Comparison all-cause mortality between individuals with COVID-19 and propensity-score-matched individuals without COVID-19 in South Korea

Tak Kyu Oh1,§, In-Ae Song1,§, Kyoung-Ho Song2, and Young-Tae Jeon1,3,*

1 Department of Anesthesiology and Pain Medicine, Seoul National University Bundang Hospital, Seongnam, South Korea

2 Division of Infectious Diseases, Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, South Korea

3 Department of Anesthesiology and Pain Medicine, College of Medicine, Seoul National University, Seoul, South Korea

§ These authors made an equal contribution as co-first authors.

* Corresponding author: Young-Tae Jeon

Department of Anesthesiology and Pain Medicine, Seoul National University Bundang Hospital, Department of Anesthesiology and Pain Medicine, College of Medicine, Seoul National University

Gumi-ro 173 Beon-gil, Bundang-gu, Seongnam 13620, South Korea

Tel: +82-31-787-7499

E-mail: ytjeon@snubh.org

© The Author(s) 2021. Published by Oxford University Press on behalf of Infectious Diseases Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Authors’ contributions: Tak Kyu Oh and In-Ae Song designed the study, analyzed the data, interpreted the data, and drafted the manuscript; Kyoung-Ho Song contributed to the acquisition of data; Young-Tae Jeon contributed to the study conceptualization, acquisition of data, and review of manuscript.; All authors have given final approval for the final version of the manuscript.

Key points: All-cause mortality in the South Korean adult population was twice as high among COVID-19 patients compared with individuals without COVID-19.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest: None declared

Prior presentation: None
Abstract

Background

We compared all-cause mortality between individuals in South Korea with and without coronavirus disease (COVID-19), using propensity score (PS)-matching.

Methods

This population-based cohort study used data from the National Health Insurance Service COVID-19 cohort database. In the database, we included individuals (COVID-19 patients, control population, and test-negative individuals) aged 20 years or older, regardless of hospitalization. The primary endpoint was all-cause mortality between January 1, 2020 and August 27, 2020.

Results

A total of 328,374 adults were included in the study: 7,713 and 320,660 in the COVID-19 group and the control group, respectively. After PS-matching, a total of 15,426 individuals (7,713 per group) were included in the analysis. All-cause mortality 3.2% (248/7,713) and 1.6% (126/7,713) in the COVID-19 group and the control group, respectively. In Cox regression analysis after PS-matching, the risk of death in the COVID-19 group was twice as high (hazard ratio: 2.00; 95% confidence interval: 1.61 to 2.48; \(P<0.001\)) than that in the control group. Among patients aged ≥60 years, the COVID-19 group had a 2.32-fold higher all-cause mortality compared with the control group, while statistically statistical differences were not observed in the age groups 20–39 years \(P=0.339\) and 40–59 years \(P=0.562\).
Conclusions

In South Korea, all-cause mortality was twice as high among individuals with COVID-19 than among those with similar underlying risks, primarily because of the elevated COVID-19-associated mortality in those aged ≥60 years. Our results highlight the need for prevention of COVID-19 with respect to mortality as a public health outcome.

Keywords: Viruses; Infections; Public Health; Population
Introduction

On March 11, 2020, the World Health Organization declared coronavirus disease-2019 (COVID-19) a pandemic. As of October 16, 2020, 38,825,968 cases of COVID-19 and 1,096,833 COVID-19 related deaths were reported globally, and there is no effective and safe vaccine for COVID-19. Therefore, it is still a global public health crisis.

The COVID-19 death rate is reported to vary by country, and the case fatality ratio was estimated to be 0.02 (or 2%) by a previously conducted meta-analysis. Moreover, the infection fatality rate has been reported as 0.95% in United States, and a recent study reported that about 6% of the global population had died due to COVID-19. However, most studies have focused on the mortality rates among COVID-19 patients, and previous studies have not compared all-cause mortality among persons diagnosed with COVID-19 to propensity score (PS)-matched controls who were not diagnosed with COVID-19. In addition to individuals diagnosed with COVID-19, the pandemic may affect health-related outcomes even among individuals without COVID-19. Elective procedures have been canceled or delayed, and access to outpatient clinics has been restricted to preserve hospital beds and intensive care unit capacity during the COVID-19 pandemic. Moreover, an unexpected decline in the number of patients seeking emergency medical care was reported during the early phase of the COVID-19 pandemic. Another study also reported that the COVID-19 pandemic has had a negative, widespread, and persistent impact on ST-elevation myocardial infarction care in United States. Thus, the risk of all-cause mortality among individuals diagnosed with COVID-19 needs to be compared to individuals who were not diagnosed with COVID-19.

