Risk Factors of SARS-CoV-2 Infections: A Global Epidemiological Study

William A Barletta

Submitted to: JMIRx Med
on: March 15, 2021

Table of Contents

Original Manuscript 4
Supplementary Files 26
Figures 27
Figure 1 28
Figure 2 29
Figure 3 30
Figure 4 31
Figure 5 32
Figure 6 33
Figure 7 34
Figure 8 35
Figure 9 36
Figure 10 37
Figure 11 38
Figure 12 39
Figure 13 40
Figure 14 41
Figure 15 42
Figure 16 43
Figure 17 44
Figure 18 45
Figure 19 46
Figure 20 47
Figure 21 48
Figure 22 49

Risk Factors of SARS-CoV-2 Infections: A Global Epidemiological Study

William A Barletta

Corresponding Author:
William A Barletta

Abstract

Since the first governmental recognitions of the pandemic characteristic of the SARS-Cov-2 infections, public health agencies have warned about the dangers of the virus to persons with a variety of underlying physical conditions, many of which are more commonly found in persons older than 50 years old. To investigate the statistical, rather than physiological basis of such warnings, this study examines correlations on a nation-by-nation basis between the statistical data concerning covid-19 fatalities among the populations of the ninety-nine countries with the greatest number of SARS-Cov-2 infections plus the statistics of potential co-morbidities that may influence the severity of the infections. It examines reasons that may underlie of the degree to which advanced age increases the risk of mortality of an infection and contrasts the risk factors of SARS-Cov-2 infections with those of influenzas and their associated pneumonias.

(JMIR Preprints 15/03/2021:28843)
DOI: https://doi.org/10.2196/preprints.28843

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
   Please make my preprint PDF available to anyone at any time (recommended).
   Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
   Only make the preprint title and abstract visible.
✓ **No, I do not wish to publish my submitted manuscript as a preprint.**
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
   Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).
   Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
✓ **Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a h**

Original Manuscript

Risk Factors of SARS-CoV-2 Infections: A Global Epidemiological Study

William A. Barletta
Department of Physics, Massachusetts Institute of Technology
Cambridge MA

Revised May 20, 2021

Abstract

Background: Since the first recognition of the pandemic characteristic of SARS-CoV-2 infections and before considerable case fatality data were available worldwide, public health agencies warned the public about the increased dangers of SARS-CoV-2 to persons with a variety of underlying physical conditions many of which are more commonly found in persons over 50 years old or in certain ethnic groups.

Objective: To investigate the statistical, rather than physiological basis in support of such warnings, this study examines correlations globally on a nation-by-nation basis between the statistical data concerning COVID-19 fatalities and the statistics of potential co-morbidities that may influence the severity of the infections.

Methods: This study considers the statistics describing the populations of the ninety-nine countries with the greatest number of SARS-CoV-2 infections. As national compilation of direct measures of immune system strength are not publicly available, it uses the frequency of fatalities in those countries due to a variety of serious diseases as a proxy for the susceptibility of those populations to those same diseases.

Results: The analysis produces plots and calculations of correlations and cross correlations of COVID-19 case fatality rates and the risks of other potential co-factors. It exposes some reasons that may underlie of the degree to which advanced age increases the risk of mortality of an infection by SARS-CoV-2. In contrast with the strong influences of co-morbidities on the seriousness of consequences of influenzas and their associated pneumonias, the correlations of the same set of risk factors with SARS-Cov-2 infections are considerably weaker. The general characteristics of the observed correlations have strengthened through three cycles of analysis starting in September 2020. The strongest correlations at ~28% and ~20% are with chronic kidney disease and coronary disease respectively.

Conclusions: This study confirms early clinical observation that infection by the SARS-Cov-2 virus presents an increased risk to persons over the age of 65. It does not support the suggestions presented by government agencies early in the pandemic that the risks are much greater for persons with certain common, potential co-morbidities.

1. Background

In only a few months after its disclosure by Chinese health authorities, the SARS-CoV-2 virus had spread around the world. By late winter of 2020, the World Health Organization had designated the COVID-19 infection caused by the virus to be a worldwide epidemic. As seen from the effects of the roughly 80 million infections reported by the end of 2020, COVID-19 can manifest in mild, flu-like symptoms or far more seriously as a severe and often deadly respiratory disease with pneumonia.

From the outset of the COVID-19 pandemic, the public has been treated to numerous speculations about the degree to which age and/or various underlying morbidities may amplify the risk of intensifying the severity of infection by the

SARS-CoV-2. Authoritative sources such as the U.S. Center for Disease Control (CDC) [1] have issued warnings. Conditions cited by the CDC as increasing risk include cancer, chronic kidney disease, obesity, coronary disease, Type 2 diabetes mellitus, and sickle cell disease. The CDC also warns that asthma, hypertension, and liver disease among others *might* subject a person to increased risk. In some countries such as the United States, the incidence of COVID-19 has been more prevalent in some ethnic groups than others leading to speculations that that disparity may be biologically rather than behavior based. Such differences are not unknown; for example, sickle cell disease is most commonly found among persons whose ancestors come from Africa and Mediterranean countries where malaria is a prevalent affliction.

As many of the diseases cited by the CDC are more common in persons in late middle age and older, a common warning early in the pandemic was that SARS-CoV-2 presented particular danger to persons over 50 years old. In the initial wave of cases in China [2] and in the strong wave of cases in Italy, the probability of death due to COVID-19 was judged to be a strong function of a patient's age, being only a few percent for those under 50 and rising to nearly 20% for patients over 80. The large number of fatalities [3] in care homes in New York, the United Kingdom and elsewhere have fueled speculations about the risks of co-morbidities frequently seen in the elderly to make contracting COVID-19 fatal. An alternative explanation is the decrease in immune functions with aging. [4]

Why is COVID-19 more dangerous to the elderly than to younger persons? Complicating the answer to this question, the actual mortality rate of COVID-19 remains highly uncertain, as the prevalence of asymptomatic and unreported infections has been estimated to be from 2 to 5 times more than infections with clearly defined symptoms. An early exemplary source of testing-based data was provided by the passengers aboard the Princess Line cruise ship, the Diamond Princess on which half of the passengers who tested positive for COVID-19 were asymptomatic or at least pre-symptomatic. [5] To some degree that uncertainty might explain the very wide distributions of reported (or apparent) rate of mortality (case fatality rate) of COVID-19 in countries ranging from < 0.03% (Singapore) to almost 30% (Yemen). Moreover, in most (but not all) countries by December 2020, the integrated average case fatality rate had declined significantly from its high levels seen in March and April of 2020.

2. Objective

For a less anecdotal (and less speculative) assessment of risk factors for serious consequences of COVID-19, a data-driven examination of worldwide national statistics seems to be in order with the goal of identifying strong correlations of mortality due to COVID-19 with other potential co-morbidities and even with ethnically specific biological factors and economic. Based on investigating the statistical correlations globally on a nation-by-nation basis between the statistical data concerning reported COVID-19 fatalities and potential co-morbidities, this manuscript presents a set of calculations of such linear and multivariate correlations that may influence the severity of the infections.

3. Methods of Analysis

The analysis that follows has not been based on clinical or physiological considerations but rather on national epidemiological statistics as reported to international authorities. Unless otherwise indicated, the following assumptions underlie the subsequent calculations:

1. *The apparent mortality outcomes (case fatality rates, CFR),* defined in Equation (1) serve as a *reliable proxy* for actual rates of infection, death, and correlation with co-morbidities;

$$National\ apparent\ mortality\ rate = \frac{deaths\ atributed\ to\ COVID-19}{reported\ cases\ of\ COVID-19} \qquad (1)$$

The apparent mortality and case number data used in the following analysis are accurate as of 30 December 2020. This analysis does not and cannot account for any uncertainty due to differing national practices in distinguishing between deaths with COVID-19 and deaths due to COVID-19.

2. The sample of 99 countries across all continents *is representative* of potential correlations between COVID-19 case fatality rates and potential co-morbidities or ethnicity. The number of COVID-19 cases in the countries not included is not statistically significant. Nevertheless, outliers with relatively small statistical significance can skew calculated correlations.

3. *Linear correlations* are examined on the basis of averaged national data for the year 2020. The sources that describe the prevalence of disease are from the World Health Organization [6], Worldometer [7], and for economic data the World Bank as reported by Trading Economics [8]. This analysis *assumes* that the published WHO data concerning the fatalities ascribed to diseases in a given country constitute valid proxies for the prevalence of those maladies in national populations. In the case of obesity, the reported number is the percentage of the population with a BMI exceeding a WHO established standard for a person of that sex.

The study examined the following factors:

Demographics: geographical region, population, national median age;

SARS-CoV-2: number of COVID-19 tests, confirmed cases of COVID-19 as reported by government authorities, and the apparent case fatality rate;

Medical factors: incidence of flu, lung disease, asthma, obesity, heart disease, common cancers, hypertension, chronic kidney disease, diabetes, and malnutrition;

Economics: GDP-PPP, average household size, % population living in slums, health expenditures per capita, and WHO universal health coverage index;

One random (or pseudo-random) variable in the range from 0 to 100.

Examination of the data begins with computing linear correlations between variables. The evaluation of the linear correlation herein uses the Pearson "product moment correlation" of Equation (2) to evaluate linear relationships between data sets:

$$r=\frac{\sum\left(x_i-x_{average}\right)\left(y_i-y_{average}\right)}{\sqrt{\sum\left(x_i-x_{average}\right)^2\cdot\sum\left(y_i-y_{average}\right)^2}} \qquad (2)$$

One may estimate the *statistical significance* of calculated correlations by computing *r* for two variables that are *uncorrelated by construction*; i.e., apparent COVID-19 mortality and a random variable in the range from 1 to 100. Once linear correlations have been computed, the next step is evaluating cross-correlations among variables and performing a multivariate analysis.

The 99 countries sampled in this study have been selected as those reporting the largest number of COVID-19 infections. The countries listed in Figure 1 represent five geographical regions: the Americas, Asia, Europe, Africa, and Middle East plus Central Asia. Regional populations are included. The combined population of nearly 5.5 billion persons accounts for the strong preponderance of all cases reported worldwide. The data cutoff date was December 30, 2020.

The SARS-CoV-2 related data are sex-aggregated because many countries still do not report sex-disaggregated data (or made these data available publicly). Therefore, the frequently reported sex-based disparities in contagion and in case fatality ratio could not be examined with respect to sex-based differences in occurrences of potential co-morbidities.

Americas	**Asia**	**Europe**	**Africa**	**Middle East**
pop = 977 M	*pop = 2504 M*	*pop = 725 M*	*pop = 768 M*	*pop = 487 M*
Argentina	*Australia*	*Albania*	*Algeria*	*Afghanistan*
Bolivia	*Bangladesh*	*Armenia*	*Cameroon*	*Bahrain*
Brazil	*China*	*Austria*	*Congo*	*Egypt*
Canada	*India*	*Azerbaijan*	*Ethiopia*	*Iran*
Chile	*Indonesia*	*Belarus*	*Ghana*	*Iraq*
Columbia	*Japan*	*Belgium*	*Ivory Coast*	*Israel*
Costa Rica	*Kazakhstan*	*Bosnia*	*Kenya*	*Lebanon*
Dominican Republic	*Kyrgyzstan*	*Bulgaria*	*Libya*	*Kuwait*
Ecuador	*Korea*	*Croatia*	*Madagascar*	*Oman*
El Salvador	*Malaysia*	*Czechia*	*Mali*	*Qatar*
Guatemala	*Nepal*	*Denmark*	*Morocco*	*Saudi Arabia*
Honduras	*New Zealand*	*Estonia*	*Nigeria*	*Turkey*
Mexico	*Pakistan*	*Finland*	*South Africa*	*UAE*
Panama	*Philipines*	*France*	*Sudan*	*Uzbekistan*
Paraguay	*Singapore*	*Germany*	*Uganda*	*Yemen*
Peru	*Thailand*	*Greece*	*Zambia*	
USA	*Taiwan*	*Hungary*		
Venezuela		*Ireland*		
		Italy		
		Macedonia		
		Moldova		
		Netherlands		
		Norway		
		Poland		
		Portugal		
		Romania		
		Russia		
		Serbia		
		Spain		
		Sweden		
		Switzerland		
		Ukraine		
		UK		

Figure 1. Countries sampled grouped into five regions. As Yemen is a statistical outlier in apparent mortality, many plots omit its data point for visual clarity.

Figure 2 plots the case fatality ratio and random numbers that are uncorrelated by construction. The Pearson coefficient for this set of 99 values is -5.3%.

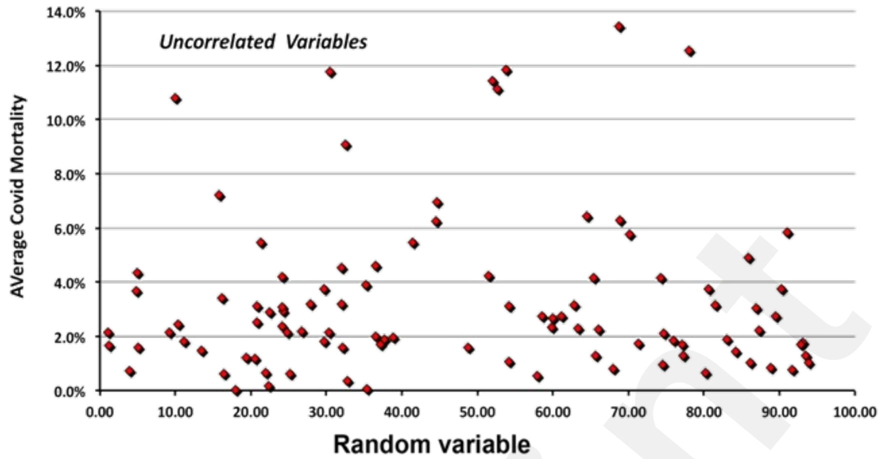


Figure 2. A plot of the variables uncorrelated by construction.

A potential limitation of this approach is that all mortality data have equal weight in the calculation of correlation. One check of whether this Ansatz introduces a bias is the correlation between apparent national mortality rates and national populations. Doing so, yields -1.4%, a value close to the Pearson coefficient for uncorrelated variables. Another possible way to attribute a rational weighting to plot the variation of COVID-19 deaths per capita against the possible risk factor. However, the number fatalities per capita depends strongly on national public health policies, on national efforts to prevent the spread SARS-CoV-2, on GDP, and other non-medical considerations. Differences in COVID-19 statistics between Norway and Sweden [7] are cases in point.

4. Results

4.a. Examination of linear correlations

To gain confidence in this statistical approach one can plot two variables for which one may expect to see a correlation (Figure 3). Here, the linear correlation is quite high, 62.5%. Closer examination of Figure 3 suggests a limitation of considering only linear correlations. The countries circled in red show a strong correlation while those in the green ellipse show scarcely any correlation of a nation's wealth with the age of its population. Clearly, a refinement of the statistical approach is needed. Identifying the data underlying each point with each country's region in Figure 4 reveals that median age and national wealth are essentially uncorrelated for European nations but strongly correlation for countries in Africa and Asia. *Regional grouping* was thus adopted throughout this study.

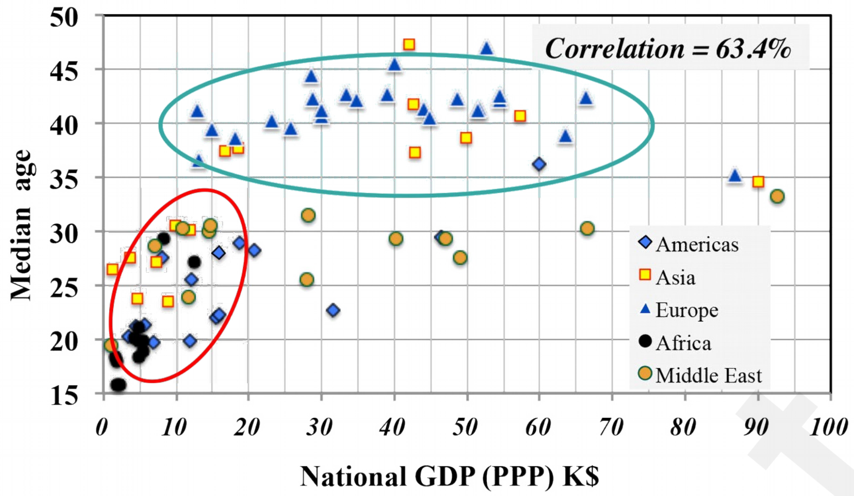


Figure 3. The plot of GDP-PPP corrected for purchasing power versus median age in countries from the five regions under study

To illustrate the utility of this refinement, Figure 4 shows the relation between deaths per 100k persons due to malnutrition as a function of national wealth measured by GDP corrected for purchasing power. The relatively strong (45.5%) correlation is driven by the high rates of malnutrition in Africa, Central America, and the poorer countries of Asia. No such effect is apparent in Europe.

