The European Research Journal 2022;8(5):678-684 Original Article

DOI: 10.18621/eurj.1140623 *Infectious Diseases*

Evaluation of the effect values of risk factors by clustering method in patients who died due to COVID-19 disease

Salih Metin[1], Suna Eraybar[2], Hüseyin Aygün[3], İlker Ercan[4], Canan Demir[5]

[1]Department of Family Medicine, Bursa Provincial Health Directorate Public Hospitals Services Presidency, Bursa, Turkey; [2]Department of Emergency, University of Health Sciences Turkey, Bursa City Hospital, Bursa, Turkey; [3]Department of Emergency, University of Health Sciences Turkey, Bursa Yüksek Ihtisas Training and Research Hospital, Bursa, Turkey; [4]Department of Biostatistics, Uludağ University School of Medicine, Bursa, Turkey; [5]Department of Infectious Diseases, University of Health Sciences Turkey, Bursa City Hospital, Bursa, Turkey

ABSTRACT

Objectives: The aim of this study is to determine the factors that may be associated with mortality in patients who died due to COVID-19 and to determine the effect sizes of the factors that make a statistically significant difference.

Methods: The patients who died due to COVID-19 between 01.03.2020 and 01.03.2021 in Bursa province were evaluated retrospectively. In addition to demographic information such as age, gender, nationality, existing chronic diseases of the patients, COVID- PCR test results, length of hospital stay, intensive care unit follow up times, intubation application times were recorded. The effect size of the variables on mortality were evaluated.

Results: Total of 3,510 deaths due to COVID-19 were evaluated. Of these, 2107 (60%) were male and 1403 (40%) were female. Three thousand three hundred and seventy-four (96.12%) patients are 50 years or older. In both sexes, the highest number of deaths were in the age range of 70-79. The most common comorbidities were hypertension (HT) (n = 1,182; 34.16%) and diabetes mellitus (DM) (n = 776; 22.43%). HT and DM had a strong effect value between the groups ($p < 0.001$ and $p < 0.001$, phi effect values: 0.661 and 0.681, respectively). Although there was a statistically significant difference for the age variable, it had an insignificant effect value ($p = 0.008$, $\delta = 0.074$).

Conclusions: Risk factors frequently reported for COVID-19 deaths but there are no studies showing the true effect values. In this study, HT and DM had a strong effect separately, gender and coronary artery disease (CAD) variables were moderate, chronic obstructive pulmonary disease (COPD), lung cancer and other chronic disease variables had weak effect values, age and non-lung cancers had insignificant effect.

Keywords: Chronical disease, COVID-19, COVID-19 related deaths, effect value, risk factors

Coronavirus disease 2019 (COVID-19), a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, was first reported in December 2019 in Wuhan, Hubei Province, China, and has spread rapidly since then [1]. A global pandemic was declared after its uncontrolled spread around the world. After the rapid spread of COVID-19 infection globally, more than 505 million confirmed cumulative

Received: *July 4, 2022;* ***Accepted:*** *July 24, 2022;* ***Published Online:*** *July 25, 2022*

How to cite this article: *Metin S, Eraybar S, Aygün H, Ercan İ, Demir C. Evaluation of the effect values of risk factors by clustering method in patients who died due to COVID-19 disease. Eur Res J 2022;8(5):678-684. DOI: 10.18621/eurj.1140623*

e-ISSN: 2149-3189

Address for correspondence: *Salih Metin, MD., Family Medicine Specialist, Bursa Provincial Health Directorate Public Hospitals Services Presidency,Dikkaldırım Mah., Hat Cad., No:4, Osmangazi, Bursa, Türkiye. E-mail: slhmtn@hotmail.com, GSM: +90 505 825 46*

cases of coranavirus disease (COVID-19) were diagnosed and more than 6 million people died worldwide [2]. Turkey was one of the countries most affected by COVID-19; Starting with the first case in Turkey on March 11, 2020, and at the end of the last 25 months, as of April 2022, over 14.7 million cumulative confirmed cases were detected and 97,666 deaths occurred [3].

Hospitalization and death rates for COVID-19 patients differ between populations and studies. Mortality rates vary between countries between 1% and 10% [4]. A systematic study conducted in China found the mortality rate to be 3-4% [5]. According to reports from Italy and the United Kingdom, death rates range from 10% to 20% [6, 7]. The death rate in Turkey has been estimated to be around 1% [3].