Therefore, this study aimed to compare all-cause mortality between individuals in South Korea with COVID-19 and individuals without COVID-19, using PS-matching. We
hypothesized that individuals with COVID-19 might have higher all-cause mortality than individuals without COVID-19, because COVID-19 might increase the risk of death among individuals with COVID-19.

Methods

Study design and ethical statement

This population-based observational study was conducted and reported according to the Reporting of Observational Studies in Epidemiology guidelines\textsuperscript{15}.

Patient Consent Statement

The study protocol was approved by the Institutional Review Board of Seoul National University Bundang Hospital (X-2004-604-905) and the Health Insurance Review and Assessment Service (NHIS-2020-1-291). Informed consent was waived because the data analyses were performed retrospectively using deidentified data derived from the South Korean NHIS database.

NHIS-COVID-19 cohort database and study population

The NHIS-COVID-19 cohort database was developed for medical research purposes in cooperation between the NHIS and the Korea Centers for Disease Control and Prevention (KCDC). The KCDC provides data on patients diagnosed with COVID-19 between January 1, 2020 and June 4, 2020, such as COVID-19 diagnosis confirmation date, treatment results, and demographic information. The COVID-19 patients in the NHIS COVID-19 database
included all patients who were confirmed as positive in the COVID-19 test regardless of hospitalization; therefore, COVID-19 patients who were admitted to the hospital with severe symptoms as well as COVID-19 patients with no or mild symptoms were included in the database. In South Korea, patients who were diagnosed with COVID-19 were admitted to the hospital if they had severe symptoms or conditions such as pneumonia. However, if they had mild or no symptoms, they were isolated and closely monitored in government-managed centers. The COVID-19 patients who are currently undergoing hospital treatment were not included in this database because their treatment outcomes have not yet been determined. Using the data on COVID-19 patients, the NHIS extracted the control population using stratification methods regarding age, sex, and place of residence as of February 2020. The NHIS-COVID-19 cohort database contains disease diagnoses according to the International Classification of Diseases (ICD)-10 codes and prescription information concerning drugs and/or procedures from 2015 to 2020. Finally, the NHIS-COVID-19 database also provides data regarding individuals who had been tested for COVID-19 but were found to be negative. Therefore, the NHIS-COVID-19 database comprised three groups: COVID-19 patients, control population, and test negative individuals. We included all individuals (COVID-19 patients, control population, and test-negative individuals) aged 20 years or older, and excluded those with incomplete medical records. The control population and test-negative individuals were defined as the control group in this study, because a larger population is needed to identify our main outcome robustly, and the PS-matching option served as a method of adjustment to ensure that the characteristics of the two groups (COVID-19 patients and control group) were similar. For this study, an independent medical record technician at the NHIS center, unaffiliated with the study, extracted the data on June 26, 2020.
Exposure variable: Confirmation of COVID-19 diagnosis

The exposure variable in this study was confirmation of COVID-19 diagnosis between January 1, 2020 and June 4, 2020. In South Korea, patients who were diagnosed with COVID-19 were admitted to the hospital if they had severe symptoms such as pneumonia. However, if they had mild or no symptoms, they were isolated and closely monitored in certain government-managed centers.

Endpoints

The primary endpoint of this study was all-cause mortality among all populations in the NHIS-COVID-19 database. It was evaluated from January 1, 2020 to August 27, 2020. All-cause mortality was defined as death due to any reason.

Covariates

The variables extracted as potential confounders included demographic characteristics (age and sex), annual income level during 2020, place of residence (Seoul, Gyeonggi-do, Daegu, Gyeongsangbuk-do, and other areas), the degree of underlying disability in 2020 (mild and moderate to severe), and Charlson Comorbidity Index (CCI), which was calculated based on the registered ICD-10 diagnostic codes (Table S1) from January 1, 2015 to December 31, 2019. Study subjects were categorized into seven groups according to age: 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, and ≥80 years.
Statistical analyses

The baseline characteristics of the participants are reported as frequencies with percentages for categorical variables, and means and their standard deviation for continuous variables. First, we performed propensity score (PS) matching, used to reduce confounders in observational studies, using the nearest neighbor method with a 1:1 ratio, without replacement, and a caliper width of 0.25\textsuperscript{16}. Logistic regression analysis was performed to calculate PSs in a logistic model, and all covariates were included in the propensity score model. The absolute standardized mean difference (ASD) was used to determine the balance between the COVID-19 group and the control group before and after PS matching. ASDs between the two groups were set below 0.1 to determine whether the two groups were well-balanced through PS matching. After confirming adequate balance between the two groups, we performed Cox proportional hazards regression analysis for all-cause mortality in the PS-matched cohort. In this Cox regression analysis, the event was defined as any mortality between January 1, 2020 and August 27, 2020, and survival time was calculated as the time from January 1, 2020 until the date of death or until August 27, 2020 for survivors.