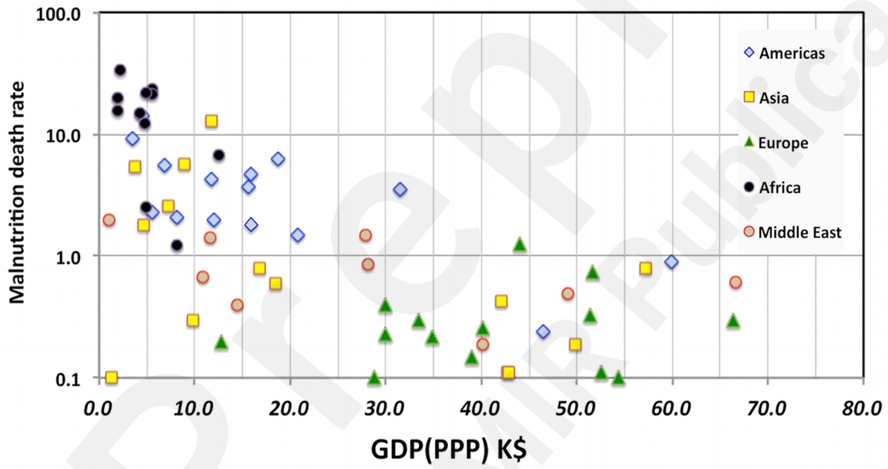


Figure 4. It is not surprising that poverty is correlated with malnutrition.

From the outset of the pandemic, national health authorities have warned the public about the increased risk of mortality for persons 60 years old and older. Figure 5 shows an early example of the basis for such warnings in the data provided by the UK Office of National Statistics in September 2020 [9] and in reference [10]. The UK government website notes several caveats: 1) The figures include deaths of non-residents of the UK; 2) they are based on the date that a death was registered rather than when it occurred; 3) they are provisional and use the tenth edition of the International Classification of Diseases, (ICD-10) for definitions for the coronavirus (COVID-19). Again one asks why should the severity of a COVID-19 infection be a function of age?

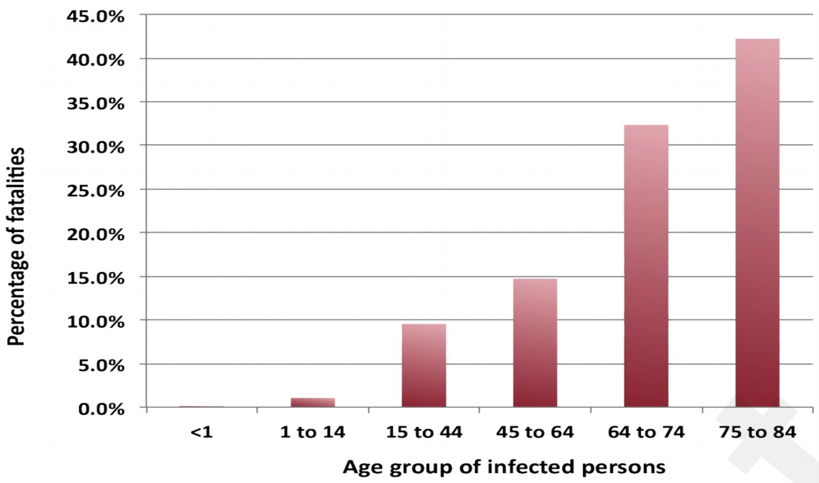


Figure 5. Deaths attributed to COVID-19 by the UK Office of National Statistics.

From these data, one might expect a strong correlation between the apparent national case fatality ratio and the median age of a country's population. Even accepting the hypothesis of universality for the data of Figure 5, one should first multiply these rates by the demographics of a nation's population normalized to the U.K. population grouped into the same age bins. Such a plot (in Figure 6a) shows a surprising result. The overall linear correlation is negative, -18.1%, partially due to the disparity among the regions: -25.8% for the Americas, 5.2% for Asia, 14.1% for Europe, 2% for Africa and -60.8% for the middle East and Central Asia.

Rather than plotting the COVID-19 case fatality ratio versus national median age, one might examine the dependence on the percentage of the population of age 65 or greater. In that case the overall correlation (- 8.1%) is negative, consistent with reference [11], but is mostly the result of regional variations with a larger, but still relatively low correlation (~19%) in Europe and Africa.

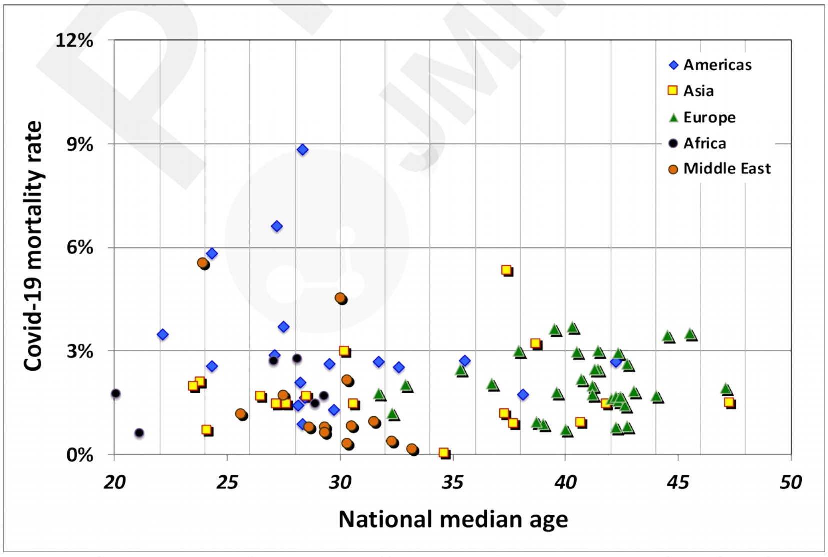


Figure 6 National median age versus CFR for the five regions

As a measure of the influence of the age of a population on SARS-CoV-2 contagion, the national rate of confirmed cases of COVID-19 per 1 million persons with respect to the percentage of population older than 65 (Figure 7) displays a moderate correlation of 44.7%.

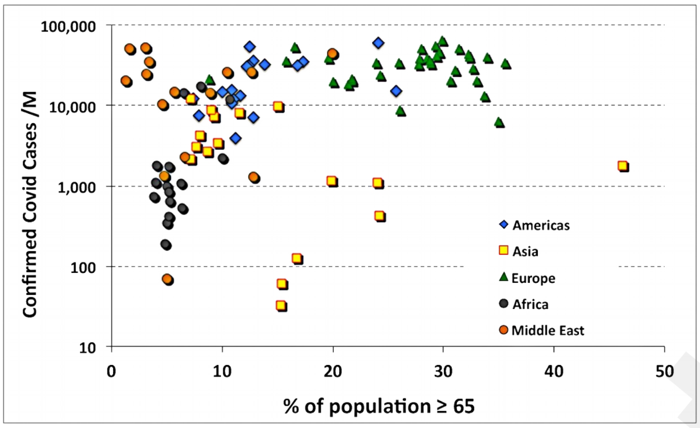


Figure 7. Confirmed COVID-19 cases as a function percent of population older than 65.

One may hypothesize that the "care home effect," i.e., the large numbers of deaths seen in nursing homes in Italy, the U.K. and N.Y. was more the result of overcrowding and poor hygienic practices compounded with the general infirmity and the reduced immune function of nursing homes residents than by any extreme dependence of the lethality of COVID-19 on specific, underlying medical disorders. The linear correlations of age with potential causal factors, showed in Figure 8 suggest the strength of candidate co-factors to explain the "care home effect." In addition to specific co-factors, the "care home effect" also reflects a generally very weakened physical condition of many occupants of care homes that could render any pneumonia-inducing disease potentially lethal. The data of Figure 8 show no evidence that age alone influences the probability of a person becoming infected by the SARS-Cov-2 virus.

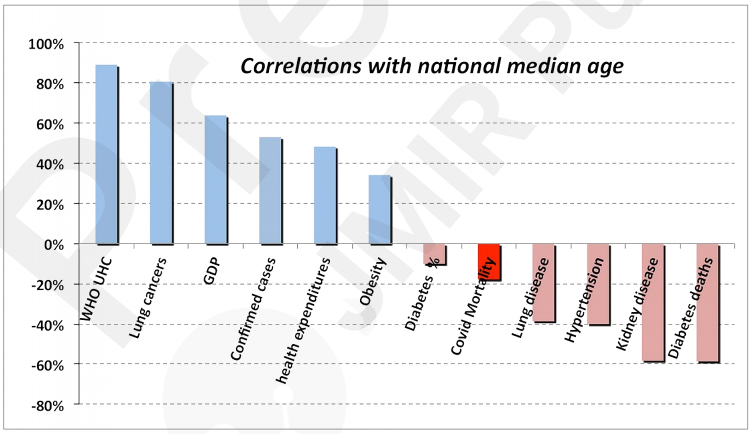


Figure 8. Correlations of potential risk factors with median age

Figure 8 may explain what seems like a startling result, namely the globally negative correlation of COVID-19 case fatality ratio with the age of national populations. The negative value is due to the strong correlations between national median age and the adjusted GDP (64%), national health care expenditures (48%), and the WHO Universal Health Care Index. Nations with the oldest populations are generally those that are the wealthiest and in which health care services are the largest, thus reducing the level of mortality.

In contrast with infections due to SARS-Cov-2, fatalities from influenza-induced pneumonia (Figure 9) are highly correlated (-65.2%) with the median age of the population. The correlation also displays a strong regional dependence. The correlation is negative for the same reasons previously explained for COVID-19.

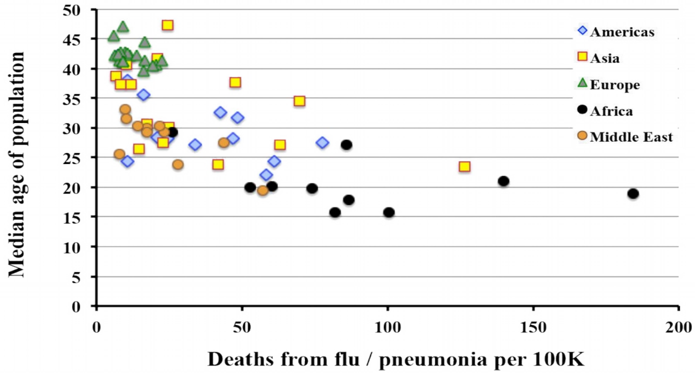


Figure 9. Incidence of flu related pneumonia deaths as a function of national median age.

This result for influenza suggests that as COVID-19 typically presents as a severe respiratory disease the severity of COVID-19 infections might correlate with incidence of asthma. The global value (Figure 10) is small but not negligible, 16.5%, largely driven by the strong correlation (68.1%) in the Middle East.

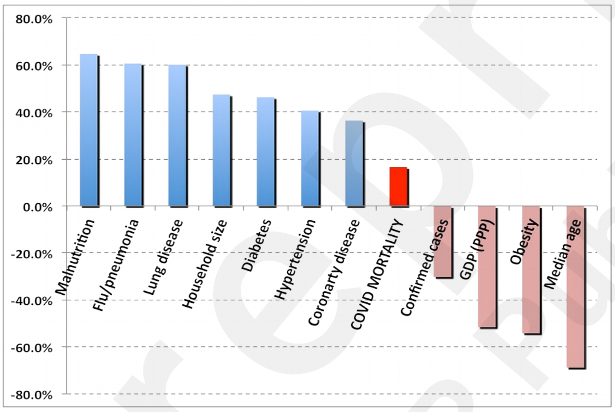


Figure 10. Correlation of severe asthma with COVID-19 case fatality ratio

For asthma as a co-factor, the contrast with influenza related pneumonia is striking. The relatively high, overall correlation of 59.4% for influenza is seen in all regions. Hence any reference to COVID-19 as a “flu-like” infection or as a “superflu” is grossly misleading.

U.S. Centers for Disease Control issued a warning early in 2020 that obesity represented a co-morbidity that could lead potentially to severe consequences of a COVID-19 infection. However, once again actual the national data of (Figure 12) display essentially no correlation (-1.7%) of COVID-19 case fatality rate with the percentage of a country’s population considered obese. A better metric of national obesity may be the average Body Mass Index (BMI in kg/m^2) of its population. With BMI as the metric national prevalence of obesity the correlation increases to 5.2%, still very small. Moreover, that figure may be misleading when comparing region to region as the correspondence between BMI and Body Fat percentage varies considerably (10% to 20%) from country to country.

The contribution of obesity to the *outcome* of other pulmonary disorders is significantly different to that of COVID-19 as is displayed in Figure 11. Obesity does have significant correlation (51.6%) with the risk of contracting infection from SARS-Cov-2, although not with the apparent outcome of the infection. Observation of increased risk of infection (although not its outcome) was previously reported in [10]. Reference [11] reports an increased risk of infection (32.9%) for people with chronic kidney disease. That correlation is not seen in the statistics of this study, which has consistently found a temporally increasing negative correlation.

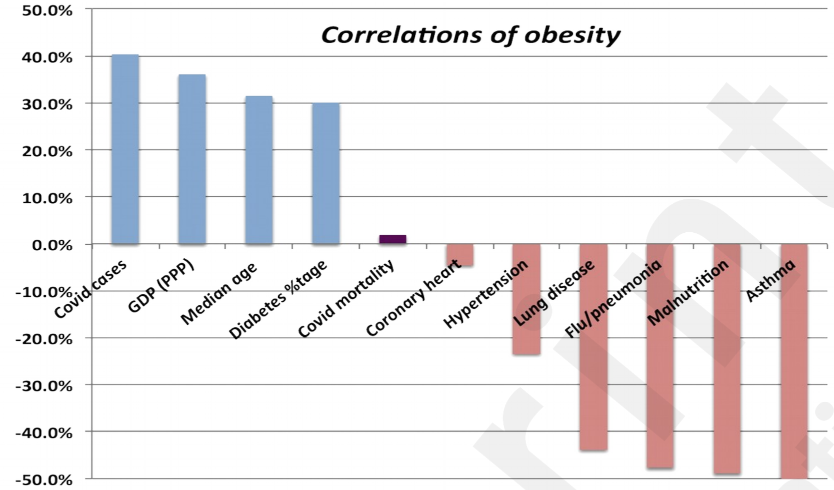


Figure 11. Correlations of obesity rates with COVID-19 mortality and other conditions. For most of conditions, rates are based on deaths per 100K persons.

One might speculate that as a chronic respiratory disorder involving pulmonary airways, asthma may increase the seriousness of consequences of COVID-19 and its induced pneumonias; this analysis shows no such significant correlation (5.3%). Examining the correlation of COVID-19 mortality with other lung diseases also showed a tiny correlation (1.3%). In contrast, the relationship of influenza-induced pneumonias with asthma and other lung diseases presents a correlation that is quite high, 59.4% and 34.8% respectively. With respect to their effects on patients with underlying conditions, influenza and COVID-19 are very different diseases.

Another early warning to persons with underlying conditions concerned diabetes mellitus. That suspicion is echoed by the strong increase of incidence of diabetic conditions with age. Whether one measures the incidence of diabetes by deaths due to diabetes or to the reported national rates of diabetes in adults (20 to 79 years of age), the correlation with COVID-19 mortality is similarly low (10.9%). In *otherwise healthy persons*, diabetes does not appear to be a significant risk factor with respect to the serious complications of infection by SARS-CoV-2.

Figure 12 and Figure 13 summarize the linear correlations and their time variation respectively of COVID-19 case fatality ratio with underlying medical and economic conditions (in green) considered herein. As the percentage of population over 65 years of age correlates at best weakly with apparent COVID-19 case fatality rate. One may surmise that poor health care management played a very large role in the “care-home effect.”

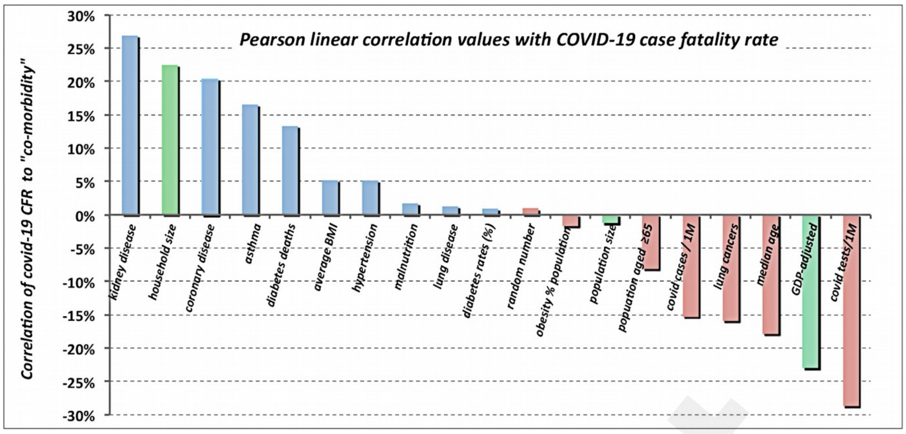


Figure 12. Summary of linear correlations with national COVID-19 case fatality ratio data.

