COVID19

Luis Zapata

3/19/2020

Abstract

A novel coronavirs SARS-CoV-2 has impacted the world creating a global pandemic. Up to date (March 19th 2020), 242713 confirmed cases with 9843 fatalities has been reported, bringing the global case fatality rate (CFR) around 4%. Multiple reports have focused on modelling the future of the viral infection, the source of origin, or the genetic differences between the strains. However, to our knowledge no study has focused on understanding the variable case fatality rate between countries. As an example, Italy has reached a CFR of 7% while Germany has a CFR of 0.2%. This widespread difference is striking and here we explore the likely causes for this heterogenety. In our first analysis, we show that the size of the population, the percentage of above 65 adults respect to the total population, and the number of hospital beds per people are the most predictive variables. This study sheds light into how large countries are at a higher risk when the majority of the population is old and there is a limited number of hospital beds.

Introduction

- Numbers globally about COVID19 (https://coronavirus.jhu.edu/map.html, https://www.who.int/ through https://ourworldindata.org/coronavirus)
- The origin of the coronavirus (https://www.nature.com/articles/s41591-020-0820-9)
- Study focused on the persistence of the virus in different surfaces (https://www.nejm.org/doi/10.1056/NEJMc2004973)
- Model to predict outcome and policy-making strategies (https://www.nature.com/articles/nature04795)
- Lessons from Italy (https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30627-9/fulltext#seccestitle30)
- Candidate targets for immunity (https://www.cell.com/cell-host-microbe/fulltext/S1931-3128(20)30166-9)

Results

Analysis of COVID-19 Case fatality rate probable causes across countries

To explore the possible causes behind the different case fatality rates (CFR) observed across countries we performed multiple regression analysis. We downloaded datasets from ourworldindata.org (https://ourworldindata.org/coronavirus), a specilized webiste that put together data from the World Health Organization daily. For the number of hospital beds per country we obtained the data from the API of the worldbank.org on the 15th of March (http://api.worldbank.org/v2/en/indicator/SH.MED.BEDS.ZS?downloadformat=csv).

We first explore different variables that might affect CFR independently. We applied a univariate linear regression of each variables using countries where there are more than 150 cases reported and more than 1 death.

Characteristic	**N = 32**
freqdeath	0.014 (0.008, 0.031)
total_cases	994 (346, 4307)
deaths	10 (4, 66)
Population	37753000 (10499000, 86104250)
logPop	7.58 (7.02, 7.93)
PopDensity	117 (69, 233)
Old	0.17 (0.07, 0.20)
Unknown	3
loga65	6.58 (6.22, 7.11)
Unknown	3
Vegetable	262 (214, 364)
Meat	74 (52, 84)
GDP	37272 (18228, 45748)
GDPc	0.001 (0.000, 0.003)
Beds	3.84 (2.96, 5.87)
Unknown	9
TotalTestsPM	6.26 (5.06, 7.20)
Unknown	16

We observed that the CFR varies widely across countries, the median and mean CFR among the 32 countries with more than 150 cases and one death was 1.4% and 2.3%, respectively (Table 1). On the extremes, we observed a 7% CFR in Italy versus a 0.2% CFR in Germany. As time progresses these number might fluctuate. We first checked the relationship between the number of cases reported and the CFR in each country. Countries with low number of cases have a less reliable CFR that seems to go down once more cases are detected and then countries with large number of cases have an increasing CFR (Fig 1A). We then observed that CFR was also correlated to the population size of the country, but several exceptions were prominent (Fig 1B). On one hand Italy, Iran, Iraq and the Philipines had more deaths than expected, on the other Germany and Malaysia have fewer.