Considering that Cox regression analysis was a time-to-event analysis, we also performed logistic regression analysis for all-cause mortality in the PS-matched cohort, as the primary sensitivity analysis. By performing logistic regression analysis as the primary sensitivity analysis, we determined the odds of all-cause mortality among the COVID-19 group compared to the control group without considering survival time.

As a secondary sensitivity analysis, we fit a multivariable Cox regression model for all-cause mortality for the entire NHIS-COVID-19 cohort in order to: (1) determine whether the results obtained from the PS-matched cohort were generalizable to the entire cohort, and (2) determine the risk of all-cause mortality among the COVID-19 group with other important covariates in context, not in isolation. All covariates were included in the multivariate Cox
model for adjustment, and the CCI and comorbidities that were used to calculate it were included in a different model to avoid multicollinearity. Finally, we performed subgroup analyses according to age and CCI because age and comorbidities were expected to affect mortality in the COVID-19 group\textsuperscript{17}. It was confirmed that there was no multicollinearity in all multivariable models involving the entire cohort, with a variance inflation factor of <2.0. The results of the Cox regression are presented as hazard ratios (HRs) with 95% confidence intervals (CIs), and those of the logistic regression analysis were presented as odds ratios (ORs) with 95% CIs. C-statistics were used to identify the C-index of the multivariable Cox regression model. All statistical analyses were performed using R software (version 3.6.3 with R packages, the R Project for Statistical Computing, Vienna, Austria). \textit{P}<0.05 was considered statistically significant.

Results

Study population

The NHIS-COVID-19 cohort comprised 8,070 individuals diagnosed with COVID-19, 222,257 test negative individuals, and 121,050 individuals in the control population; thus, 351,377 individuals were initially screened. Then, 23,003 individuals were excluded since they were <20 years old. Of the remaining 328,374 adult individuals, 7,713 belonged to the COVID-19 group and 320,660 to the control group. After PS matching, a total of 15,426 individuals (7,713 individuals in each group) were included in the analysis (Fig.1). The results of the comparison of baseline characteristics between the COVID-19 group and control group before and after PS matching are presented in Table 1. All ASDs between the two groups were below 0.1 after PS matching, reflecting adequate balance between the two
groups through PS matching. The distribution of PSs also became similar through PS matching (Fig. S1)

**Survival analysis**

Table 2 shows the results of the survival analysis before and after PS matching. After PS matching, all-cause mortality in the COVID-19 group was 3.2% (248 of 7,713), while that in the control group was 1.6% (126 of 7,713). Cox regression analysis found that the risk of all-cause mortality was twice as high in the COVID-19 group compared with the control group (HR: 2.00; 95% CI: 1.61 to 2.48; \( P<0.001 \)). Logistic regression analysis showed a similar tendency (OR: 1.97; 95% CI: 1.59 to 2.44; \( P<0.001 \)). In the multivariable Cox regression model for all-cause mortality in the entire NHIS-COVID-19 cohort, the COVID-19 group showed a 2.11-fold higher all-cause mortality compared with the control group (HR: 2.11; 95% CI: 1.85 to 2.40; \( P<0.001 \); Table 3). The C-index of the multivariable model was 0.90 (95% CI: 0.89 to 0.90), showing high predictability for all-cause mortality in the multivariable model.

**Subgroup analyses**

Table 4 shows the results of the subgroup analyses according to age and CCI. The COVID-19 group showed a 2.32-fold higher all-cause mortality compared with the control group among individuals aged \( \geq 60 \) years (HR: 2.32; 95% CI: 2.03 to 2.65; \( P<0.001 \)), while the difference in the risk of all-cause mortality was not statistically significant among individuals in the age groups of 20–39 years (\( P=0.339 \)) and 40–59 years old (\( P=0.562 \)) compared to the control group.
Discussion

This population-based cohort study showed that there was a two-fold increase in the risk of all-cause mortality among COVID-19 patients, compared to a PS-matched COVID-19-negative control group. Notably, this higher all-cause mortality is almost entirely attributable to the elevated all-cause mortality in those aged ≥60 years. To our knowledge, this is the first study to report an increase in all-cause mortality associated COVID-19 in direct comparison to non-COVID-19 causes of all-cause mortality.

