medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

# Face mask wearing rate predicts country's COVID-19 death rates

Daisuke Miyazawa[1], Gen Kaneko[2]

[1]Miyazawa Clinic, 1-6-5 Akuraminami Takarazuka, Hyogo 665-0823, Japan

[2]School of Arts & Sciences, University of Houston-Victoria. 3007 Ben Wilson, Victoria, TX 77901, United States.

Corresponding author: Daisuke Miyazawa

Email: kusami1@ybb.ne.jp

ORCID (Daisuke Miyazawa): https://orcid.org/0000-0001-9743-1242

ORCID (Gen Kaneko): https://orcid.org/0000-0003-4295-5123

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

**Abstract**

Identification of biomedical and socioeconomic predictors for the number of deaths by COVID-19 among countries will lead to the development of effective intervention. While previous multiple regression studies have identified several predictors, little is known for the effect of mask non-wearing rate on the number of COVID-19-related deaths possibly because the data is available for limited number of countries, which constricts the application of traditional multiple regression approach to screen a large number of potential predictors. In this study, we used the hypothesis-driven regression to test the effect of limited number of predictors based on the hypothesis that the mask non-wearing rate can predict the number of deaths to a large extent together with age and BMI, other relatively independent risk factors for hospitalized patients of COVID-19. The mask non-wearing rate, percentage of age ≥ 80 (male), and male BMI showed Spearman's correlations up to about 0.8, 0.7, and 0.6 with the number of deaths per million from 22 countries from mid-March to mid-June, respectively. The observed number of deaths per million were significantly correlated with the numbers predicted by the lasso regression model including four predictors, age ≥ 80 (male), male BMI, and mask non-wearing rates from mid-March and late April to early May (Pearson's coefficient = 0.918). The multiple linear regression models including the mask non-wearing rates, age, and obesity-related predictors explained up to 79% variation of the number of deaths per million. Furthermore, 56.8% of the variation of mask non-wearing rate in mid-March, the strongest predictor of the number of deaths per million, was predicted by age ≥ 80 (male) and male BMI, suggesting the confounding role of these predictors. Although further verification is needed to identify causes of the national differences in COVID-19 mortality rates, these results highlight the importance of the mask, age, and BMI in predicting the COVID-19-related deaths, providing a useful strategy for future regression analyses that attempt to contribute to the mechanistic understanding of COVID-19.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

## Introduction

There have been considerable differences in the number of deaths by COVID-19 among countries. Western countries (Europe and the United States), in particular, have been suffered from the high number of deaths and the high death rate compared to East Asian countries. Several regression studies have identified predictors, or candidates for true risk factors, for the difference among countries such as age, obesity, and previous BCG vaccination (Jeet et al., 2020; Bellali, Chtioui, & Chahed, 2020; Leung, Bulterys, & Bulterys, 2020; Squalli, 2020; Knittel & Ozaltun, 2020). However, it is important to continue characterizing novel and known predictors under various biomedical and socioeconomic contexts for potential future intervention to counteract with the ongoing COVID-19 pandemic.

Leffler et al. (2020) used the issued dates for each government's mandates or recommendations for face mask wearing as a predictor in their analysis. By their multivariable linear regression, duration of infection in the country, duration of wearing masks, and percentage of the population over age 60 significantly predicted each country's COVID-19 mortality rate. Also, a systematic review and meta-analysis identifying 172 observational studies indicated that face mask use could result in a large reduction in risk of COVID-19 (Chu et al., 2020). These findings, along with the big difference in face mask wearing rates between Western and other East Asian countries, suggest that the importance of mask wearing rate can be a significant predictor for the number of deaths by COVID-19. However, to our knowledge, there has been no analytical studies including the face mask wearing rate in regression models. This is possibly because the common regression approach uses a large number of samples and predictors to achieve the high prediction rates. The mask wearing rate is available only for limited number of countries.

Other studies identified age, BMI, and male sex as provisional and relatively independent risk factors for hospitalized patients of COVID-19 (Palaiodimos et al., 2020). In 5279 patients with laboratory confirmed COVID-19 patients in New York City, age is the strongest risk factor associated with hospital admission, while the critical illness after admission was associated with age, heart failure, BMI >40, and male sex (Petrilli et al., 2020). Docherty et al. (2020) reported using data from 20133 patients admitted to hospital with COVID-19 in the United Kingdom that independent risk factors are increasing age, male sex, and chronic comorbidity, including obesity.

In this study, we employed the hypothesis-driven regression approach to identify predictors for the number of deaths by COVID-19 across countries based on the hypothesis that the mask non-wearing rate can predict the number of deaths to a large extent together with age and BMI, other relatively independent risk factors for hospitalized patients of COVID-19. This approach uses limited number of a priori determined predictors based on a certain hypothesis. The hypothesis-driven regression is especially useful in providing biological insights related to the association with small number of samples, while the implication is restricted to the hypothesis, and has been used in various fields of research (Klein et al., 2013; Sabuncu et al., 2016; Carriedo et al., 2020). Based on the results we discussed the cultural factors for the observed difference among countries, which may have lowered

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

the number of COVID-19-related deaths in East Asian countries.

## Methods

### Data collection

All data was collected from publicly available secondary sources. Analyses in the present study included UK, France, Italy, USA, Spain, Mexico, Germany, Malaysia, Canada, Sweden, China, Saudi Arabia, India, Indonesia, Philippines, Norway, Finland, Japan, Denmark, Australia, Singapore, and Thailand. These countries were chosen because of the availability for mid-March mask-wearing data. The face mask wearing rates in mid-March (3/9-18) 2020 and late April to early May (4/26-5/1) 2020 across countries were derived from "% of people in each country who say they are: Wearing a face mask when in public spaces" in Smith (2020). Total COVID death per million were obtained from Ritchie (2020). BMI data was obtained from Global Status Report on Noncommunicable Diseases 2014. 5 Oct. 2015, www.who.int/nmh/publications/ncd-status-report-2014/en/. Population percent by age data was obtained from Kose et al. (2020).

### Data analysis

R version 3.6.2 was used for all statistical analyses. Pearson's and Spearman's correlations were calculated by cor and cor.test fuctions. Lasso regression was used to find potential effects of predictors on the number of deaths per million (Tibshirani, 2011). Multiple linear regression models were built using the lm function.

