https://doi.org/10.32394/pe.75.43

Seyyed Amir Yasin Ahmadi[1], Farhad Shahsavar[2], Arash Tehrani-Banihashemi[3], Mitra Elmi[4], Parisadat Ahmadi[5], Saeed Sohrabpour[5], Hossein Nazarian[1], Mehrzad Jafarzadeh[6], Arman Karimi Behnagh[7], Khatereh Anbari[8]*

AN ECOLOGICAL MODELING ON THE ADJUSTED EFFECTS OF SOCIOECONOMIC DETERMINANTS AND HLA-DRB1 ALLELES IN FATALITY OF COVID-19 DURING THE EARLY PHASE OF EPIDEMICS IN A GROUP OF COUNTRIES

[1]Student Research Committee, Lorestan University of Medical Sciences, Khorramabad, Iran
[2]Department of Immunology, Lorestan University of Medical Sciences, Khorramabad, Iran
[3]Department of Community Medicine, Preventive Medicine and Public Health Research Center, Iran University of Medical Sciences, Tehran, Iran
[4]School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin Health Innovation Research Institute, Perth, WA, Australia
[5]Otolaryngology Research Center, Amir Alam Hospital, Tehran University of Medical Sciences, Tehran, Iran
[6]Endocrine Research Center, Institute of Endocrinology and Metabolism, Iran University of Medical Sciences, Tehran, Iran
[7]Student Research Committee, Iran University of Medical Sciences, Tehran, Iran
[8]Department of Community Medicine, Social Determinants of Health Research Center, Lorestan University of Medical Sciences, Khorramabad, Iran,
[*]Corresponding author

ABSTRACT

INTRODUCTION. Socioeconomic determinants along with genetic status may affect fatality rate of COVID-19. We intend to investigate the adjusted effects of the HLA-DRB1 alleles and socioeconomic determinants including gross domestic product per capita (GDP cap) and health expenditure per capita (HE cap) in fatality of COVID-19 during the early phase of epidemic in a group of countries.

METHODS. As an ecological study, early exposure to epidemics was defined as having more than 5000 confirmed cases of COVID-19 from 1 March 2020 to 1 April 2020. Poisson regression was used to report adjusted incidence rate ratio (IRR) for case fatality rate in this constant time period.

RESULTS. Fourteen countries were eligible. Among the alleles, DR7 showed the strongest risk factor (IRR=112.535, P<0.001). Having GDP cap more than 40000$ or having HE cap more than 3000$ was a protecting factor (IRR=0.899, P<0.001, adjusted with allele DR7). Having GDP cap more than 40000$ along with having HE cap more than 3000$ was a protecting factor (IRR=0.471, P<0.001, adjusted with allele DR7).

CONCLUSION. Socioeconomic status of the countries may compensate the probable harmful effect of some HLA-DRB1 alleles. This conclusion was limited to a period that all the selected countries had almost similar governmental intervention.

Keywords: *coronavirus, COVID-19, genetic epidemiology, gross domestic product, health expenditure, HLA-DRB1*

INTRODUCTION

Background. In the field of virology, the coronavirus family is called "Coronaviridae". This single-stranded RNA virus is the second most common cause of common cold syndrome after rhinoviruses (1). So far, several new mutations have been identified among this virus family. The first famous mutation in this family occurred in year 2002 that caused the severe acute respiratory syndrome (SARS) (2). After that, another mutation happened in 2012 that led to the Middle East respiratory syndrome (MERS) (3). The

recent mutation of this family occurred in 2019 (severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)) that became the cause of coronavirus disease in year 2019 (COVID-19) (4). The world statistics of SARS and MERS have previously been published, and COVID-19 statistics is also under publication and updating. SARS globally infected 8096 cases in total that finally had 9.6% fatality rate (5). MERS globally infected 1789 people with a fatality rate of 31.1% in the end (6). One of the distinct characteristics of COVID-19 compared with the two previous mutations, is its high rate of contagion and infectivity in spite of lower case fatality rate (CFR) that has so far infected millions cases. Until now, almost all world countries have been involved, and as a consequence, the world health organization (WHO) confirmed that it was now a pandemic. The statistics of world countries show significant differences in terms of CFR (7).

Human leukocyte antigens (HLA) are important molecules for the function of the immune system. These molecules are categorized in two major subgroups of class I (including A, B, C types, etc.) and class II (including DRB1, DQB1, etc.) (8). HLA class II is located on immune system cells, and its role is identifying and presenting antigens. Antigen-presenting cells (APC) present epitope of the antigen to T helpers (Th) *via* this molecule so that through the Th differentiation toward their own species, the final destiny of the pathway continuation of immune system response will become apparent (9). In the human genome, *HLA* genes are next to each other as a haplotype in a complex. Each gene of this complex on its own is polymorphic. These polymorphisms lead to ethnic diversity and consequently a different function of the immune system in various people. This complex is the most polymorphic gene group in the human genome (10). Some statistical models have been suggested for analysis of HLA related data such as logistic regression (11).

Apart from the immune system, social factors also have important roles in infectious diseases. The economic health of a society is associated with the social determinants of health (SDH). One of the examples of economic health indices is gross domestic product (GDP) (12). The ratio of health expenditure per capita to GDP per capita is also one of the crucial indices in health (13).

Rationale and objectives. In nationwide epidemics, economic and social factors have very important roles. Moreover, the role of genetics besides environmental factors, is very prominent. As an example for the importance of genetics, in our case of study, regarding the *HLA* genes, whenever antigen epitopes are identified and presented to Th cells by the HLA class II, a cytokine storm may start. This response may be different in different individuals according to their genetic profiles. In some viral infections, systemic immune response syndrome (SIRS) occurs after the cytokine storm (14). Up to this point, no ecological studies have been performed based on the genetic and social databases in regard to COVID-19. Therefore, considering the importance of the genetics and social aspects, in this study we intend to investigate the ecological role of *HLA-DRB1* allele frequency using the international database of *HLA* alleles and adjust the impacts with socioeconomic indices including, population, area, median age, GDP per capita and health expenditure per capita. In addition, early exposure of countries to the early phase of epidemics seems to be accompanied by equal performance of the governments, and therefore, this condition was chosen to remove the bias of the governmental performance.

MATERIAL AND METHODS

Study design. The present work was an ecological study based on public information of online databases. All the countries which had more than 5000 confirmed cases of COVID-19 from 1 March 2020 to 1 April 2020 were initially selected. This period was regarded as the primary exposure status of the countries at the early phase of epidemics. For China, 1 February 2020 to 1 March 2020 was regarded. Both the time and this minimum number of cases were selected based on referring to expert opinion.