In South Korea, the annual mortality was reported as 0.57% in 2019 by Statistics Korea (http://kostat.go.kr/portal/eng/index.action). However, the all-cause mortality rate among the control group after PS-matching was 1.6% (126/7,713) until August 27, 2020 in this study. This suggests that the COVID-19 patients were a sick group to begin with, and so the individuals in the PS-matched control group were also in poorer health than people in the general population of South Korea.

Many previous studies reported the factors associated with increased mortality among COVID-19 patients, and the underlying comorbid status was associated with a higher risk of all-cause mortality among COVID-19 patients 18,19. In this study, we adjusted for many confounders, including CCI and comorbidities, using PS matching or multivariable Cox regression modeling to estimate the independent effect of COVID-19 on all-cause mortality in the South Korean population. According to the NHIS-COVID-19 database, the all-cause mortality in among those confirmed as having COVID-19 up until June 4, 2020 was 3.2% (28 of 7,713; Table 2). It was lower than the potential COVID-19-related global mortality of 6%7. Despite the lower all-cause mortality among South Korean COVID-19 patients, COVID-19 increases the risk of all-cause mortality; thus, our results justify the effective prevention of COVID-19 infection in the future by vaccination or by wearing a mask.
The results of subgroup analyses according to age are important in this study. A previous study reported that older age is one of the known risk factors for increased mortality among COVID-19 patients, in addition to other risk factors such as male sex, smoking, and underlying comorbid diseases\(^\text{20}\). Thus, the prevention, isolation, and treatment of COVID-19 infection in elderly individuals has become an important public health issue in the COVID-19 pandemic\(^\text{21}\). In this study, we showed that COVID-19 infection increases all-cause mortality among elderly people, with an HR of 2.16 (95% CI: 1.88 to 2.48).

Although all-cause mortality was not elevated in COVID-19 infection in the younger age group (20–59 years old) compared to PS-matched control younger group, prevention of COVID-19 infection in this group is still important for two reasons. First, since rapid asymptomatic transmission of COVID-19 during the incubation period demonstrated strong infectivity among young COVID-19 patients\(^\text{22}\), older people can be affected by transmission from young people. Second, clinical sequelae are commonly reported among COVID-19 survivors, including young people\(^\text{23}\). Third, there might be some cases who died due to COVID-19 without being diagnosed with COVID-19 among the PS-matched control group.

Although we did not report the hospitalization rate for all individuals in this study, there were a few cases of pre-hospital death due to COVID-19 prior to June 4, 2020 in South Korea. Thus, the impact of hospitalization in the COVID-19 patients and the PS-matched controls might be limited in this study. Therefore, lack of elevated mortality among those aged under 60 years is not the only consideration for future prevention and management.

The impact of CCI on the association between all-cause mortality and COVID-19 infection was also notable in this study because underlying comorbidities are well-known risk factors for increased mortality among COVID-19 patients\(^\text{18,19}\). Our study showed that COVID-19 infection increased the risk of all-cause mortality regardless of comorbid status (HR: 1.75 in the CCI 0–2 group, and HR: 2.16 in the CCI ≥3 group). This suggests that COVID-19 is
associated with a higher risk of all-cause mortality in both healthy and unhealthy adult populations. Therefore, our results suggest that the prevention of COVID-19 infection should be emphasized regardless of comorbid status.

This study has some limitations. First, some important variables, including body mass index and lifestyle factors such as history of smoking and alcohol consumption, were not included in the analysis because they were not available in the NHIS COVID-19 database. Second, both PS matching and multivariable adjustment are known to reduce the number of known and measured confounders. Therefore, there may be some residual confounding that might have affected the study results. Third, to calculate the CCI, we defined comorbidities using ICD-10 codes. However, the diseases specified by the ICD-10 codes may differ from the actual underlying diseases in our study population. Fourth, since the NHIS-COVID-19 database did not provide the cause of death, the proportion of COVID-19-related deaths among COVID-19 patients were not evaluated in this study. Lastly, the all-cause mortality was evaluated from January 1, 2020 to August 27, 2020; therefore, the results may have been different if we had been able to assess mortality over a longer period.