Potential co-factor	Date of COVID-19 CFR statistics		
	30-Dec	20-Nov	16-Oct
Kidney disease	26.9%	28.9%	17.6%
household size	22.5%	22.8%	12.6%
heart disease	20.4%	19.4%	9.9%
Asthma	16.5%	16.8%	9.1%
diabetes deaths	13.3%	14.8%	5.0%
covid deaths per 1M	9.2%	7.0%	17.0%
%slums	9.0%	7.2%	5.9%
hypertension	5.1%	4.9%	-1.1%
flu/pneumonia	3.4%	4.0%	-2.0%
malnutrition	1.7%	0.2%	-3.7%
lung disease	1.3%	2.4%	-11.2%
Random number	0.9%	-2.4%	2.6%
Diabetes (% population)	0.9%	4.6%	-4.3%
population	-1.3%	0.0%	-1.4%
obesity % population	-1.7%	-0.6%	1.4%
% population ≥65	-8.1%	-10.3%	2.8%
covid cases /1M	-15.3%	-16.3%	-8.6%
health care expenditure	-15.5%	-14.3%	-2.0%
lung cancers	-15.9%	-17.9%	-9.8%
life expectancy	-16.3%	-15.2%	-5.5%
median age	-17.9%	-19.1%	-7.4%
WHO UHC index	-19.7%	-16.8%	-7.6%
% population in cities	-19.7%	-17.7%	-12.1%
adjusted GDP	-23.0%	-21.5%	-11.9%
covid tests per 1M	-28.7%	-25.7%	-11.1%

Figure 13. Correlation with national values of apparent COVID-19 case fatality ratio

One notes in figure 12 the strong negative correlation of case fatality ratio with both COVID tests per million and with the number of cases per million. More tests mean earlier detection, more detection of mild and weakly symptomatic cases, better triage followed by earlier and more effective clinical treatments.

4.b. Cross-correlations and multivariate analysis

Before investigating cross-correlations for root causes, one should perform a multivariate analysis of COVID-19 case fatality ratio against a common trio of risk factors commonly found in patients in nursing and convalescent homes– namely diabetes mellitus, hypertension, and coronary disease (DHC). For that trio, the coefficient of multiple correlation is 17.1%, not negligible but unlikely to be the root cause of the “care home effect.” Computing the correlation of DHC with deaths due to influenza and its associated pneumonia yields a stronger correlation of 35.9%. Replacing hypertension with asthma in the DHC trio reduces the coefficient of multivariate correlation for COVID-19 mortality to 12.1%. In contrast, analogous analysis for influenza increases the multiple correlation coefficient to 62.7%, demonstrating once again (see Table 1) that influenza and COVID-19 are very different diseases.

Table 1. Multivariate correlations for a trio of input variables: namely, diabetes mellitus, hypertension, and coronary disease.

Output Variable	**Regression coefficient**	**Pearson r values**
COVID-19	12.3%	3.5%, 5.3%, -4.1%
Influenza / pneumonia	43.9%	38.6%, 14.7%, 24.7%

Other calculations of multivariate correlations with the apparent national mortality rates of COVID-19 are presented in Table 2.

Table 2. Multivariate correlations with national COVID-19 mortality data

Multiple Variables	**Regression coefficient**	**Pearson r values**
GDP and household size	7.0%	-5.9%, 5.6 %
Obesity and diabetes	3.5%%	3.5%, - 7.1%
Influenza and lung disease	11.7%	-6.4%, -14.8%
Diabetes, heart and hypertension	12.3%	3.5%, 5.3%, -4.1%
Median age and # of cases	13.8%	0.4%, -13.7%
Flu deaths and diabetes	10.7%	-6.4%, 3.5%
Flu deaths and hypertension	6.8%	-6.4%, -4.1%
Obesity, asthma and diabetes	14.2 %	3.5%, 5.3%, 3.5%

4.b.1. Cross-correlations

The previous section argues and Figure 13 illustrates the striking contrast between the correlations of COVID-19 with those of influenza/pneumonia with respect to other potential underlying conditions.

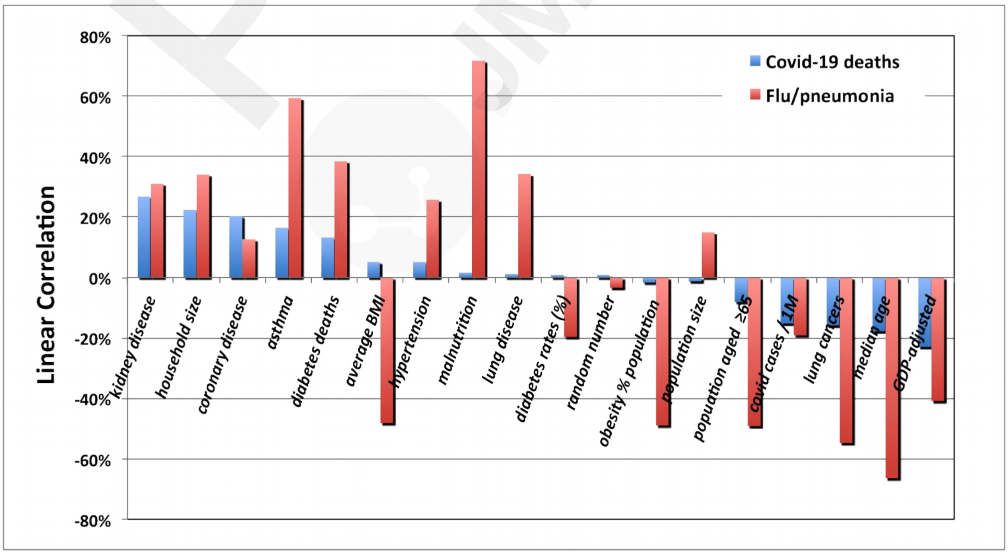


Figure 14. Contrast between correlations of COVID-19 with of flu-induced pneumonia

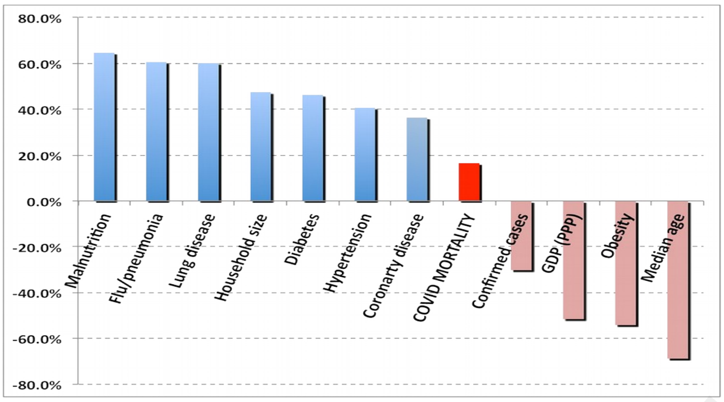


Figure 15. Cross-correlations of asthma with several disorders that might affect the outcome of COVID-19 cases.

Although obesity appears correlated with SARS-CoV-2 contagion, it appears uncorrelated with the outcome of COVID-19 infections contrary to the findings of reference [12]. Understanding the correlations of obesity calls for a deeper look at the relation of obesity with the conditions that show the most influence. Already in the case of contagion, regional differences make for a substantial fraction of the apparent effect. The regional differences could be due to factors such as national median age or it may be influenced by national wealth reckoned in terms of per capita GDP-PPP as showed in Figure 16.

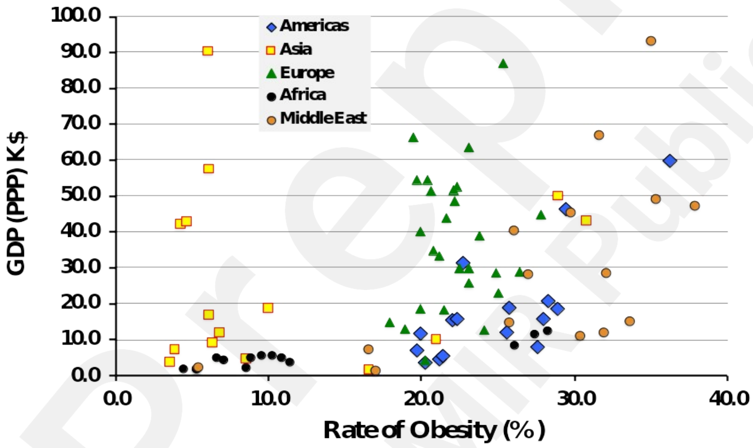


Figure 16. Or is wealth of the nation a driver of obesity?

As is the case with asthma, diabetes mellitus shows (Figure 17) significant correlations with several medical and economic conditions such as age, household size and mortality due to influenza/pneumonia. Once again no correlation with COVID-19 mortality is evident.

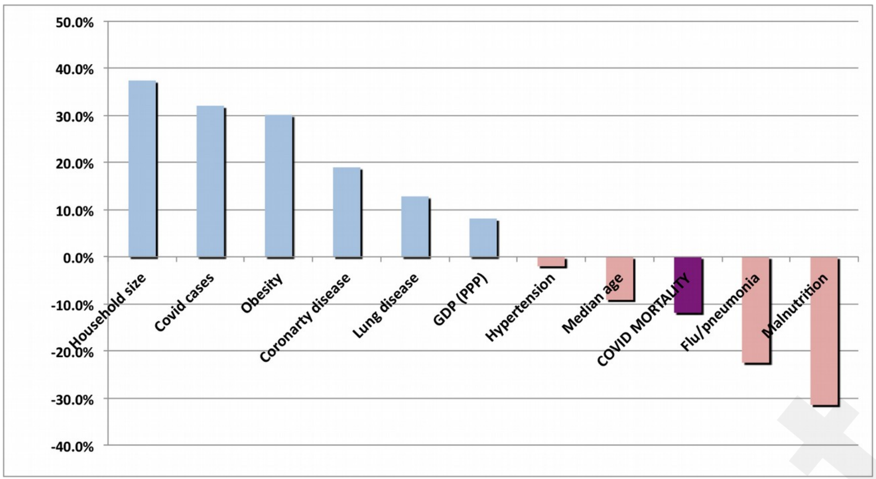


Figure 17. Correlations of national rates of diabetes mellitus with other medical and economic conditions. The correlation with COVID-19 mortality is the purple bar.

4.b.2 Regional analysis

A key assumption of this study is the high degree of country dependence of the COVID-19 case fatality rate. Even though the case fatality ratio has fallen dramatically in many countries with rates originally greater than 10%, after nearly one year of pandemic, the disparity by country and by region remains large, ranging over an order of magnitude as illustrated in Figure 18.

The size of regional data sets is obviously much smaller and the uncertainties in computed correlations higher than the aggregated world data. Nonetheless, examining the regional dependence of COVID-19 case fatality ratio on the most commonly cited co-morbidities is instructive. (See Table 1 of the Appendix.)

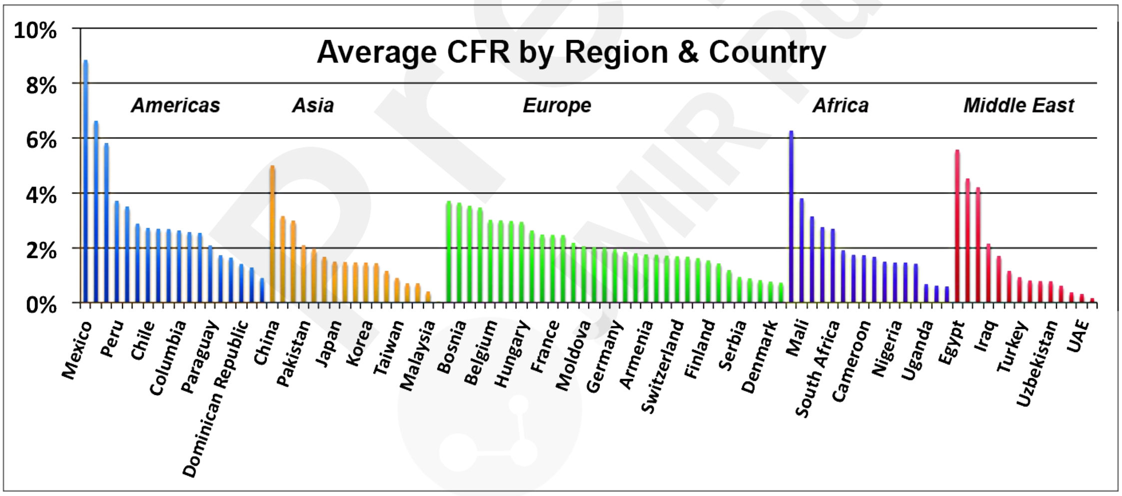


Figure 18. Average case fatality rate by country and region

4.c. Factors related to national economics and public health policies

The differences in the magnitude, outcomes, and characteristics of the waves of infections among national sub-regions with roughly equivalent medical factors indicates that economics and public health policies makes a significant difference in the severity of SARS-Cov-2 infections. This section examines dependencies on GDP-PPP, average household size, percentage of population living in slums, percentage of urban population, health expenditures per capita,

and the WHO Universal Health Coverage (UHC) index.
Figure 4 has already shown an example of economic impact on medical outcomes; the per capita GDP-PPP has a strong influence (-44.6%) on the rate of deaths due to malnutrition. That observation is hardly surprising. One may ask the same question with respect to mortality due to COVID-19 infections. The distribution of COVID-19 mortality with national wealth shows essentially no correlation (-5.9%). The politics of poverty does not, of itself, explain the observed national rates of COVID-19 mortality.
The distribution of contagion of SARS-Cov-2 over the global data set is noticeable and positive (29.9%). However, as shown in Figure 19 that value is entirely driven by strong dependence of rising contagion with rising income in African countries. If one removes the African countries from the sample, the correlation disappears (2.8%).

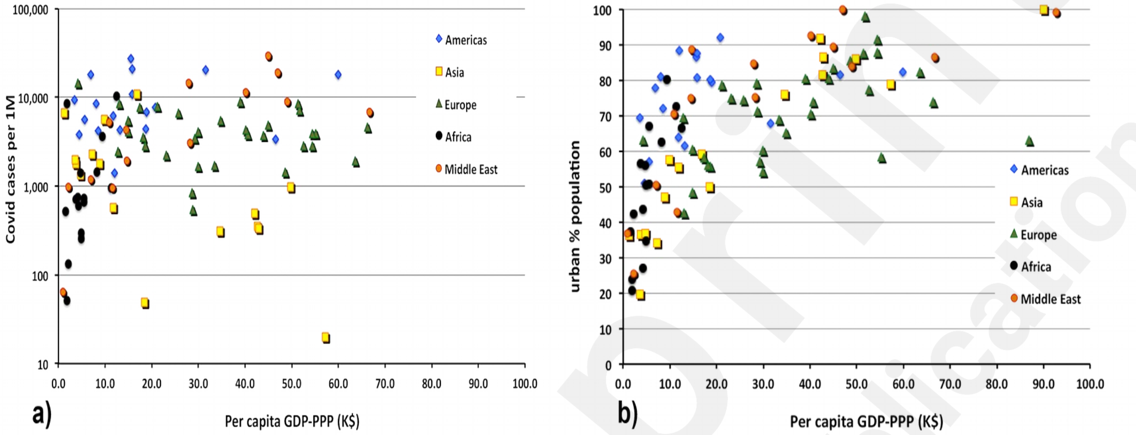


Figure 19. a) The distribution of COVID-19 cases with national GDP-PPP. b) The degree of urbanization with increasing GDP

The strange behavior in Africa in Figure 19a may be due to the increase in urbanization with increasing national wealth. One might further suspect that the increase in urbanization is also likely to increase the fraction of the national population living in slums. In fact, the percentage of people living in slums actually decreases with the urbanized fraction of the population.
The correlation of economic and policy factors with contagion (measured in confirmed COVID-19 cases per 1 million of population and apparent COVID-19 mortality is presented in Table 3. As the mortality rate varies in time and seems to decline as the pandemic progresses (at least in the Northern Hemisphere) the mortality rate has been benchmarked on December 30, 2020. The surprising negative correlation in contagion with the percentage of the urban population living in slums is likely due to the trend in Africa that the smaller the fraction of the population living in cities, the more likely it is that they live in slums [World Bank data].
An examination by region of the impact of economic co-factors in the COVID-19 case fatality rate is shown in Figure 20. The negative correlations with national wealth and with national health care expenditures are to be expected. Nonetheless these effects are weaker in Africa than in other regions. More detailed investigation of these effects would require examination of underlying conditions on a country-by-country basis.