Eligibility criteria. All the countries with the above conditions were primarily selected. Then, the countries with single race (not necessarily single ethnicity) or studying a central population for *HLA-DRB1* alleles instead (with real proportions of their subpopulation) were eligible for the study.

Variables. The variables used in this study were *HLA-DRB1* alleles (from the bioinformatics database allelefrequencies.net) (15), GDP per capita (named GDP cap) (16), health expenditure per capita (named HE cap) (17), median age of the country, population of the country, area of the country, number of confirmed cases and number of confirmed deaths. For *HLA-DRB1* alleles, each country had some populations in the database. In the cases of existing more than one *HLA* typing study, we chose a central population with more sample size and sum of allele frequencies nearer to 100% as much as possible. Nevertheless, the sum of allele frequencies were not exactly 100% in some countries. Low resolution alleles were used including *01 (DR1), *03 (DR3), *04 (DR4), *07 (DR7), *08 (DR8), *09 (DR9), *10 (DR10), *11 (DR11), *12 (DR12), *13 (DR13), *14 (DR14), *15 (DR15) and *16 (DR16). In addition, a new categorical variable was considered to be created if necessary for combination

of GDP cap and HE cap (called as economy score). Another generated variable was called HLA factor in which a factor analysis was performed for the alleles.

Statistical analysis. Poisson regression analysis was used in Stata 14 (StataCorp LLC, US) to adjust effects of the variables. The outcome variables were death count (per 1000 confirmed cases-one month as the exposure variable) and confirmed case count (per one million of population-one month as the exposure variable). The units of the primary covariates (other than allele frequencies) were changed to a suitable unit for better imagination of incidence rate ratio (IRR) with 95% confidence interval (CI), but the constants (β0) were not suppressed in our models but they were not mentioned. P <0.05 was considered as the significance level. "Spmap" command (18) was used to design ecological map. We ran three different models. The first one was a semi-saturated modeling (with regression degrees of freedom [DF] n-1 in with the total DF was n). The second one was a modeling to adjust the effects of some specific alleles with economy score. The third one was a modeling to adjust the effects of HLA factor with socioeconomic determinants. Heatmap was designed for clustering of the countries based on frequencies of *HLA DRB1* alleles using R 3.6.3 software (R foundation for statistical computing, Austria) (19).

Ethical considerations. This work was approved by the ethics committee of Lorestan University of Medical sciences with ethical registration number # IR.LUMS.REC.1399.012. The public information of the data bases was used according to their copyright policies (referring and acknowledgement and not using for primary aims other than research or education).

RESULTS

Primary findings. A total of 14 countries were eligible for this ecological analysis. The countries were China, South Korea, Italy, Spain, Germany, France, Switzerland, United Kingdom (UK), Netherlands, Austria, Belgium, Portugal, Turkey and Iran. Australia, Canada, Israel and United States of America had more than one race/population without any pooled allele frequency report of *HLA-DRB1* alleles as a central population (according to the allele frequency database). Therefore, they were excluded from the study in spite of having more than 5000 confirmed cases in the mentioned time period. The genetic, demographic and socioeconomic determinants of the countries were summarized (Tables 1 and 2). Clustering of the countries based on frequencies of *HLA DRB1* alleles is shown (Figure 1). Distribution of CFR among the countries and its comparison with Poisson and normal distributions is also shown (Figure 2).

Multiple regression models

– ***Semi-saturated models.*** To achieve the adjusted effect of the genetic, demographic and socioeconomic variables on death count (per case-time period), all the covariates were subjected for Poisson regression analysis. Nevertheless, variable

Table 1. The frequencies (%) of HLA-DRB1 alleles.

Country	DR1	DR3	DR4	DR7	DR8	DR9	DR10	DR11	DR12	DR13	DR14	DR15	DR16
China	1	6.5	13	5	7.8	13.1	1.4	5.8	15.4	4.6	6.2	15.6	5.1
Korea	6.3	2.1	19.5	6.8	9.8	10.2	1.7	4.5	8	10.7	8.1	11.2	1.1
Italy	7.4	6.1	8.2	15.8	2	0.1	0.6	27	1.7	10.5	7	6.3	2.2
Spain	11.25	12.89	12.11	16.97	29.4	0.65	1.4	12.77	1.41	13.22	2.87	9.31	2.08
Germany	11.11	10.55	13.16	12.59	3.25	0.95	0.85	12.53	1.86	12.87	3.09	14.21	2.43
France	10.8	11.3	14.4	13.1	3.5	0.8	1	11.8	1.2	13.9	3.7	11.9	2.7
Switzerland	10.66	9.03	12.39	14.16	3.6	0.73	0.59	13.04	1.56	14.56	4.16	11.6	3
UK	10.53	16.54	17.29	10.15	3.67	1.88	0.75	9.4	1.13	10.9	3.01	13.16	1.5
Netherland	7.41	11.82	15.12	8.4	2.97	1.1	0.65	15.3	1.55	18.14	4.07	10.28	22.7
Austria	13.9	11.8	9.5	13.4	2.4	0.8	0.2	14.2	1.2	14	4.4	13.5	2.7
Belgium	11	15.7	13.1	11.1	4	1	1.5	10.6	1	11.1	1.5	14.2	3
Portugal	12.1	11.2	13.1	14	4	0.8	1.9	9.8	1.3	12.5	2.5	7	2.6
Turkey	3.6	9.6	13.1	11.7	0.2	6.8	0.6	35.4	0.4	6.2	4.3	5.4	2.7
Iran	5.39	11.83	13.17	10.43	2.38	0.7	2.66	20.8	0.98	10.85	4.34	12.61	3.85

Table 2. Demographic and socioeconomic determinants of the countries with the units used in the regression models.

Country	Case (1000 person)	Death (person)	CFR	Population (million person)	Area (1000 mile^2)	Median age (year)	GDP cap (1000$)	HE cap (100$)	Economy score*
China	68.147	2614	0.038	1336.72	3705.39	37.4	10.747	3.9833	1
Korea	6.151	147	0.024	48.7547	38.023	48.1	34.024	20.4386	1
Italy	104.664	12401	0.118	61.0168	116.305	45.5	34.575	27.3871	1
Spain	94.372	8189	0.087	46.7548	194.896	42.7	32.02	23.8989	1
Germany	67.309	732	0.011	81.4718	137.846	47.1	49.617	47.1426	3
France	51.377	3512	0.068	65.3122	211.208	41.4	44.062	42.6336	2
Switzerland	16.09	373	0.023	7.63996	15.942	42.4	85.585	98.3596	3
UK	25.131	1789	0.071	62.6984	94.525	40.5	43.118	39.582	3
Netherland	12.588	1039	0.083	16.847	16.033	42.6	55.73	47.4203	3
Austria	10.172	128	0.013	8.21728	32.382	44	53.482	46.8828	3
Belgium	12.774	705	0.056	10.4315	11.787	41.4	47.782	41.4939	3
Portugal	7.443	160	0.021	10.7603	35.672	42.4	24.509	18.0086	1
Turkey	13.531	214	0.016	78.7855	301.382	30.9	9.599	4.6865	1
Iran	44.013	2855	0.065	77.8912	636.293	30.3	5.902	4.1539	1

* 1: GDP cap <40 and HE cap <30; 2: GDP cap >40 or HE cap >30; 3: GDP cap >40 and HE cap >30.