## Results

### Spearman's correlations

We first calculated Spearman's correlations between the number of deaths per million from various dates and parameters related to the mask non-wearing rate, age, and obesity (Fig. 1). The correlation coefficients between the number of deaths per million and mask non-wearing rates (both mid-March and late April to early May) showed time-dependent increases, reaching the apparent plateau in early May. The mask non-wearing rate in mid-March generally showed higher correlation coefficients than those that determined in late April to early May. The highest correlation of 0.79 was observed between the mid-March mask non-wearing rate and the number of deaths per million on June 6, 2020 ($P = 6.865e-06$).

Age-related predictors were also highly correlated with the number of deaths per million with Spearman's correlation coefficients of up to ~0.68. In both sexes, correlation coefficients were lower in younger age groups (age 65–69) compared to those of higher age groups, and the highest

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

correlations were found in age ≥ 80 groups. In contrast to the mask non-wearing rate, correlations between the number of deaths and age-related predictors were highest in mid-May and decreased thereafter.

In general BMI showed lower correlations with the number of deaths per million with a clear sex-dependent difference. Male BMI was more closely associated with the number of deaths per million than female BMI.

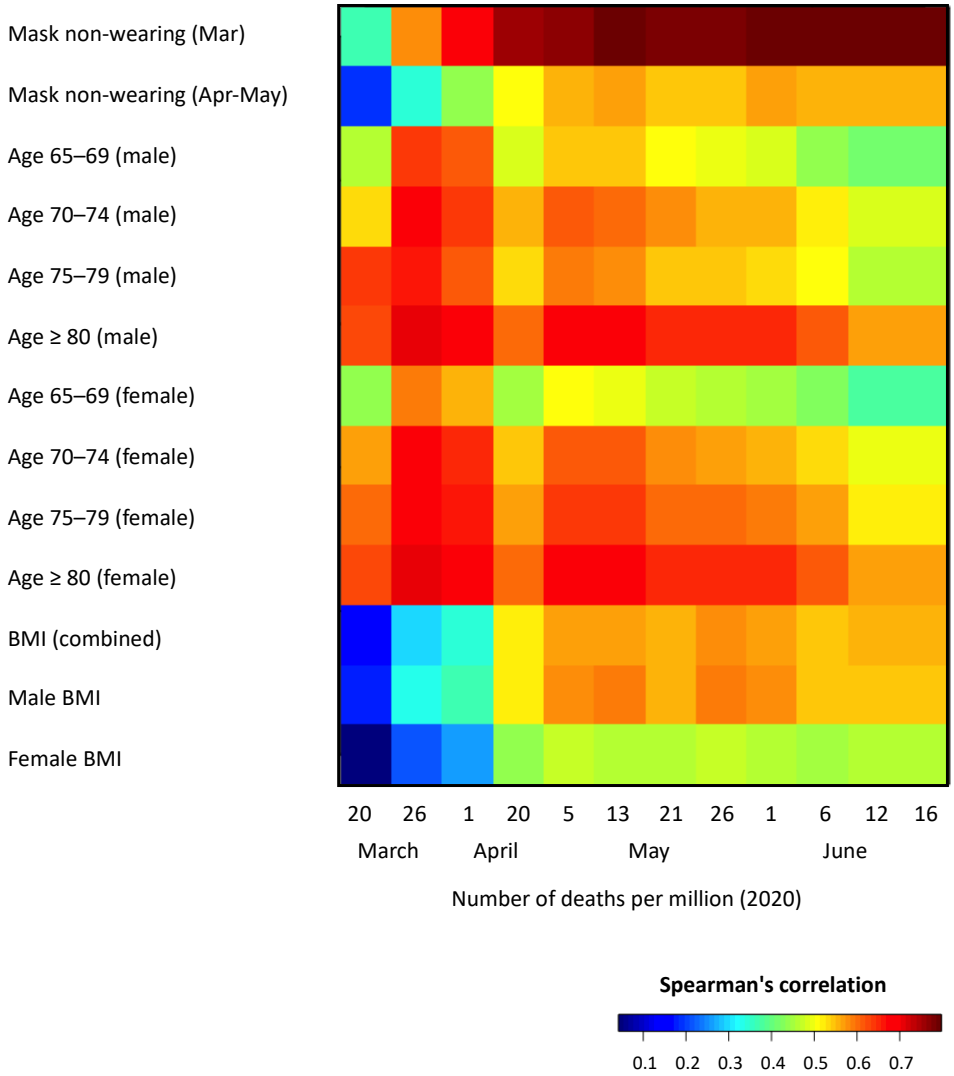


Fig. 1. Spearman's correlation (over time) among parameters related to COVID19 deaths, obesity, age, and mask non-wearing rate.

Time course profiles for Spearman's correlation coefficients of several highly correlated parameters demonstrated the unique feature of age ≥ 80 (male) among parameters used (Fig. 2). Spearman's correlation coefficients for this parameter remained almost constant throughout the study period,

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

whereas those for other three parameters showed time-dependent increase and reached the apparent plateau around day 50 (mid-May).