Fever, cough, respiratory difficulty, muscle soreness, and exhaustion are the most common symptoms of COVID-19 infection [8]. Although the majority of COVID-19 patients recover in a short time, it can lead to fatal results in some patients [9].

The severity and mortality of the COVID-19 disease have been attributed to a variety of demographic variables and co-morbidities. Advanced age is the factor most closely associated with COVID-19 mortality [10, 11]. Other factors linked to mortality include the presence of comorbidities such as diabetes mellitus (DM), cardiovascular and chronic respiratory system diseases, hypertension (HT), obesity, kidney and liver disease, cancer, chronic obstructive pulmonary disease (COPD), asthma [12-15]. In a retrospective study of 828 confirmed cases of COVID-19 from open-access individual-level data worldwide, male, older age, hypertension, diabetes, and USA patients identified as independent risk factors for death among COVID-19 patients [16]. It is stated that neurological and chronic respiratory system diseases are also associated with COVID-19 death [17]. In several studies patients with comorbidities; has been shown that there is a relationship between the severity of the disease and the occurrence of death in those with obesity and also smoking history [18-22].

The aim of this study is to determine the factors that may be associated with mortality in patients who died due to COVID-19 and to determine the effect sizes of the factors that make a statistically significant difference.

METHODS

In this study, the patients who died due to COVID-19, which occurred between 01.03.2020 and 01.03.2021 in Bursa province, were evaluated retrospectively. Study data includes data from a total of 14 hospitals, including 2 training and research hospitals, two central state hospitals and 12 district state hospitals across Bursa. A total of 3,510 deaths were examined and 139 (3.96%) of these patients were foreign nationals. During the planning phase of the study, the Ministry of Health COVID-19 scientific studies permission and the approval of the ethics committee in our hospital were obtained (2019- KAEK-140). The medical records of patients were evaluated retrospectively from the patient files registered in the public health management system and hospital information management system. Patients with at least one positive COVID-PCR sample between March 2020 and March 2021 were considered as COVID positive.

Demographic information of patients like gender, age, nationality, existing chronic diseases of the patients, COVID-PCR test results of the patients, length of hospital stay, intensive care follow-up times, intubation application times and hospitals where they were followed were recorded. Presence of chronic disease of the patients were classified as HT, DM, COPD, malignancy (lung malignant neoplasm, extrapulmonary malignancy) and other chronic diseases. The patients' existing chronic conditions were obtained from hospital information management systems and reported using International Classification (ICD 10) codes.

Statistical Analysis

Cluster analysis was used to show the relationships in a multivariate way, as the data set includes

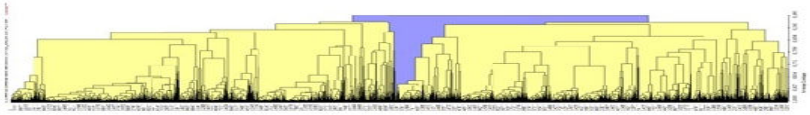

Fig. 1. Hierarchical cluster analysis dendrogram.

both continuous and nominal type variables. Hierarchical clustering methods were designed as part of multivariate statistical analysis to generate homogeneous groups of cases or entities called clusters. The Gower similarity coefficient was used to calculate the similarities between the deaths caused by COVID-19. Clusters were combined with the full linkage method and a dendrogram was created showing the clustering of cases according to their similarity levels (Fig. 1). Clustan Graphics V.8.00 software was used for cluster analysis. Clusters (Cluster-1 and Cluster-2) were statistically tested to identify variables that were influential in clustering.

The conformity of the variables to the normal distribution was examined with the Shapiro Wilk test and the Kolmogorov-Smirnov test. Because the data did not show normal distribution, Mann Whitney U test was used for comparisons between the two groups. Descriptive statistics are given as median (minimum-maximum) because the data are not normally distributed and nonparametric tests are used. Pearson Chi Square test was used to compare categorical variables between groups. Categorical variables were expressed as n (%). Because the sample size was too large and the p-value was statistically significant, the effect size was calculated. In the case of non-parametric Cliff's Delta effect size and Phi effect size for categorical data are given. Statistical analyzes were made with SPSS v22.0 program and R v3.6.3 program "effsize" package. In statistical analyzes, $\alpha = 0.05$ was taken as the level of significance. In the interpretation of Cliff's Delta effect size values; $\delta < 0.147$ unimportant, $\delta < 0.330$ small, $\delta < 0.474$ medium, $\delta \geq 0.474$ intervals are taken into account.The intervals specified in Table 1 were taken into account in the interpretation of the Cliff's Delta effect size and Phi effect size values [23].