Fig 1A CFR versus Total Cases

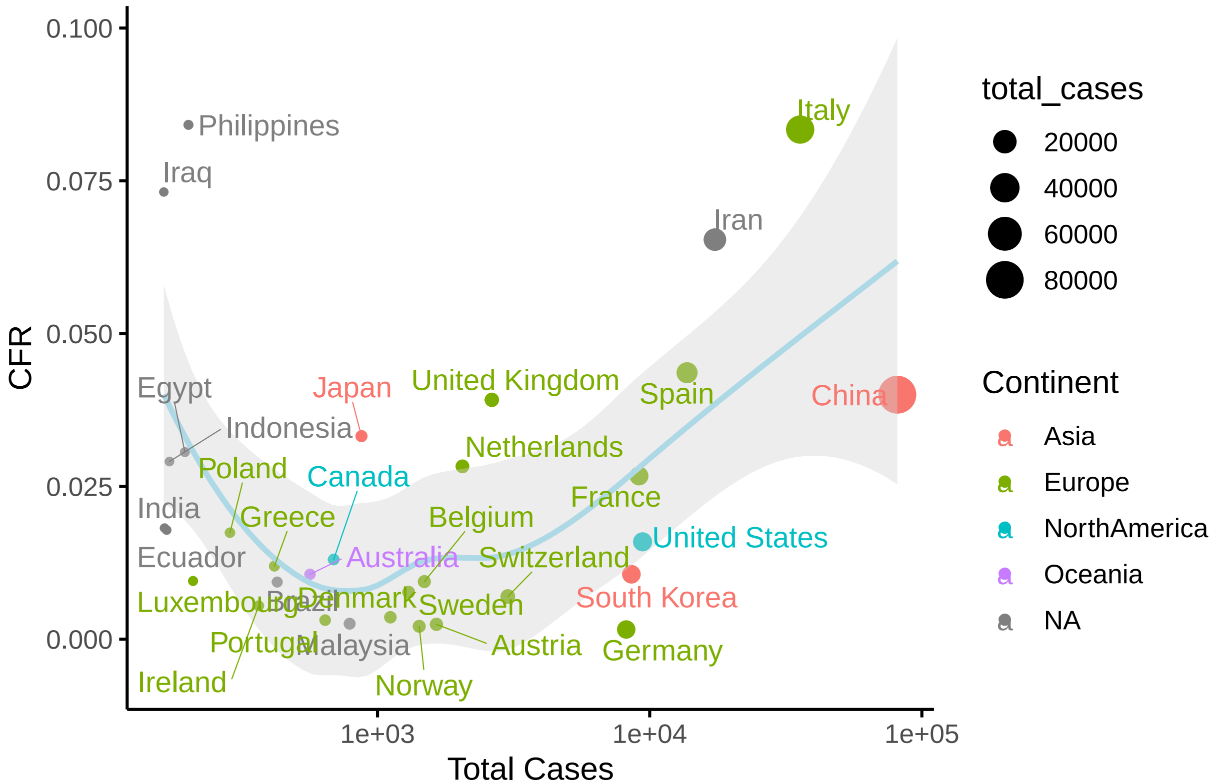
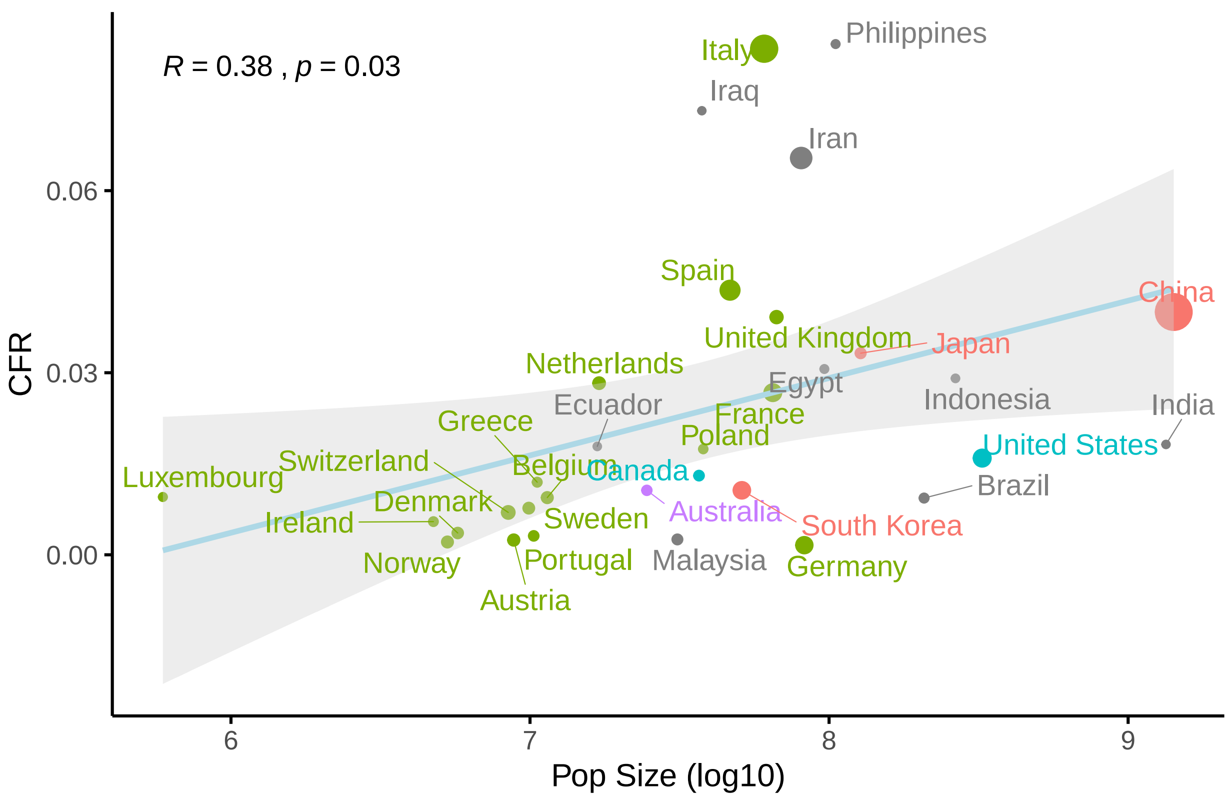
