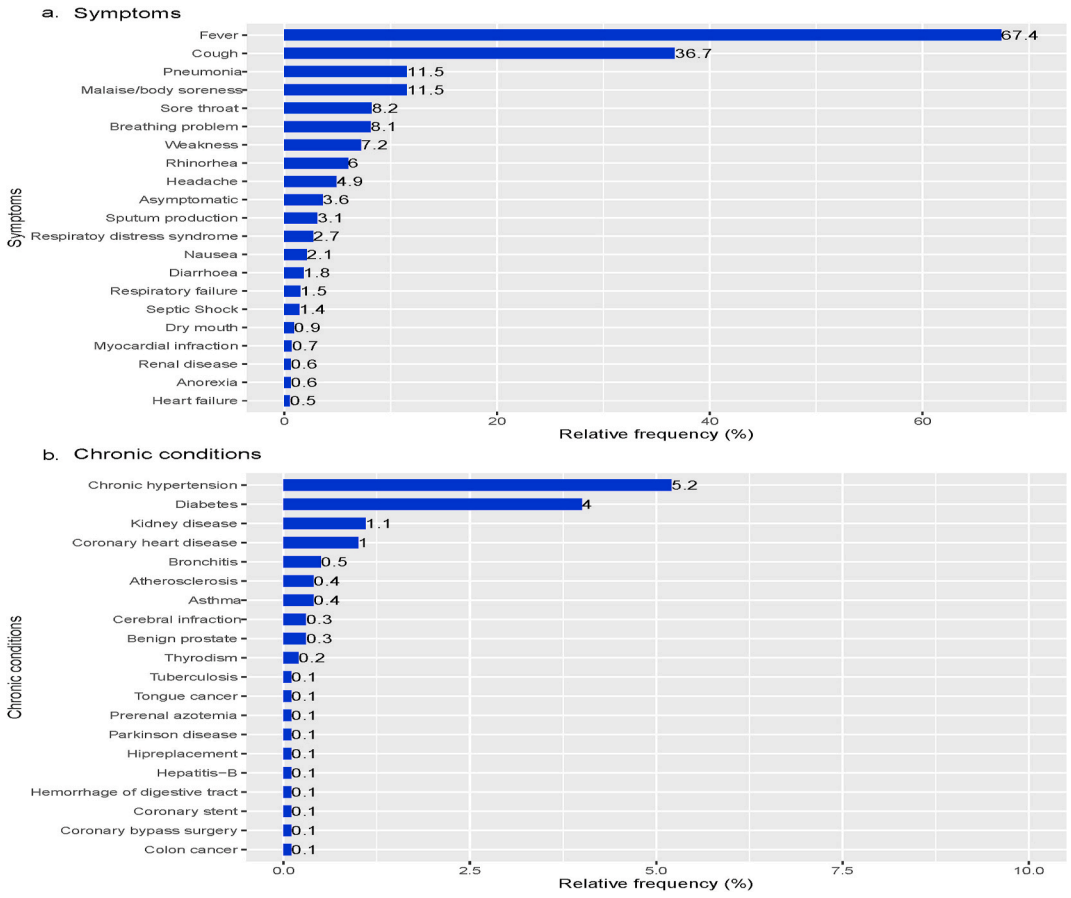
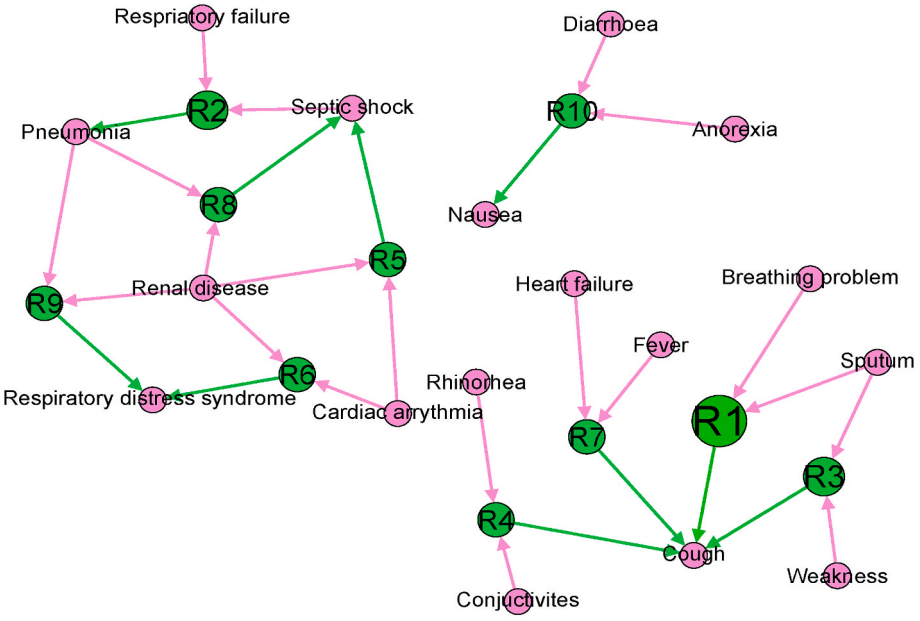


Fig. 3. Relative frequency of symptom and chronic disease in COVID-19 patients.