In conclusion, using the NHIS-COVID-19 database, we showed that all-cause mortality among patients with COVID-19 was twice that of those with similar underlying risks, regardless of hospitalization. This higher all-cause mortality almost entirely attributable to the elevated mortality in those aged ≥60 years. Our results highlight the need for prevention of COVID-19 with respect to mortality as a public health outcome.

**Acknowledgement:** None
References

1. Bedford J, Enria D, Giesecke J, et al. COVID-19: towards controlling of a pandemic. 2020;395(10229):1015-1018.

2. COVID C. Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2020.

3. Moore JP, Klasse PJ. COVID-19 Vaccines: "Warp Speed" Needs Mind Melds, Not Warped Minds. J Virol. Aug 17 2020;94(17).

4. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. JAMA. May 12 2020;323(18):1775-1776.

5. Khalili M, Karamouzian M, Nasiri N, Javadi S, Mirzazadeh A, Sharifi H. Epidemiological characteristics of COVID-19: a systematic review and meta-analysis. Epidemiol Infect. Jun 29 2020;148:e130.

6. Mahajan S, Carballo C, Li S-X, et al. SARS-CoV-2 Infection Hospitalization Rate and Infection Fatality Rate among the Non-Congregant Population in Connecticut. 2020.

7. Grech V. Unknown unknowns - COVID-19 and potential global mortality. Early Hum Dev. May 2020;144:105026.

8. Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection. Lancet Infect Dis. Jul 2020;20(7):773.

9. Spinelli A, Pellino G. COVID-19 pandemic: perspectives on an unfolding crisis. Br J Surg. Jun 2020;107(7):785-787.

10. De Simone B, Chouillard E, Di Saverio S, et al. Emergency surgery during the COVID-19 pandemic: what you need to know for practice. Ann R Coll Surg Engl. May 2020;102(5):323-332.

11. Wen X, Li Y. Anesthesia Procedure of Emergency Operation for Patients with...
Suspected or Confirmed COVID-19. *Surg Infect (Larchmt)*. Apr 2020;21(3):299.

12. De Filippo O, D'Ascenzo F, Angelini F, et al. Reduced Rate of Hospital Admissions for ACS during Covid-19 Outbreak in Northern Italy. *N Engl J Med*. Jul 2 2020;383(1):88-89.

13. Kansagra AP, Goyal MS, Hamilton S, Albers GW. Collateral Effect of Covid-19 on Stroke Evaluation in the United States. *N Engl J Med*. Jul 23 2020;383(4):400-401.

14. Garcia S, Stanberry L, Schmidt C, et al. Impact of COVID-19 pandemic on STEMI care: An expanded analysis from the United States. *Catheter Cardiovasc Interv*. Aug 7 2020.

15. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JPJAom. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. 2007;147(8):573-577.

16. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association*. 1984;79(387):516-524.

17. Iaccarino G, Grassi G, Borghi C, et al. Age and Multimorbidity Predict Death Among COVID-19 Patients: Results of the SARS-RAS Study of the Italian Society of Hypertension. *Hypertension*. Aug 2020;76(2):366-372.

18. Tian W, Jiang W, Yao J, et al. Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis. *J Med Virol*. May 22 2020.

19. Zhang JJY, Lee KS, Ang LW, LeoYS, Young BE. Risk Factors of Severe Disease and Efficacy of Treatment in Patients Infected with COVID-19: A Systematic Review, Meta-Analysis and Meta-Regression Analysis. *Clin Infect Dis*. May 14 2020.

20. Zheng Z, Peng F, Xu B, et al. Risk factors of critical & mortal COVID-19 cases: A
systematic literature review and meta-analysis. *J Infect.* Aug 2020;81(2):e16-e25.

21. Daoust JF. Elderly people and responses to COVID-19 in 27 Countries. *PLoS One.* 2020;15(7):e0235590.

22. Huang L, Zhang X, Zhang X, et al. Rapid asymptomatic transmission of COVID-19 during the incubation period demonstrating strong infectivity in a cluster of youngsters aged 16-23 years outside Wuhan and characteristics of young patients with COVID-19: A prospective contact-tracing study. *J Infect.* Jun 2020;80(6):e1-e13.