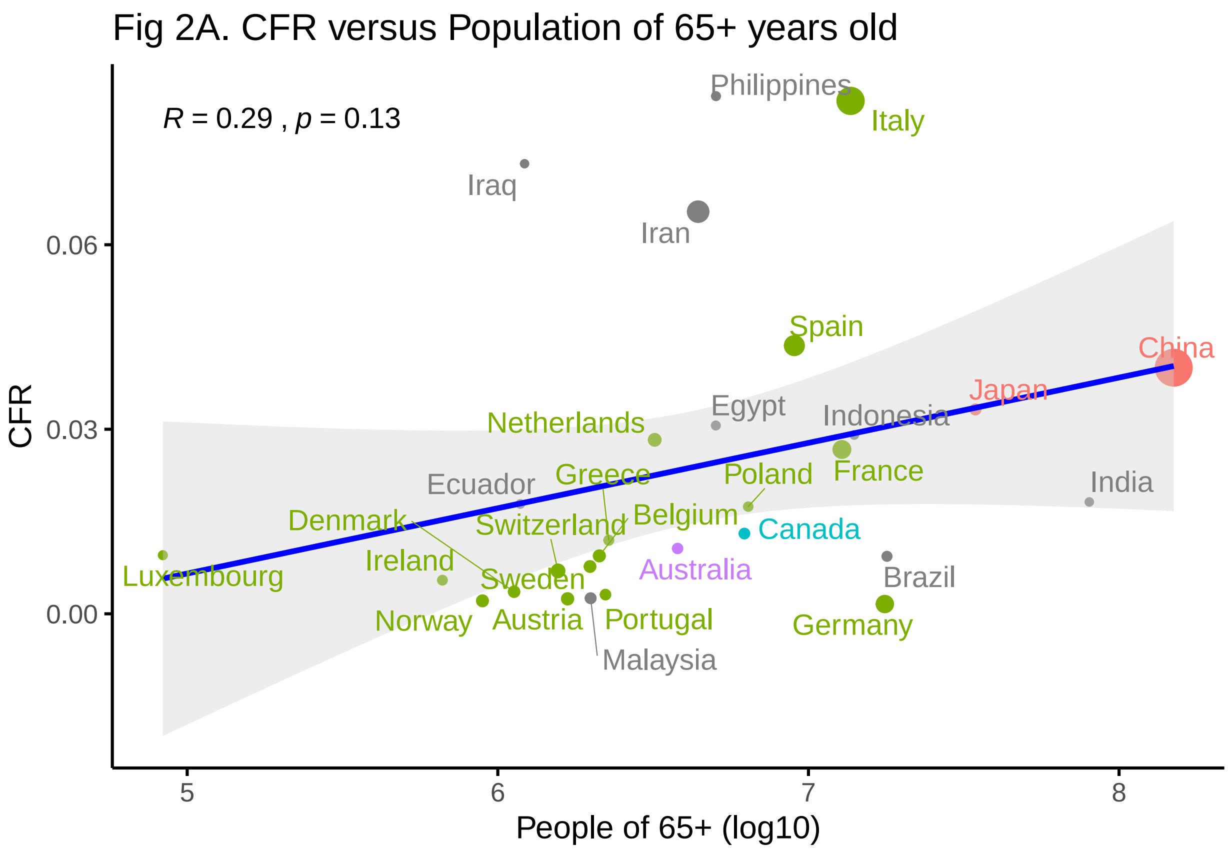
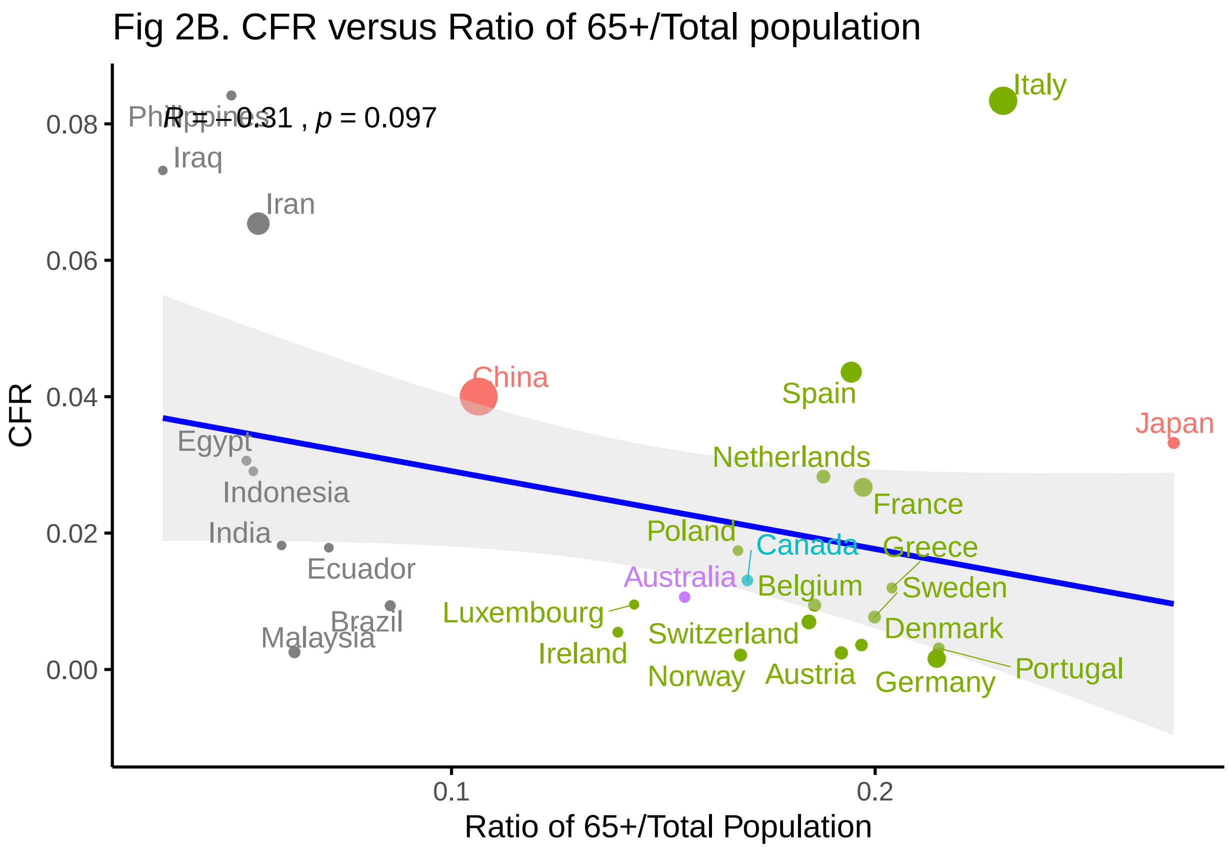