Note:
Green circles (nodes) represent top 10 rules by support, R1= rule1 and so on. Higher the support value, the larger the green nodes.
Pink circles (nodes) represent symptom.
Symptom with pink arrows towards rules (e.g,sputum -> R1) are antecedents.
Symptom with green arrow outwards rules (e.g, cough <- R1) are consequents.

Fig. 4. Graphical presentation of symptom rules.

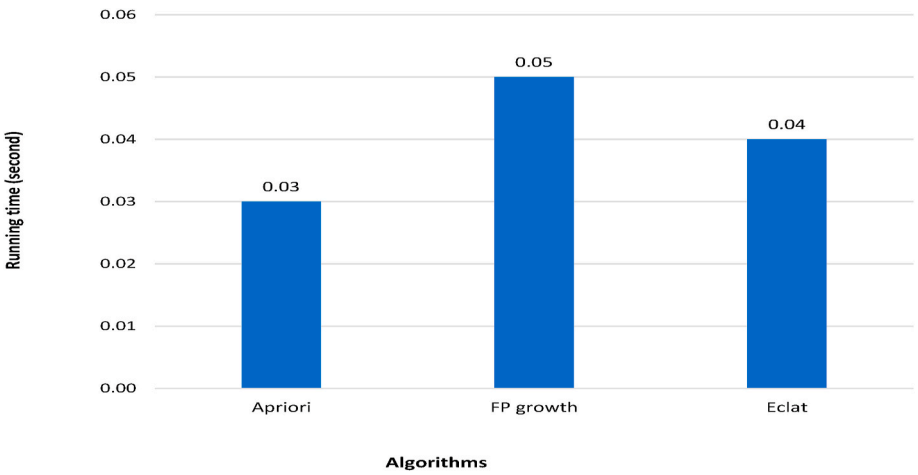


Fig. 5. Run-time comparison for different rule mining algorithms.

survived COVID-19 and those who died. The top 10 symptom rules showed that if a patient had breathing problems and sputum production, there was high confidence that they would present a cough. Likewise, septic shock and respiratory distress syndrome were consequents for COVID-19 patients presenting with cardiac arrhythmia, renal disease, and pneumonia. Patients with chronic conditions and patients who died of COVID-19 showed more severe symptom rules, such as cardiac arrhythmia, hypertension, respiratory failure, septic shock, heart failure, myocardial infarction, and pneumonia, accompanied by fever and cough.

The most important future work that can stem from this research is the application of the same idea in dynamic data sets. COVID-19 web data statistics are frequently updated. In the current setting, our approach to extract the symptom patterns relied on a static data sets, hence, they are not applicable in a dynamic setting. Thus, dynamic algorithms are needed to extract the patterns from the database. Though some work in dynamic rule mining has been done [56,57], we would like to extend the same approach applying these algorithms in COVID-19 data sets. However, the main challenge surrounding COVID-19 web data is that they are noisy. Hence, it is worth investigating the quality of the results produced by these algorithms in future studies.

CRediT authorship contribution statement

Meera Tandan: Concept, data curation, data management and analysis, drafting manuscript and finalization. **Yogesh Acharya:** Clinical discussion, revision and finalizing the manuscript. **Suresh Pokharel:** support in application of association rule mining, finalizing the manuscript. **Mohan Timilsina:** Concept, data cleaning, support in application of association rule mining, finalizing the manuscript.

Funding

No funding available for this research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104249.