23. Xiong Q, Xu M, Li J, et al. Clinical sequelae of COVID-19 survivors in Wuhan, China: a single-centre longitudinal study. *Clin Microbiol Infect.* Sep 23 2020.
Figure legend

Fig.1. Flowchart depicting selection process for study subjects
Table 1. Comparison of baseline characteristics between COVID-19 group and control group before and after propensity score matching

| Variable                      | Before propensity score matching | After propensity score matching |
|-------------------------------|---------------------------------|---------------------------------|
|                               | COVID-19 (n=7,713) | Control (n=320,660) | ASD | COVID-19 (n=7,713) | Control (n=7,713) | ASD |
| Sex, male                     | 3048 (39.5) | 142710 (44.5) | 0.102 | 3048 (39.5) | 3084 (40.0) | 0.010 |
| Age, year                     |                  |                  |      |                  |                  |      |
| 20-29                         | 2057 (26.7) | 68361 (21.3) |      | 2057 (26.7) | 2052 (2.7) |      |
| 30-39                         | 832 (10.8) | 50457 (15.7) | 0.160 | 832 (10.8) | 835 (10.8) | 0.002 |
| 40-49                         | 1036 (13.4) | 46901 (14.6) | 0.035 | 1036 (13.4) | 1048 (13.6) | 0.005 |
| 50-59                         | 1567 (20.3) | 52598 (16.4) | 0.097 | 1567 (20.3) | 1552 (20.1) | 0.012 |
| 60-69                         | 1199 (15.5) | 43884 (13.7) | 0.051 | 1199 (15.5) | 1200 (15.6) | 0.007 |
| 70-79                         | 617 (8.0) | 31342 (9.8) | 0.065 | 617 (8.0) | 625 (8.1) | 0.002 |
| 80+                           | 405 (5.3) | 27117 (8.5) | 0.143 | 405 (5.3) | 401 (5.2) | 0.001 |
| Annual income level in 2020   |                  |                  |      |                  |                  |      |
| Q1 (Lowest)                   | 2439 (31.6) | 74362 (23.2) |      | 2439 (31.6%) | 2333 (30.2%) |      |
| Q2                            | 1445 (18.7) | 62782 (19.6) | 0.181 | 1445 (18.7) | 1495 (19.4) | 0.017 |
| Q3                            | 1577 (20.4) | 77265 (24.1) | 0.021 | 1577 (20.4) | 1643 (21.3) | 0.021 |
| Q4 (Highest)                  | 2135 (27.7) | 100814 (31.4) | 0.091 | 2135 (27.7) | 2112 (27.4) | 0.007 |
|                      | 117 (1.5) | 5437 (1.7) | 0.084 | 117 (1.5) | 130 (1.7) | 0.014 |
|----------------------|-----------|------------|-------|-----------|-----------|-------|
| Residence at 2010    |           |            |       |           |           |       |
| Seoul                | 510 (6.6) | 53559 (16.7) | 510 (6.6) | 541 (7.0) |
| Gyeonggi-do          | 431 (5.6) | 57348 (17.9) | 0.535 | 431 (5.6) | 479 (6.2) | 0.027 |
| Daegu                | 5036 (65.3) | 96291 (30.0) | 0.741 | 5036 (65.3) | 4964 (64.4) | 0.020 |
| Gyeongsangbuk-do     | 933 (12.1) | 26338 (8.2) | 0.119 | 933 (12.1) | 960 (12.4) | 0.011 |
| Other area           | 803 (10.4) | 87124 (27.2) | 0.549 | 803 (10.4) | 769 (10.0) | 0.014 |
| Underlying disability|           |            |       |           |           |       |
| Mild degree          | 318 (4.1) | 16790 (5.2) | 0.056 | 318 (4.1) | 288 (3.7) | 0.020 |
| Moderate to severe   | 293 (3.8) | 12073 (3.8) | 0.002 | 293 (3.8) | 277 (3.6) | 0.011 |
| Charlson comorbidity index | 2.7 (2.7) | 3.5 (3.4) | 0.275 | 2.7 (2.7) | 2.8 (2.8) | 0.016 |
| Hypertension         | 1889 (24.5) | 99768 (31.1) | 0.154 | 1889 (24.5) | 1943 (25.2) | 0.016 |
| Myocardial infarction| 222 (2.9) | 12162 (3.8) | 0.055 | 222 (2.9) | 225 (2.9) | 0.002 |
| Congestive heart failure | 503 (6.5) | 36058 (11.2) | 0.191 | 503 (6.5) | 497 (6.4) | 0.003 |
| Peripheral vascular disease | 1314 (17.0) | 64626 (20.2) | 0.083 | 1314 (17.0) | 1353 (17.5) | 0.013 |
| Cerebrovascular disease | 899 (11.7) | 49746 (15.5) | 0.120 | 899 (11.7) | 890 (11.5) | 0.004 |
| Peptic ulcer disease | 3180 (41.2) | 150619 (47.0) | 0.117 | 3180 (41.2) | 3229 (41.9) | 0.013 |
| DM without chronic complication | 1893 (24.5) | 94401 (29.4) | 0.114 | 1893 (24.5) | 1929 (25.0) | 0.011 |
| DM with chronic complication | 591 (7.7) | 32084 (10.0) | 0.088 | 591 (7.7) | 599 (7.8) | 0.004 |
| Condition                      | COVID-19 Cases | Coronavirus Disease-2019 Cases | p-value | COVID-19 Cases | Coronavirus Disease-2019 Cases | p-value |
|-------------------------------|----------------|-------------------------------|---------|----------------|-------------------------------|---------|
| Renal disease                 | 183 (2.4)      | 16134 (5.0)                   | 0.175   | 183 (2.4)      | 185 (2.4)                     | 0.002   |
| Hemiplegia or paraplegia      | 140 (1.8)      | 7095 (2.2)                    | 0.030   | 140 (1.8)      | 150 (1.9)                     | 0.010   |
| Rheumatic disease             | 745 (9.7)      | 35463 (11.1)                  | 0.047   | 745 (9.7)      | 731 (9.5)                     | 0.006   |
| Mild liver disease            | 3376 (43.8)    | 150617 (47.0)                 | 0.065   | 3376 (43.8)    | 3451 (44.7)                   | 0.020   |
| Moderate to severe liver disease | 33 (0.4)    | 3057 (1.0)                    | 0.081   | 33 (0.4)       | 36 (0.5)                      | 0.006   |
| Chronic pulmonary disease     | 3961 (51.4)    | 186747 (58.2)                 | 0.138   | 3961 (51.4)    | 4072 (52.8)                   | 0.029   |
| Any cancer                    | 602 (7.8)      | 48899 (15.2)                  | 0.278   | 602 (7.8)      | 633 (8.2)                     | 0.015   |
| Metastatic solid tumor        | 70 (0.9)       | 9198 (2.9)                    | 0.207   | 70 (0.9)       | 62 (0.8)                      | 0.011   |
| HIV/AIDS                      | 9 (0.1)        | 647 (0.2)                     | 0.025   | 9 (0.1)        | 8 (0.1)                       | 0.004   |