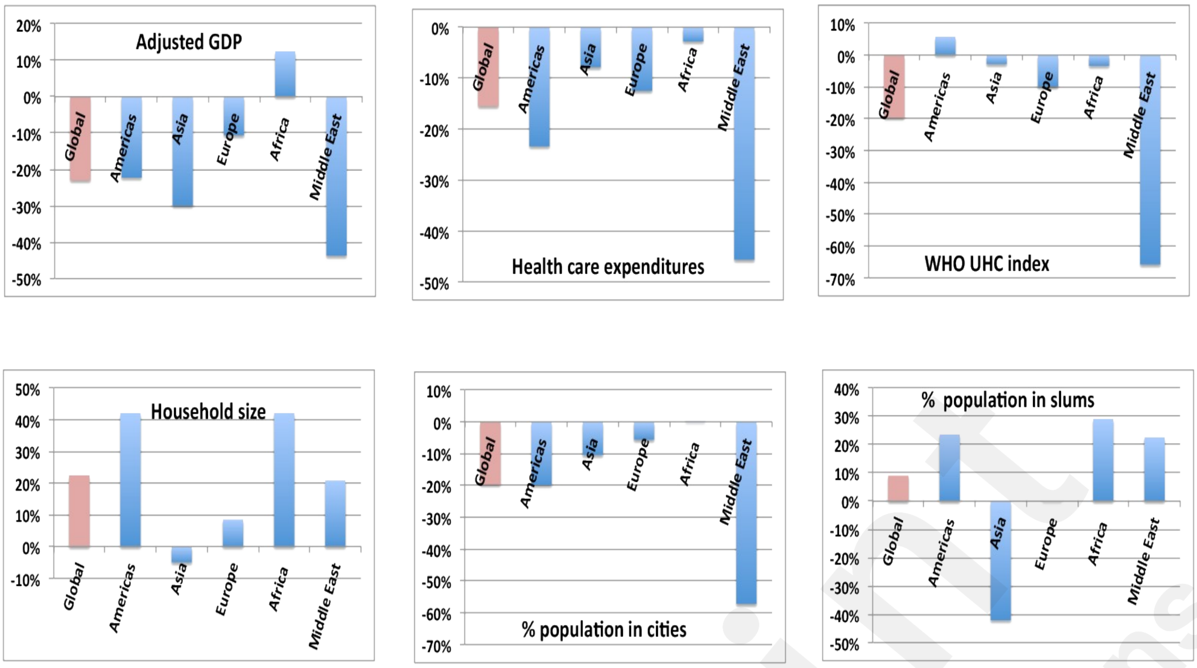


Figure 20. The impact of economic co-factors on case fatality ratio varies strongly by region

Table 3. Correlations of economic and political factors with number of cases of COVID-19 infection (contagion) and apparent COVID-19 mortality.

	Correlation	
Factor	**Conta-gion**	**Mortality**
% in cities	47.60%	-8.50%
Testing for COVID-19	43.80%	-9.90%
GDP-PPP	32.00%	-5.90%
UHC INDEX	30.30%	-2.00%
Health spending	24.00%	4.60%
Household size	17.20%	5.60%
Urban % in slums	-40.70%	4.10%

The correlation with respect to GDP is explained by the correlation of GDP with percentage of population over 65. The substantial correlation of contagion with testing results from the obvious fact that the more one looks, the more one sees. The correlation of contagion with percentage of urban population is due to the cross-correlation of GDP with percentage of urban population (64.8%) and the high cross-correlation of urban population with testing for COVID-19 (49.7%). The values for average health care expenditures and the UHC index of the WHO are similarly explained. The data that underlie the value of case fatality rate versus the percentage of the urban population that live in slums appears below in Figure 21.

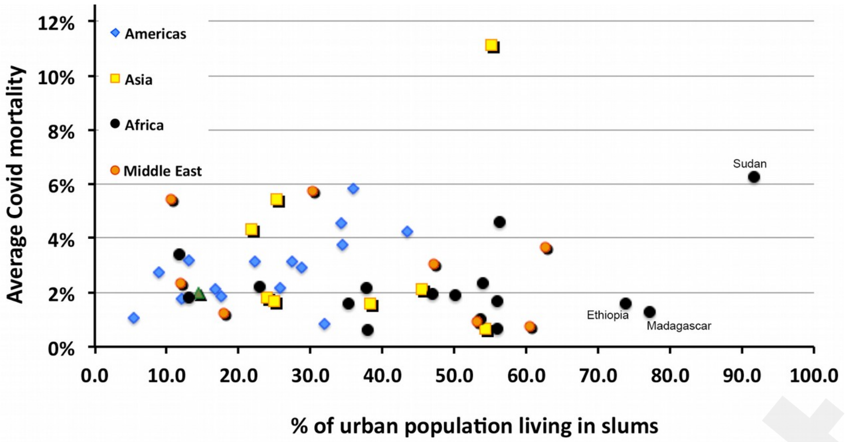


Figure 21. Correlation of COVID-19 mortality with the percentage of the urban population living in slums. The three outlying nations are identified.

5. Summary and conclusions

While this study covering statistics from countries with ~70% of the world's population confirms the early clinical observation that infection by the SARS-CoV-2 virus presents an increased risk to persons over the age of 65, it does not support the suggestions offered by government agencies early in the pandemic that risks are much greater for persons with certain common, potential co-morbidities. Many of the early deaths of elderly patients early in the pandemic occurred in circumstances that likely promoted rather than impeded contagion among person already in a generally poor state of health, likely accompanied by compromised immune functions.

Reference [4] and also the analysis of Koff and Williams in [13] provide plausible explanations. Namely, the virulence of COVID-19 in the elderly is strongly driven by the decrease in adaptive and innate immune responses with aging. Koff and Williams go on to recommend more longitudinal studies in aging populations including assessing the potential of a decrease in the efficacy of vaccines as being "critical to the future of global health."

Many persons who object to strict measures to prevent the spread of the SARS-CoV-2 virus commonly claim that COVID-19 is similar to influenza, is only slightly more lethal, and should be treated in the same manner as influenza as a matter of public policy. In fact, comparing the severity of medical outcomes of COVID-19 with those caused by influenza strains and their resulting pneumonias displays dramatic differences. Promulgating the idea that COVID-19 a "flu-like disease" spreads gross misinformation to the detriment of the public health worldwide.

A broader comparative assessment SARS-CoV-2 and influenza strains against an overall measure of immune system responsiveness to infection would need a global database of an appropriate metric. One candidate could be the as "Wellness Index" proposed by J. Han [14]; however, that metric would require genetic sequencing of large, representative samples of individuals in a broad range of countries. Consequently, at present the possibility of such a database of immune system readiness seems highly doubtful.

Governmental actions can reduce the consequences of SARS-CoV-2infections. Comparing the cases of Germany and Italy may be instructive in this regard. By mid-October 2020, Italy had 150% the numbers of confirmed cases of COVID-19 as Germany. Yet the case fatality ratio in Italy was roughly triple that in Germany. In early 2020, Germany had put in place an extensive network of triage and early treatment centers outside of hospitals. Germany also moved quickly to secure adequate supplies of personal protective equipment. Hence, infected patients were identified early in the course of the disease and were treated in a manner that did not overwhelm the central intensive care facilities in hospitals as happened in the Italian region of Lombardia.

Perhaps a similar lesson comes from comparing the experience in the United States in California and New York through the fall of 2020. The early lockdown in California more than doubled the first duration of the first wave of infections as compared with New York leading to 60% more cases in California yet half the death rate of New York in which medical resources were badly stressed.

Presently, authoritative data on a worldwide, country-to-country basis are not publicly available to evaluate the effectiveness of either prevention and/or treatment modalities. Also unavailable over the full range of those countries included in this analysis are the full range of statistics related to COVID-19 disaggregated with respect to sex. When and if such data become available, expanding the analysis with respect to sex-based differences in testing, contagion, and mortality would prove useful.

The roll out of large-scale vaccination programs during a time when the vaccines are in short supply necessitated schemes for prioritizing recipients. If probability of severe illness is a primary consideration, then the early guess about the risks connected with potential co-morbidities should be replaced with data such as presented here along with detailed clinical evaluations accumulated throughout 2020.

A word of caution: Data used in this study were accumulated before the UK (B.1.1.7), South African (B.1.351), Brazilian (P.1) and Indian (B.1.617) strains of SARS-CoV-2 began to propagate. Initial evidence suggests that these new strains are somewhat more virulent than the original strain. Examining the national case fatality ratio averaged over the duration of the pandemic during early 2021 shows a troubling slight but statistically significant increase in several countries including the United States. Indeed, based on [7], over the period from March 1, 2021 to June 15, 2021, the apparent case fatality ratio in the U.S. looks significantly higher (Figure 22) than that before the appearance of the B.1.1.7 strain.

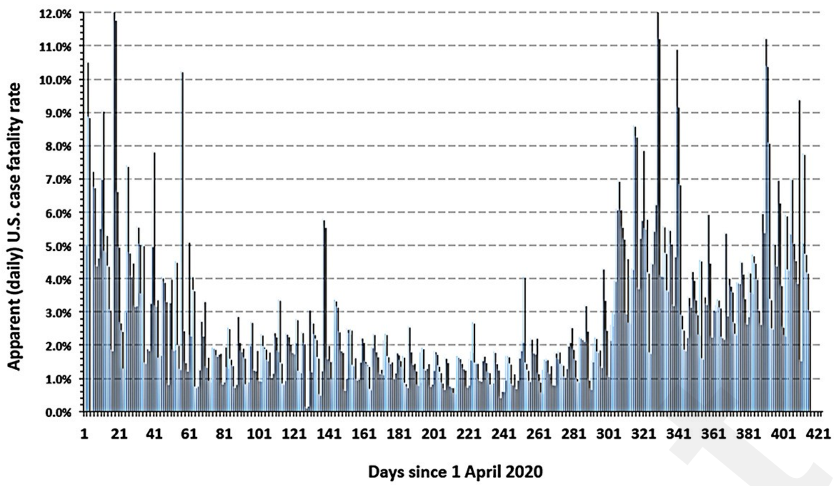


Figure 22. Apparent daily case fatality ratios in the US shows a disturbing trend

Admittedly the data are much noisier than earlier data due to the marked decrease of the daily reports of the number of new cases and deaths. Despite the reduced statistical significance, the trend is troubling. It is too soon to judge whether the increase is a reflection of increased virulence in variants of SARS-CoV-2, whether it is an indication of increased susceptibility and physical and psychological stress on so-called “essential workers,” or whether it is a result of some form of COVID weariness among large portions of national populations. Differences in virulence of the several variant strains now circulating will complicate the interpretation of national data collected in 2021.

Acknowledgements

The author acknowledges his colleagues in the World Federation of Scientists for their encouragement to continue, expand, and report this research. The author’s work has been completely self-supported without any outside funding or other material support.

References

1. “People with Certain Medical Conditions,” https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html, August 14, 2020
2. H. Li, S. Wang, F., Zhong, W. Bao, Y. Li, L. Liu, H. Wang and Y. He, “Age-Dependent Risks of Incidence and Mortality of COVID-19 in Hubei Province and Other Parts of China, Front. Med. (Lausanne) v.7, 190; 2020
3. U.S. CDC, “COVID-19 Hospitalization and Death by Age,“ https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html, August 18, 2020, 3.
4. E. Montecino-Rodriguez, B. Berent-Maoz, and K. Dorshkind, Causes, consequences, and reversal of immune system aging,” **J Clin Invest. 2013 Mar 1; 123(3): 958–965.**
5. M. Kenji , K. Katsushi , Z. Alexander , and C. Gerardo . “Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan,” 2020. Euro Surveill. 2020;25 (10):pii=2000180. https://doi.org/10.2807/1560-7917.
6. World Health Rankings, https://www.worldlifeexpectancy.com/world-health-rankings
7. Worldometer, https://www.worldometers.info/
8. Trading Economics, https://tradingeconomics.com/
9. Https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19roundup/2020-03-26.
10. J. P.A. Ioannidis, C. Axfors, D. G. Contopoulos-Ioannidis, “Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters,” Environmental Research, ISSN: 0013-9351, Vol: 188, 109890
11. S. de Lusignan, J. Dorward, Ana Correa, et al., “Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study,” **Lancet Infect Dis., 2020 Sep; 20(9): 1034–1042. Published online 2020 May 15. doi: 10.1016/S1473-3099(20)30371-6**
12. A. Tamara and D. L.Tahapary, “Obesity as a predictor for a poor prognosis of COVID-19: A systematic review,” Diabetes & Metabolic Syndrome: Clinical Research & Reviews, V.14, Issue 4, July–August 2020, 655-659
13. W.C. Koff and M. A. Williams, “COVID-19 and Immunity in Aging Populations— A New Research Agenda,” N Engl J Med 383; 9 (Aug. 27, 2020)
14. Jian Han, “Wellness Index: measuring the readiness of our immune system,” J Immunol May 1, 2016, 196 (1 Supplement) 54.33, Abstract

Table A.1 Regional variation of COVID-19 case fatality ratio correlations

	World	**Amer-icas**	**Asia**	**Europe**	**Africa**	**Middle East**
kidney disease	26.9%	19.7%	-8.0%	15.6%	28.3%	45.9%
household size	22.5%	42.1%	-5.0%	8.6%	42.2%	20.9%
heart disease	20.4%	-15.7%	5.7%	-14.2%	39.0%	55.0%
asthma	16.5%	-23.3%	3.8%	-20.9%	17.7%	69.1%
diabetes deaths	13.3%	55.9%	9.7%	20.4%	-4.0%	-10.1%
covid deaths per 1M	9.2%	38.0%	6.5%	72.6%	7.7%	-24.6%
%slums	9.0%	23.7%	-42.1%	0.0%	29.1%	22.7%
hypertension	5.1%	-4.8%	38.7%	10.1%	-13.4%	2.4%
flu/pneumonia	3.4%	50.2%	-31.2%	7.4%	-38.4%	62.6%
malnutrition	1.7%	46.9%	28.9%	-5.8%	-15.4%	63.0%
lung disease	1.3%	1.3%	12.4%	20.9%	41.4%	75.6%
random number	0.9%	-21.4%	-15.8%	26.9%	7.2%	-13.2%
diabetes % of population	0.9%	19.5%	-2.3%	-1.1%	75.3%	-45.7%
population	-1.3%	2.9%	12.8%	10.8%	-12.2%	16.0%
obesity % of population	-1.7%	-17.5%	5.9%	24.1%	-1.2%	-41.1%
% of population ≥65	-8.1%	-21.2%	2.1%	19.2%	19.1%	-4.2%
covid cases /1M	-15.3%	-12.6%	-34.3%	13.1%	-5.3%	-47.1%
health expenditures	-15.5%	-23.4%	-7.9%	-12.5%	-2.8%	-45.7%
lung cancers	-15.9%	-40.5%	23.1%	27.0%	-2.6%	-34.7%
life expectancy	-16.3%	-18.5%	-10.1%	13.8%	6.2%	-60.8%
median age	-17.9%	-25.9%	7.5%	14.1%	2.8%	-61.1%
WHO UHC index	-19.7%	5.8%	-2.9%	-10.0%	-3.4%	-65.9%
% of population in cities	-19.7%	-19.9%	-10.5%	-5.7%	0.2%	-57.3%
adjusted GDP	-23.0%	-22.3%	-30.0%	-10.7%	12.5%	-43.8%
covid tests	-28.7%	-6.9%	-27.6%	-32.4%	-38.9%	-49.8%