References

[1] WHO, Coronavirus Disease (COVID-19) Pandemic, May 2020. Accessed: 2020-07-26.
[2] Worldometer, COVID-19 Coronavirus Pandemic Reported Cases and Deaths by Country, Territory, or Conveyance, 2020. Accessed: 2020-07-26.
[3] RECOVERY Collaborative Group,Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, Linsell L, Staplin N, Brightling C, Ustianowski A, Elmahi E, Prudon B, Green C, Felton T, Chadwick D, Rege K, Fegan C, Chappell LC, Faust SN, Jaki T, Jeffery K, Montgomery A, Rowan K, Juszczak E, Baillie JK, Haynes R, Landray MJ., Dexamethasone in hospitalized patients with covid-19-preliminary report, N. Engl. J. Med. Jul 17:NEJMoa2021436 (2020).
[4] WHO, Solidarity" Clinical Trial for COVID-19 Treatment, 2020. Accessed: 2020-07-27.
[5] Jonathan Grein, Norio Ohmagari, Daniel Shin, Diaz George, Erika Asperges, Antonella Castagna, Torsten Feldt, Gary Green, Margaret L. Green, FranÃ§ois-Xavier Lescure, et al., Compassionate use of remdesivir for patients with severe covid-19, N. Engl. J. Med. 382 (24) (2020) 2327–2336.
[6] Bauchner Howard, Phil B. Fontanarosa, Randomized clinical trials and covid-19: managing expectations, JAMA 323 (22) (2020) 2262–2263.
[7] Jeff Craven, COVID-19 Therapeutic Tracker, 2020. Accessed: 2020-12-20.
[8] Michael A. Matthay, B Taylor Thompson, Dexamethasone in hospitalised patients with covid-19: addressing uncertainties, Lancet Respir. Med. 8 (12) (2020) 1170–1172.
[9] Gyanshankar P. Mishra, Jasmin Mulani, Corticosteroids for covid-19: the search for an optimum duration of therapy, Lancet Respir. Med. 9 (1,E8) (2020).
[10] Sui-Lee Wee Carol Zimmer, Jonathan Corum, Coronavirus Vaccine Tracker, Updated on December 18, 2020, 2020. Accessed: 2020-12-20.
[11] Bloomberg. More, Than 1.6 Million People Have Been Vaccinated- Covid-19 Tracker, Updated December 19, 2020, 2020. Accessed: 2020-12-20.
[12] Kristine A. Moore, Marc Lipsitch, John M. Barry, Michael T. Osterholm, Part 1: the Future of the Covid-19 Pandemic: Lessons Learned from Pandemic Influenza. COVID-19: the CIDRAP Viewpoint, Center for Infectious Disease Research and Policy, 2020.
[13] Gangqiang Guo, Lele Ye, Kan Pan, Yu Chen, Xing Dong, Kejing Yan, Zhiyuan Chen, Ning Ding, Wenshu Li, Hong Huang, et al., New insights of emerging sars-cov-2: epidemiology, etiology, clinical features, clinical treatment, and prevention, Front. Cell Dev. Biol. 8 (2020) 410.
[14] Ã„lgÃ¥ Andreas, Oskar Eriksson, Martin Nordberg, Analysis of scientific publications during the early phase of the covid-19 pandemic: topic modeling study, J. Med. Internet Res. 22 (11) (2020), e21559.
[15] A. Sayed, Y. Acharya, K.C.V. Long, L. Lynam, M. Tandan, Estimation of clinical comorbidities in covid-19 patients: a systematic re-view and meta-analysis, Ann. Microbiol. Res. 4 (1) (2020) 105–111.
[16] Jean-Marc Adamo, Data Mining for Association Rules and Sequential Patterns: Sequential and Parallel Algorithms, Springer Science & Business Media, 2001.
[17] Shreshth Tuli, Shikhar Tuli, Gurleen Wander, Praneet Wander, Sukhpal Singh Gill, Schahram Dustdar, Rizos Sakellariou, Omer Rana, Next generation technologies for smart healthcare: challenges, vision, model, trends and future directions, Internet Technol. Lett. 3 (2) (2020) e145.
[18] Adrien Depeursinge, Ann N Leung Anne S Chin, Donato Terrone, Michael Bristow, Glenn Rosen, Daniel L. Rubin, Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution ct, Invest. Radiol. 50 (4) (2015) 261.
[19] Maxwell W. Libbrecht, William Stafford Noble, Machine learning applications in genetics and genomics, Nat. Rev. Genet. 16 (6) (2015) 321–332.
[20] Shreshth Tuli, Shikhar Tuli, Rakesh Tuli, Sukhpal Singh Gill, Predicting the Growth and Trend of Covid-19 Pandemic Using Machine Learning and Cloud Computing, Internet of Things, 2020, p. 100222.

[21] Mohammad Jamshidi, Lalbakhsh Ali, Jakub Talla, ZdenÄ)k Peroutka, Farimah Hadjilooei, Pedram Lalbakhsh, Morteza Jamshidi, Luigi La Spada, Mirhamed Mirmozafari, Mojgan Dehghani, et al., Artificial intelligence and covid-19: deep learning approaches for diagnosis and treatment, IEEE Access 8 (2020) 109581–109595.