Fig 1B CFR versus Population size



Next, as reported previously that the older than 65 people (+65) was at more danger, we hypothesized that the proportion of elderly had an infleunce on the reported CFR. Although we did not observe a significant correlation either for the total number of +65 adults (Fig 2A) or the ratio of +65 respect to the total population (Fig 2B), the trend was as expected.





Another interesting variable to look was the GDP per country. Here we found that both total GDP and GDP per capita were significantly negative correlated (Fig 3C and 3D), which is expected given the higher capacity of expenditure of rich countries to increase their helth system when needed. Again the exception were Italy, Iran, Iraq and Phillipines.

Fig 3A. CFR versus GDP

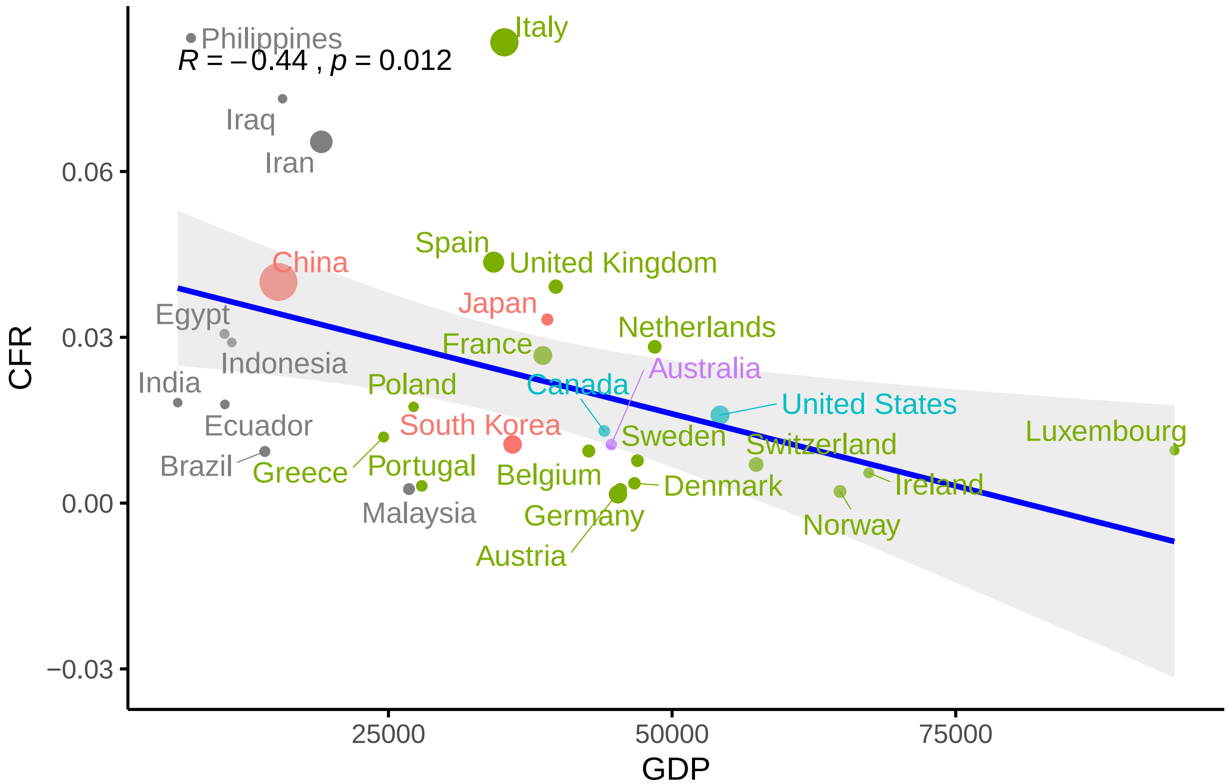
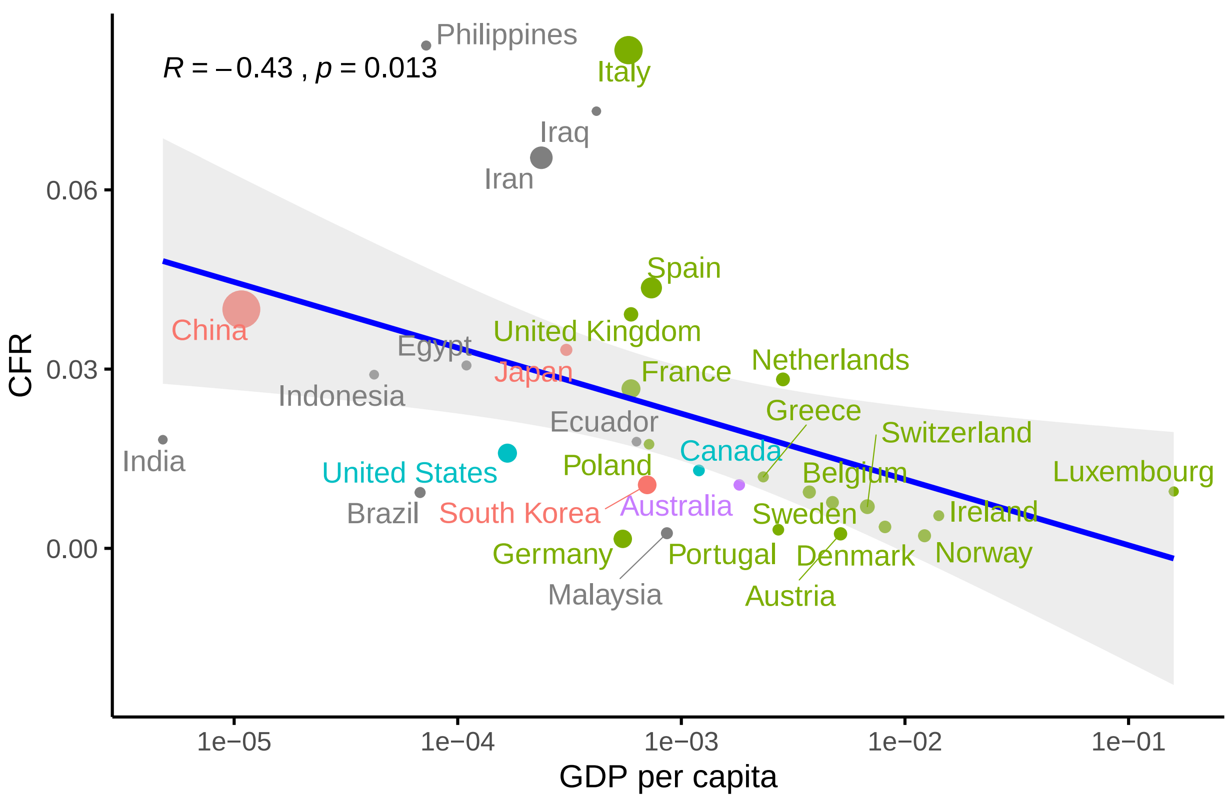
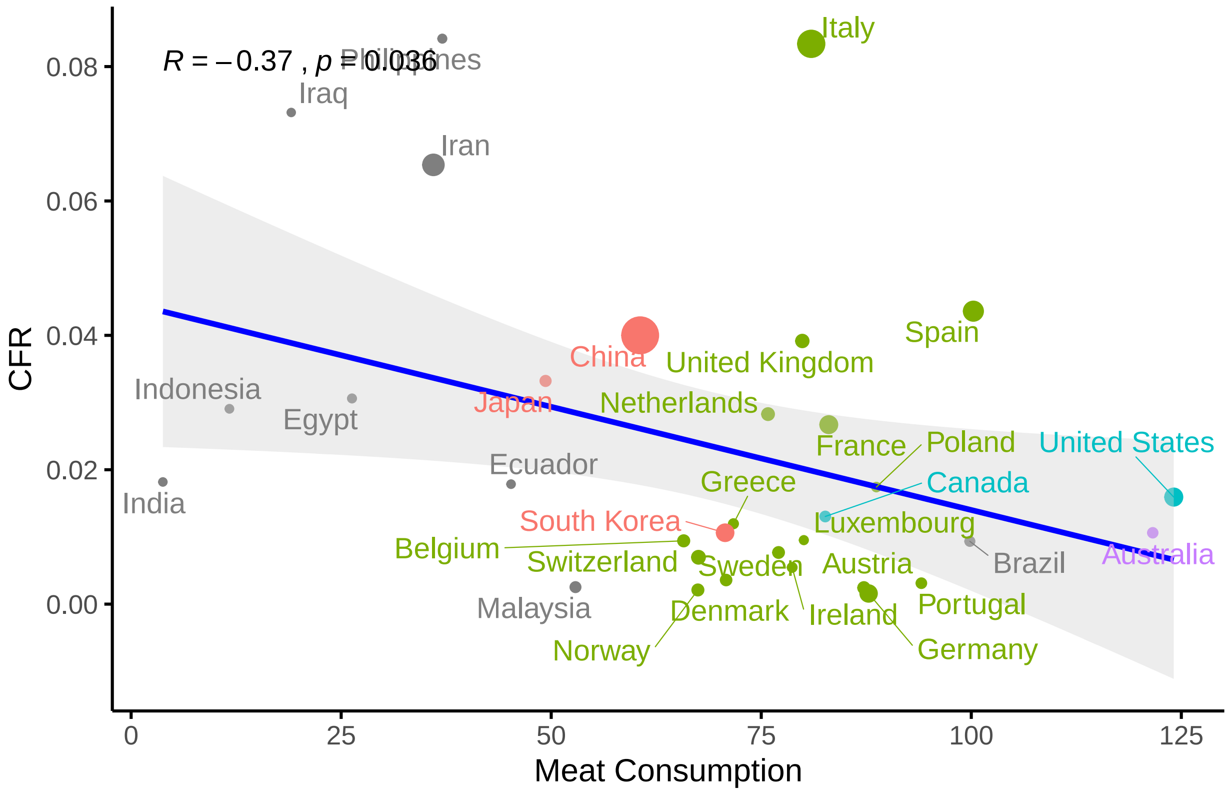


