

1

1 **TITLE PAGE**

2

3 **FREQUENCY OF TESTING FOR COVID 19 INFECTION AND THE PRESENCE OF HIGHER NUMBER OF**
4 **AVAILABLE BEDS PER COUNTRY PREDICT OUTCOMES WITH THE INFECTION, NOT GDP OF THE**
5 **COUNTRY – A DESCRIPTIVE STATISTICAL ANALYSIS**

6 Authors: Binayak Sinha¹, Sumit Sengupta², Samit Ghosal³

7 1. Consultant Endocrinologist. AMRI Hospitals, Kolkata, India.

8 2. Consultant Pulmonologist. AMRI Hospitals, Kolkata, India.

9 3. Consultant Endocrinologist. Nightingale Hospital, Kolkata, India.

10 Corresponding author: Samit Ghosal. Email: ramdasghosal@gmail.com. Phone number: +91-
11 9674328281

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33 **ABSTRACT:**

34 ***Introduction:***

35 The novel coronavirus epidemic which originated in late 2019 from China has wreaked havoc on
36 millions across the world with illness, death and socioeconomic recession. As of now no valid
37 treatment or preventative strategy has evolved worldwide and governments across the world have
38 been forced to take the draconian step of social isolation in communities by enforcing “lockdowns”.

39 ***Aim of this Study:***

40 This study aims to correlate the rates of infection with the novel coronavirus and total deaths as the
41 primary output variable. In addition the strength of association between infection rates and total
42 death in comparison to GDP share of the respective countries, physicians, hospital beds and rates of
43 testing for COVID 19 infection per thousand patients, is being assessed, in a bid to develop a model
44 which would help to develop tools to reduce the impact of this disease.

45 ***Material & Methods***

46 Data relating to number of cases, severity, cases recovered and deaths worldwide and specifically for
47 the top six countries affected was collected from the WHO COVID-19 situation report which is being
48 updated on a daily basis till 22nd March 2020, the date of analysis. Additional data related to GDP,
49 physician and hospital bed per 1000 patients were procured from the World Bank database. All data
50 were collected in a file in CSV format. Analysis was conducted in Jupyter notebook with Python 3.8.2
51 software and also with XL-Stat statistical software for excel. The analytical strategy was descriptive
52 with no inferential overtones.

53 ***Results:***

54 COVID 19 infection strongly correlates with total deaths ($r: 0.89$), with a predicted death rate of 25
55 patients per 1000 affected. There was no correlation between the GDP growth of the country and
56 number of treating physicians/1000 patient population with any COVID 19 related outcome.
57 However there was a negative correlation between COVID 19-related deaths and the number of
58 beds available per 1000 population [$r=-0.34$]. Importantly there is an inverse correlation between
59 the number of tests conducted per million population with the rates of active infections [$r=-0.12$],
60 new cases [$r=-0.38$] and new deaths [$r=-0.28$] in COVID 19.

61 ***Conclusion:***

62 This is the first study to assess parameters other than age and sex and sets out a robust dataset
63 which indicates an increased risk of worsening outcomes with lesser number of beds and testing,
64 suggesting that the need of the hour is to increase available bed numbers and to increase rates of
65 testing.

66

67 **Keywords:** COVID-19; mortality; recovery rates; testing; hospital beds; correlation.

68

69

70

71

72 **1. Introduction:**

73 The outbreak of coronavirus pandemic which started from late 2019 has created a havoc on human
74 civilization. We are not only faced with the disease-related morbidity and mortality, but also
75 extensive economic, social and psychological upheaval. In view of the rapidity of spread of infection
76 and a high mortality rate the present day health-care system is firmly focussed on both prevention
77 and treatment. Research required to arrive at a definitive therapeutic intervention is going to take
78 some time, but there are encouraging reports from non-randomised trials with drugs like
79 chloroquine/hydroxychloroquine, remdesivir, azithromycin.[1,2] Governments across the globe have
80 taken on this calamity of herculean proportions head on with lockdown of cities in an attempt to
81 contain further spread of the disease.

82 However, it is also extremely important to analyse the evolving data at hand to get an understanding
83 of the direction we are headed. Such an analysis might also help us prepare more effectively to plan
84 for the future and fortify preventive strategies giving researchers ample space and time to come up
85 with potent and effective cure for COVID-19 pneumonia.