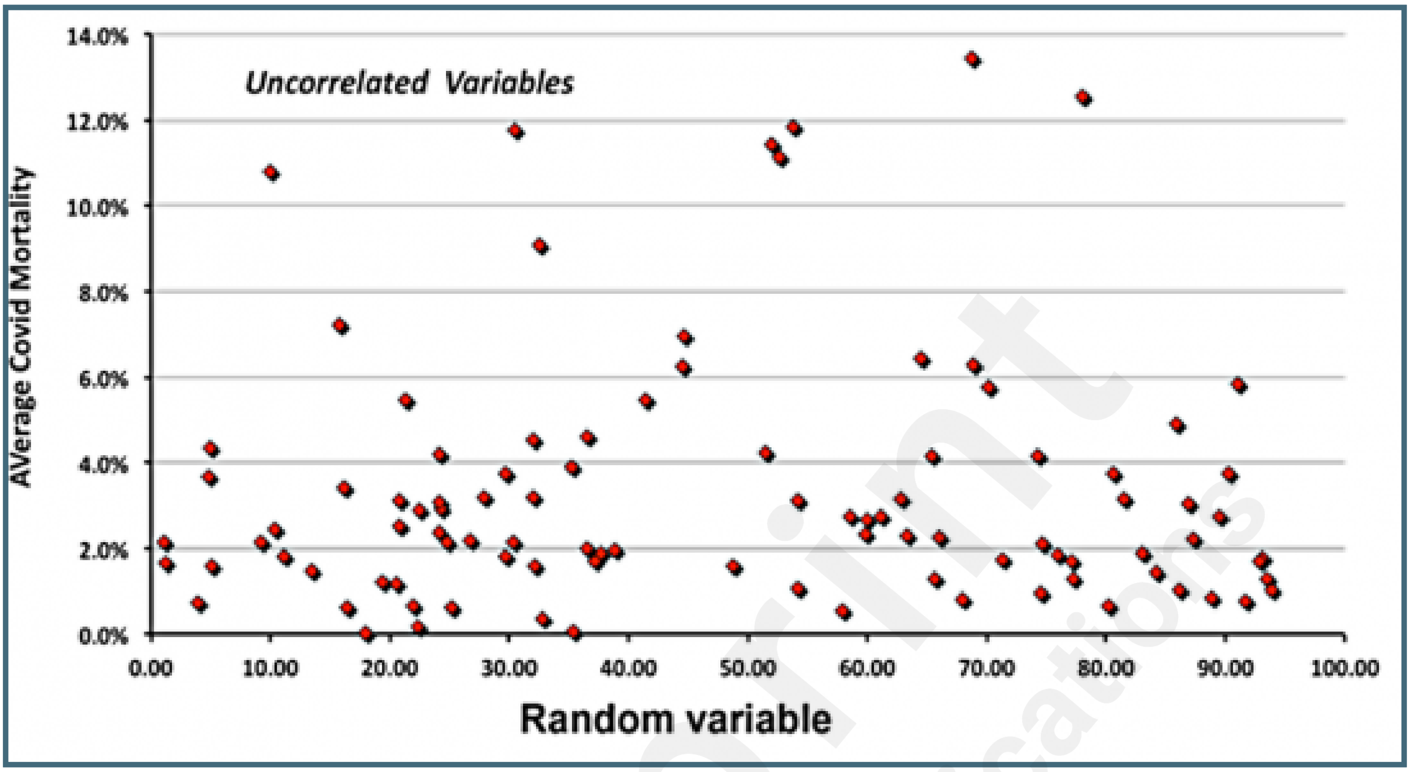
Supplementary Files

Figures

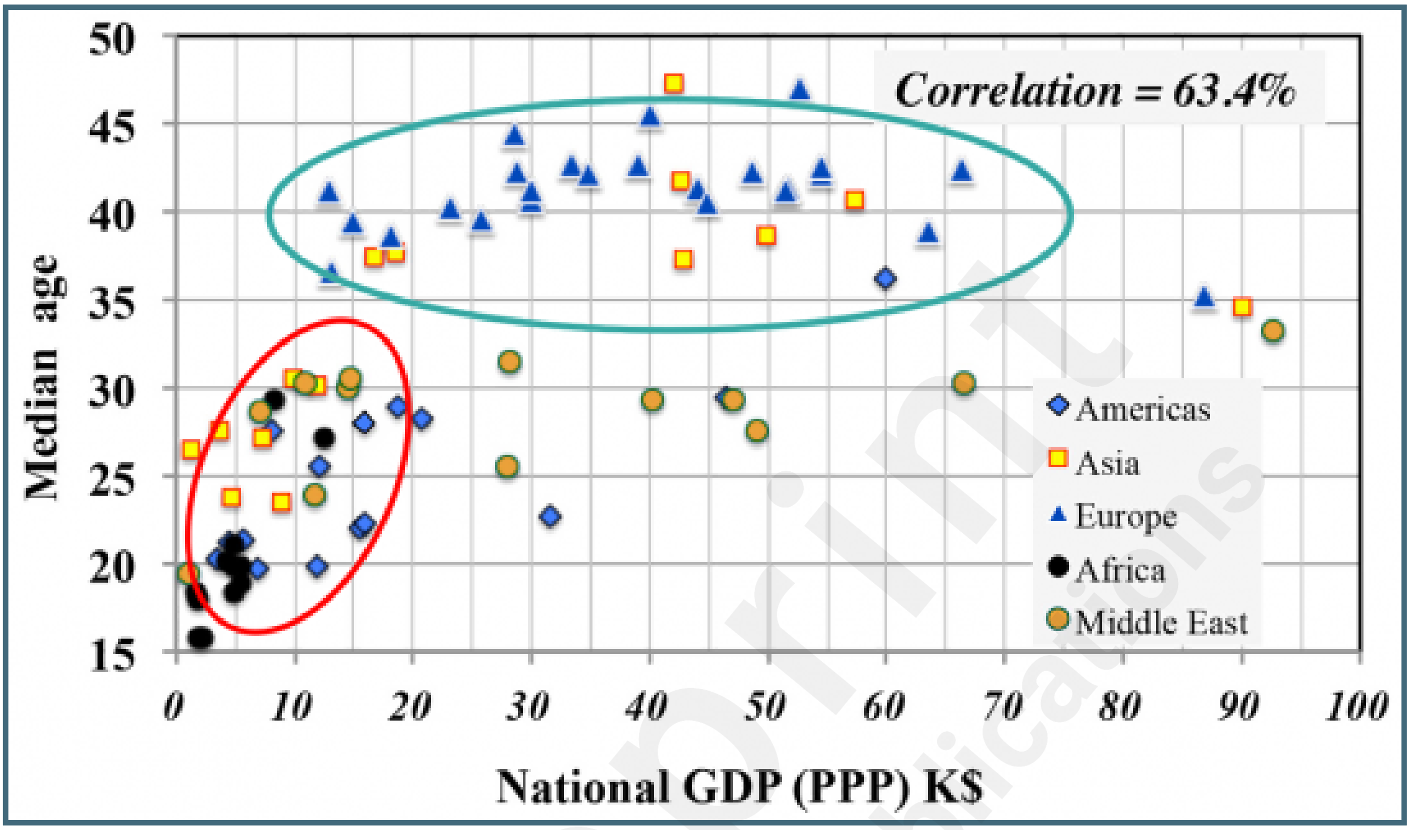
Countries sampled grouped into five regions. As Yemen is a statistical outlier in apparent mortality, many plots omit its data point for visual clarity.

Americas	Asia	Europe	Africa	Middle East
pop = 977 M	*pop = 2504 M*	*pop = 725 M*	*pop = 768 M*	*pop = 487 M*
Argentina	*Australia*	*Albania*	*Algeria*	*Afghanistan*
Bolivia	*Bangladesh*	*Armenia*	*Cameroon*	*Bahrain*
Brazil	*China*	*Austria*	*Congo*	*Egypt*
Canada	*India*	*Azerbaijan*	*Ethiopia*	*Iran*
Chile	*Indonesia*	*Belarus*	*Ghana*	*Iraq*
Columbia	*Japan*	*Belgium*	*Ivory Coast*	*Israel*
Costa Rica	*Kazakhstan*	*Bosnia*	*Kenya*	*Lebanon*
Dominican Republic	*Kyrgyzstan*	*Bulgaria*	*Libya*	*Kuwait*
Ecuador	*Korea*	*Croatia*	*Madagascar*	*Oman*
El Salvador	*Malaysia*	*Czechia*	*Mali*	*Qatar*
Guatemala	*Nepal*	*Denmark*	*Morocco*	*Saudi Arabia*
Honduras	*New Zealand*	*Estonia*	*Nigeria*	*Turkey*
Mexico	*Pakistan*	*Finland*	*South Africa*	*UAE*
Panama	*Philipines*	*France*	*Sudan*	*Uzbekistan*
Paraguay	*Singapore*	*Germany*	*Uganda*	*Yemen*
Peru	*Thailand*	*Greece*	*Zambia*	
USA	*Taiwan*	*Hungary*		
Venezuela		*Ireland*		
		Italy		
		Macedonia		
		Moldova		
		Netherlands		
		Norway		
		Poland		
		Portugal		
		Romania		
		Russia		
		Serbia		
		Spain		
		Sweden		
		Switzerland		
		Ukraine		
		UK		

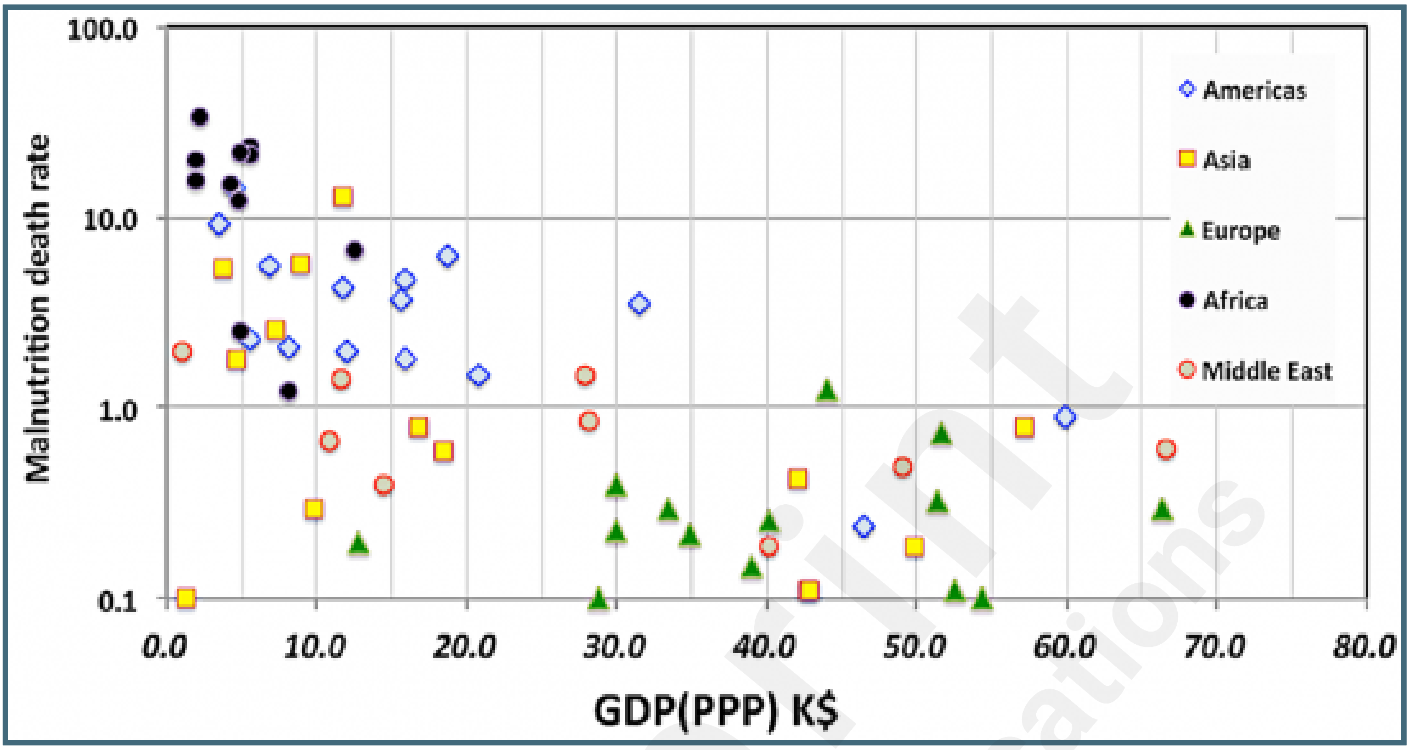
A plot of the variables uncorrelated by construction.



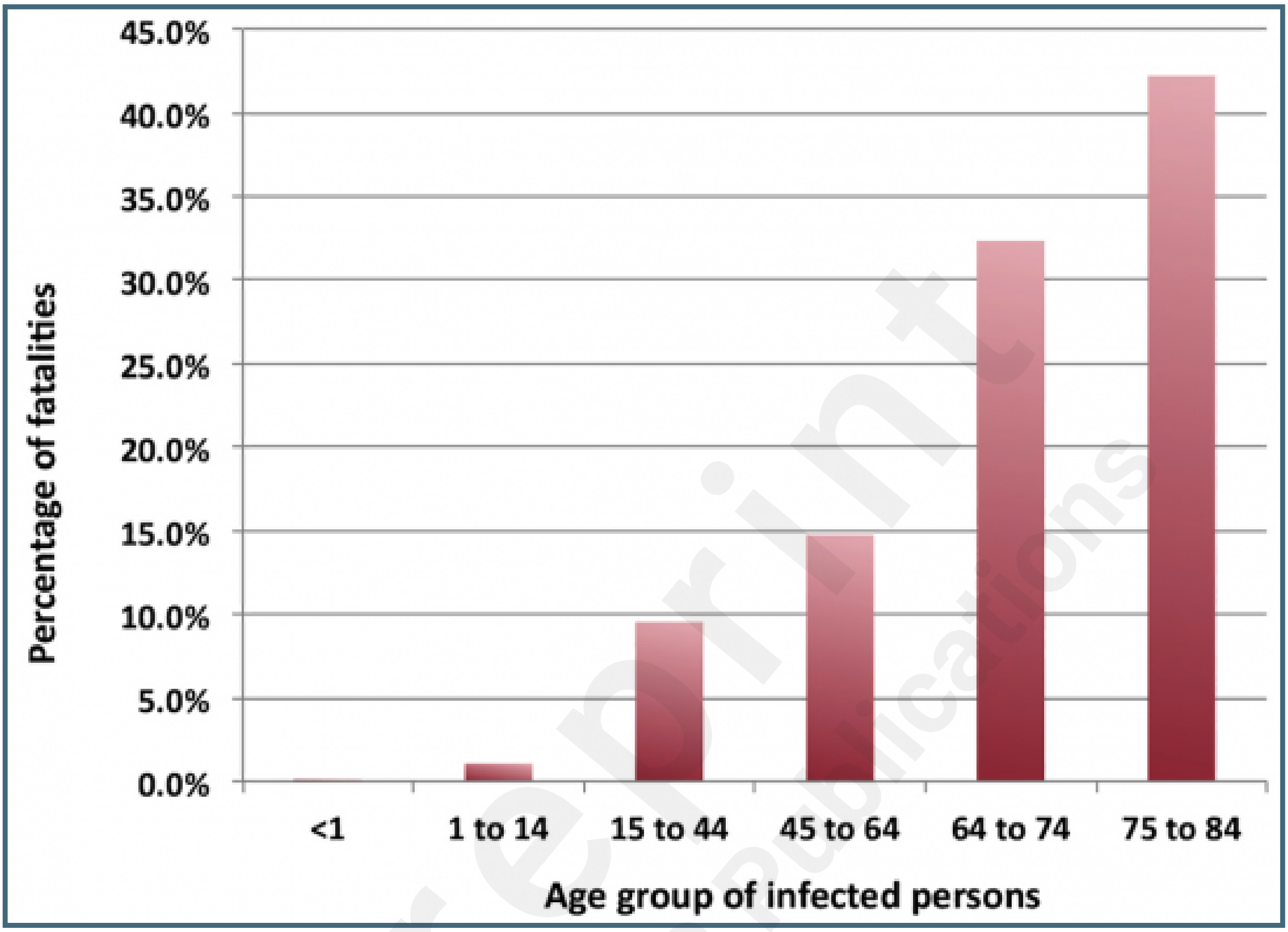
The plot of GDP-PPP corrected for purchasing power versus median age in countries from the five regions under study.



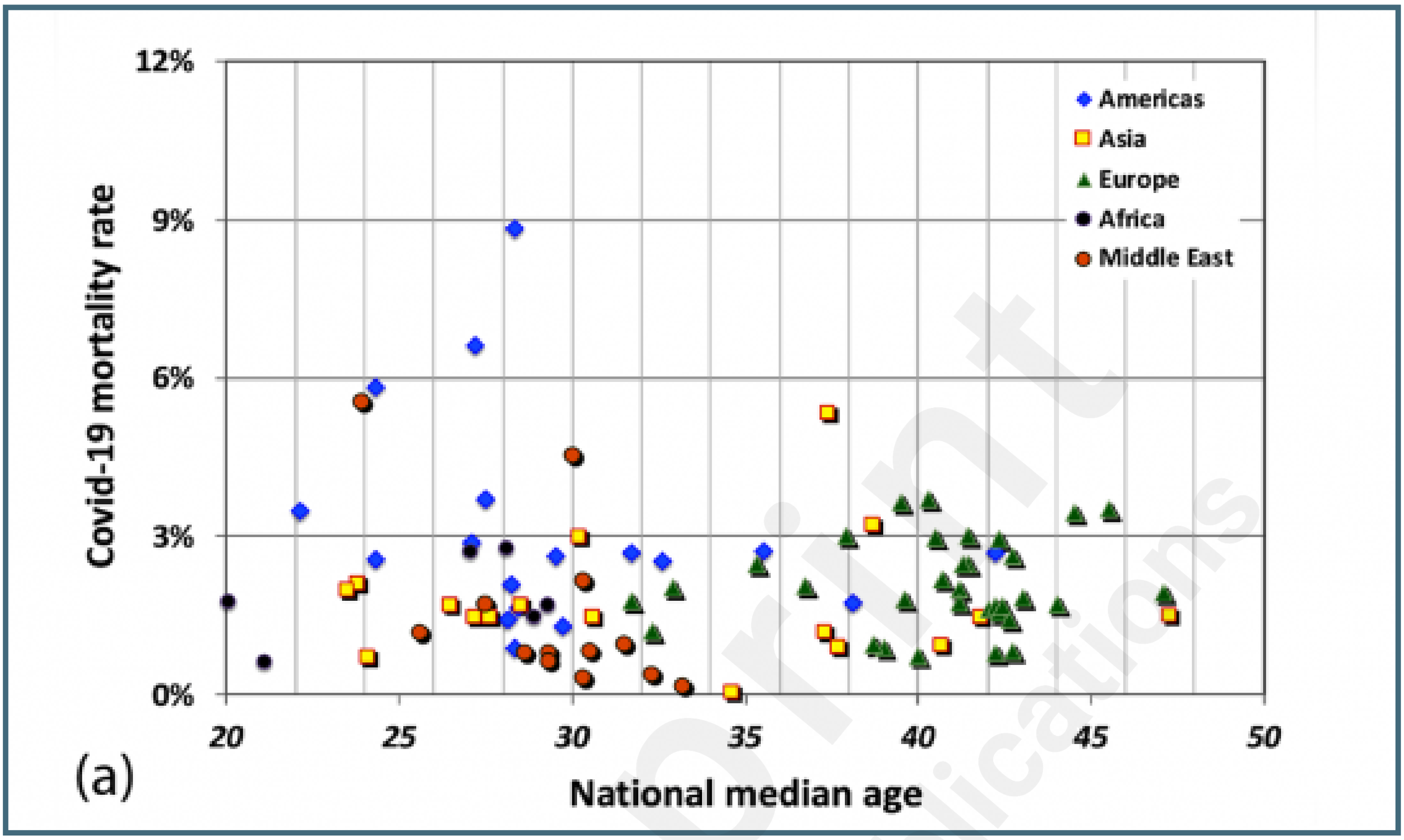
It is not surprising that poverty is correlated with malnutrition.