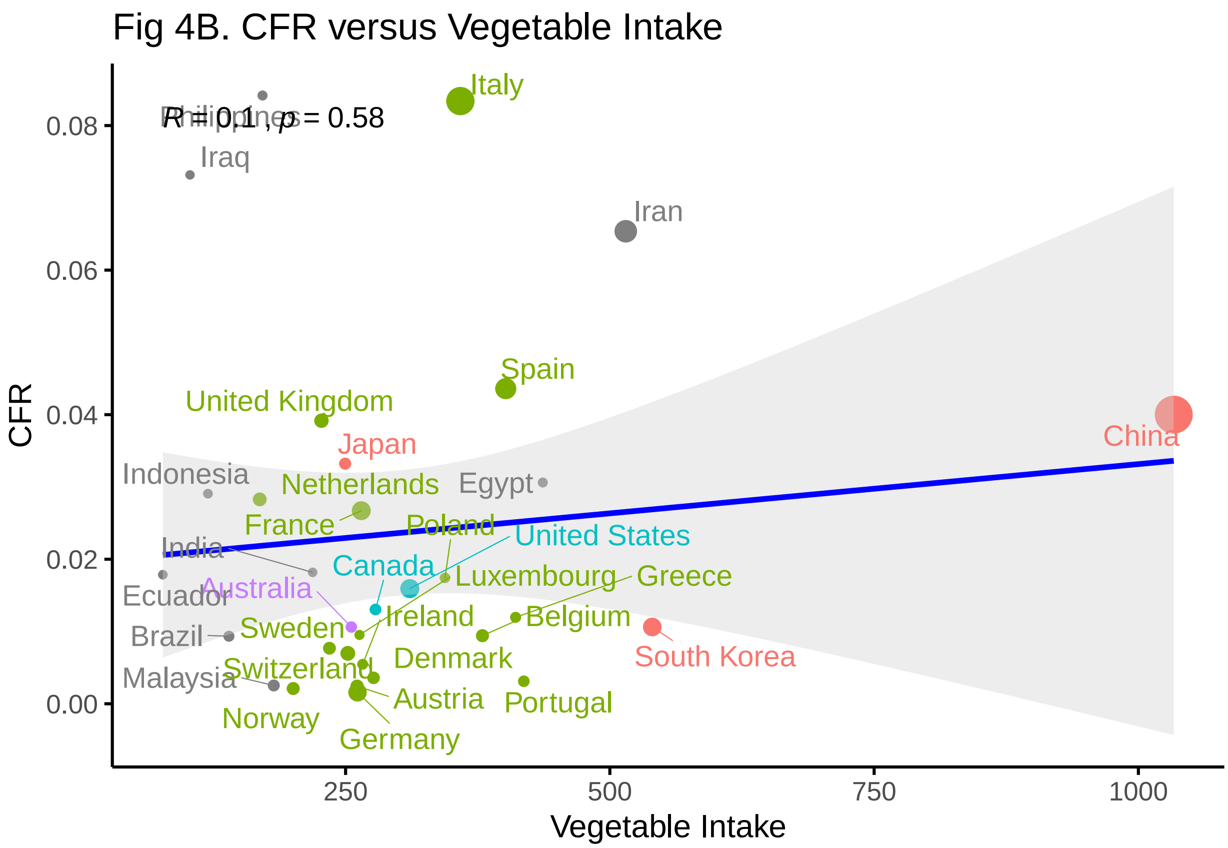
Fig 3B. CFR versus GDP per capita

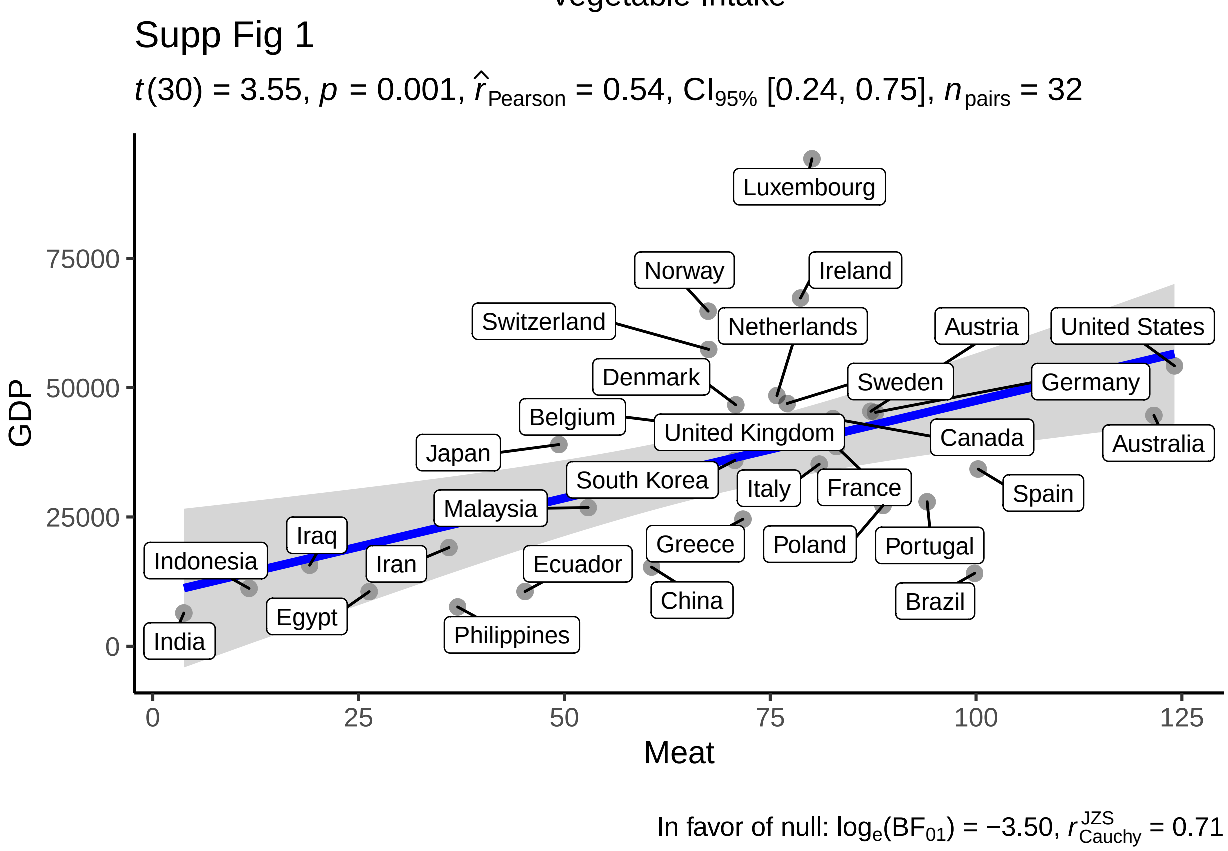


To contrast with other variables associated to quality of life or as a surrogate of iron consumption we compared the proportion of Meat consumption per capita and the vegetable consumption index to the CFR (Fig 4A,4B). We found that meat consumption and not vegetable intake was associated but likely because meat consumption is also associated to a higher GDP (Supp. Fig. 1).