Table 1. Interpretation of Cliff's delta effect size and Phi effect size

Effect Size	Interpretation
Cliff's delta effect size	
$\delta < 0.147$	Negligible
$\delta < 0.330$	Small
$\delta < 0.474$	Medium
$\delta \geq 0.474$	Large
Phi effect size	
0.0-0.1	Negligible association
0.1-0.2	Weak association
0.2-0.4	Moderate association
0.4-0.6	Relatively strong association
0.6-0.8	Strong association
0.8-1.0	Very strong association

Interpretation of Cliff's delta effect size and Phi effect size

RESULTS

Total of 3,510 deaths due to COVID-19 were evaluated in this study. When the historical and monthly distributions of deaths were examined, it was seen that the highest values were reached in November with 32.99% (n = 1,158) cases, followed by December with 29.66% (n = 1,041) cases. The lowest mortality rates were recorded in March 2020, at 0.14% (n = 5), and in June of 1.08% (n = 38).

When the gender distribution of our study population is examined, 2,107 (60%) were male and 1,403 (40%) were female (male/female ratio 1.5). Three thousand three hundred and seventy-four (96.12%) patients are 50 years or older (Fig. 2). In both sexes, the highest number of deaths were in the age range of 70-79. In this age group, 34.08% of men (n = 718) and 33.14% of female (n = 465) died due to COVID-19. In the age range of 60-69 years for males (n = 613; 29.09%) and 80-89 years for female (n =349; 24.88%) was the second highest number of deaths were recorded. Fig. 3 shows death distributions by age and gender.

The most common comorbidities were HT (n = 1,182; 34.16%) and DM (n =776; 22.43%). Cancer was present in 130 patients, with 48 (1.39%) having lung cancer and 82 (2.43%) having extra-pulmonary cancer. Five hundred and sixteen (14.91%) patients did not have any concomitant chronic disease. The distributions of chronic diseases accompanying the patients are given in Fig. 4.

When Cluster-1 and Cluster-2 groups are compared, a statistically significant difference was found in terms of age and time to death after COVID-19 diagnosis ($p < 0.008$ and $p < 0.001$, respectively). The effect size was found to be insignificant between the two groups according to Cliff's Delta. There was no statistically significant difference between the groups

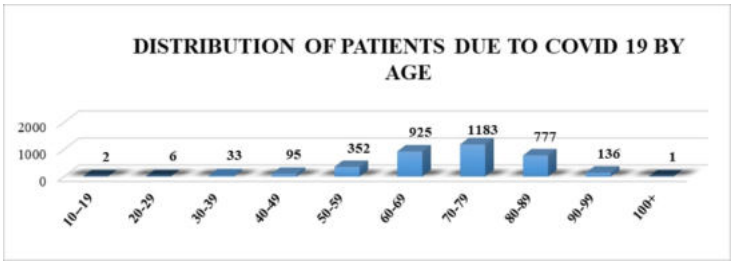

Fig. 2. Age distribution of study population.

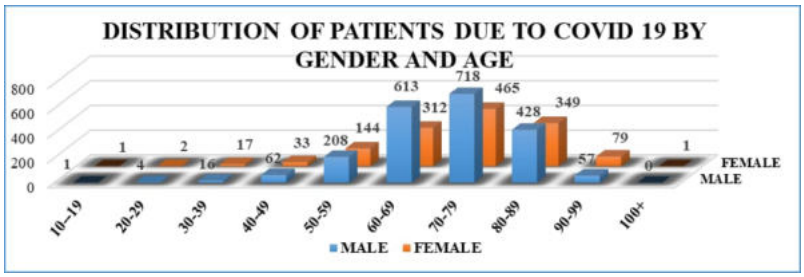

Fig. 3. Age distribution of study population by gender.