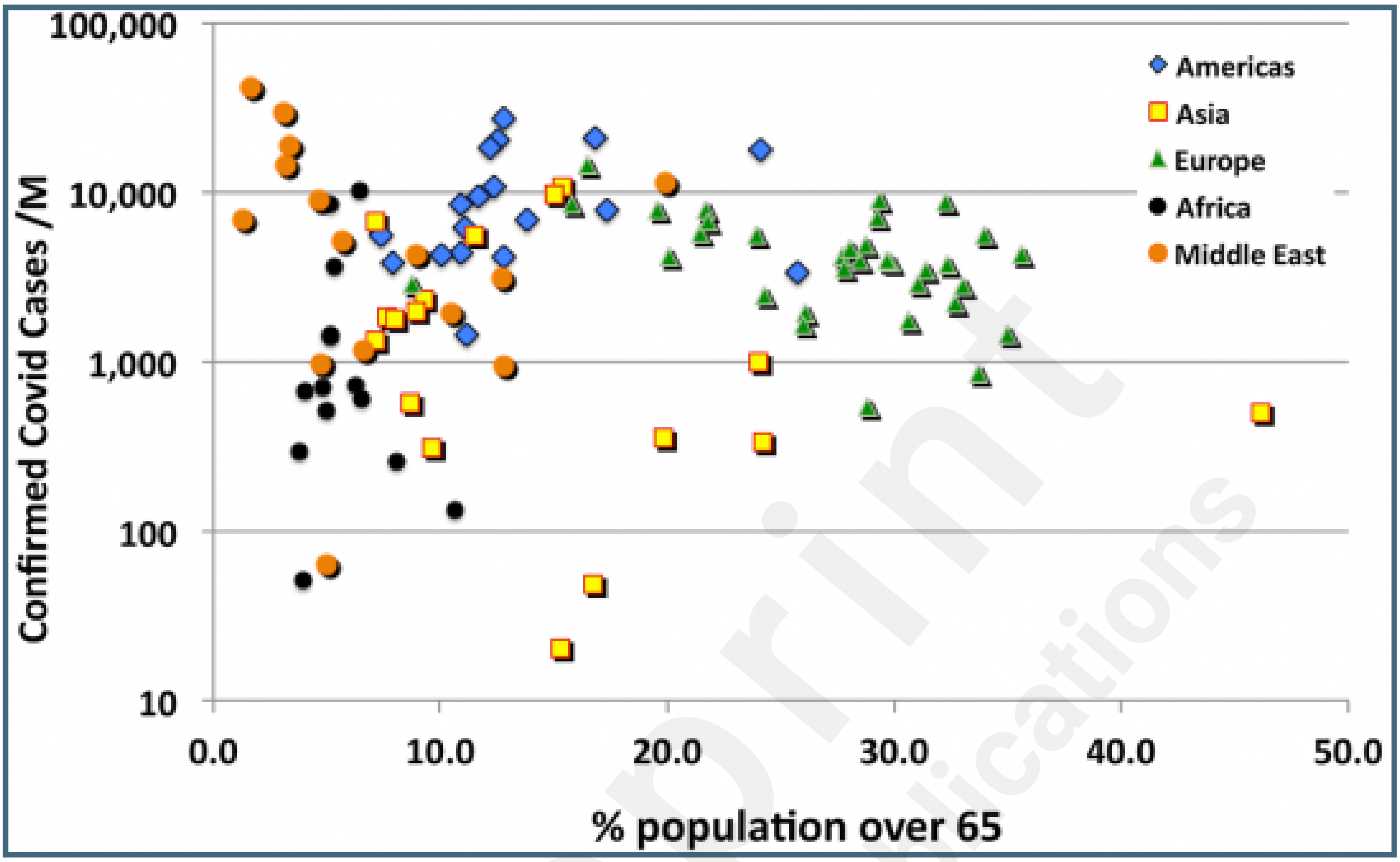
Deaths attributed to COVID-19 by the UK Office of National Statistics.



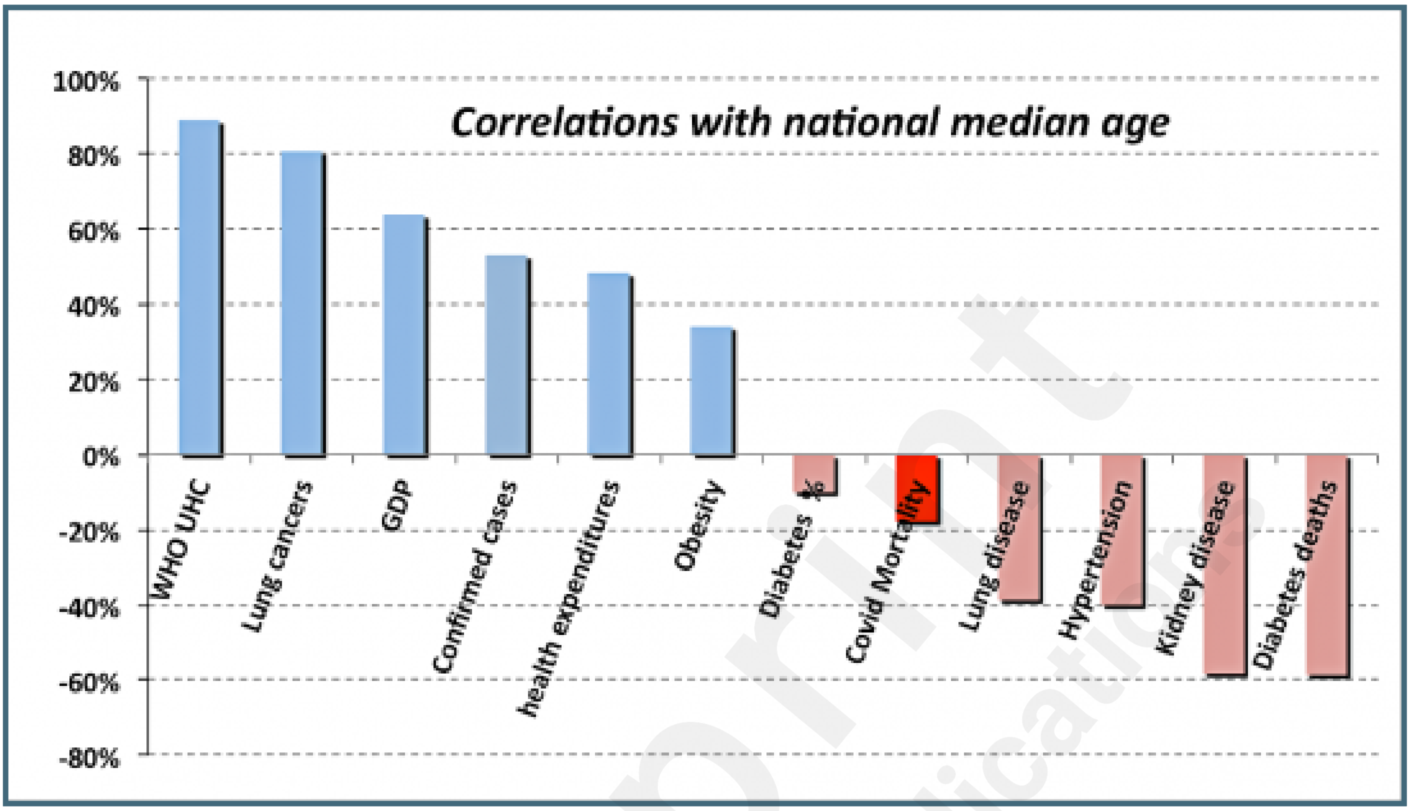
National median age versus CFR for the five regions.



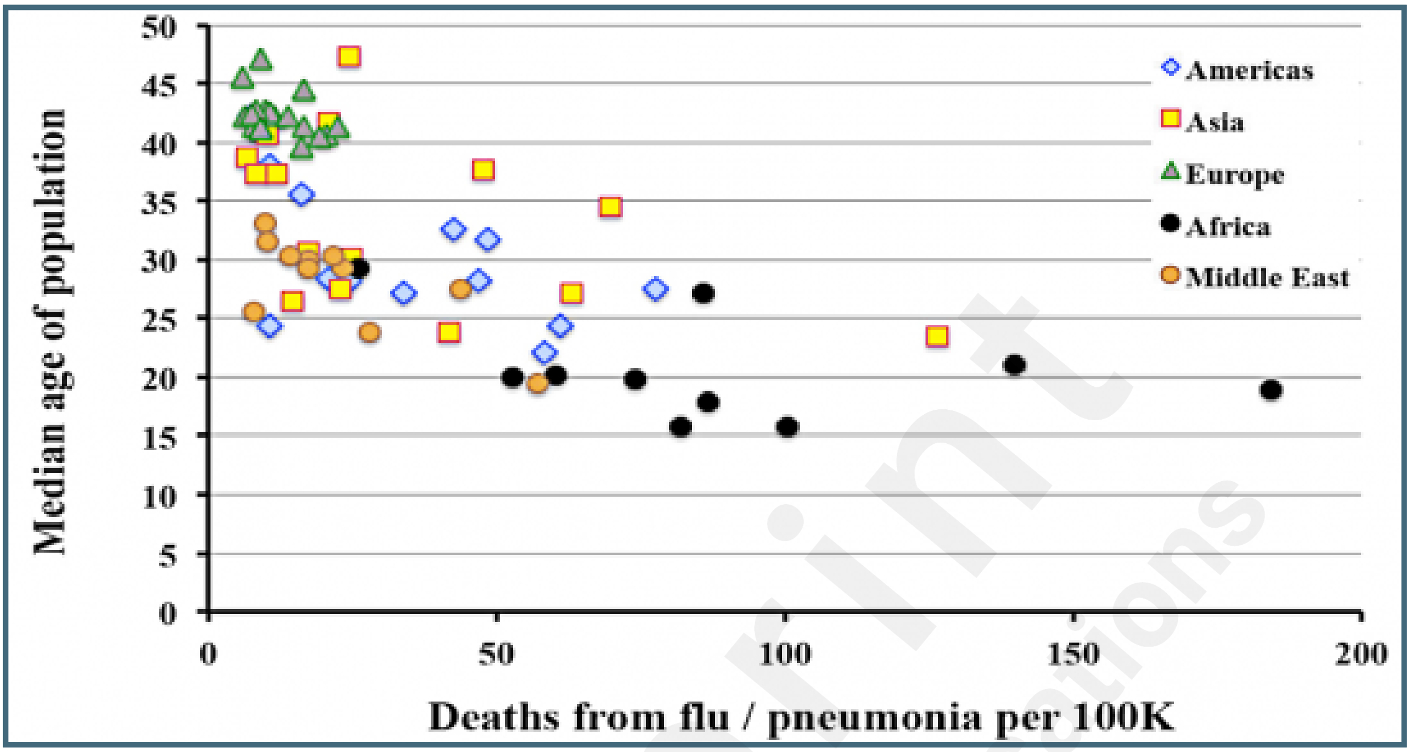
Confirmed COVID-19 cases as a function percent of population older than 65.



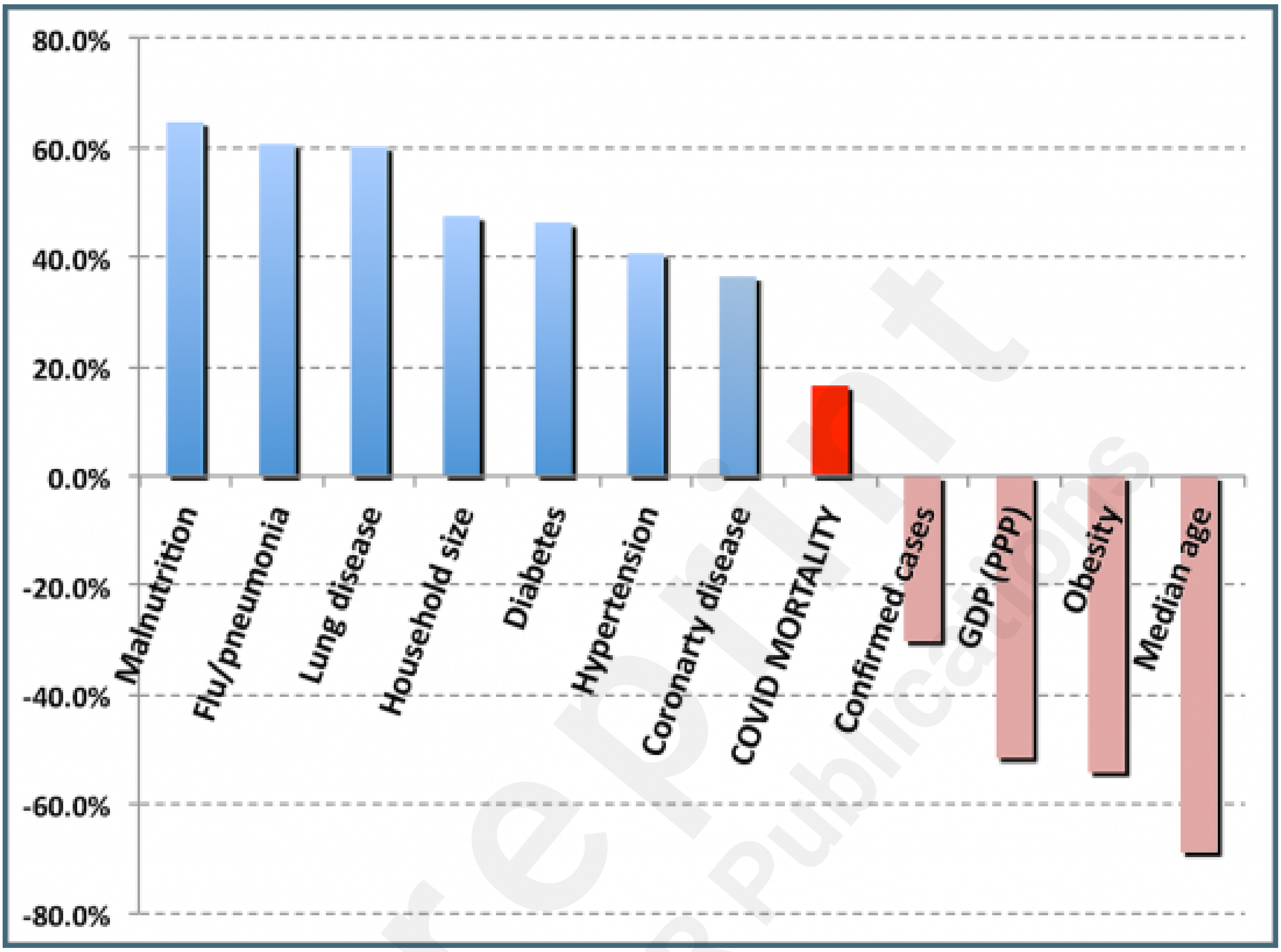
Correlations of potential risk factors with median age.