[22] Talha Burak Alakus, Ibrahim Turkoglu, Comparison of deep learning approaches to predict covid-19 infection, Chaos, Solit. Fractals 140 (2020) 110120.

[23] Adi L. Tarca, Vincent J. Carey, Xue-wen Chen, Roberto Romero, Sorin DrÄƒghici, Machine learning and its applications to biology, PLoS Comput. Biol. 3 (6) (2007) e116.

[24] Chunming Xu, Scott A. Jackson, Machine Learning and Complex Biological Data, 2019.

[25] Wolfgang Huber, Vincent J. Carey, Robert Gentleman, Simon Anders, Marc Carlson, Benilton S. Carvalho, Hector Corrada Bravo, Sean Davis, Laurent Gatto, Thomas Girke, et al., Orchestrating high-throughput genomic analysis with bioconductor, Nat. Methods 12 (2) (2015) 115–121.

[26] Martin Kircher, Udo Stenzel, Janet Kelso, Improved base calling for the illumina genome analyzer using machine learning strategies, Genome Biol. 10 (8) (2009) R83.

[27] Geoffrey R. Oxnard, Cloud P. Paweletz, Lynette M. Sholl, Genomic analysis of plasma cell-free dna in patients with cancer, Jama Oncol. 3 (6) (2017) 740–741.

[28] Rosanna Upstill-Goddard, Diana Eccles, Joerg Fliege, Andrew Collins, Machine learning approaches for the discovery of gene-gene interactions in disease data, Briefings Bioinf. 14 (2) (2013) 251–260.

[29] Xiangxiang Zeng, Yuanlu Liao, Yuansheng Liu, Quan Zou, Prediction and validation of disease genes using hetesim scores, IEEE ACM Trans. Comput. Biol. Bioinf 14 (3) (2016) 687–695.

[30] Timilsina Mohan, Haixuan Yang, Ratnesh Sahay, Dietrich Rebholz-Schuhmann, Predicting links between tumor samples and genes using 2-layered graph based diffusion approach, BMC Bioinf. 20 (1) (2019) 462.

[31] Daniela Nitsch, Joana P. GonÃ§alves, Fabian Ojeda, Bart De Moor, Yves Moreau, Candidate gene prioritization by network analysis of differential expression using machine learning approaches, BMC Bioinf. 11 (1) (2010) 460.

[32] Pokharel Suresh, Zhenkun Shi, Zuccon Guido, Li Yu, Discriminative features generation for mortality prediction in icu, in: International Conference on Advanced Data Mining and Applications, Springer, 2020.

[33] Chandra Prasetyo Utomo, Hanna Kurniawati, Xue Li, Suresh Pokharel, Personalised medicine in critical care using bayesian reinforcement learning, in: International Conference on Advanced Data Mining and Applications, Springer, 2019, pp. 648–657.

[34] Qiuling Suo, Fenglong Ma, Ye Yuan, Mengdi Huai, Weida Zhong, Jing Gao, Aidong Zhang, Deep patient similarity learning for personalized healthcare, IEEE Trans. NanoBioscience 17 (3) (2018) 219–227.

[35] Yu-Chen Lo, Stefano E. Rensi, Torng Wen, Russ B. Altman, Machine learning in chemoinformatics and drug discovery, Drug Discov. Today 23 (8) (2018) 1538–1546.

[36] Jessica Vamathevan, Dominic Clark, Czodrowski Paul, Ian Dunham, Edgardo Ferran, George Lee, Bin Li, Anant Madabhushi, Parantu Shah, Michaela Spitzer, et al., Applications of machine learning in drug discovery and development, Nat. Rev. Drug Discov. 18 (6) (2019) 463–477.

[37] Antonio Lavecchia, Machine-learning approaches in drug discovery: methods and applications, Drug Discov. Today 20 (3) (2015) 318–331.

[38] Natalie Stephenson, Emily Shane, Jessica Chase, Jason Rowland, David Ries, Nicola Justice, Jie Zhang, Leong Chan, Renzhi Cao, Survey of machine learning techniques in drug discovery, Curr. Drug Metabol. 20 (3) (2019) 185–193.

[39] Brant W. Chee, Richard Berlin, Bruce Schatz, Predicting adverse drug events from personal health messages, in: AMIA Annual Symposium Proceedings, American Medical Informatics Association, 2011, p. 217, 2011.

[40] Aurel Cami, Alana Arnold, Manzi Shannon, Ben Reis, Predicting adverse drug events using pharmacological network models, Sci. Transl. Med. 3 (114) (2011), 114ra127-114ra127.