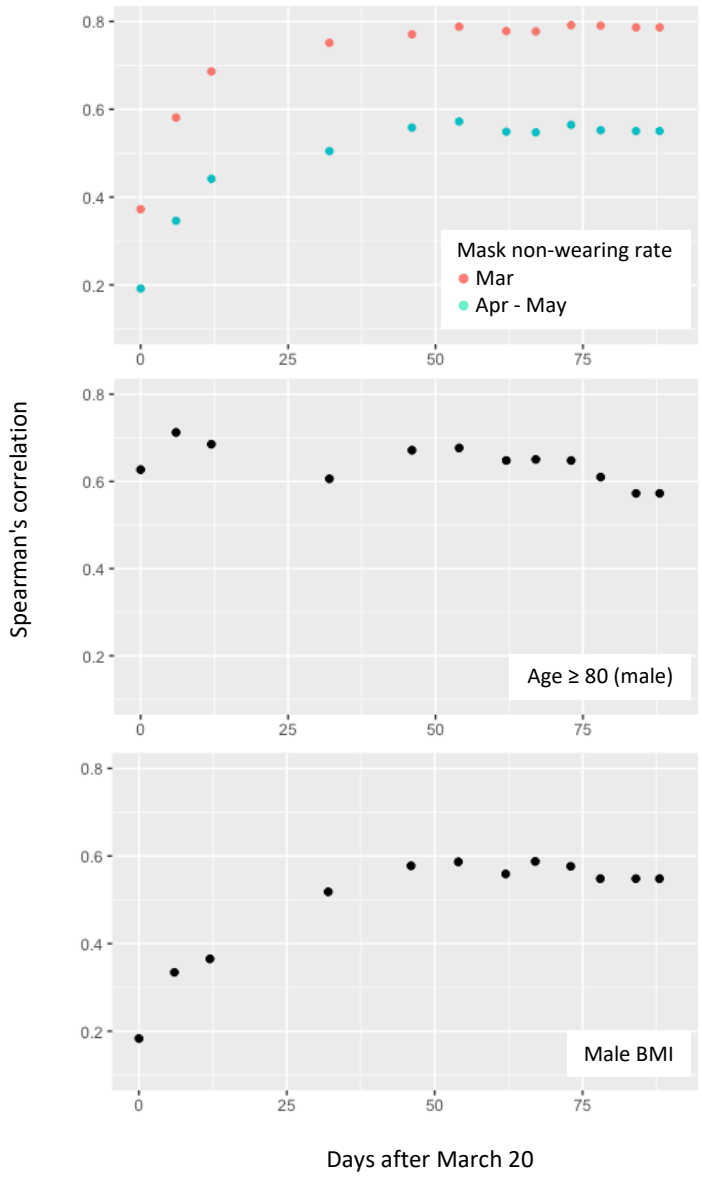


Fig. 2. Time-dependent changes in Spearman's correlations between the number of deaths per million on May 13 and mask non-wearing rates (mid-March and late April to early May), age ≥ 80, and male BMI.

**Scatter plots**

We subsequently made scatter plots showing the relationship between the number of deaths per million and predictors that showed high Spearman's correlation (Fig. 3). The number of deaths per million on May 13 was used because it showed the highest correlations with most important

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

predictors including the mask non-wearing rates (mid-March), age ≥ 80 (male), and male BMI. As expected, the number of deaths per million did not show a clear correlation with the mask non-wearing rate (late April to early May). The number of deaths per million showed exponential type increases in most plots, in particular with the mask non-wearing rate (mid-March).

We therefore applied the logarithmic transformation to the number of deaths per million. The transformed number of deaths per million on May 13 showed significant linear correlation with the mask non-wearing rate in mid-March (r = 0.865, P = 9.739e-08), mask non-wearing rate in late April to early May (r = 0.530, P = 0.009296), age ≥ 80 (male) (r = 0.658, P = 0.0008717), and BMI (male) (r = 0.682, P = 0.0004762). The relationship between the number of deaths per million (May 13) and age ≥ 80 (female) appeared to be similar in the scatter plot (data not shown), but their linear correlation did not reach the statistical significance with a lower Pearson correlation coefficient compared to that of age ≥ 80 of males (r = 0.416, P = 0.05381).

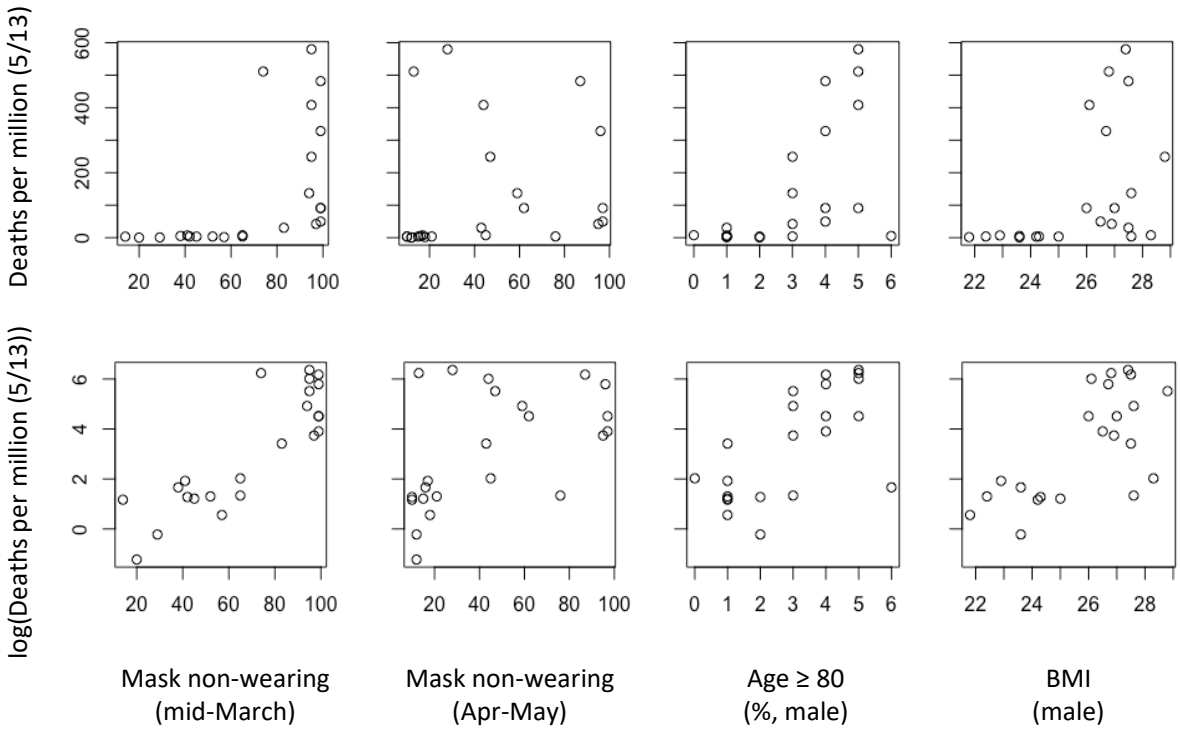


Fig. 3. Scatter plots showing the correlations among parameters related to COVID19 deaths, obesity, age, and mask non-wearing rate. The number of deaths per million on May 13 was used with (bottom panels) and without (top panels) logarithmic transformation.

## Lasso regression

Next, we used the least absolute shrinkage and selection operator (lasso) regression to find the potential effects of these predictors on the log-transformed number of deaths per million. The lasso regression estimates the regression coefficients with the lasso penalty terms, retaining important

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

explanatory variables, while less important variables are removed from the regression (Tibshirani, 2011). When all parameters in Fig. 2 were included together with age ≥ 80 (female), the lasso regression selected BMI (male) (beta = 0.21610188), age ≥ 80 (male) (0.34906818), the mask non-wearing rate (mid-March) (0.06019861), and the mask non-wearing rate in late April to early May (-0.02008978) with an intercept of -6.65720405. Age ≥ 80 (female) was omitted from the regression. The observed number of deaths per million showed high correlation with the number predicted by the lasso regression model (Fig. 4; Pearson's coefficient = 0.918, P = 1.635e-09).