Presented as mean value with standard deviation or number with percentage

COVID-19, coronavirus disease-2019; DM, diabetes mellitus; HIV, human immunodeficiency virus; AIDS, acquired immune deficiency syndrome
Table 2. Mortality from January 1, 2020 to August 27, 2020

| Variable       | Event (%)       | Cox or Logistic regression | P-value |
|----------------|-----------------|----------------------------|---------|
|                |                 | Hazard ratio (95% CI)      |         |
| Before PSM     |                 |                            |         |
| Control        | 11,318 of 320,660 (3.5) | 1                           |         |
| COVID-19       | 248 of 7,713 (3.2)   | 0.90 (0.79, 1.02)           | 0.097   |
| After PSM      |                 |                            |         |
| Control        | 126 of 7,713 (1.6)   | 1                           |         |
| COVID-19       | 248 of 7,713 (3.2)   | 2.00 (1.61, 2.48)           | <0.001  |
| Sensitivity analysis (LR) |                |                            |         |
| Control        | 126 of 7,713 (1.6)   | 1                           |         |
| COVID-19       | 248 of 7,713 (3.2)   | 1.97 (1.59, 2.44)           | <0.001  |

CI, confidence interval; PSM, propensity score matching; COVID-19, coronavirus disease-2019; LR, logistic regression
Table 3. Multivariable Cox regression model for mortality in 2020 among entire NHIS-COVID-19 cohort