[41] Sanjoy Dey, Heng Luo, Achille Fokoue, Jianying Hu, Ping Zhang, Predicting adverse drug reactions through interpretable deep learning framework, BMC Bioinf. 19 (21) (2018) 1–13.

[42] Timilsina Mohan, Meera Tandan, Mathieu d'Aquin, Haixuan Yang, Discovering links between side effects and drugs using a diffusion based method, Sci. Rep. 9 (1) (2019) 1–10.

[43] Pokharel Suresh, Zuccon Guido, Li Xue, Chandra Prasetyo Utomo, Li Yu, Temporal Tree Representation for Similarity Computation between Medical Patients. Artificial Intelligence in Medicine, 2020, p. 101900.

[44] Pokharel Suresh, Li Xue, Xin Zhao, Anoj Adhikari, Li Yu, Similarity Computing on Electronic Health Records, PACIS, 2018, p. 198.

[45] Andreas Holzinger, Georg Langs, Helmut Denk, Kurt Zatloukal, Heimo MÃ¼ller, Causability and explainability of artificial intelligence in medicine, Wiley Interdiscipl. Rev.: Data Min. Knowl. Discov. 9 (4) (2019), e1312.

[46] Meera Tandan, Timilsina Mohan, Cormican Martin, Akke Vellinga, Role of patient descriptors in predicting antimicrobial resistance in urinary tract infections using a decision tree approach: a retrospective cohort study, Int. J. Med. Inf. 127 (2019) 127–133.

[47] Erico Tjoa, Cuntai Guan, A Survey on Explainable Artificial Intelligence (Xai): towards Medical Xai, 2019 arXiv preprint arXiv:1907.07374.

[48] Jean-Baptiste Lamy, Boomadevi Sekar, Gilles Guezennec, Jacques Bouaud, Brigitte SÃ©roussi, Explainable artificial intelligence for breast cancer: a visual case-based reasoning approach, Artif. Intell. Med. 94 (2019) 42–53.

[49] Rakesh Agrawal, Ramakrishnan Srikant, Mining sequential patterns, in: Proceedings of the Eleventh International Conference on Data Engineering, IEEE, 1995, pp. 3–14.

[50] Rakesh Agrawal, Tomasz ImieliÅ„ski, Arun Swami, Mining association rules between sets of items in large databases, in: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 1993, pp. 207–216.

[51] Rakesh Agarwal, Ramakrishnan Srikant, et al., Fast algorithms for mining association rules, in: Proc. Of the 20th VLDB Conference, 1994, pp. 487–499.

[52] Jiawei Han, Jian Pei, Mining frequent patterns by pattern-growth: methodology and implications, ACM SIGKDD Explor. Newslett. 2 (2) (2000) 14–20.

[53] M. Zaki, S. Parthasarathy.M. ogihara, w. li, New algorithms for fast discovery of association rules, Proc. of the 3st KDD (1997).

[54] Pelin YILDIRIM TAŞER, Kökten Ulaş BİRANT, Derya BİRANT, Multitask-based association rule mining, Turk. J. Electr. Eng. Comput. Sci. 28 (2) (2020) 933–955.

[55] Gabriela Czibula, Istvan Gergely Czibula, Diana-Lucia Miholca, Liana Maria Crivei, A novel concurrent relational association rule mining approach, Expert Syst. Appl. 125 (2019) 142–156.

[56] Dang Nguyen, Wei Luo, Dinh Phung, Svetha Venkatesh, Ltarm: a novel temporal association rule mining method to understand toxicities in a routine cancer treatment, Knowl. Base Syst. 161 (2018) 313–328.

[57] Iyad Aqra, Norjihan Abdul Ghani, Carsten Maple, JosÃ© Machado, Nader Sohrabi Safa, Incremental algorithm for association rule mining under dynamic threshold, Appl. Sci. 9 (24) (2019) 5398.

[58] Xiangyu Liu, Xinzheng Niu, Philippe Fournier-Viger, Fast top-k association rule mining using rule generation property pruning, Appl. Intell. (2020) 1–17.

[59] Jesmin Nahar, Tasadduq Imam, Kevin S. Tickle, Yi-Ping Phoebe Chen, Association rule mining to detect factors which contribute to heart disease in males and females, Expert Syst. Appl. 40 (4) (2013) 1086–1093.

[60] Sangita Khare, Deepa Gupta, Association rule analysis in cardiovascular disease, in: 2016 Second International Conference on Cognitive Computing and Information Processing (CCIP), IEEE, 2016, pp. 1–6.