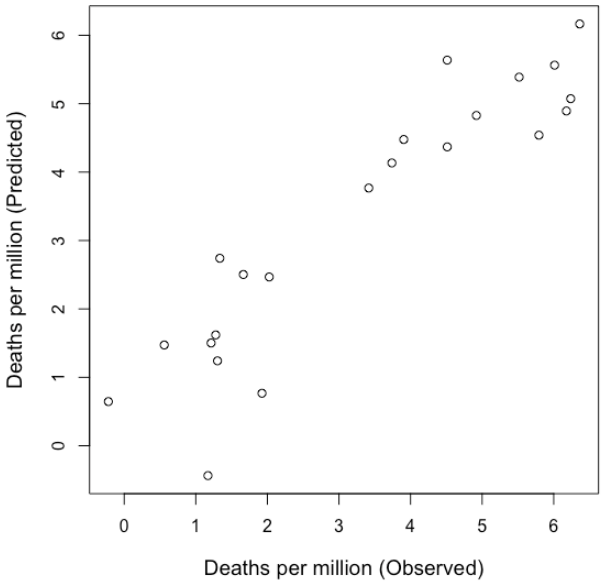


Fig. 4. Pearson's correlation between observed and predicted values of log-transformed number of deaths per million. The minimum lambda of 0.02511886 was used for the lasso regression. Pearson's coefficient = 0.918, P = 1.635e-09.

**Multiple linear regression**

We performed a hypothesis-driven linear regression using three or two of predictors selected by the lasso regression. Because the mask non-wearing rate is available only for 22 countries (Supplementary File 1), we first selected BMI (male), age ≥ 80 (male), and the mask non-wearing rate in mid-March for this regression (Table 1, model 1). This multiple regression model explained 76.0% variation of the number of deaths per million on May 13 (P = 2.075e-06), identifying the significant effects of mask non-wearing rate in mid-March and age ≥ 80 (male). BMI (male) also showed a trend for association, but the association was not statistically significant.

Other combinations of three predictors also resulted in significant associations between the number of deaths per million on May 13 and the selected predictors with ~0.7 adjusted R square values (Table 1). We found several important features in the multiple regressions. First, the effect of mask non-wearing rate in mid-March was significant in all regressions tested with low P values, and exclusion

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

of this predictor markedly decreased the adjusted R square (model 2). Second, BMI (male) showed a relatively low association with the number of deaths on May 13, being significant only when the strongest predictor, mask non-wearing rate in mid-Match, was not included the regression (model 3). Third, the highest adjusted R square of 0.7934 was obtained when we include predictors age ≥ 80 (male) and mask non-wearing rates of mid-March and late April to early May (model 4). All regression coefficients were statistically significant in this model.

Regressions including two predictors also underscored the importance of the mask non-wearing rate in mid-March as the predictor for the number of deaths per million on May 13. Namely, the mask non-wearing rate in mid-March was always a significant predictor (models 5–7) with adjusted R square of up to ~0.75. While BMI (male) and age ≥ 80 (male) were also significantly associated with the number of deaths per million on May 13, exclusion of the mask non-wearing rate in mid-March dropped the adjusted R square to ~0.65.

Table 1. Multiple linear regressions for predicting the number of deaths per million (May 13) using three or two predictors

| Model | Predictors | Beta (standard error) | P value | Adjusted R2 (overall P) |
|---|---|---|---|---|
| 1 | BMI (male) | 0.21212 (0.16368) | 0.21137 | 0.7603 (2.075e-06) |
| | Age ≥ 80 (male) | 0.39160 (0.15572) | 0.02163* | |
| | Mask non-wearing rate (Mar) | 0.04163 (0.01330) | 0.00579** | |
| | Intercept | -6.26822 (3.68145) | 0.10584 | |
| 2 | BMI (male) | 0.539400 (0.175061) | 0.00644** | 0.631 (9.438e-05) |
| | Age ≥ 80 (male) | 0.600767 (0.176575) | 0.00318** | |
| | Mask non-wearing rate (Apr-May) | 0.002544 (0.010958) | 0.81900 | |
| | Intercept | -12.468384 (4.223399) | 0.00853** | |
| 3 | BMI (male) | 0.19276 (0.16627) | 0.2615 | 0.751 (2.908e-06) |
| | Mask non-wearing rate (Mar) | 0.07963 (0.01567) | 7.77e-05*** | |
| | Mask non-wearing rate (Apr-May) | -0.02649 (0.01138) | 0.0318* | |
| | Intercept | -6.13726 (3.74871) | 0.1190 | |
| 4 | Age ≥ 80 (male) | 0.33405 (0.14491) | 0.03328* | 0.7934 (5.529e-07) |
| | Mask non-wearing rate (Mar) | 0.07551 (0.01354) | 2.71e-05*** | |

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

| | | | | |
|---|---|---|---|---|
| | Mask non-wearing rate (Apr-May) | -0.02292 (0.01043) | 0.04124* | |
| | Intercept | -1.99613 (0.62727) | 0.00516** | |
| 5 | Age ≥ 80 (male) | 0.371983 (0.157729) | 0.0292* | 0.7517 |
| | Mask non-wearing rate (Mar) | 0.053365 (0.009917) | 3.42e-05*** | (6.898e-07) |
| | Intercept | -1.570567 (0.654079) | 0.0267* | |
| 6 | BMI (male) | 0.17211 (0.18432) | 0.362153 | 0.6931 |
| | Mask non-wearing rate (Mar) | 0.05656 (0.01347) | 0.00048*** | (5.163e-06) |
| | Intercept | -5.16465 (4.13573) | 0.226911 | |
| 7 | Mask non-wearing rate (Mar) | 0.08907 (0.01351) | 2.6e-06*** | 0.7465 |
| | Mask non-wearing rate (Apr-May) | -0.02579 (0.01147) | 0.0366* | (8.407e-07) |
| | Intercept | -1.86463 (0.69196) | 0.0144* | |
| 8 | BMI (male) | 0.5608 (0.1450) | 0.00104** | 0.6494 |
| | Age ≥ 80 (male) | 0.6091 (0.1685) | 0.00185** | (1.832e-05) |
| | Intercept | -12.9298 (3.6328) | 0.00209** | |

* P < 0.05, ** P < 0.01

**Single regression**

Because the mask non-wearing rate in mid-March showed the strong association with the number of deaths on May 13, we also built several single linear regression models for predicting the number of deaths from other dates using this predictor (Table 2). Overall, changes in adjusted R square showed the similar tendency with those of the Spearman's correlation coefficients (Fig. 1), in which the value gradually increased during March and April, reaching the plateau in mid-May. Overall, the mask non-wearing rate in mid-March explained up to 72% of variations of the number of deaths per million.