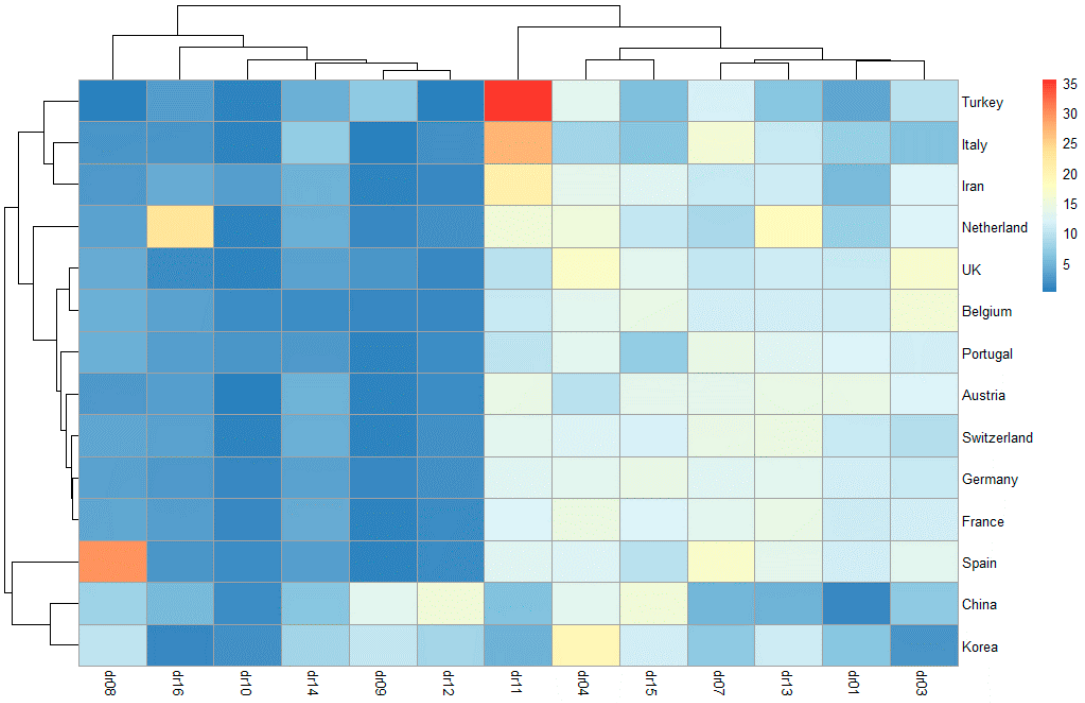

Figure 1. Heatmap of HLA-DRB1 allele frequncy for the studied countries.

removal was detected by the software due to low total DF. Among the alleles, *DR7* showed the strongest independent risk factor (IRR=112.535, 95% CI=84.496-149.878, P<0.001, pseudo R^2 =0.990), of course merely in comparison to other alleles, because of the removal of other covariates for the mentioned reason (Table 3). Similar problem occurred for case count outcome. After saturation of the total DF with only *HLA-DRB1* alleles, *DR1* showed the strongest significant protecting factor (IRR=0.572, 95% CI=0.385-0.851, P=0.006, pseudo R^2 =0.934) (Table 4).

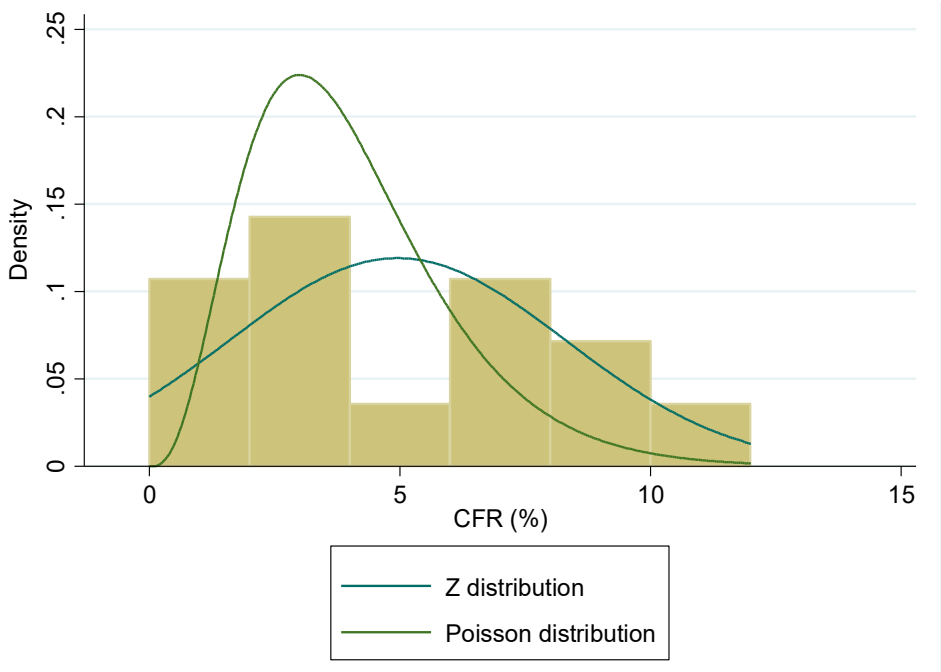

Figure 2. Distribution of CFR for the studied countries (the Poisson distribution is based on lambda =3)

Table 3. Poisson regression analysis for the possible effects of all covariates on death count (in confirmed cases-one month).