Fig 4A. CFR versus Meat Consumption

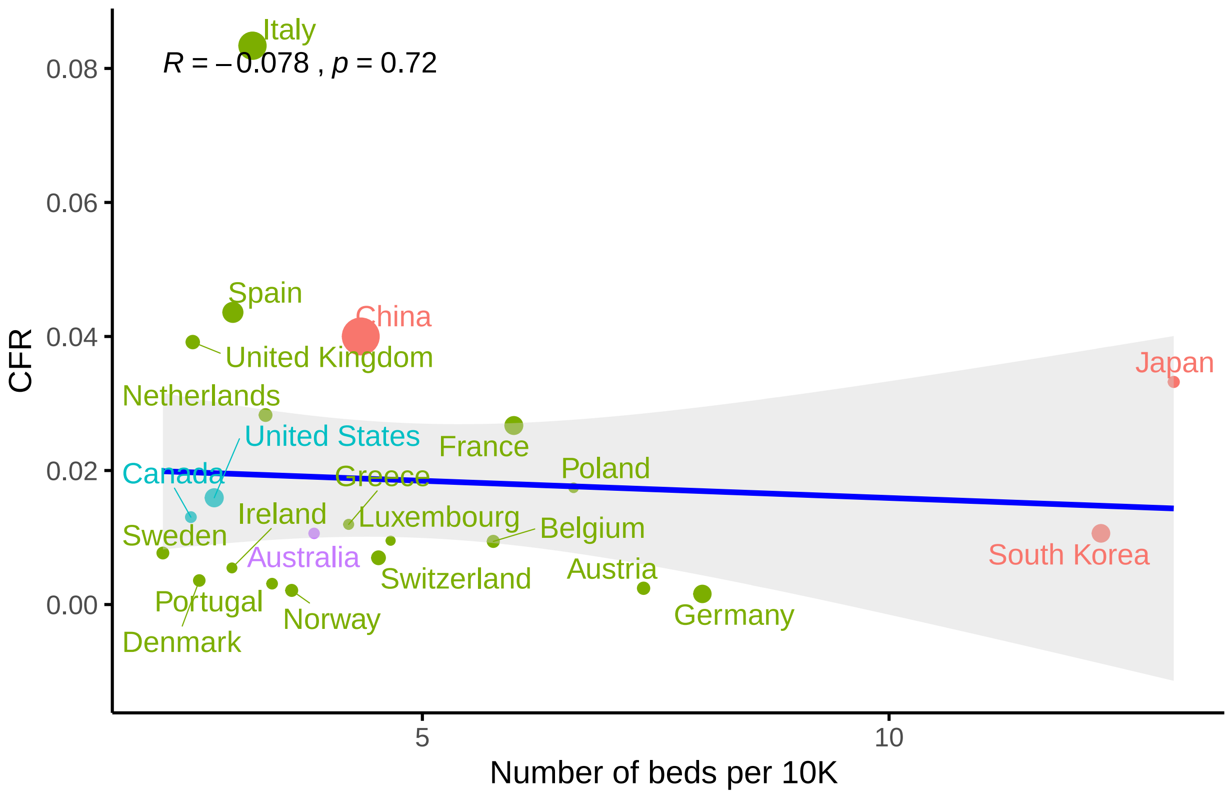


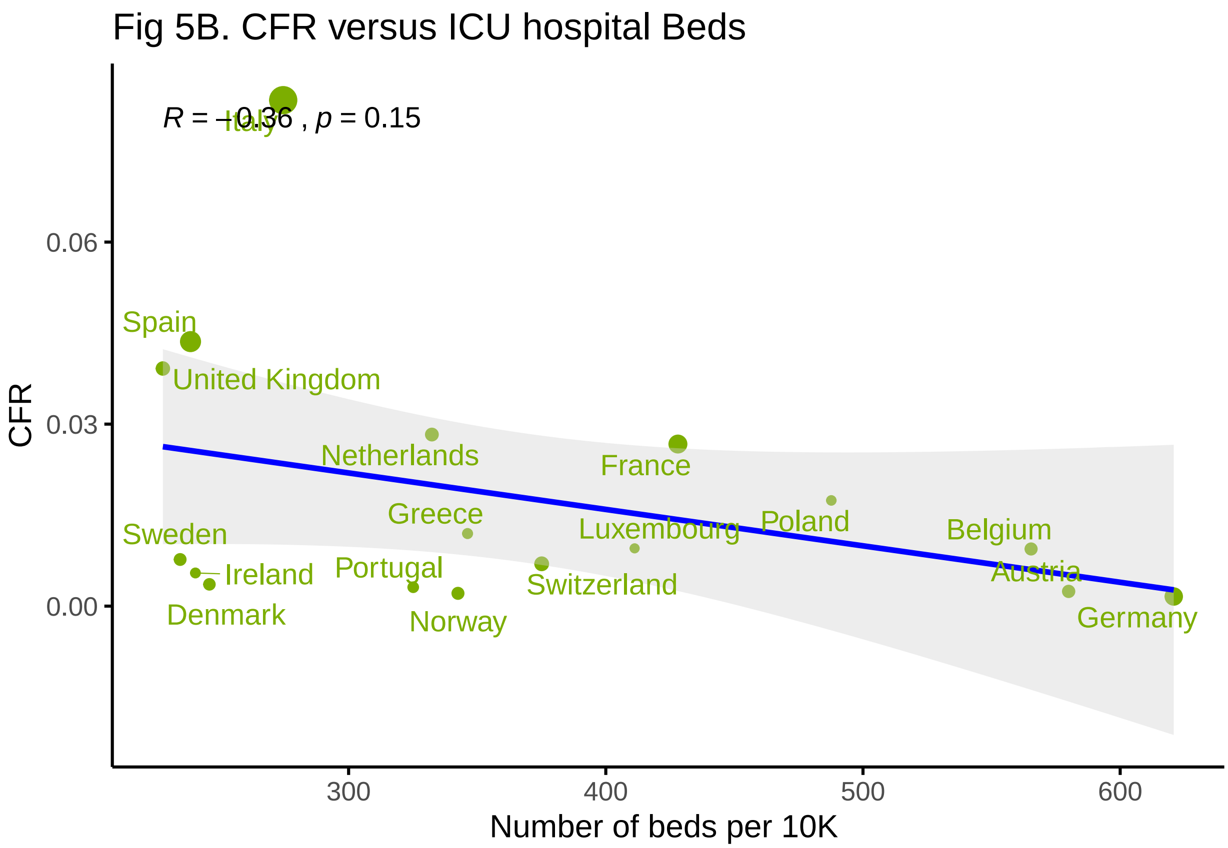


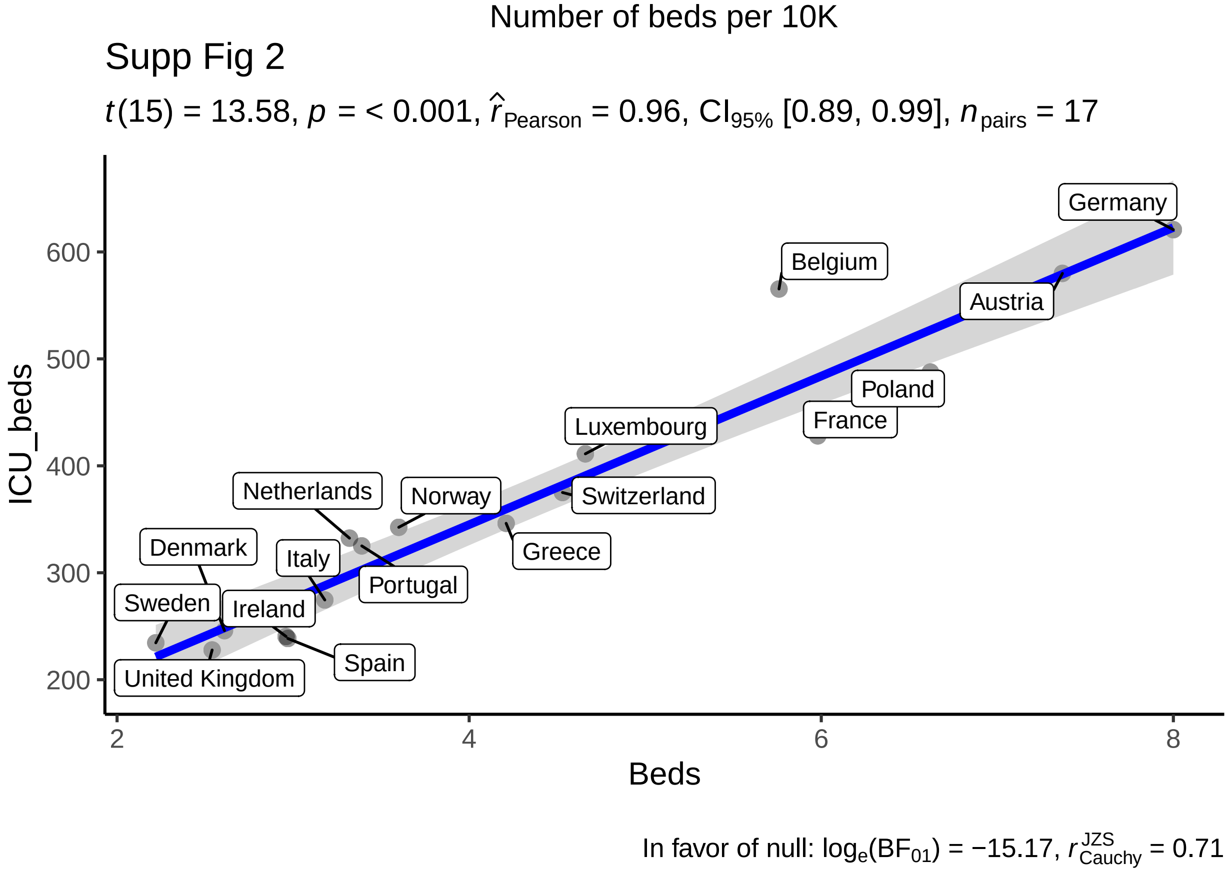


We also check if there is an association between number of hospital beds to the CFR of COVID-19. Surprisingly, we found no direct correlation between these measures when looking independently (Fig 5A, 5B). We observed that the number of hospital bed strongly correlates with the number of intensive care unit beds (Supp Fig 2), and given we had more information on the number of total beds, we used this parameter for the regression analysis.