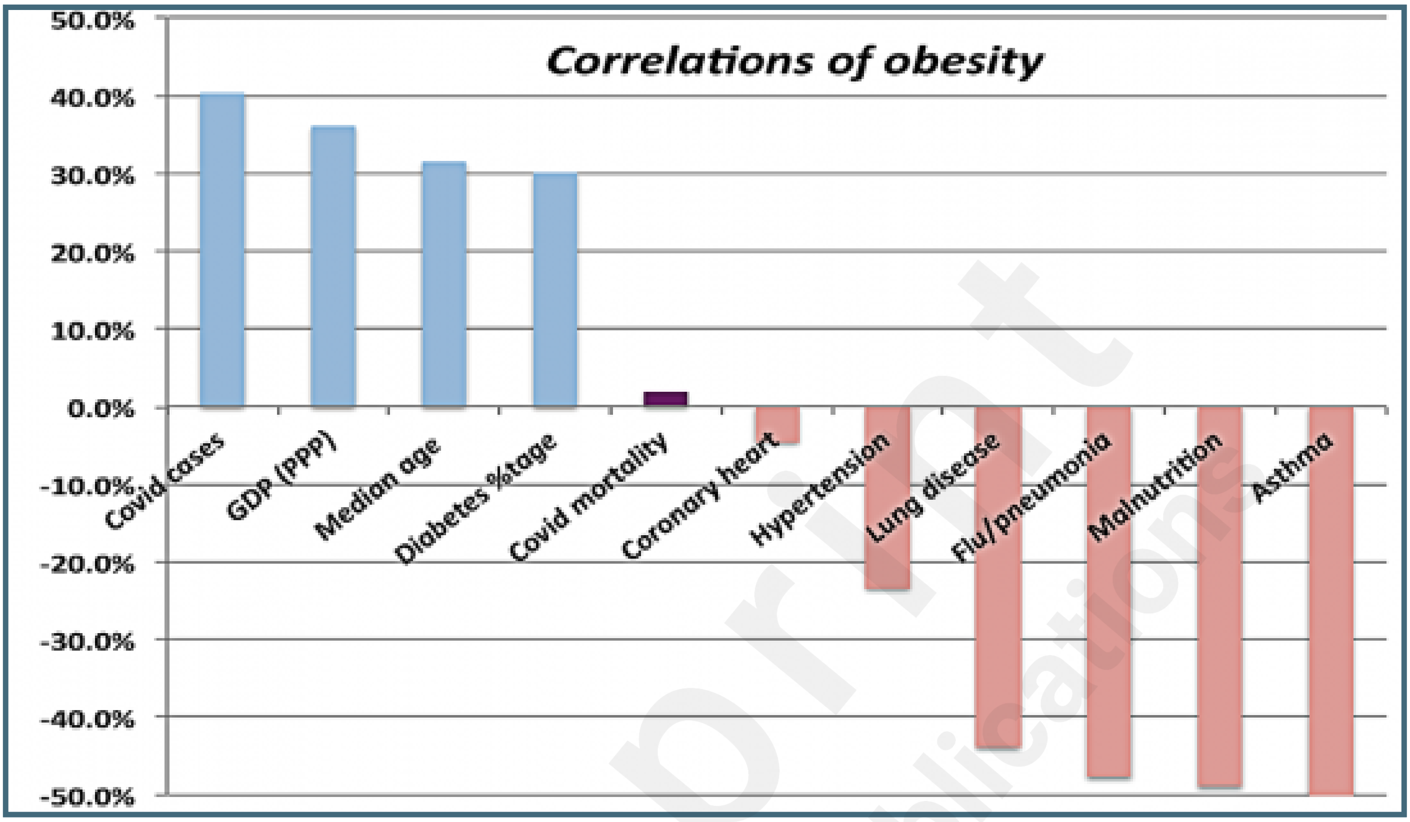
Incidence of flu related pneumonia deaths as a function of national median age.



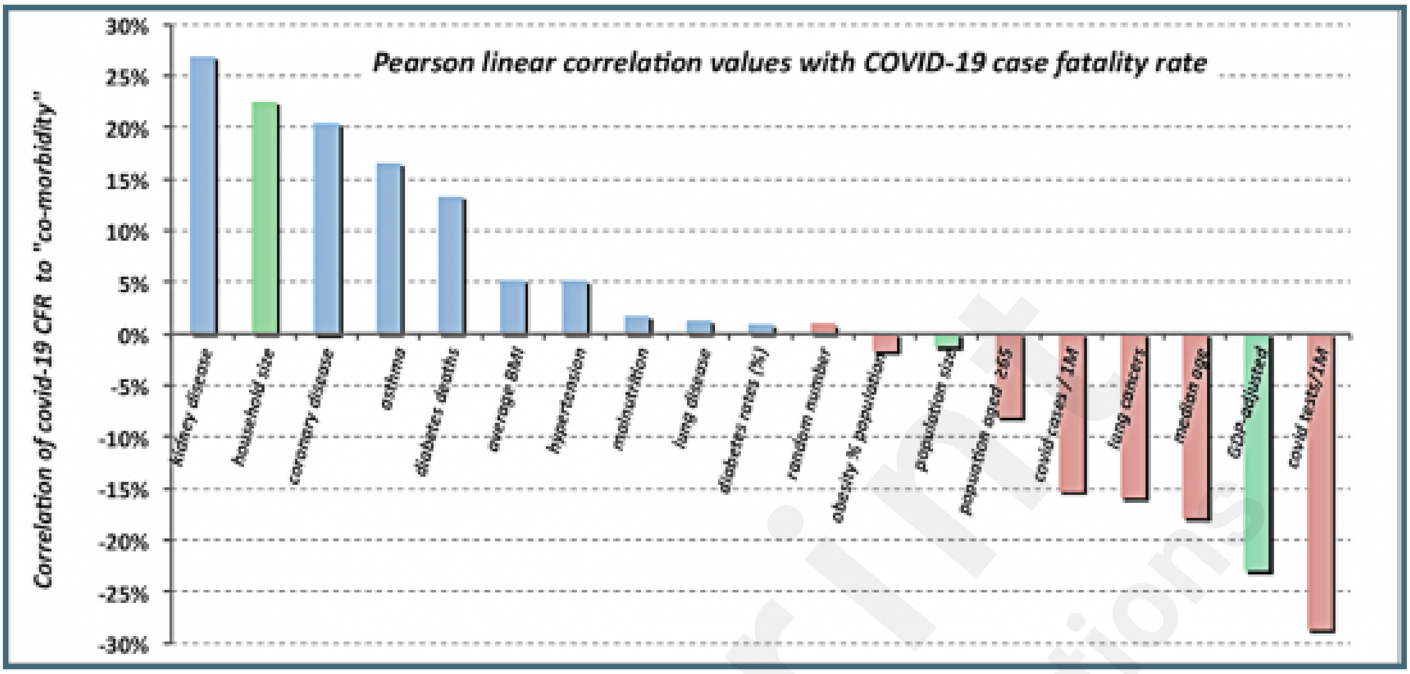
Correlation of severe asthma with COVID-19 CFR.



Correlations of obesity rates with COVID-19 mortality and other conditions. For most of conditions, rates are based on deaths per 100K persons.



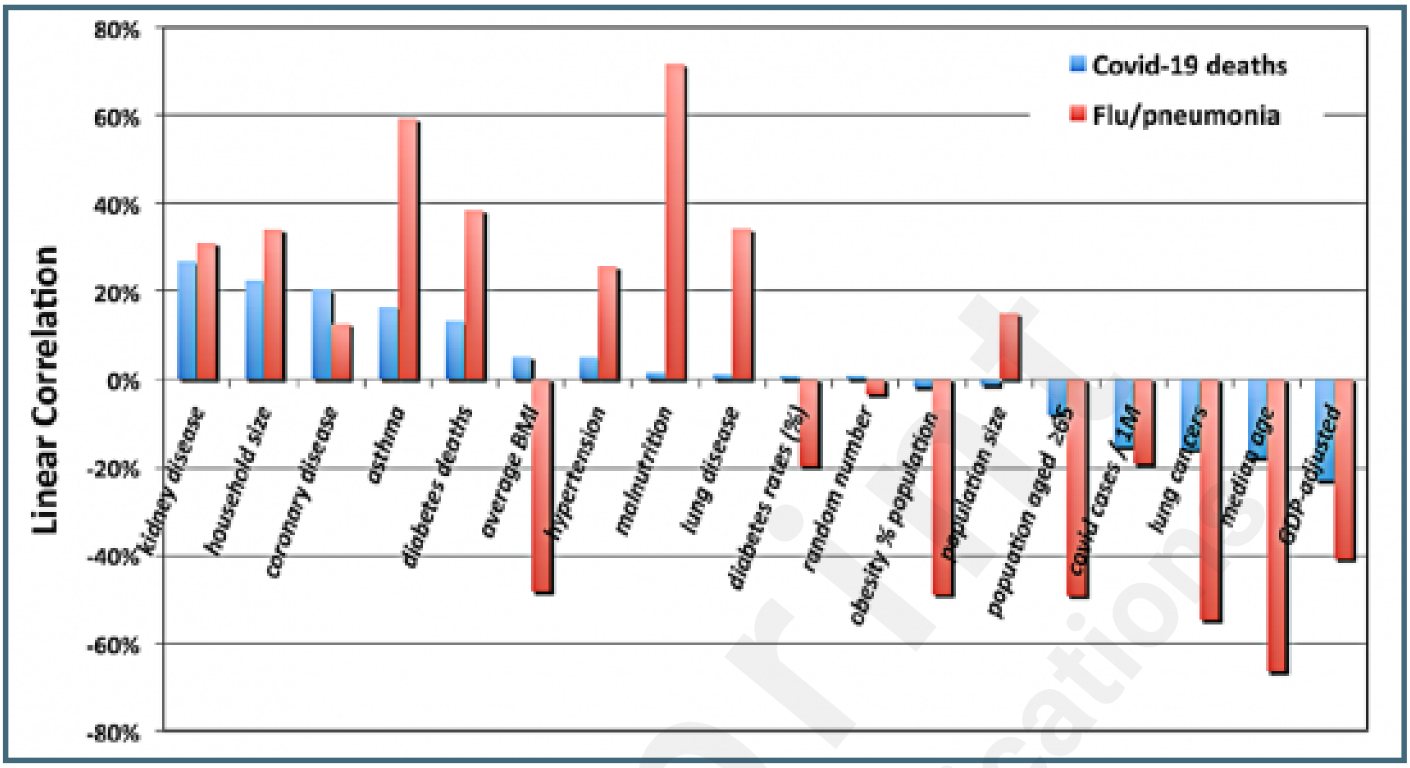
Summary of linear correlations with national COVID-19 CFR data.



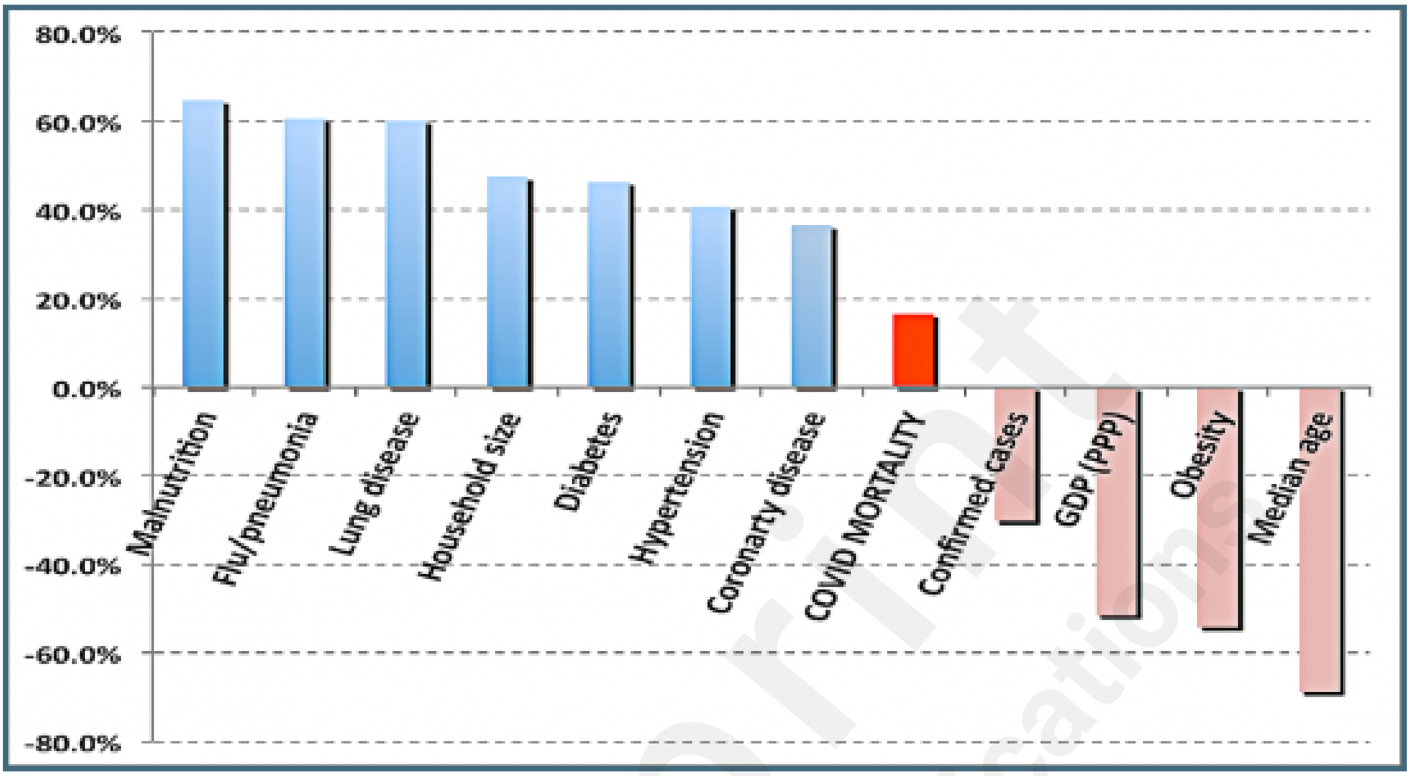
Correlation with national values of apparent COVID-19 CFR.

Potential co-factor	Date of COVID-19 CFR statistics		
	30-Dec	20-Nov	16-Oct
Kidney disease	26.9%	28.9%	17.6%
household size	22.5%	22.8%	12.6%
heart disease	20.4%	19.4%	9.9%
Asthma	16.5%	16.8%	9.1%
diabetes deaths	13.3%	14.8%	5.0%
covid deaths per 1M	9.2%	7.0%	17.0%
%slums	9.0%	7.2%	5.9%
hypertension	5.1%	4.9%	-1.1%
flu/pneumonia	3.4%	4.0%	-2.0%
malnutrition	1.7%	0.2%	-3.7%
lung disease	1.3%	2.4%	-11.2%
Random number	0.9%	-2.4%	2.6%
Diabetes (% population)	0.9%	4.6%	-4.3%
population	-1.3%	0.0%	-1.4%
obesity % population	-1.7%	-0.6%	1.4%
% population ≥65	-8.1%	-10.3%	2.8%
covid cases /1M	-15.3%	-16.3%	-8.6%
health care expenditure	-15.5%	-14.3%	-2.0%
lung cancers	-15.9%	-17.9%	-9.8%
life expectancy	-16.3%	-15.2%	-5.5%
median age	-17.9%	-19.1%	-7.4%
WHO UHC index	-19.7%	-16.8%	-7.6%
% population in cities	-19.7%	-17.7%	-12.1%
adjusted GDP	-23.0%	-21.5%	-11.9%
covid tests per 1M	-28.7%	-25.7%	-11.1%

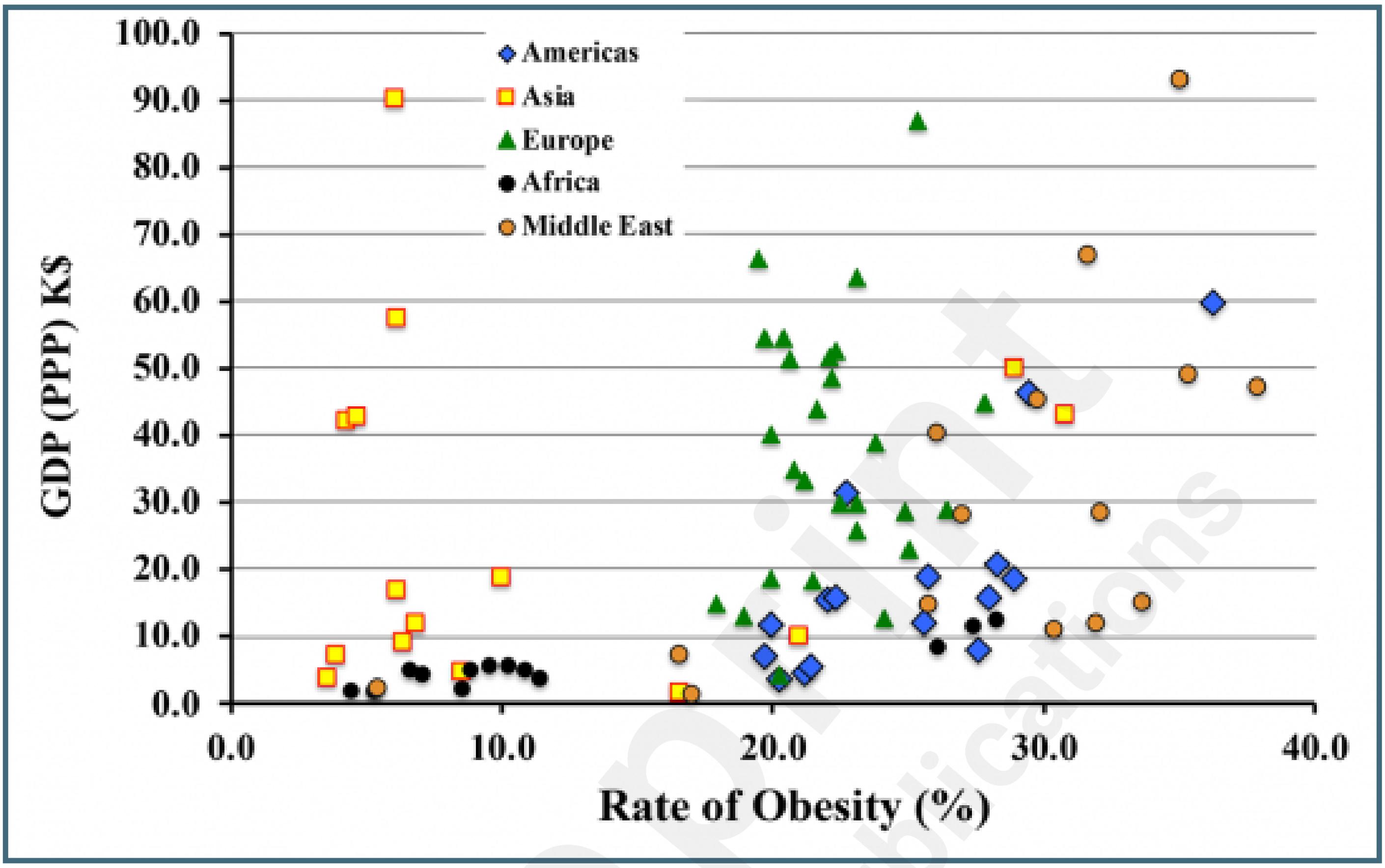
Contrast between correlations of COVID-19 with of flu-induced pneumonia.