Table 2. Linear regressions for predicting the log-transformed number of deaths per million from various dates using the mask non-wearing date in mid-March as the single predictor

| Model | Number of deaths per million | Beta (standard error) P value | Intercept (standard error) P value | Adjusted R2 (overall P) |
|---|---|---|---|---|
| 1 | March 26 | 0.04089 (0.01706) | -2.89183 (1.30653) | 0.1843 |

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

| | | | | |
|---|---|---|---|---|
| | | P = 0.0264* | P = 0.0387* | (0.02641) |
| 2 | April 1 | 0.05149 (0.01513)<br>P = 0.00282** | -2.68139 (1.15891)<br>P = 0.03144* | 0.335<br>(0.002824) |
| 3 | April 20 | 0.06303 (0.01127)<br>P = 1.79e-05*** | -1.84714 (0.86336)<br>P = 0.0449* | 0.5903<br>(1.794e-05) |
| 4 | May 5 | 0.065157 (0.009786)<br>P = 1.76e-06*** | -1.492568 (0.749481)<br>P = 0.0603 | 0.6736<br>(1.755e-06) |
| 5 | May 13 | 0.065560 (0.009377)<br>P = 8.75e-07*** | -1.362001 (0.718190)<br>P = 0.0724 | 0.6951<br>(8.754e-07) |
| 6 | May 21 | 0.066014 (0.00915)<br>P = 5.61e-07*** | -1.290532 (0.701396)<br>P = 0.0807 | 0.7082<br>(5.609e-07) |
| 7 | May 26 | 0.065582 (0.008961)<br>P = 4.48e-07*** | -1.195657 (0.686308)<br>P = 0.0968 | 0.7145<br>(4.484e-07) |
| 8 | June 1 | 0.065597 (0.008839)<br>P = 3.64e-07*** | -1.130428 (0.676922)<br>P = 0.111 | 0.7203<br>(3.641e-07) |
| 9 | June 6 | 0.065558 (0.008759)<br>P = 3.21e-07*** | -1.072095 (0.670825)<br>P = 0.126 | 0.7238<br>(3.21e-07) |
| 10 | June 12 | 0.065304 (0.008685)<br>P = 2.99e-07*** | -1.000452 (0.665144)<br>P = 0.148 | 0.7256<br>(2.995e-07) |
| 11 | June 16 | 0.06495 (0.00866)<br>P = 3.11e-07*** | -0.94215 (0.66326)<br>P = 0.171 | 0.7246<br>(3.111e-07) |

* P < 0.05, ** P < 0.01, *** P < 0.001

**Prediction of mask non-wearing rates**

Lastly, we examined the relationship between the mask non-wearing rate, age ≥ 80 (male), and male BMI (Fig. 5). The mask non-wearing rate in mid-March showed significant correlations with both age ≥ 80 (male) and male BMI. Multiple linear regressions found both age ≥ 80 (male) and male BMI as significant predictors of the mask non-wearing rate with regression coefficients of 5.224 (P = 0.042559) and 8.377 (P = 0.000682), respectively, with an interception of -160.020 (P = 0.006040). This multiple regression model explained 56.8% variation of the mask non-wearing rate in mid-March

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

(P = 0.0001325).

Compared to the mask non-wearing rate in mid-March, that in late April to early May showed a weak association with age ≥ 80 (male) and male BMI (Fig. 5). Accordingly, in the following multiple linear regression analysis, age ≥ 80 (male) and male BMI explained only 29.4% variation of the mask non-wearing rate although the regression was still statistically significant (P = 0.01419), The regression coefficients of men BMI and age ≥ 80 (male) were 8.422 and 3.270, respectively (P = 0.0141 and 0.3776, respectively).

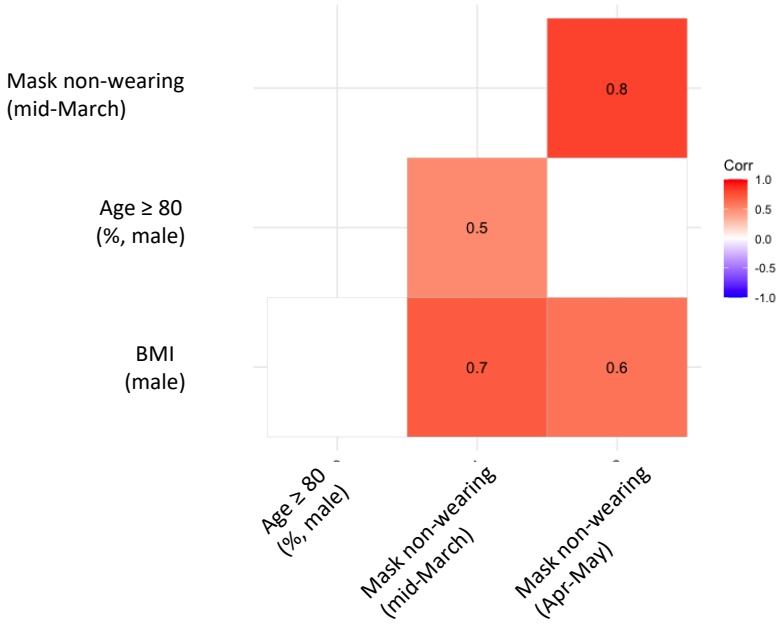


Fig. 5. Pearson's correlation matrix between mask non-wearing rate, age ≥ 80 (male), and mae BMI. Only significant correlations are shown (P < 0.05).

## Discussion

The number of deaths per million on a certain date highly depends on when the epidemic began in the country, in particular in the early stages of the pandemic. In this study, we have tentatively determined March 20 as the start date of a global pandemic. It is likely that this assumption is a reasonable approximation for our hypothesis because all employed predictors (mask non-wearing rates, age and BMI) tended to show weak correlations with the number of deaths per million at the onset of the pandemic followed by marked increases as the progress of the pandemic. These features enabled us to select a specific date (May 13), at which most parameters showed the high correlations with the number of deaths per million, and to build regression models that well explain the variation of the response variable with limited number of predictors. On the other hand, it is noteworthy that the age-related predictors were highly correlated to the number of deaths per million from the very beginning

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

of the pandemic and then remained constant or gradually declined. Below we discuss the potential biomedical and socioeconomic mechanisms by which each predictor affects the number of deaths by COVID-19.