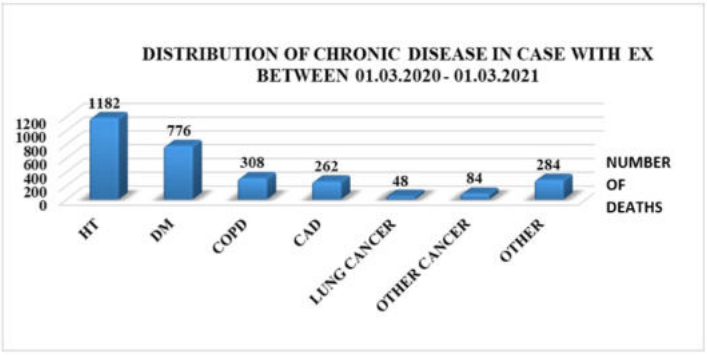

Fig. 4. Distribution of existing chronic diseases in the study population. HT = Hypertension, DM = Diabetes Mellitus, CAD = Coronary artery disease, COPD = Chronic obstructive pulmonary disease.

Table 2. Comparisons of Cluster-1 and Cluster-2 groups

	Cluster-1 (n = 854)	Cluster-2 (n = 842)	*p* value	Cliff's Delta	
				δ	%95 CI.
Age (years)	71 (14-97)	72 (35-102)	**0.008**	0.074	(0.019-0.129)
Intubation time (time from hospitalization to intubation) (hours)	0 (0-49)	0 (0-46)	0.144	-	-
Death time (time from hospitalization to death) (days)	7 (0-90)	8 (0-132)	**< 0.001**	0.097	(0.042-0.151)

Data are given as mean (minimum-maximum)

in terms of intubation time ($p = 0.144$) (Table 2). When Cluster-1 and Cluster-2 groups were compared, there was a statistically significant difference in terms of gender, HT, DM, CAD, COPD, lung cancer, extra-pulmonary cancer and other chronic disease variables ($p < 0.001$, $p < 0.001$, $p < 0.001$, $p < 0.001$, $p < 0.001$, $p < 0.001$, $p = 0.006$ and $p < 0.001$, respectively). When the effect sizes were evaluated, insignificant correlation between the two groups in terms of extra-pulmonary cancer. Weak correlation between the two groups in terms of COPD, lung cancer and other chronic disease; moderate correlation between the two

groups in terms of gender and CAD; A strong correlation was observed between the two groups in terms of HT and DM (Table 3). The rate of Gender-1 observation was higher in the Cluster-1 group than in the Cluster-2 group ($p < 0.001$) (Table 3). The incidence of HT, DM, CAD, respiratory and other chronic diseases was found to be lower in the Cluster-1 group than in the Cluster-2 group ($p < 0.001$, $p < 0.001$, $p < 0.001$, $p < 0.001$ and $p < 0.001$, respectively). Lung cancer and extrapulmonary cancer observed rates were higher in the Cluster-1 group than in the Cluster-2 group ($p < 0.001$ and $p = 0.006$, respectively) (Table 3).

DISCUSSION

During the pandemic process, many studies have been conducted investigating the mortality and morbidity relationships of patients due to COVID-19 [9-11]. Many studies have found that severe COVID-19 disease and variables linked to death, such as HT, DM, Obesity, COPD, CAD, advanced age, and other chronic diseases, are statistically significant. However, the literature on the actual effect values could not be reached. In this study, we examined the effect values in COVID-19 patients who died, in addition to the statistical significance of these variables. HT and DM had a significant effect value between groups, whereas gender and CAD variables had a medium effect value, COPD, lung cancer, and other chronic illness variables had weak effect values, and age and non-lung cancer variables had no effect value.

In the meta-analysis examining on the intensity of underlying diseases in hospitalized 76993 patients for COVID-19, HT, cardiovascular diseases, DM, smoking, COPD, malignancy and chronic kidney disease were detected respectively [24]. The most common comorbidities in COVID-19 patients, according to various studies, are HT, DM, and CAD, with HT being the most common. However, no research has been carried to establish that HT and DM have an independent effect on mortality [25, 26]. In this study, it was determined that HT and DM have a strong effect on deaths related to COVID-19 and pose a statistically significant risk in deaths caused by COVID-19.

In their meta-analysis, Nakanishi *et al*. [27] they

Table 3. Comparisons of Cluster-1 and Cluster-2 groups

		Cluster-1 (n = 854)	Cluster-2 (n = 842)	*p* value	Phi (*p*)
Gender, n (%)	**Male**	607 (71.1)	403 (47.9)	**< 0.001**	0.237 (**< 0.001**)
	Female	247 (28.9)	439 (52.1)		
HT (available), n (%)		93 (10.9)	643 (76.4)	**< 0.001**	0.661 (**< 0.001**)
DM (available), n (%)		12 (1.4)	552 (65.6)	**< 0.001**	0.681 (**< 0.001**)
CAD (available), n (%)		111 (13)	306 (36.3)	**< 0.001**	0.271 (**< 0.001**)
COPD (available), n (%)		69 (8.1)	166 (19.7)	**< 0.001**	0.168 (**< 0.001**)
Lung cancer (available), n (%)		30 (3.5)	4 (0.5)	**< 0.001**	0.108 (**< 0.001**)
Extrapulmonary cancer (available), n (%)		67 (7.8)	39 (4.6)	**0.006**	0.066 (**0.006**)
Other chronical diseases (available), n (%)		114 (13.3)	236 (28)	**< 0.001**	0.181 (**< 0.001**)