[61] KM Mehedi Hasan Sonet, Md Mustafizur Rahman, Pritom Mazumder, Abid Reza, Rashedur M. Rahman, Analyzing patterns of numerously occurring heart diseases using association rule mining, in: 2017 Twelfth International Conference on Digital Information Management (ICDIM), IEEE, 2017, pp. 38–45.

[62] Vladimir Ivančević, Ivan Tušek, Jasmina Tušek, Marko Knežević, Salaheddin Elheshk, Ivan Luković, Using association rule mining to identify risk factors for early childhood caries, Comput. Methods Progr. Biomed. 122 (2) (2015) 175–181.

[63] Anindita Borah, Bhabesh Nath, Identifying risk factors for adverse diseases using dynamic rare association rule mining, Expert Syst. Appl. 113 (2018) 233–263.

[64] Sugam Sharma, Concept of association rule of data mining assists mitigating the increasing obesity, in: Healthcare Policy and Reform: Concepts, Methodologies, Tools, and Applications, IGI Global, 2019, pp. 518–536.

[65] Yoshihiro Noguchi, Anri Ueno, Manami Otsubo, Hayato Katsuno, Ikuto Sugita, Yuta Kanematsu, Aki Yoshida, Hiroki Esaki, Tachi Tomoya, Hitomi Teramachi, A new search method using association rule mining for drug-drug interaction based on spontaneous report system, Front. Pharmacol. 9 (2018) 197.

[66] Ruichu Cai, Mei Liu, Yong Hu, Brittany L. Melton, Michael E. Matheny, Hua Xu, Lian Duan, Lemuel R. Waitman, Identification of adverse drug-drug interactions through causal association rule discovery from spontaneous adverse event reports, Artif. Intell. Med. 76 (2017) 7–15.

[67] S. Ramasamy, K. Nirmala, Disease prediction in data mining using association rule mining and keyword based clustering algorithms, Int. J. Comput. Appl. 42 (1) (2020) 1–8.

[68] K Hema Prasanna Murari Devakannan Kamalesh, B. Bharathi, R. Dhanalakshmi, R Aroul Canessane, Predicting the risk of diabetes mellitus to subpopulations using association rule mining. Proceedings of the International Conference on Soft Computing Systems, Springer, 2016, pp. 59–65.

[69] Pokharel Suresh, Zuccon Guido, Li Yu, Representing ehrs with temporal tree and sequential pattern mining for similarity computing, in: International Conference on Advanced Data Mining and Applications, Springer, 2020.

[70] Annie Lau, Siew Siew Ong, Ashesh Mahidadia, Achim Hoffmann, Johanna Westbrook, Tatjana Zrimec, Mining patterns of dyspepsia symptoms across time points using constraint association rules, in: Pacific-asia Conference on Knowledge Discovery and Data Mining, Springer, 2003, pp. 124–135.

[71] SriJyothsna Yeleswarapu, Aditya Rao, Thomas Joseph, Vangala Govindakrishnan Saipradeep, Rajgopal Srinivasan, A pipeline to extract drug-adverse event pairs from multiple data sources, BMC Med. Inf. Decis. Making 14 (1) (2014) 13.

[72] Parul Arora, Himanshu Kumar, Bijaya Ketan Panigrahi, Prediction and analysis of covid-19 positive cases using deep learning models: a descriptive case study of India, Chaos, Solit. Fractals 139 (2020) 110017.

[73] Lei Qin, Qiang Sun, Yidan Wang, Ke-Fei Wu, Mingchih Chen, Ben-Chang Shia, Szu-Yuan Wu, Prediction of number of cases of 2019 novel coronavirus (covid-19) using social media search index, Int. J. Environ. Res. Publ. Health 17 (7) (2020) 2365.

[74] Anuradha Tomar, Neeraj Gupta, Prediction for the spread of covid-19 in India and effectiveness of preventive measures, Sci. Total Environ. (2020) 138762.

[75] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, U Rajendra Acharya, Automated detection of covid-19 cases using deep neural networks with x-ray images, Comput. Biol. Med. (2020) 103792.

[76] Ioannis D. Apostolopoulos, Tzani A. Mpesiana, Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks, Phys. Eng. Sci. Med. 1 (2020).