**Obesity**

We have proposed possible mechanisms by which obesity could be a risk factor for the severity of COVID-19 and H1N1 influenza infections (Miyazawa, 2020). Namely, respiratory failure is the most important pathology that contributes to the severity of both COVID-19 and H1N1 influenza infections. Since obese patients generally show a restrictive breathing pattern and reduced lung volumes, the obesity-hypoventilation syndrome can lead to respiratory failure in COVID-19 patients, being a risk factor especially for patients with severe symptoms. Also, obesity has been reported to be a risk factor for the development of acute respiratory distress syndrome (Zhi et al, 2016), which is a serious clinical manifestation of COVID-19 (Simonnet et al., 2020). According to Moriconi et al. (2020), in patients of COVID-19, inflammatory markers were higher in obese group than non-obese group at admission, and obese group showed a worse pulmonary clinical picture, with lower PaO2. Thus, obesity could be a risk factor for already hospitalized COVID-19 patients or even for those with serious symptoms.

These potential mechanisms may explain the reason why the association of BMI (male) and the number of deaths increased May to June 2020. Typical COVID-19 patients are hospitalized a few days to a week after infection and in some cases the symptom become severe a few weeks after hospitalization. If the obesity affects the last phase of the transition of the disease, there must be some delay in the increase in correlations between BMI and the number of deaths per million compared to the spread of COVID-19. However, we cannot rule out the possibility that obese people have high chance to be infected because obese adults are known to inhale air average 50% more per day than non-obese adults (Brochu, 2014), which may lead to increase the chance to inhale the virus.

**Age**

In contrast to BMI, age ≥ 80 (male) was correlated with the number of deaths from the onset of the global pandemic. This may reflect the common etiological feature that old people are more susceptible to infection and have higher risk of death after infection compared to young people because of the dysfunction of immunity.

The precise reason why older age contributes to the number of COVID-19-related deaths remains to be elucidated, while the immune dysfunction has been proposed as a potential mechanism (Mueller, McNamara, & Sinclair, 2020). Meanwhile, it is controversial whether the reduced immunity simply contributes to the higher mortality after infection. Excessive immunity could cause the cytokine storm and ARDS in COVID-19 in some cases (Ye, Wang, & Mao, 2020). Furthermore, two patients with X-

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

linked agammaglobulinemia have recovered from COVID-19, suggesting that B-cell response may not be necessary for the recover from this disease (Soresina et al., 2020). While other dysfunction of immunity, such as age-related T-cell dysfunction, may be involved in the severity of COVID-19 (Minato, Hattori, & Hamazaki, 2020), the mechanism should be carefully investigated.

Along with immune dysfunction, old individuals tend to provoke unwanted inflammation, which may contribute to the severity of COVID-19 (Mueller, McNamara, & Sinclair, 2020). This tendency can be seen in SARS-CoV-infected old nonhuman primate (Smits et al., 2010). Also, To et al. (2020) reported that older age is correlated with the higher salivary viral load, which is highest during the first week after symptom onset. Furthermore, as in the case of obese patients, decreased respiratory function of elderly could also be the reason. In elder individual, both forced expiratory volume in one second (FEV1) and forced vital capacity (FVC) decrease dramatically (Falaschetti et al., 2004), to almost half of their lifetime maximum values, especially in males (Leem et al., 2019).

**Sex differences**

We identified a clear sex-dependent difference in the Spearman's correlation coefficients between the number of deaths per million and obesity-related parameters. This is in line with the fact that males are more susceptible to severe outcomes from COVID-19 than females (Montopoli et al., 2020; Grandi , Facchinetti, & Bitzer, 2020). It has been proposed that estrogen may be protective against COVID-19 in females (Grandi , Facchinetti, & Bitzer, 2020), while androgens may worsen COVID-19 in males (Montopoli et al., 2020). It is also possible that smoking, which is more common in males globally, may play a role in worsening COVID-19 (Grandi , Facchinetti, & Bitzer, 2020). Also, females generally have more fear to COVID-19 and tend to take more preventive behaviors than males (Yıldırım, 2020)

**Face mask**

Face masks are considered to be effective in preventing transmission of COVID-19 to others, and also have a protective effect in preventing transmission from others. A recent simulation study showed that universal masking at 80% adoption suppresses COVID-19 deaths significantly more than maintaining a lockdown. Comparison between the validated modeling results and empirical data from Asian regions showed an almost perfect correlation between early universal masking and successful suppression of COVID-19 outbreaks (Kai et al., 2020). Also, Zhang et al. (2020) reported that face mask alone could have reduced the number of infections by over 78,000 in Italy from April 6 to May 9 and over 66,000 in New York City from April 17 to May 9, claiming that the failure in containing the propagation of COVID-19 pandemic worldwide is largely attributed to the unrecognized importance of airborne virus transmission.

Likewise, the mask non-wearing rate in mid-March was found to be the strongest predictor for the

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

number of deaths per million in this study, even the single regression by which explained more than 70% of the variation of the response variable.

There was a big difference in face mask wearing rates between Western countries and Asian countries (especially East Asian countries). The county's policy for wearing face mask alone cannot explain this big difference because, for example, face mask has never been mandated in Japan, despite its high face mask wearing rate. We speculate that the cultural may be the major reason for the difference. Many Japanese wear surgical masks on a daily basis not for shedding infections or pollens, but also for achieving anonymity just like westerners wearing sunglasses, which is referred to as "mask dependency" in excessive cases (Li, 2017). However, while people may hope to achieve anonymity, most of them want to avoid making others uncomfortable, and there is a regional difference in what people feel uncomfortable. Jack, Caldara, & Schyns (2012) mentioned that "whereas Western Caucasian internal representations predominantly featured the eyebrows and mouth, East Asian internal representations showed a preference for expressive information in the eye region" . This tendency may be the major reason why it is more considered rude wearing sunglasses among East Asia (Gesteland, 2020), whereas it is considered more suspicious to wear face masks in Western countries. Yamanaka (2020) wonders if Japan has an "X-factor" that led to the low rates of COVID-19 deaths. Although the present study is not intended to show the causal role of face mask wearing rate, future attempts for intervention may consider the face mask wearing rate as the major candidate for the "X-factor".

Additionally, most Japanese keep silent while using public transportation, because loud chatter in the public transportation is considered rude in Japan (Baseel, 2020). This may also be able to reduce the case of COVID-19 because more aerosols are exhaled from asymptomatic individuals during speaking than breathing (Buonanno, Stabile, & Morawska, 2020) and it is considered to contribute largely to spread of COVID-19 (Prather, Wang, & Schooley, 2020). Also aerosols' viral density should be different between speaking and breathing, because origin of the aerosols should be different, mainly from lung or mouth. Aerosols from mouth than lung may contain more virus in asymptomatic individuals. Long et. al (2020) noted that although virus shedding does not equate with viral infectivity, the asymptomatic group had a significantly longer duration of viral shedding than the symptomatic group.