Covariate	Adjusted IRR	St. Err.	t-value	p-value	95% CI of IRR
DR1	0.879	0.037	-3.04	0.002	0.809-0.955
DR3	1.427	0.023	21.59	<0.001	1.381-1.473
DR4	1.995	0.077	17.86	<0.001	1.85-2.152
DR7	112.535	16.453	32.31	<0.001	84.496-149.878
DR8	0.683	0.008	-32.23	<0.001	0.668-0.699
DR9	7.251	0.49	29.33	<0.001	6.352-8.277
DR10	24.461	3.066	25.51	<0.001	19.133-31.273
DR11	0.384	0.011	-32.56	<0.001	0.362-0.407
DR12	0.065	0.006	-29.52	<0.001	0.054-0.078
DR13	0.049	0.005	-29.4	<0.001	0.04-0.06
DR14	35.826	3.674	34.9	<0.001	29.303-43.802
DR15	5.508	0.341	27.57	<0.001	4.879-6.218
DR16	7.433	0.493	30.27	<0.001	6.528-8.464
Population	Omitted*				
Area	Omitted				
GDP cap	Omitted				
HE cap	Omitted				
Age	Omitted				
In (case)	Exposure				
Pseudo R^2	0.990				

* Automatically omitted because of low degrees of freedom.

Table 4. Poisson regression analysis for the possible effects of all covariates on confirmed case count (in population-one month).

Covariate	Adjusted IRR	St. Err.	t-value	p-value	95% CI of IRR
DR1	0.572	0.116	-2.76	0.006	0.385-0.851
DR3	0.816	0.065	-2.56	0.010	0.699-0.953
DR4	0.603	0.124	-2.45	0.014	0.402-0.904
DR7	1.011	0.888	0.01	0.990	0.181-5.655
DR8	1.050	0.078	0.66	0.507	0.909-1.214
DR9	1.005	0.416	0.01	0.990	0.447-2.262
DR10	0.439	0.322	-1.12	0.261	0.104-1.846
DR11	0.690	0.12	-2.13	0.033	0.491-0.971
DR12	0.369	0.206	-1.79	0.074	0.124-1.101
DR13	0.824	0.495	-0.32	0.748	0.254-2.675
DR14	1.049	0.623	0.08	0.936	0.327-3.362
DR15	1.047	0.383	0.13	0.900	0.511-2.143
DR16	1.043	0.41	0.11	0.916	0.482-2.253
Area	Omitted*				
GDP cap	Omitted				
HE cap	Omitted				
Age	Omitted				
In (population)	Exposure				
Pseudo R^2	0.934				

* Automatically omitted because of low degrees of freedom.

Table 5. Poisson regression analysis for the possible effect of economy score on death count (in confirmed cases-one month) adjusted with DR7.

Covariate	Adjusted IRR	St.Err.	t-value	p-value	95% CI of IRR
DR7	1.073	0.002	45.53	<0.001	1.07-1.076
Economy score					
1	1				Reference
2	0.899	0.016	-5.89	<0.001	0.868-0.932
3	0.471	0.008	-46.92	<0.001	0.456-0.486
In (case)	Exposure				
Pseudo R^2	0.480				

– ***Economy score combination models.*** In order to solve the problem of the saturation due to low total DF we tried to limit the number of covariates. Hence, we decided to define a new categorical score for combination of the status GDP cap and HE cap. This score was named economy score. Score 1 belonged to GDP cap less than 40000$ and HE cap less than 3000$. Score 2 belonged to GDP cap more than 40000$ or HE cap more than 3000$. Score 3 belonged to GDP cap more than 40000$ along with HE cap more than 3000$ (Table 2). For death count model, we adjusted economy score with allele *DR7* as the most effective allele. Accordingly, economy score 2 showed a significant protecting effect in comparison to score 1 as the baseline (IRR=0.899, 95% CI=0.868-0.932, P<0.001) and also economy

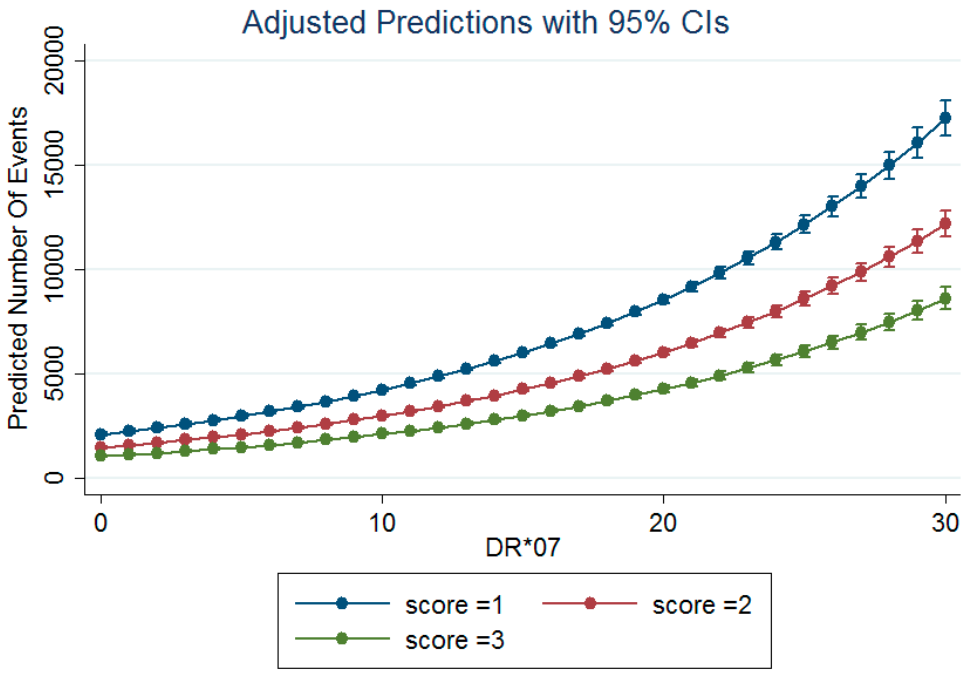

Figure 3. Marginal analysis of DR7 for prediction of death number per 1000 confirmed cases at different economy scores (the regression model consisted of DR7 and economy score)

Table 6. Poisson regression analysis for the possible effect of economy score on case count (in population-one month) adjusted with DR1, DR3 and DR4.

Covariate	Adjusted IRR	St.Err.	t-value	p-value	95% CI of IRR
DR1	1.319	0.028	12.94	<0.001	1.265-1.375
DR3	1.041	0.029	1.46	0.144	0.986-1.099
DR4	0.803	0.024	-7.21	<0.001	0.757-0.853
Economy score					
1	1				Reference
2	0.762	0.142	-1.45	0.147	0.529-1.1
3	0.576	0.074	-4.3	<0.001	0.448-0.74
In (population)	Exposure				
Pseudo R^2	0.842				

Table 7. Unadjusted effects of socioeconomic determinants on death count (in confirmed case-one month) and case count (in population-one month).