HT = Hypertension, DM = Diabetes Mellitus, CAD = Coronary artery disease, COPD = Chronic obstructive pulmonary disease

reported the effect of age-related severity and mortality of the main genetic risk factor for COVID-19. They stated that the risk of mortality increased in those carrying the RS10490770 allele gene on the chromosome 3 locus. They also stated that the effect size is more common than other similar clinical risk factors and is associated with an increased risk of morbidity and mortality, which is more pronounced among individuals aged 60 years and younger [27]. In this study, we found that although the age variable was statistically significant in deaths due to COVID-19, the effect value was insignificant. Therefore, we think that the age variable can be reported as a mortality risk due to strong risk factors such as risky allele, HT, and DM.

Limitations

The main limitation of this study was that when evaluating chronic diseases in our patient population, data on drug use, treatment compliance and disease duration could not be recorded. The other limitation is that treatment differences or additional treatments were not taken into account, since the treatments received by our patients were applied in accordance with standard treatment algorithms.

CONCLUSION

It's critical to understand the disease's epidemiology in order to predict what to expect in the event of a worldwide pandemic. Knowing who is at risk of morbidity and mortality can help with patient management. Risk factors frequently reported for COVID-19 deaths in the literature are HT, DM, CAD, COPD, advanced age, and cancer. Although these variables are statistically significant in the studies, there are no studies showing the true effect values on mortality from COVID-19. In this study, in which 3,510 COVID-19-related deaths were examined, HT and DM had a strong effect separately, gender and CAD variables were moderate, COPD, lung cancer and other chronic disease variables had weak effect values, age and non-lung cancers had insignificant effect.

Understanding the epidemiology of the disease is essential to assist the effort to manage the 2019 coronavirus pandemic and to determine what to look out for in a global new outbreak. We need to know which individuals are at high risk of SARS-CoV-2 infection, as well as the morbidity and mortality risks if infected. The number of articles describing these aspects is increasing, almost similar to the epidemic, revealing the uncertainty about this disease.

Author' Contribution

Study Conception: SM; Study Design: SM, HA; Supervision: SE; Funding: CD; Materials: HA, İE; Data Collection and/or Processing: HA, CD; Statistical Analysis and/or Data Interpretation: SM, İE; Literature Review: SE, CD; Manuscript Preparation: SM and Critical Review: SE, İE.

Ethical disclosure

Ethical committee approval was obtained from Bursa City Hospital Ethical committee during the study planning phase. An informed consent form has been signed by the parents of the cases involved.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during conduction or writing of this study.

REFERENCES

1. Shao A, Zhou Y, Tu S, Sheng J. A novel scoring system in mortality prediction of severe patients with COVID-19. EClinicalMedicine 2020;24:100450.
2. WHO Coronavirus Disease (COVID-19) Dashboard. Available at 'https://covid19.who.int/'. Accessed on Apr 24, 2022
3. Republic of Turkey Ministry of Health. COVID-19 Information Page. Available at 'https://covid19.saglik.gov.tr/TR-66935/genel-koronavirus-tablosu.html'. Accessed on Apr 24, 2022.
4. Cobos-Siles M, Cubero-Morais P, Arroyo-Jiménez I, Rey Hemandez M, Hernández-Gómez L, Vargas-Parra DJ, et al. Cause-specific death in hospitalized individuals infected with SARS-CoV-2: more than just acute respiratory failure or throm-