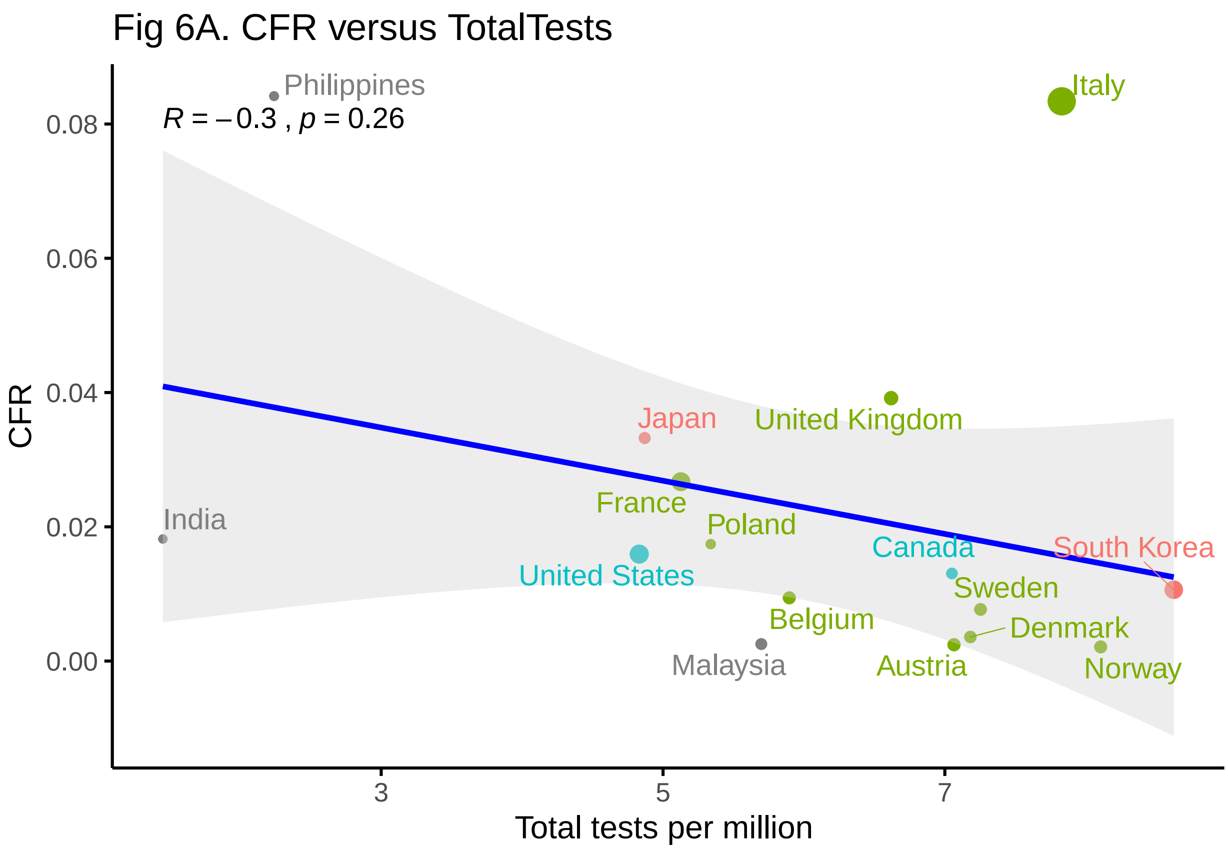
Fig 5A. CFR versus hospital Beds





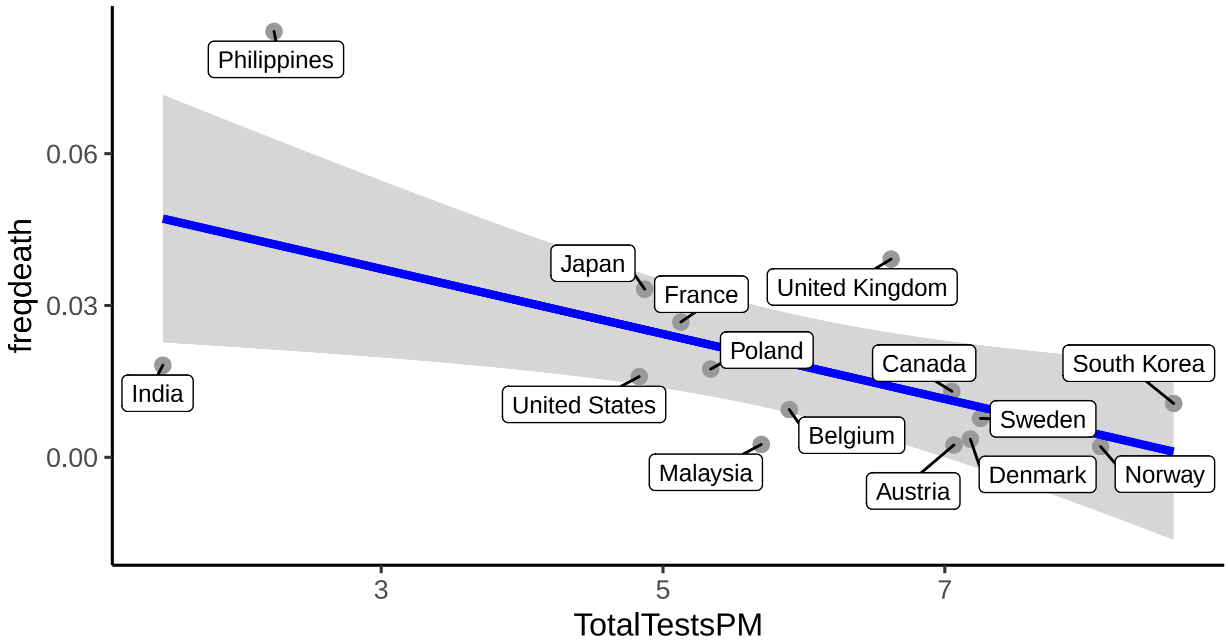


A lot of speculation has been placed in the importance of testing, thus we also explored this variable and found no association between the variable and the number of total tests performed per million inhabitants (Fig 6A). However, if we remove Italy there is a significant correlation between this variable and CFR (Supp Fig 3).