Input/outcome variable	Unadjusted IRR	95% CI	p-value
GDP cap/death	0.997	0.996-0.997	<0.001
HE cap/death	0.995	0.994-0.995	<0.001
GDP cap/case	1.049	1.046-1.053	<0.001
HE cap/case	1.039	1.037-1.043	<0.001

score 3 showed a significant protecting effect in comparison to score 1 as the baseline (IRR=0.471, 95% CI=0.456-0.486, P<0.001) (pseudo R^2=0.480) (Table 5). Marginal analysis showed a possible negative interaction of economy score with *DR7* to predict death counts per confirmed cases (Figure 3). For confirmed case count, we adjusted economy score with alleles *DR1*, *DR3* and *DR4* as the most significant alleles of the complete model. Accordingly, economy score 3 showed a significant protecting effect in comparison to score 1 as the baseline (IRR=0.576, 95% CI=0.448-0.740, P<0.001) (pseudo R^2 =0.842) (Table 6). Unadjusted effects of GDP cap and HE cap on death count and confirmed case count is also shown (Table 7). The ecological maps of CFR, frequency of allele *DR7*, GDP cap and HE cap are shown (Figures 4 and 5).

– ***HLA factor model.*** Factor analysis weighted with one million of population unit was performed on the alleles. Accordingly, we performed the analysis to give three components with Eigenvalue cutoff point 1. The alleles with correlation coefficient less than 0.6 with factor 1 were removed in the next step, in which two factors remained. Factor score was calculated for factor 1 as the new variable called HLA factor (Kaiser-Meyer-Olkin value =0.678). Accordingly, HLA factor was a risk factor (IRR=1.094, 95% CI=1.057-1.133, P<0.001). The effects of the other covariates are shown (Table 8).

DISCUSSION

Summary of evidence. In the present ecological study, we tried to solve ecological fallacy using regression modeling as well as showing the modifying effects of genes and socioeconomic status on each other. It was important for us to choose a period, when all the countries had a similar status. Therefore, 14 countries with more than 5000 confirmed cases in this time period being single race were selected. Due to the limitation of low DF semi-saturated modeling was performed only for *HLA DRB1* alleles. Then, we analyzed the adjusted effect of the most important alleles with economy score. In addition, a factor analysis was performed for the alleles.

In the semi-saturated models, socioeconomic determinants were automatically removed due to low total DF. Both models of death count outcomes and confirmed case count outcomes were highly fitted with the individual distribution due to saturation (pseudo R^2 >90%) (Tables 3 and 4). Regarding this methodology, *DR7* increased incidence of death (the most effective allele selected for the economy score combination model), and *DR1*, *DR3*, *DR4* and *DR11* decreased incidence of the disease (the first three ones were selected for the economy score combination model). In the combination model of adjusting, the mentioned alleles with economy score, the effect of *DR7* on death remained significant with similar effect