[77] Wolfram Research, Patient Medical Data for Novel Coronavirus COVID-19" from the Wolfram Data Repository, 2020. Accessed: 2020-06-02.
[78] Geoffrey I. Webb, Discovering significant patterns, Mach. Learn. 68 (1) (2007) 1–33.
[79] Roberto J. Bayardo, Rakesh Agrawal, Dimitrios Gunopulos, Constraint-based rule mining in large, dense databases, Data Min. Knowl. Discov. 4 (2–3) (2000) 217–240.
[80] Stefan Mutter, Mark Hall, Eibe Frank, Using classification to evaluate the output of confidence-based association rule mining, in: Australasian Joint Conference on Artificial Intelligence, Springer, 2004, pp. 538–549.
[81] Tyler McCormick, Cynthia Rudin, David Madigan, A Hierarchical Model for Association Rule Mining of Sequential Events: an Approach to Automated Medical Symptom Prediction, 2011.
[82] Laszlo Szathmary, Petko Valtchev, Amedeo Napoli, Generating Rare Association Rules Using the Minimal Rare Itemsets Family, 2010.
[83] K. Pazhanikumar, S. Arumugaperumal, Association rule mining and medical application: a detailed survey, Int. J. Comput. Appl. 80 (17) (2013).
[84] Kamran Shaukat, Sana Zaheer, Iqra Nawaz, Association rule mining: an application perspective, Int. J. Contr. Syst. Instrum. (1) (2015) 29–38, 2015.
[85] Michael C. Grant, Luke Geoghegan, Marc Arbyn, Zakaria Mohammed, Luke McGuinness, Emily L. Clarke, Ryckie Wade, The Prevalence of Symptoms in 24,410 Adults Infected by the Novel Coronavirus (Sars-cov-2; Covid-19): A Systematic Review and Meta-Analysis of 148 Studies from 9 Countries, 2020. Available at: SSRN 3582819.
[86] Melina Michelen, Nicholas Jones, Charitini Stavropoulou, In patients of covid-19, what are the symptoms and clinical features of mild and moderate cases, Centre Evid. Based Med. 16 (2020). https://www.cebm.net/covid-19/in-patients-of-covid-19-what-are-the-symptoms-and-clinical-features-of-mild-andmoderatecase/.
[87] Karl-Heinz Kuck, Arrhythmias and Sudden Cardiac Death in the Covid-19 Pandemic, Herz, 2020, p. 1.
[88] Jun Wu, Jushuang Li, Geli Zhu, Yanxia Zhang, Zhimin Bi, Yean Yu, Bo Huang, Shouzhi Fu, Yiqing Tan, Jianbin Sun, et al., Clinical features of maintenance hemodialysis patients with 2019 novel coronavirus-infected pneumonia in wuhan, China, Clin. J. Am. Soc. Nephrol. 15 (8) (2020) 1139–1145.
[89] Dahai Zhao, Feifei Yao, Lijie Wang, Ling Zheng, Yongjun Gao, Jun Ye, Feng Guo, Hui Zhao, Rongbao Gao, A Comparative Study on the Clinical Features of Covid-19 Pneumonia to Other Pneumonias, Clinical Infectious Diseases, 2020.
[90] Natalia Ridao, José Luño, Soledad García de Vinuesa, Francisco Gómez, Alberto Tejedor, Fernando Valderrábano, Prevalence of hypertension in renal disease, Nephrol. Dial. Transplant. 16 (suppl_1) (2001) 70–73.
[91] Kai Liu, Ying Chen, Ruzheng Lin, Kunyuan Han, Clinical features of covid-19 in elderly patients: a comparison with young and middle-aged patients, J. Infect. 80 (6) (2020) e14–e18.
[92] Yang Liu, Bei Mao, Shuo Liang, Jia-wei Yang, Hai-wen Lu, Yan-hua Chai, Lan Wang, Li Zhang, Qiu-hong Li, Lan Zhao, et al., Association between ages and clinical characteristics and outcomes of coronavirus disease 2019, Eur. Respir. J. 55 (5) (2020).
[93] Jerome R. Lechien, Carlos M. Chiesa-Estomba, Sammy Place, Yves Van Laethem, Pierre Cabaraux, Quentin Mat, Kathy Huet, Plzak Jan, Mihaela Horoi, Stéphane Hans, Maria Rosaria Barillari, Giovanni Cammaroto, Nicolas Fakhry 1, Delphine Martiny, Tareck Ayad, Lionel Jouffe, Claire Hopkins, Sven Saussez, et al., COVID-19 Task Force of YO-IFOS, Clinical and epidemiological characteristics of 1,420 european patients with mild-to-moderate coronavirus disease 2019, J. Intern. Med. 288 (3) (2020) 335–344.
[94] Garima Sharma, Annabelle Santos Volgman, Erin D. Michos, Sex Differences in Mortality from Covid-19 Pandemic: Are Men Vulnerable and Women Protected? *JACC*, Case Reports, 2020.