Supp Fig 3

$t(13) = -2.71$, $p = 0.018$, $\hat{r}_{\text{Pearson}} = -0.60$, $\text{CI}_{95\%}$ $[-0.85, -0.13]$, $n_{\text{pairs}} = 15$



In favor of null: $\log_e(\text{BF}_{01}) = -1.47$, $r^{\text{JZS}}_{\text{Cauchy}} = 0.71$

Now, understanding that CFR is a multifactorial process, we performed a multivariate regression analysis to determine the best predictors of CFR with the current available data. We first checked for colinearity of all predicted variables by plotting a multi correlation (Fig. 7). To select for the best model, we perform multiple regression using different subset of the independent variables. We then selected the one with the lowest AIC. By looking at the performance of all different models we observed that the best model had an R2 adjusted of 0.48 and AIC of -106. This model included the population size, the proportion of 65+ adults respect to the rest of the population, the GDP per capita, and the number of hospital beds (Fig. 8). The final model then was CFR = 0.029*(log10(Populatio)) + 0.34*(GDP per capita) + 0.33*(65+/65- population) - 0.0044*(Number of hospital beds).

Fig 7. Correlation of all tested variables

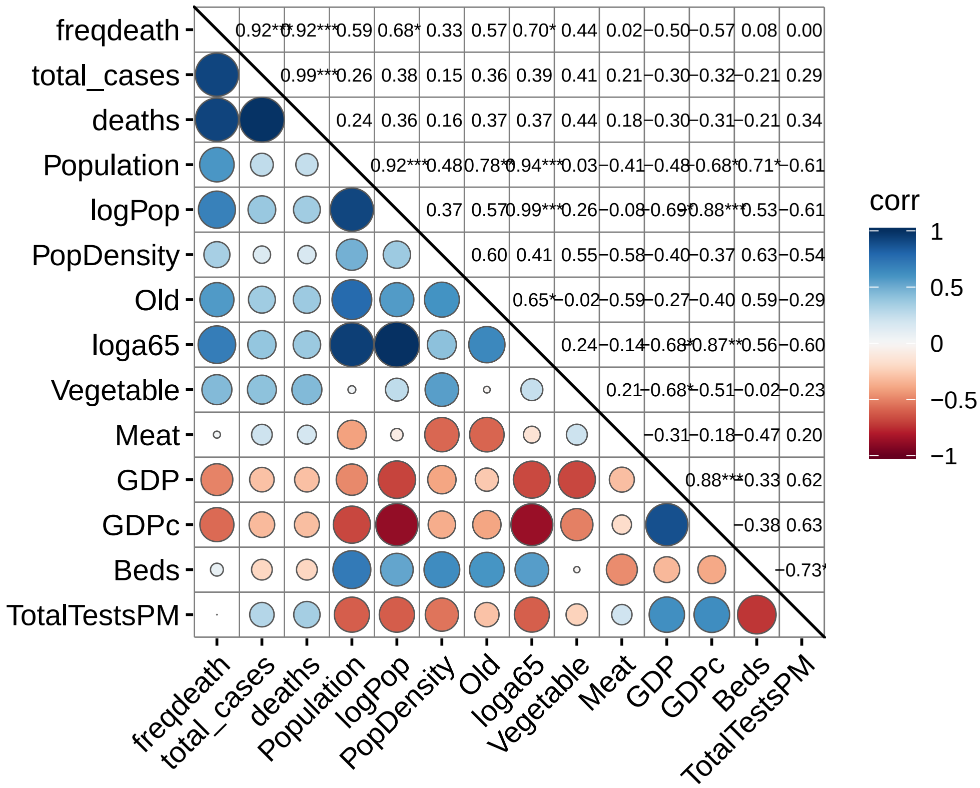
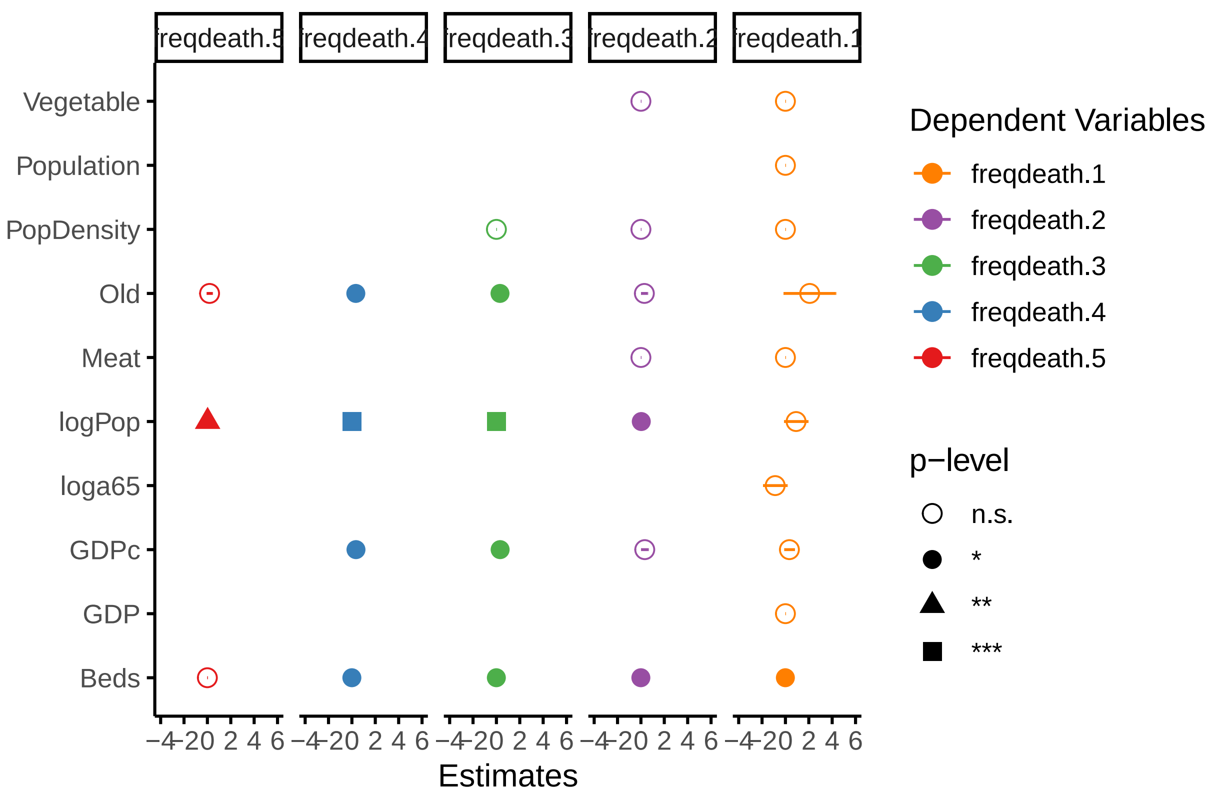


Fig 8. Coefficient estimates and significance value for each variab



Discussion

Here we have presented an up to date analysis (March 19th) of the likely factors contributing to the difference of case fatality rate (CFR) observed across countries during the SARS-CoV-2 pandemic outbreak. We performed multiple comparisons of different variables that can explain CFR. We identified four factors including the population size of the country, the proportion of elderly people, the GDP per capita and the number of hospital beds per country that explain CFR. Given that the first three can not be modify in the short term, we propose that the number of available beds, specifically the number of intensive care unit bebds, should be increase as soon as possible to tackle the current pandemic. As an example for our model, if we increase the current value in spain from 3 beds per 1000 people to 4 beds per 1000 people, we can reduce the CFR from an estimated of 3.3% to 2.9%, a 0.3% difference that can save the lives of approximately 50 people out of the currently confirmed cases of ~18000 people.