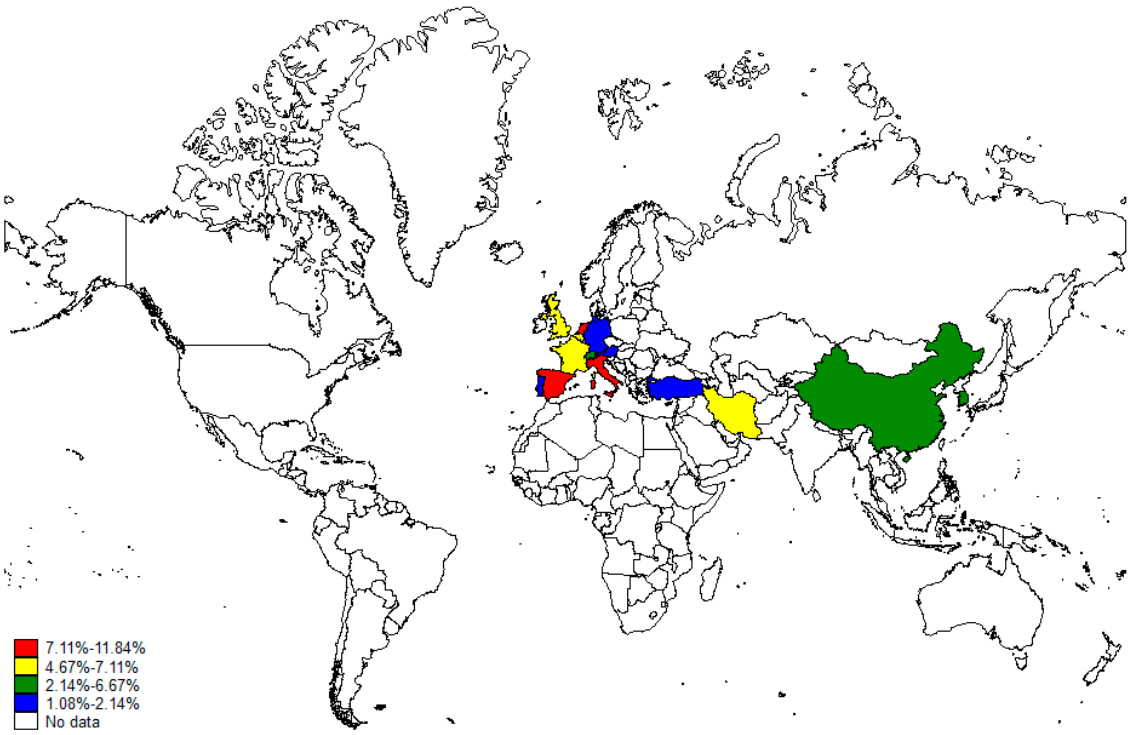

Figure 4. CFR of the countries. The classification thresholds are automatic

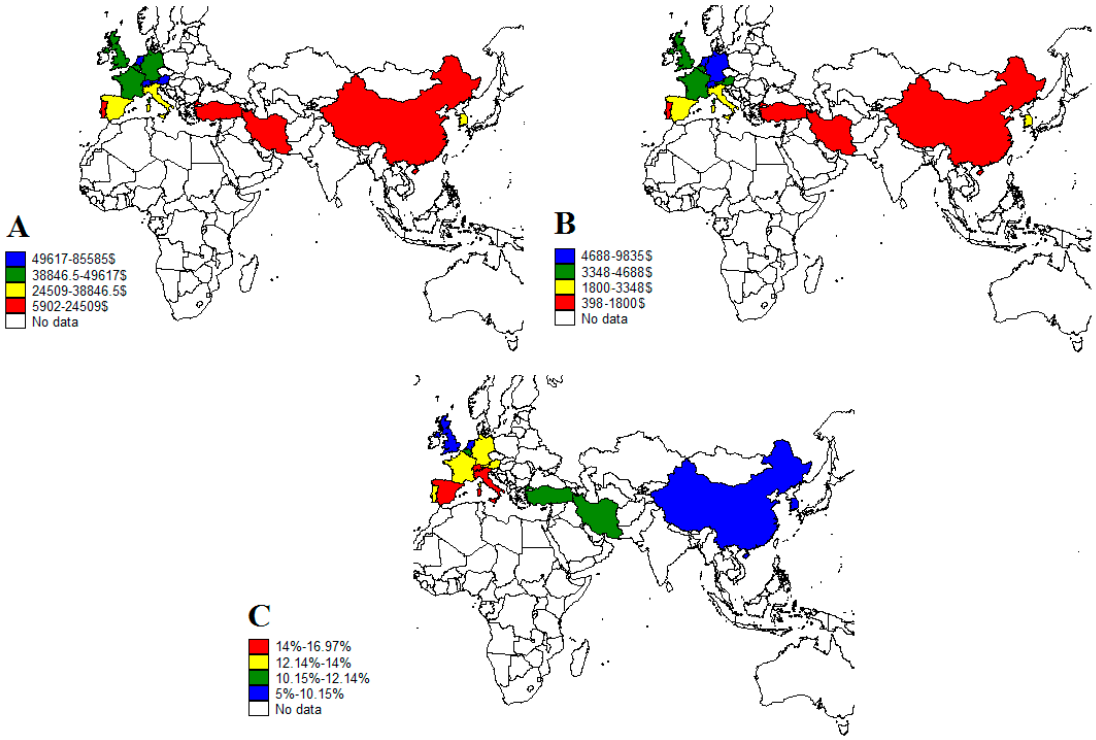

Figure 5. A) GDP cap of the countries; B) HE cap of the countries; C) frequency of DR7 allele in the countries. The classification thresholds are automatic. Comparision of Italy, Germany and Switzerland in the above conditions is notable

Table 8. Poisson regression analysis for the possible effect of economy score on death count (in confirmed cases-one month) adjusted with the HLA factor, area, population and age.

Covariate	Adjusted IRR	St.Err.	t-value	p-value	95% CI of IRR
HLA factor	1.094	0.020	5.05	<0.001	1.057-1.133
Population	1.002	<0.001	8.19	<0.001	1.002-1.003
Area	0.999	<0.001	-10.06	<0.001	0.999-0.999
Economy score					
1	1				
2	0.750	0.014	-15.62	<0.001	0.723-0.777
3	0.328	0.006	-65.73	<0.001	0.317-0.339
age	1.004	0.003	1.58	0.115	0.999-1.009
In (case)	Exposure				
Pseudo R^2	0.518				

direction (though with lower magnitude of effect) (Table 5). However, the effect of *DR3* on taking the disease did not remain significant and the effect direction of *DR1* was changed (Table 6). Unadjusted analysis of GDP cap and HE cap showed inconclusive results, because magnitude of effects was not enough. The unadjusted risk role of GDP cap and HE cap on case count outcome might be due to better screening of the wealthier countries (Table 7). The HLA factor modeling showed that some *HLA* alleles increase the risk of death in spite of adjusting with socioeconomic determinants (Table 8).

As expected, each country has a variety of alleles neutralizing the harmful effect of some other alleles like *DR7*, ecologically (Figure 1). This point can be understood visually *via* the heatmap. As it is seen, countries with higher frequency of *DR7* do not have necessarily higher CFR. It seems that the protecting interactions of the alleles as well as better economic status may overcome the effect of *DR7*.

Previously, an *in silico* study for vaccine design of SARS-CoV-2 has shown that among *HLA-DRB1* alleles, *04:01 and *07:01 had more affinity to the epitopes of SARS-CoV-2 spike protein (20). This result

was similar to our study. According to our findings, both of them were associated with increased incidence of death whereas *DR4* was associated with decreased incidence of getting symptomatic disease. It seems that *DR4* results in better recognition of epitopes for better antiviral reaction whereas *DR7* results in triggering cytokine storm in infected patients. Of course, it is just a hypothesis which may be affected by ecological fallacy.

Until now, some studies have been conducted about the role of HLA in diseases of coronaviruses. Lin et al. (2003) in Taipei examined the role of HLA class I and II in susceptibility to SARS. They found that HLA-B *4601 was associated with SARS susceptibility and disease severity (21). In a similar study in Japan (2009), HLA-DRB1 *1202 behaved as a risk factor (22). Har et al. (2016) in Saudi Arabia studied the relationship between HLA class II and MERS susceptibility. The results showed that DRB1 *1101 and DQB1 *1202 were associated with MERS, but were not associated with disease outcome (23). It was very interesting that the HLA disease association studies on coronaviruses were few in spite of their very important anthropological role in susceptibility to or severity of different infectious diseases.

Disease association of HLA DR7. Previously, disease association of HLA DR7 with other diseases has been studied at individual level. Bavinck et al. (2000) in a population-based study in Saba Island showed that in 124 white individuals, the subjects who had *HLA-DR7* allele had higher risk of the development of basal cell carcinoma (odds ratio [OR]=3.8) (24). Newman et al. (2004) studied the association of *HLA-DRB1* alleles in 507 patients with Crohn's disease. They reported that the ileal involvement in Crohn's disease was associated with *HLA-DRB1* alleles. According to their study, subjects with *07:01 allele, had higher risk of ileal involvement in Crohn's disease (OR=1.9) (25). Chu et al. (2005) showed that among different *HLA-DRB1* alleles, the frequency of *07 was the highest in Chinese patients with chronic HBV. Moreover, the presence of such allele was associated to lower response rate to interferon treatment (26). Noble et al. (2011) showed that different *HLA* haplotypes containing *DRB1*07:01* allele could play a paradoxical role in the patients with type 1 diabetes. The European-derived haplotype *DRB1*07:01-DQA1*02:01-DQB1*02:01g* was associated with lower risk (OR=0.34) of type 1 diabetes, on the other hand, the African-derived haplotype *DRB1*07:01-DQA1*03:01-DQB1*02:01g* (OR=3.96) increased the risk of type 1 diabetes (27). Paradowska-Gorycka et al. (2016) showed that among Polish patients with mixed connective tissue disease, *HLA-DRB1* genotypes were highly associated with the development of disease. Interestingly, the type of association varied among the different members of genotypes. Regarding such association, they reported that *07:01 allele was found to be protective (OR= 0.50); however, other members including *15:01 and *09:01 increased the risk of mixed connective tissue disease (28). Shen et al. (2018) studied the association of between *HLA* alleles and body mass index (BMI). They examined the correlation of 72 low-resolution *HLA* alleles and 163 high-resolution *HLA* alleles in 1.3 million individuals and found that the *HLA DRB1*07* was among those *HLA* alleles with strongest evidence of association with higher BMI (29). Leonard et al. (2019) showed that type I diabetes patients and their first-degree relatives were in high risk of developing celiac disease autoimmunity. Based on their study, the risk of such celiac disease autoimmunity was associated with the presence of DR7-DQ2 and DR4-DQ8 haplotypes in affected subjects (30). According to the above literature, HLA DR7 disease association was different in different diseases. These associations may be causal or confounded by other alleles and factors. The limitation of such studies was lack of adjusting the effects of the alleles with each other. According to the word of Noble et al., HLA DR7 was ethnicity associated. In our study, high prevalence of HLA DR7 in European countries might result in higher fatality rate. In Germany and Switzerland, better GDP cap and HE cap covered this detrimental effect.