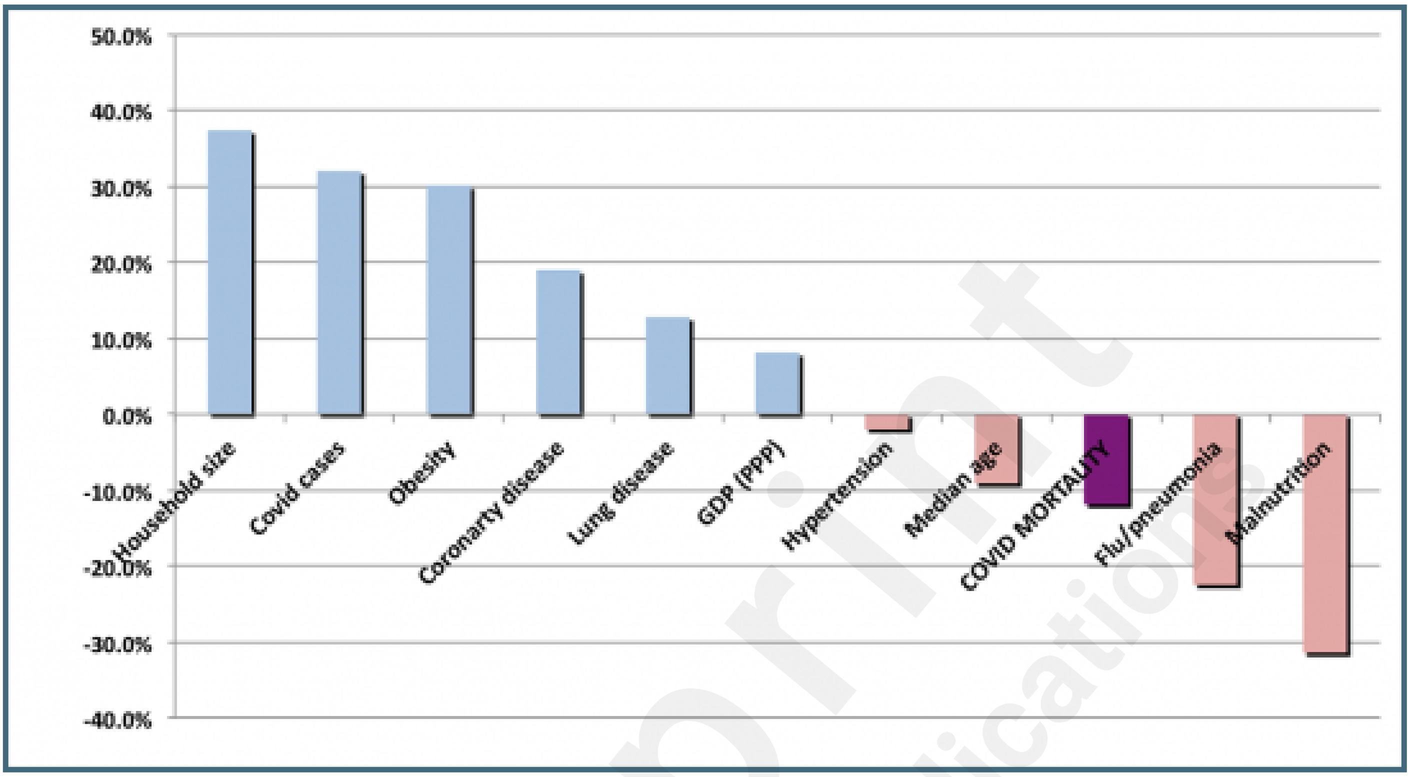
Cross-correlations of asthma with several disorders that might affect the outcome of COVID-19 cases.



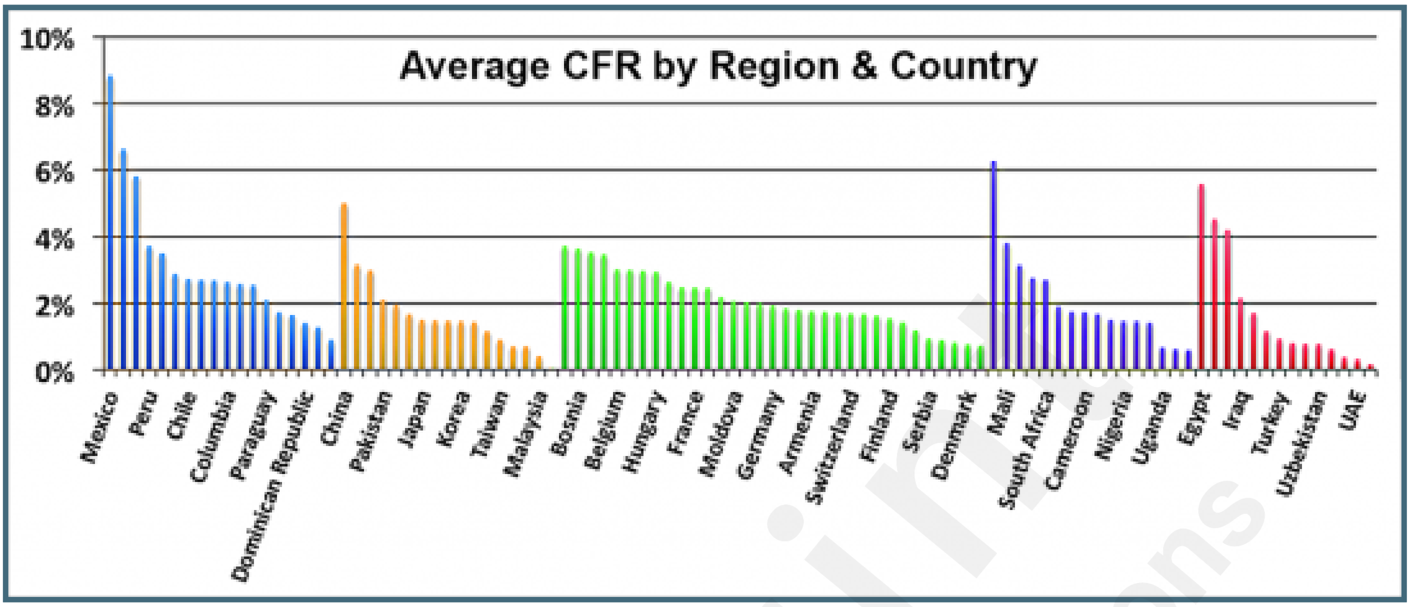
Or is wealth of the nation a driver of obesity?



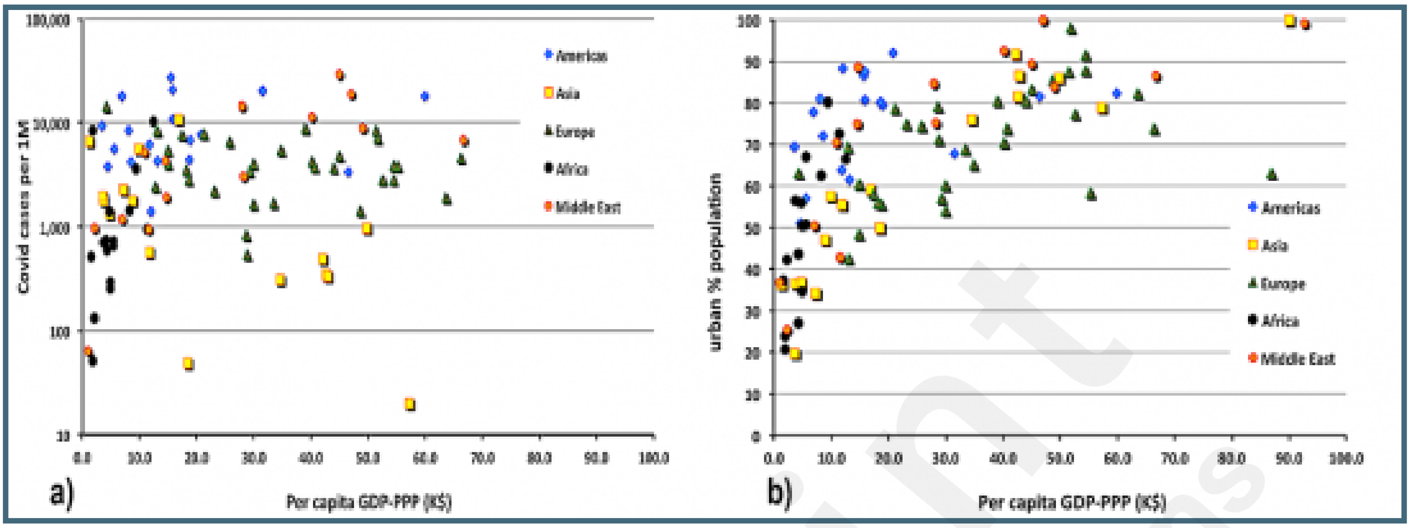
Correlations of national rates of diabetes mellitus with other medical and economic conditions. The correlation with COVID-19 mortality is the purple bar.



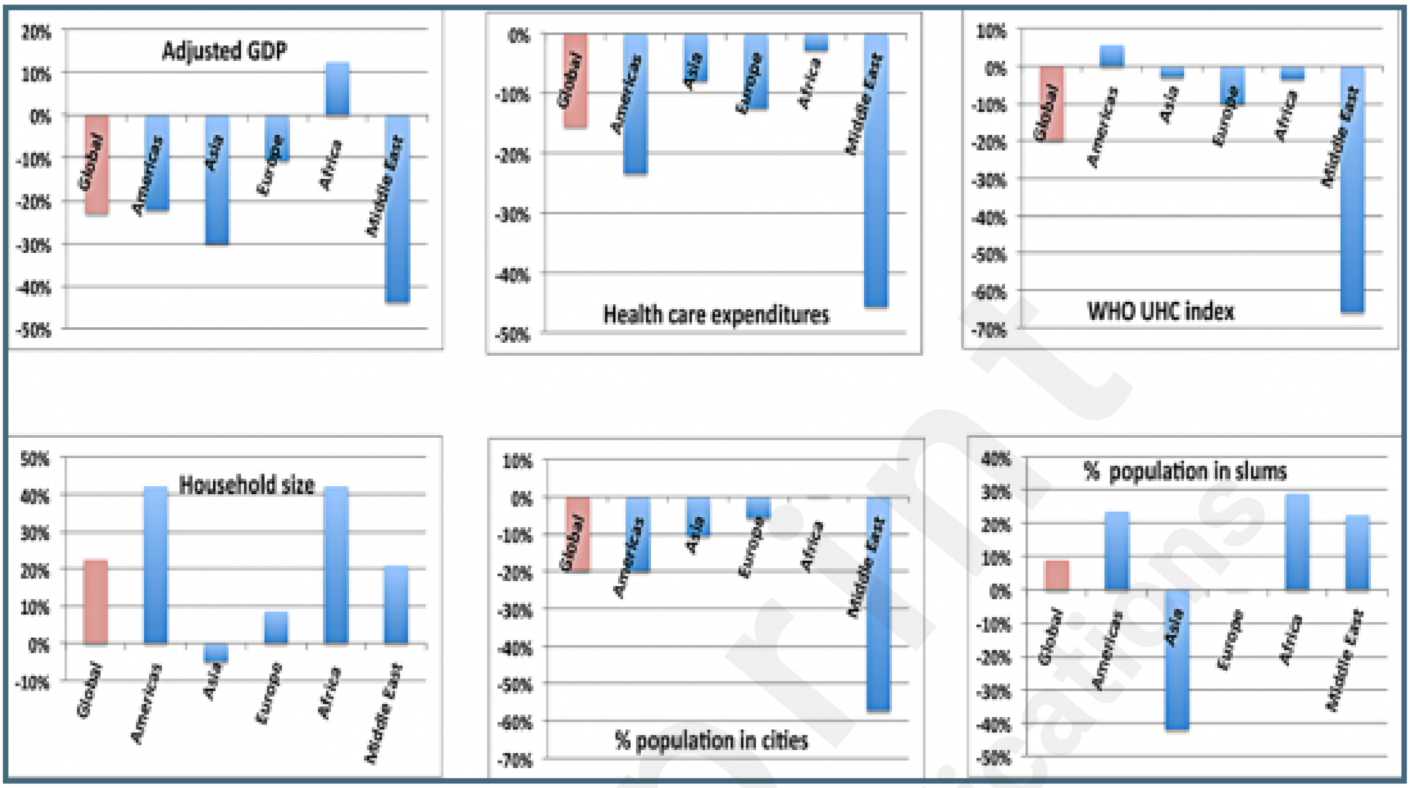
Average case fatality rate by country and region.



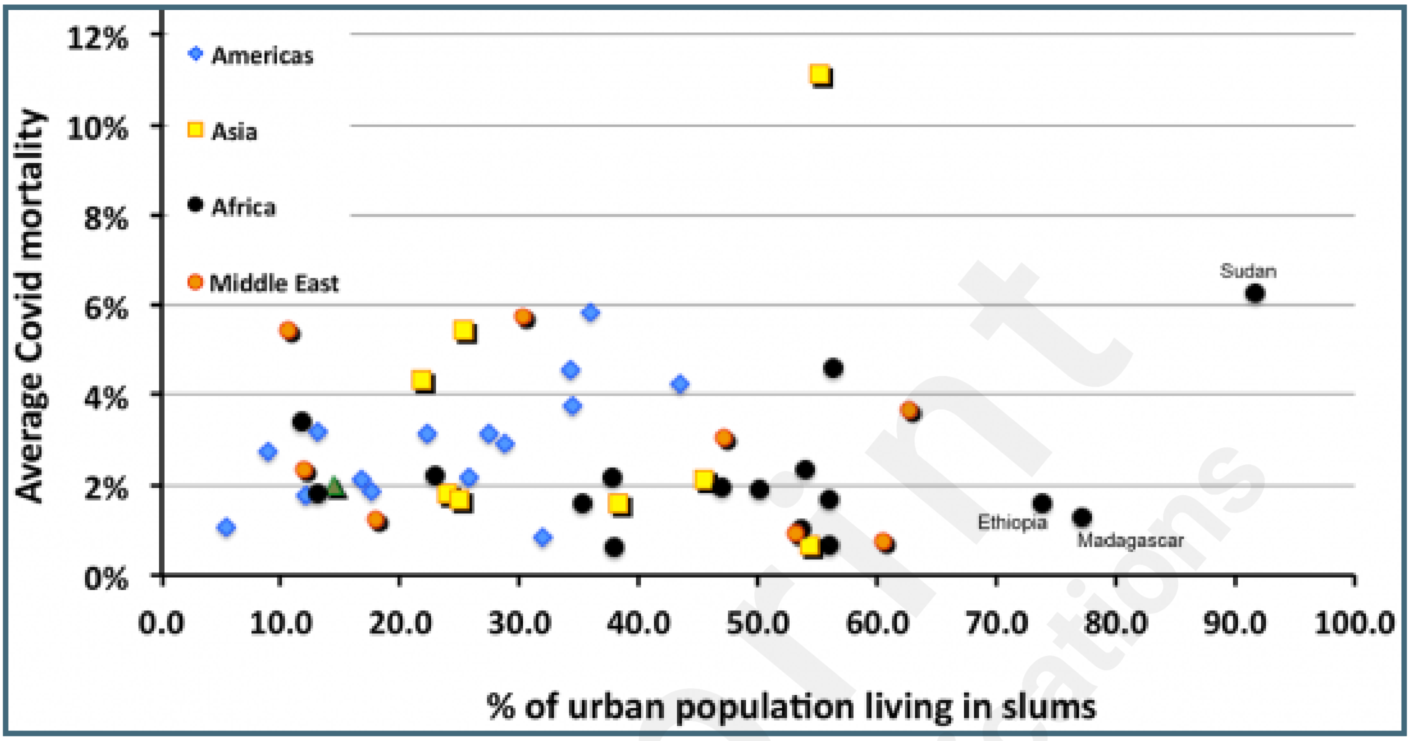
a) The distribution of COVID-19 cases with national GDP-PPP. b) The degree of urbanization with increasing GDP.



The impact of economic co-factors on CFR varies strongly by region.



Correlation of COVID-19 mortality with the percentage of the urban population living in slums. The three outlying nations are identified.



Apparent daily case fatality ratios in the US shows a disturbing trend.