[95] Yogesh Acharya, Suman Pant, Pradip Gyanwali, Ganesh Dangal, Priyanka Karki, Navindra Raj Bista, Meera Tandan, Gender disaggregation in covid-19 and increased male susceptibility, J. Nepal Health Res. Council 18 (3) (2020) 345–350.
[96] Hong Liu, Shiyan Chen, Min Liu, Hao Nie, Hongyun Lu, Comorbid chronic diseases are strongly correlated with disease severity among covid-19 patients: a systematic review and meta-analysis, Aging Dis. 11 (3) (2020) 668.
[97] Véronique L. Roger, Alan S. Go, Donald M. Lloyd-Jones, Emelia J. Benjamin, Jarett D. Berry, William B. Borden, Dawn M. Bravata, Shifan Dai, Earl S. Ford, Caroline S. Fox, Heather J. Fullerton, Cathleen Gillespie, Susan M. Hailpern, John A. Heit, Virginia J. Howard, Brett M. Kissela, Steven J. Kittner, Lackland, Judith H. Lichtman, Lynda D. Lisabeth, Diane M. Makuc, Gregory M. Marcus, Ariane Marelli, David B. Matchar, Claudia S. Moy, Dariush Mozaffarian, Michael E. Mussolino, Graham Nichol, Nina P. Paynter, Elsayed Z. Soliman, Paul D. Sorlie, Nona Sotoodehnia, Tanya N. Turan, Salim S. Virani, Nathan D. Wong, Daniel Woo, Melanie B. Turner, et al., American Heart Association Statistics Committee and Stroke Statistics Subcommittee, Executive summary: heart disease and stroke statistics-2012 update: a report from the american heart association, Circulation 125 (1) (2012) 188–197.
[98] DeLisa Fairweather, Leslie T. Cooper Jr., Lori A. Blauwet, Sex and gender differences in myocarditis and dilated cardiomyopathy, Curr. Probl. Cardiol. 38 (1) (2013) 7–46.
[99] Fei Zhou, Yu Ting, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, et al., Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, China: a retrospective cohort study, Lancet 395 (10229) (2020) 1054–1062.
[100] Sayed Abida, Yogesh Acharya, Vi Long, Lynam Laura, Meera Tandan, Estimation of clinical comorbidities in covid-19 patients: a systematic review and meta-analysis, Ann. Microbiol. Res. 4 (07 2020).
[101] Iaccarino Guido, Grassi Guido, Claudio Borghi, Claudio Ferri, Massimo Salvetti, Massimo Volpe, Age and multimorbidity predict death among covid-19 patients: results of the sars-ras study of the Italian society of hypertension, Hypertension 76 (2) (2020) 366–372.
[102] Wei-jie Guan, Wen-hua Liang, Yi Zhao, Heng-rui Liang, Zi-sheng Chen, Yi-min Li, Xiao-qing Liu, Ru-chong Chen, Chun-li Tang, Tao Wang, et al. Comorbidity and its impact on 1590 patients with covid-19 in China: a nationwide analysis. Eur. Respir. J., 55(5), 2020.
[103] GÃ¶rkem SarÄ±yer, Ceren Ã-cal TaÅŸar, Highlighting the rules between diagnosis types and laboratory diagnostic tests for patients of an emergency department: use of association rule mining, Health Inf. J. 26 (2) (2019) 1177–1193, 1460458219871135.
[104] Sengul Dogan, Ibrahim Turkoglu, Diagnosing hyperlipidemia using association rules, Math. Comput. Appl. 13 (3) (2008) 193–202.
[105] Shantakumar B. Patil, Y.S. Kumaraswamy, Extraction of significant patterns from heart disease warehouses for heart attack prediction, IJCSNS 9 (2) (2009) 228–235.
[106] M. Harahap, A.M. Husein, S. Aisyah, F.R. Lubis, B.A. Wijaya, Mining association rule based on the diseases population for recommendation of medicine need, in: Journal of Physics: Conference Series, vol. 1007, IOP Publishing, 2018, 012017.
[107] Xueyan Mei, Hao-Chih Lee, Kai-yue Diao, Mingqian Huang, Bin Lin, Chenyu Liu, Zongyu Xie, Yixuan Ma, Philip M. Robson, Michael Chung, et al., Artificial intelligence-enabled rapid diagnosis of patients with covid-19, Nat. Med. 26 (8) (2020) 1224–1228.
[108] Feng Shi, Jun Wang, Jun Shi, Ziyan Wu, Qian Wang, Zhenyu Tang, Kelei He, Yinghuan Shi, Dinggang Shen, Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19, IEEE Rev. Biomed. Eng. 14 (2020) 4–15.
[109] Sajid Mahmood, Muhammad Shahbaz, Aziz Guergachi, Negative and positive association rules mining from text using frequent and infrequent itemsets, Sci. World J. 2014 (2014).