Controversies on fatality rate. The countries around the world applied different approaches to detect COVID-19 and estimate CFR. Various factors likely effect on CFR estimation, and identifying them will help reduce CFR that one of these factors is GDP per capita of each country. The countries with more health benefits are likely able to test non-symptomatic carriers with none or mild symptoms in addition to whom are ill and dying consequently. CFR estimation is probably more accurate in such countries than in countries where tests are performed only on critically ill and dying patients. Therefore, they are more successful in diagnosing and controlling the disease and its treatment. There is a large difference in GDP per capita and its impact on the current CFR estimation among the different countries. Thus, with increasing prevalence, in those with lower GDP per capita CFR can significantly increase due to the infection fatality rate (IFR) consistence for COVID-19. Assuming that the number of deaths reported worldwide is true, the actual number of infected people is likely to be much higher than reported. Therefore, the CFR will increase significantly, which could have adverse consequences, including imposing more force on the health system. Finally, due to the increasing the trend of epidemics in all countries of the world, it is necessary to conduct

experiments beyond the clinical priority on a large number of people in the community so that timely diagnosis and control of the disease can significantly reduce CFR (31).

COVID-19 and socioeconomic status. It is known that socioeconomic indicators are associated with COVID-19 related outcomes. Many countries use household prevention models. However, these models are fragile among the low socioeconomic status populations (32). A study in New York on the association of human mortality and sociodemographic factors showed that these factors could be associated with increased risk of COVID-19 infection and failure in social distancing interventions (33). In previous studies, socioeconomic gradient could affect the associations of other variables and result on controversial results (34-35).

Ecological studies on coronaviruses. So far, some ecological studies have been conducted for diseases of coronaviruses. The first ecological map of HLA in COVID-19 was shown by Nguyen et al. (2020). The focus of this study was mainly on the affinity of HLA class I with *in silico* approach (36). Tan et al. (2005) published an ecological study on the effect of air on SARS outbreaks. The results showed that temperature changes in China cities were associated with the SARS outbreak. However, the authors acknowledged the possibility of ecological fallacy (37). In another spatial map study, the distribution of SARS infection risk in different parts of a hospital was examined (38). Elkind et al. (2020) explained the role of the American Heart Association in COVID-19, which was one of their activities in the field of social factors affecting health (39). Zhang et al. designed an ecological map of the COVID-19 transmission in China. They demonstrated the trend using regression models (40). So far, no ecological studies have been published about the combination role of genetic and environmental factors.

Limitations. The most important limitation in each ecological study is ecological fallacy. Our study is not optimal, although we tried to solve the fallacy *via* these ways. First, we have chosen a time period in which all the studied countries had more similar status in screening, case detection and treatment. After this time, the policies of the countries were changed. Second, multiple regression models helped us to adjust the confounding effects of *HLA-DRB1* alleles and socioeconomic determinants; however, we ran into saturation of the models. Third, we have chosen countries with one race or evaluation of *HLA* alleles in a central population instead with combination of ethnicities to reduce the bias of race and ethnicity.

CONCLUSION

According to our models, *HLA DR7* was the most important *HLA-DRB1* allele associated with increased fatality rate. After adjustment with economy score, it remained as a risk factor for increased fatality rate. *HLAs DR1*, *DR3* and *DR4* were associated with decreased number of symptomatic cases. After adjustment with economy score, *DR1* became a risk factor and *DR4* remained a protecting factor. It seems that socioeconomic status of a country may affect genetic susceptibility of individuals. This study should be repeated after the end of this pandemic using panel regression models. The key points and suggestions of this study are summarized as bellow.

- Both socioeconomic determinants and genetic status affect fatality rate of COVID-19. There may be some potential gene-environment interactions.
- Countries with higher GDP per capita along with higher health expenditure per capita may have better screening and treatment at early phases of epidemics (not necessarily till the end of pandemics).
- This ecological study was performed at the early stage of the outbreak as a period, when all the involved countries had similar status.
- HLA-DRB1 plays role in antigen recognition and then immune response. Among *HLA-DRB1* alleles, *DR7* showed the strongest adjusted association with increased fatality due to COVID-19. Most *HLA-DRB1* alleles were associated with fatality rate, but not necessarily with incidence of symptomatic case.
- Higher GDP per capita along with higher health expenditure per capita may cover detrimental effect of *HLA DR7*. However, it should not be concluded as causation relationship.
- Conducting association study on the role of *HLA-DRB1* alleles in fatality of COVID-19 in different countries, ethnicities and races is strongly suggested. The researchers are supposed to perform multiple regression models. It is very important that bivariable analysis of the single associations is not enough.
- According to the semi-saturated model for the profile of *HLA DRB1* alleles and *in silico* studies, designing and trial of peptide vaccines are strongly suggested regarding ethical issues and conflicts of interest.
- Ecological fallacy should always be regarded.

Conflicts of interest. The authors declare that they have no conflict of interest. No financial or governmental secondary interest exists. We deny any interpretation in favor of epidemiologists or immunologists interests.

Funding. None.

Acknowledgement. This study has been intellectually supported by Student Research Committee, Lorestan University of Medical Sciences, Khorramabad, Iran with grant number #1374 under supervision of the corresponding author. We also acknowledge WHO website for its available data.