boembolic events. Intern Emerg Med 2020;15:1533-44.
5. Fu L, Wang B, Yuan T, Chen X, Ao Y, Fitzpatrick T, et al. Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: a systematic review and meta-analysis. J Infect 2020;80:656-65.
6. Giorgi Rossi P, Marino M, Formisno D, Venturelli F, Vincentini M, Grilli R, et al. COVID-19 Working Group. Characteristics and outcomes of a cohort of COVID-19 patients in the Province of Reggio Emilia, Italy. PLoS One 2020;15:e0238281.
7. Goodacre S, Thomas B, Lee E, Sutton L, Loban A, Waterhouse S, et al. Characterisation of 22445 patients attending UK emergency departments with suspected COVID-19 infection: Observational cohort study. PLoS One 2020;15:e0240206.
8. Noor, FM, Islam M. Prevalence of clinical manifestations and comorbidities of coronavirus (COVID-19) infection: a meta-analysis. Fortune J Health Sci 2020;3: 55-97.
9. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress dyndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med 2020;180:934-43.
10. Becerra-Muñoz VM, Núñez-Gil IJ, Eid CM, García Aguado M, Romero R, Huang J, et al. Clinical profile and predictors of in-hospital mortality among older patients hospitalised for COVID-19. Age Ageing 2021;50:326-34
11. CDC COVID-19 Response Team. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. MMWR Morb Mortal Wkly Rep 2020;69:343-6.
12. Albitar O, Ballouze R, Ooi JP, Sheikh Ghadzi SM. Risk factors for mortality among COVID-19 patients. Diabetes Res Clin Pract 2020;166:108293.
13. Liu SP, Zhang Q, Wang W, Zhang M, Liu C, Xiao X, et al. Hyperglycemia is a strong predictor of poor prognosis in COVID-19. Diabetes Res Clin Pract 2020;167:108338.
14. Caramelo F, Ferreira N, Oliveiros B. Estimation of risk factors for COVID-19 mortality- preliminary results. MedRxiv 2020. doi: https://doi.org/10.1101/2020.02.24.20027268
15. Gansevoort RT, Hilbrands LB. CKD is a key risk factor for COVID-19 mortality. Nat Rev Nephrol 2020;16:705-6.
16. Ranzani OT, Bastos LSL, Gelli JGM, Marchesi JF, Baião F, Hamacher S, et al. Characterisation of the first 250,000 hospital admissions for COVID-19 in Brazil: a retrospective analysis of nationwide data. Lancet Respir Med 2021;9:407-18.
17. de Souza FSH, Hojo-Souza NS, Batista BDO, da Silva CM, Guidoni DL. On the analysis of mortality risk factors for hospitalized COVID-19 patients: a data-driven study using the major Brazilian database. PLoS One 2021;16:e0248580.
18. Zhang J, Wang X, Jia X, Li J, Hu K, Chen G, et al. Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China. Clin Microbiol Infect 2020;26:767-72.
19. Yu C, Lei Q, Li W, Wang X, Liu W, Fan X,et al. Clinical characteristics, associated factors, and predicting COVID-19 mortality risk: a retrospective study in Wuhan, China. Am J Prev Med 2020;59:168-75
20. Cen Y, Chen X, Shen Y, Zhang XH, Lei Y, Xu C, et al. Risk factors for disease progression in patients with mild to moderate coronavirus disease 2019-a multi-centre observational study. Clin Microbiol Infect 2020;26:1242-7.
21. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395:1054-62
22. Laguna-Goya R , Utrero-Rico A., Talayero P, Lasa-Lazaro M, Ramirez-Fernandez A, Naranjo L, et al. IL-6-based mortality risk model for hospitalized patients with COVID-19. J Allergy Clin Immunol 2020;146:799-807.e9.
23. Romano J, Kromrey JD, Coraggio J, Skowronek J . Appropriate statistics for ordinal level data: should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys? Paper presented at the annual meeting of the Florida Association of Institutional Research, February 1 -3, 2006, Cocoa Beach, Florida.
24. Emami A, Javanmardi F, Pirbonyeh N, Akbari A. Prevalence of underlying diseases in hospitalized patients with COVID-19: a systematic review and meta-Analysis. Arch Acad Emerg Med 2020;8:e35.
25. Tadic M, Cuspidi C, Grassi G, Mancia G. COVID-19 and arterial hypertension: Hypothesis or evidence? J Clin Hypertens (Greenwich) 2020;22:1120-6.
26. Tadic M, Cuspidi C, Sala C. COVID-19 and diabetes: is there enough evidence? J Clin Hypertens (Greenwich) 2020;22:943-8.
27. Nakanishi T, Pigazzini S, Degenhardt F, Cordioli M, Butler-Laporte G, Maya-Miles D, et al. Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. J Clin Invest 2021;131:e152386.