| Variable                                      | Multivariable model | P-value |
|-----------------------------------------------|---------------------|---------|
|                                | Hazard ratio (95% CI) |         |
| COVID-19 (vs control)                        | 2.11 (1.85, 2.40)   | <0.001  |
| Age, 10 year increase                        | 2.01 (1.98, 2.05)   | <0.001  |
| Sex, male (vs female)                        | 1.64 (1.58, 1.70)   | <0.001  |
| Annual income level in 2020                  |                     |         |
| Q1 (Lowest)                                  | 1                   |         |
| Q2                                            | 0.96 (0.91, 1.02)   | 0.170   |
| Q3                                            | 0.90 (0.85, 0.95)   | <0.001  |
| Q4 (Highest)                                 | 0.80 (0.77, 0.84)   | <0.001  |
| Unknown                                       | 1.14 (0.98, 1.32)   | 0.091   |
| Residence at 2010                             |                     |         |
| Seoul                                         | 1                   |         |
| Gyeonggi-do                                   | 1.08 (1.03, 1.14)   | 0.004   |
| Daegu                                         | 0.34 (0.32, 0.37)   | <0.001  |
| Gyeongsangbookdo                              | 0.63 (0.58, 0.68)   | <0.001  |
| Other area                                    | 1.04 (0.99, 1.09)   | 0.132   |
| Underlying disability                        |                     |         |
| Mild degree (vs no disability)               | 1.00 (0.95, 1.06)   | 0.927   |
| Moderate to severe (vs no disability)        | 1.33 (1.25, 1.41)   | <0.001  |
| Charlson comorbidity index, 1 point increase (in other model) | 1.10 (1.10, 1.11) | <0.001  |
| Hypertension                                  | 1.12 (1.06, 1.18)   | <0.001  |
| Myocardial infarction                         | 0.98 (0.92, 1.054)  | 0.530   |
| Congestive heart failure                      | 1.19 (1.14, 1.25)   | <0.001  |
| Peripheral vascular disease                   | 0.90 (0.86, 0.93)   | <0.001  |
| Cerebrovascular disease                       | 0.94 (0.90, 0.98)   | 0.008   |
| Condition                              | Odds Ratio (95% CI) | p-Value |
|----------------------------------------|---------------------|---------|
| Peptic ulcer disease                   | 0.89 (0.85, 0.92)   | <0.001 |
| DM without chronic complication        | 1.11 (1.06, 1.16)   | <0.001 |
| DM with chronic complication           | 1.05 (1.01, 1.10)   | 0.033  |
| Renal disease                          | 1.12 (1.07, 1.19)   | <0.001 |
| Hemiplegia or paraplegia              | 1.23 (1.15, 1.32)   | <0.001 |
| Rheumatic disease                      | 0.88 (0.84, 0.93)   | <0.001 |
| Mild liver disease                     | 0.97 (0.93, 1.02)   | 0.238  |
| Moderate to severe liver disease       | 1.76 (1.60, 1.94)   | <0.001 |
| Chronic pulmonary disease              | 0.88 (0.84, 0.92)   | <0.001 |
| Any cancer                             | 1.56 (1.49, 1.63)   | <0.001 |
| Metastatic solid tumor                 | 2.85 (2.70, 3.01)   | <0.001 |
| HIV/AIDS                               | 0.96 (0.68, 1.35)   | 0.821  |

C-index : 0.90 (0.89, 0.90)

CI, confidence interval; COVID-19, coronavirus disease-2019; DM, diabetes mellitus; HIV, human immunodeficiency virus; AIDS, acquired immune deficiency syndrome
Table 4. Subgroup analyses according to age group and Charson comorbidity index

| Variable                                              | Multivariable model | P-value |
|-------------------------------------------------------|---------------------|---------|
|                                                      | Hazard ratio (95% CI) |         |
| Age: 20-39 yrs old (n=121,707, mortality=219)        |                     |         |
| COVID-19 (vs control)                                 | 0.38 (0.05, 2.74)   | 0.339   |
| Age: 40-59 yrs old (n=102,102, mortality=1,362)       |                     |         |
| COVID-19 (vs control)                                 | 1.14 (0.72, 1.81)   | 0.562   |
| Age ≥ 60 yrs old (n=104,564, mortality=9,985)         |                     |         |
| COVID-19 (vs control)                                 | 2.32 (2.03, 2.65)   | <0.001  |
| Charson comorbidity index: 0-2 (n=171,035, mortality=1,297) | |         |
| COVID-19 (vs control)                                 | 1.75 (1.25, 2.45)   | 0.001   |
| Charson comorbidity index ≥ 3 (n=157,338, mortality=10,269) | |         |
| COVID-19 (vs control)                                 | 2.16 (1.88, 2.48)   | <0.001  |

CI, confidence interval; COVID-19, coronavirus disease-2019
Figure 1

COVID-19 n=8,070

Test negative n=222,257

Control population n=121,050

Initially screened n=351,377

Exclude
Age < 20 yrs old: n=23,003

Adult (≥20) n=328,374

COVID-19 n=7,713

Control group n=320,660

1:1 PSM

COVID-19 n=7,713

Control group n=7,713