REFERENCES

1. El-Sahly HM, Atmar RL, Glezen WP, et al. Spectrum of clinical illness in hospitalized patients with "common cold" virus infections. Clin Infect Dis. 2000;31(1):96-100.
2. Vicenzi E, Canducci F, Pinna D, et al. Coronaviridae and SARS-associated coronavirus strain HSR1. Emerg Infect Dis. 2004;10(3):413-8.
3. McIntosh K, Perlman S. Coronaviruses, including severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS)2015.
4. Kurdoğlu M, Khaki A. What Is the Current State of Knowledge About the Novel Coronavirus Infection During Pregnancy? Int J Womens Health Reprod Sci. 2020;8(2):110-1.
5. Keogh-Brown MR, Smith RD. The economic impact of SARS: how does the reality match the predictions? Health policy. 2008;88(1):110-20.
6. Salamatbakhsh M, Mobaraki K, Sadeghimohammadi S, et al. The global burden of premature mortality due to the Middle East respiratory syndrome (MERS) using standard expected years of life lost, 2012 to 2019. BMC Public Health. 2019;19(1):e1523.
7. Coronavirus disease (COVID-2019) situation reports. Geneva: World Health Organization; 2020. Available from: www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.
8. Trachtenberg E, Vinson M, Hayes E, et al. HLA class I (A, B, C) and class II (DRB1, DQA1, DQB1, DPB1) alleles and haplotypes in the Han from southern China. Tissue Antigens. 2007;70(6):455-63.
9. Ahlers JD, Belyakov IM, Thomas EK, et al. High-affinity T helper epitope induces complementary helper and APC polarization, increased CTL, and protection against viral infection. J Clin Invest. 2001;108(11):1677-85.
10. Shahsavar F, Varzi A-M, Ahmadi SAY. A genomic study on distribution of human leukocyte antigen (HLA)-A and HLA-B alleles in Lak population of Iran. Genom Data. 2017;11:3-6.
11. Cordell HJ, Clayton DG. A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. American J Hum Genet. 2002;70(1):124-41.
12. Anderson M, Mossialos E. Beyond gross domestic product for New Zealand's wellbeing budget. Lancet Public Health. 2019;4(7):e320-e1.
13. Smith N, Mitton C, Kershaw P. The reallocation challenge: Containing Canadian medical care spending to invest in the social determinants of health. Can J Public Health. 2016;107(1):e130-e2.
14. Wu W, Shi Y, Gao H, et al. Immune derangement occurs in patients with H7N9 avian influenza. Critical Care. 2014;18(2):R43.
15. Gonzalez-Galarza FF, McCabe A, Santos EJMd, et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. Nucleic Acids Res. 2020;48(D1):D783-D8.
16. World Population Review 2019. Available from: https://worldpopulationreview.com/countries/countries-by-gdp/.
17. Global Health Expenditure Database Geneva2016. Available from: https://apps.who.int/nha/database.
18. Pisati M, editor. The A to Z of how to create thematic maps of Italy using spmap. Italian Stata Users Group Meeting, November; 2014.
19. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
20. Ahmed SF, Quadeer AA, McKay MR. Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. Viruses. 2020;12(3):254.
21. Lin M, Tseng H-K, Trejaut JA, et al. Association of HLA class I with severe acute respiratory syndrome coronavirus infection. BMC Med Genet. 2003;4(9):1-7.
22. Keicho N, Itoyama S, Kashiwase K, et al. Association of human leukocyte antigen class II alleles with severe acute respiratory syndrome in the Vietnamese population. Hum Immunol. 2009;70(7):527-31.
23. Hajeer AH, Balkhy H, Johani S, et al. Association of human leukocyte antigen class II alleles with severe Middle East respiratory syndrome-coronavirus infection. Ann Thoracic Med. 2016;11(3):211-3.
24. Bavinck JNB, Bastiaens MT, Marugg ME, et al. Further evidence for an association of HLA-DR7 with basal cell carcinoma on the tropical island of Saba. Arch Dermatol. 2000;136(8):1019-22.
25. Newman B, Silverberg MS, Gu X, et al. CARD15 and HLA DRB1 alleles influence susceptibility and disease localization in Crohn's disease. Am J Gastroenterol. 2004;99(2):306-15.
26. Chu R-H, Ma L-X, Wang G, et al. Influence of HLA-DRB1 alleles and HBV genotypes on

interferon-α therapy for chronic hepatitis B. World J Gastroenterol. 2005;11(30):4753-7.
27. Noble JA, Johnson J, Lane JA, et al. Race-specific type 1 diabetes risk of HLA-DR7 haplotypes. Tissue Antigens. 2011;78(5):348-51.
28. Paradowska-Gorycka A, Stypińska B, Olesińska M, et al. Association of HLA-DRB1 alleles with susceptibility to mixed connective tissue disease in Polish patients. HLA. 2016;87(1):13-8.
29. Shen J, Guo T, Wang T, et al. HLA-B* 07, HLA-DRB1* 07, HLA-DRB1* 12, and HLA-C* 03: 02 Strongly Associate With BMI: Data From 1.3 Million Healthy Chinese Adults. Diabetes. 2018;67(5):861-71.
30. Leonard MM, Camhi S, Kenyon V, et al. Targeted genotyping for the prediction of celiac disease autoimmunity development in patients with type 1 diabetes and their family members. World J Diabetes. 2019;10(3):189-99.
31. Ward D. Sampling Bias: Explaining Wide Variations in COVID-19 Case Fatality Rates. Preprint. 2020.
32. Chung RY-N, Dong D, Li MM. Socioeconomic gradient in health and the covid-19 outbreak. BMJ. 2020;369:m1329.
33. Sy KTL, Martinez ME, Rader B, et al. Socioeconomic disparities in subway use and COVID-19 outcomes in New York City. Am J Epidemiol. 2021;190(7):1234-42.
34. Mena GE, Martinez PP, Mahmud AS, et al. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. Science. 2021;372(934):1-9.
35. Little C, Alsen M, Barlow J, et al. The impact of socioeconomic status on the clinical outcomes of COVID-19; a retrospective cohort study. J Community Health. 2021;46:794-802.
36. Nguyen A, David JK, Maden SK, et al. Human leukocyte antigen susceptibility map for SARS-CoV-2. J Virol. 2020;94(13):e00510-20.
37. Tan J, Mu L, Huang J, et al. An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. J Epidemiol Community Health. 2005;59(3):186-92.
38. Qian H, Li Y, Nielsen PV, et al. Spatial distribution of infection risk of SARS transmission in a hospital ward. Build Environ. 2009;44(8):1651-8.
39. Elkind MS, Harrington RA, Benjamin IJ. The Role of the American Heart Association in the Global COVID-19 Pandemic. Circulation. 2020;141(15):e743-e5.
40. Zhang F, Zhang J, Cao M, et al. A Simple Ecological Model Captures the Transmission Pattern of COVID-19 Outbreak in China. Preprint. 2020.

Received: 02.08.2021
Accepted for publication: 20.12.2021

Address for correspondence:
Khatereh Anbari MD, MPH
Social Determinants of Health Research Center,
Lorestan University of Medical Sciences,
Khorramabad, Iran
Phone: +989161613257
Email: anbari.kh@lums.ac.ir