**Correlation between variables**

There was a significant positive correlation between male BMI and mask non-wearing rate in mid-March. Although this does not simply indicate that obese people tend not to wear masks, it is possible to speculate that when people become more obese, they feel more uncomfortable to wear masks since obese adult inhale air average 50% more per day than non-obese adults (Brochu, 2014). Also, body temperature is positively associated with obesity (Bastardot, 2019), and face masks could rise body temperature (Yip, 2005; Hayashi 2004). Therefore, obese people may feel more heat and discomfort

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

(Li, 2005) when wearing face masks depending on the temperature and humidity. The small size of universal face masks may be a simple reason for the correlation.

There was a significant positive correlation between age ≥ 80 (male) and mask non-wearing rate in mid-March. Due to the lack of data on mask wearing rates by age, this cannot be fully discussed either. It has been reported that age was correlated with fear and were related to preventive behaviors (Yıldırım, 2020). Face mask wearing in public correlates with the proportion of people who are afraid of COVID-19, except for Scandinavian countries and United Kingdom (Smith, 2020). These facts are inconsistent with the correlation found in this study. Further data is required for detailed investigation.

## Conclusion

The mask non-wearing rate in mid-March alone explained up to 72% of variations of the number of deaths per million. The fact that the face mask non-wearing rate in mid-March, the early phase of the pandemic, more strongly predicted the total number of deaths per population over the entire three-month course of the pandemic than the face mask non-wearing rate in late April to early May may suggest that face mask wearing from the early phase of the pandemic may be very important for suppressing COVID-19-related deaths, if mask wearing is assumed to be an independent risk factor.

In the lasso regression, four factors, BMI (male), age ≥ 80 (male), the mask non-wearing rate (mid-March), and the mask non-wearing rate in late April to early May were chosen, and the observed and predicted model of number of deaths per million showed high correlation (Pearson's coefficient = 0.918, P = 1.635e-09). Although this has not been proven at this time to be the cause of the differences in COVID-19 mortality by country, it may be a useful method for predicting COVID-19 mortality.

The death rate of the country in the early stages of a pandemic could be further predicted if factors such as the timing and number of infected people entering the country from abroad are taken into account.

## References


- Baseel C. 15 Rude Things Not to Do on Trains in Japan. 9 Sept. 2019, japantoday.com https://japantoday.com/category/features/lifestyle/15-rude-things-not-to-do-on-trains-in-japan
- Bastardot F, Marques-Vidal P, Vollenweider P. Association of body temperature with obesity. The CoLaus study. Int J Obes (Lond). 2019;43(5):1026-1033. doi:10.1038/s41366-018-0218-7
- Bellali H, Chtioui N, Chahed M. Factors associated with country-variation in COVID-19 morbidity and mortality worldwide: an observational geographic study medRxiv 2020.05.27.20114280; doi: https://doi.org/10.1101/2020.05.27.20114280
- Brochu P, Bouchard M, Haddad S. Physiological daily inhalation rates for health risk assessment

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

in overweight/obese children, adults, and elderly. Risk Anal. 2014;34(3):567-582. doi:10.1111/risa.12125

- Buonanno G, Stabile L, Morawska L. Estimation of airborne viral emission: Quanta emission rate of SARS-CoV-2 for infection risk assessment. Environ Int. 2020;141:105794. doi:10.1016/j.envint.2020.105794
- Carriedo A, Cecchini JA, Gonzalez C. (2020) Soccer spectators' moral functioning and aggressive tendencies in life and when watching soccer matches. Ethics & Behav published online. doi:10.17605/OSF.IO/TPA6U.
- Chu DK, Akl EA, Duda S, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis [published online ahead of print, 2020 Jun 1]. Lancet. 2020;S0140-6736(20)31142-9. doi:10.1016/S0140-6736(20)31142-9
- Docherty AB, Harrison EM, Green CA, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ*. 2020;369:m1985. Published 2020 May 22. doi:10.1136/bmj.m1985
- Falaschetti E, Laiho J, Primatesta P, Purdon S. Prediction equations for normal and low lung function from the Health Survey for England. Eur Respir J. 2004;23(3):456-463. doi:10.1183/09031936.04.00055204
- Gesteland R, Cross-cultural Business Behavior: A Guide for Global Management (Fifth Edition), Copenhagen buissiness school press, Norway 2012 p85
- Grandi G, Facchinetti F & Bitzer J (2020) The gendered impact of coronavirus disease (COVID-19): do estrogens play a role?, The European Journal of Contraception & Reproductive Health Care, DOI: 10.1080/13625187.2020.1766017
- Hartman WR, Hess AS, Connor J. Persistent viral RNA shedding after COVID-19 symptom resolution in older convalescent plasma donors [published online ahead of print, 2020 Jun 13]. Transfusion. 2020;10.1111/trf.15927. doi:10.1111/trf.15927
- Hayashi C, Tokura H. The effects of two kinds of mask (with or without exhaust valve) on clothing microclimates inside the mask in participants wearing protective clothing for spraying pesticides. Int Arch Occup Environ Health. 2004;77(1):73-78. doi:10.1007/s00420-003-0472-3
- Jack RE, Caldara R, Schyns PG. Internal representations reveal cultural diversity in expectations of facial expressions of emotion. J Exp Psychol Gen. 2012;141(1):19-25. doi:10.1037/a0023463
- Jeet G, Sharma A, Goyal A, et al. Determinants of COVID-19 incidence and mortality: A cross-country analysis
medRxiv 2020.06.11.20128561; doi: https://doi.org/10.1101/2020.06.11.20128561

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .


- Kai D, Goldstein G, Morgunov A, et al. Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations, submitted 22 Apr 2020 and released 26 Apr 2020. DOI: 10.13140/RG.2.2.21662.08001. arXiv:2004.13553
- Klein NJ, Beck AJ, Hutchins DA, Sanudo-Wilhelmy SA (2013) Regression modeling of the North East Atlantic Spring Bloom suggests previously unrecognized biological roles for V and Mo. Front Microbiol 4, 45.
- Knittel C, Ozaltun B. What does and does not correlate with COVID-19 death rates medRxiv 2020.06.09.20126805; doi: https://doi.org/10.1101/2020.06.09.20126805
- Kose, Ayhan, et al. "World Bank Open Data." Data, 15 June 2020, data.worldbank.org/.
- Leem, A.Y., Park, B., Kim, Y.S. et al. Longitudinal decline in lung function: a community-based cohort study in Korea. Sci Rep 9, 13614 (2019). https://doi.org/10.1038/s41598-019-49598-9
- Leffler CT, Ing EB, Lykins V JD, et al. Association of country-wide coronavirus mortality with demographics, testing, lockdowns, and public wearing of masks. Update June 2, 2020. medRxiv 2020.05.22.20109231; doi: https://doi.org/10.1101/2020.05.22.20109231
- Leung NY, Bulterys MA, Bulterys PL. Predictors of COVID-19 incidence, mortality, and epidemic growth rate at the country level medRxiv 2020.05.15.20101097; doi: https://doi.org/10.1101/2020.05.15.20101097
- Li T. "Mask Appeal: The Addiction of Surgical Masks in Japan." The Straits Times, Singapore Press Holdings Ltd. Co, 7 Apr. 2017, www.straitstimes.com/asia/east-asia/mask-appeal.
- Li Y, Tokura H, Guo YP, et al. Effects of wearing N95 and surgical facemasks on heart rate, thermal stress and subjective sensations. Int Arch Occup Environ Health. 2005;78(6):501-509. doi:10.1007/s00420-004-0584-4
- Long QX, Liu BZ, Deng HJ, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. Nat Med. 2020;26(6):845-848. doi:10.1038/s41591-020-0897-1

  May 16]. Metabolism. 2020;108:154262. doi:10.1016/j.metabol.2020.154262
- Long QX, Tang XJ, Shi QL, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections [published online ahead of print, 2020 Jun 18]. *Nat Med*. 2020;10.1038/s41591-020-0965-6. doi:10.1038/s41591-020-0965-6
- Minato N, Hattori M, Hamazaki Y. Physiology and pathology of T-cell aging. Int Immunol. 2020;32(4):223-231. doi:10.1093/intimm/dxaa006
- Miyazawa D. Why obesity, hypertension, diabetes, and ethnicities are common risk factors for COVID-19 and H1N1 influenza infections. Authorea. June 10, 2020. DOI: 10.22541/au.159181185.59632853

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .


- Montopoli M, Zumerle S, Vettor R, et al. Androgen-deprivation therapies for prostate cancer and risk of infection by SARS-CoV-2: a population-based study (n=4532) [published online ahead of print, 2020 May 4]. Ann Oncol. 2020;S0923-7534(20)39797-0. doi:10.1016/j.annonc.2020.04.479
- Moriconi D, Masi S, Rebelos E, et al. Obesity prolongs the hospital stay in patients affected by COVID-19, and may impact on SARS-COV-2 shedding [published online ahead of print, 2020 Jun 4]. Obes Res Clin Pract. 2020;S1871-403X(20)30401-4. doi:10.1016/j.orcp.2020.05.009
- Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older people?. Aging (Albany NY). 2020;12(10):9959-9981. doi:10.18632/aging.103344
- Palaiodimos L, Kokkinidis DG, Li W, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York [published online ahead of print, 2020
- Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ*. 2020;369:m1966. Published 2020 May 22. doi:10.1136/bmj.m1966
- Prather KA, Wang CC, Schooley RT. Reducing transmission of SARS-CoV-2 [published online ahead of print, 2020 May 27]. Science. 2020;eabc6197. doi:10.1126/science.abc6197
- Ritchie, Hannah. Coronavirus Source Data. 16 June 2020, ourworldindata.org/coronavirus-source-data
- Sabuncu MR, Ge T, Holmes AJ, Smoller JW, Buckner RL, Fischl B, the Alzheimer's Disease Neuroimaging Initiative. (2016) Morphometricity as a measure of the neuroanatomical signature of a trait. Proceedings of the National Academy of Sciences, 113, E5749-5756.
- Simonnet A, Chetboun M, Poissy J, et al. High prevalence of obesity in severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) requiring invasive mechanical ventilation [published online April 9, 2020]. Obesity doi:10.1002/oby.22831
- Smith M. International COVID-19 Tracker Update: 2 May. 2 May 2020, yougov.co.uk/topics/international/articles-reports/2020/05/01/international-covid-19-tracker-update-2-may
- Smits SL, de Lang A, van den Brand JMA, Leijten LM, van IJcken WF, Eijkemans MJC, et al. (2010) Exacerbated Innate Host Response to SARS-CoV in Aged Non-Human Primates. PLoS Pathog 6(2): e1000756. https://doi.org/10.1371/journal.ppat.1000756
- Soresina A, Moratto D, Chiarini M, et al. Two X-linked agammaglobulinemia patients develop pneumonia as COVID-19 manifestation but recover [published online ahead of print, 2020 Apr 22]. Pediatr Allergy Immunol. 2020;10.1111/pai.13263. doi:10.1111/pai.13263

medRxiv preprint doi: https://doi.org/10.1101/2020.06.22.20137745; this version posted June 23, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .


- Squalli J. Evaluating the determinants of COVID-19 mortality: A cross-country study
- Tibshirani R (2011) Regression shrinkage and selection via the lasso: a retrospective. J R Stat Soc 73(3):273–282
- To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis. 2020;20(5):565-574. doi:10.1016/S1473-3099(20)30196-1
- Yamanaka S. "Uncovering Japan's Coronavirus X-Factor Matters to the World." Nikkei Asian Review, Nikkei Inc. , 10 June 2020, https://asia.nikkei.com/Business/Science/Yamanaka-on-COVID-19/Uncovering-Japan-s-coronavirus-X-factor-matters-to-the-world
- Ye Q, Wang B, Mao J. The pathogenesis and treatment of the `cytokine storm' in COVID-19. J Infect 2020; (Epub ahead of print). https://doi.org/10.1016/j.jinf.2020.03.037
- Yip, W.L., Leung, L.P., Lau, P.F., & Tong, H.K. (2005). The Effect of Wearing a Face Mask on Body Temperature. Hong Kong Journal of Emergency Medicine, 12, 23 - 27.
- Yıldırım M, Geçer E, Akgül Ö. The impacts of vulnerability, perceived risk, and fear on preventive behaviours against COVID-19 [published online ahead of print, 2020 Jun 3]. Psychol Health Med. 2020;1-9. doi:10.1080/13548506.2020.1776891
- Zhang R, Li Y, Zhang AL, Wang Y, Molina MJ. Identifying airborne transmission as the dominant route for the spread of COVID-19 [published online ahead of print, 2020 Jun 11]. *Proc Natl Acad Sci U S A*. 2020;202009637. doi:10.1073/pnas.2009637117
- Zhi G, Xin W, Ying W, Guohong X, Shuying L. "Obesity Paradox" in acute respiratory distress syndrome: asystematic review and meta-analysis. PLoS One 2016;11(9):e0163677.


medRxiv 2020.05.12.20099093; doi: https://doi.org/10.1101/2020.05.12.20099093