86 **2. Aims of this analysis:**

87 This analysis aims at finding a correlation between the rates of infection and total death. In addition
88 we also looked at the strength of association between infection rates and death in comparison to
89 GDP share of the respective countries, physicians & hospital beds/1000 patients and rates of testing
90 for COVID 19 infection per 1 million patients. Lastly, we plan to create a regression model which
91 would help predicting the death rates in those infected with COVID-19.

92 **3. Materials & Methods:**

93 We collected the data from the WHO COVID-19 situation report which is being updated on a daily
94 basis.[3] Since we needed to do the analysis at a point in time we selected the data available as of
95 22nd March as the date of analysis. Since data relating to age and sex and their correlation to COVID
96 19 outcomes have already been published, those data were not collated [4]. Additional data related
97 to GDP, physician and hospital bed per 1000 patients were procured from the World Bank database.
98 [5,6,7]

99 All data were collected in a file in CSV file format. Analysis was conducted in Jupyter notebook with
100 Python 3.8.2 software and also with XL-Stat statistical software for excel.

101 The analytical strategy was descriptive with no inferential overtones.

102 **4. Results:**

103 **4.1 Patients infected with COVID-19 and mortality:**

104 A very strong correlation ($r: 0.89$) was found between the patients infected with COVID-19 and total
105 death. (Fig 1) There was a 89% association between death and getting infected ($P < 0.001$, 95% CI:
106 0.046-0.056).

107 For every 1000 patient the predictive death count was 25.

108 The regression model for prediction was: $y(\text{total number of death from COVID-19}) = -26.72 +$
109 $0.05 * \text{total infected population}$. (Fig 2)

110

111 **4.2 Relationship between GDP, physician and hospital bed per 1000 patient population and**
112 **infection rate & mortality:**

113 There was poor correlation between GDP growth and infection rate [$r=0.04$] or death from COVID-19
114 [$r=0.01$]. As a matter of fact the top 5 infected countries bearing most of the brunt from COVID-19
115 infection shares 48% of the Global GDP between them. (Table 1)

116 In view of the wide heterogeneity of data entry across the different countries including most of the
117 input parameters there were no correlation trend visible as far as infection rates and its relationship
118 with the physicians or hospital bed per 1000 patient population was concerned. (Fig 3) However,
119 analysing the top 15 infected countries gave a definitive trend. (Fig 3) New deaths from COVID 19
120 was inversely correlated with the number of available hospital beds per 1000 infected patients
121 admitted [$r=-0.34$]. However, there was no trend observed between the COVID-19 related deaths
122 and physicians per 1000 population [$r=-.011$]. (Fig 3)

123 Amongst the top 6 infected countries Germany had the best physician and hospital bed per 1000
124 patient population.

125

126 **4.3 Correlation between testing for COVID-19 and other variables:**

127 In view of increased frequency of testing seen predominantly in the top 15 countries, analysis was
128 done on the same to improve the yield of outcomes. Increased testing frequency was associated
129 with a positive correlation with recovery [$r=0.18$] and was negatively correlated with new cases, [$r=-$
130 0.38] new deaths, [$r=-0.28$] active cases [$r=-0.12$]. (Fig 3)

131

132 **5. Discussion:**

133 The number of people diagnosed with COVID-19 have crossed 3 lakhs with more than 12,000 deaths
134 at the time of this analysis. The figures keep climbing up steeply with each passing days. It is a crisis
135 of mammoth proportion requiring the cooperation from the individual to the global leaders in health
136 and politics. The situation came into focus when 44 cases of pneumonia of unknown cause were
137 reported from the Wuhan province of China. [8] It was the beginning of the second week of January
138 when a novel strain of coronavirus was found to be responsible. Thereafter there has been a steep
139 rise in cases not only in the Hubei province but different regions of China and adjoining countries like
140 Thailand. With increasing spread of the disease and the associated number of deaths, different
141 countries started to react in an attempt to prevent the disease. The approaches included case
142 finding, isolation of suspected cases and their close contacts and also hunting for active therapeutic
143 agents to tackle the viral-related morbidity and mortality.

144 Recent genomic studies have practically ruled out the possibility of a laboratory related experiment
145 gone out of hand. [9] The two hypothesis doing rounds are a non-pathogenic strain of coronavirus
146 which jumped into a human host from an animal intermediary and became pathogenic by the
147 process of natural selection and the other possibility being a virus becoming pathogenic by the
148 process of natural selection in an animal species and then jumps on to its human host [9]. Having
149 identified the genomic sequence and its predilection for the lungs, the next step was to find a
150 diagnostic test and discover medications to tackle the situation. [10]

151 However, a more important aspect of the disease in question has to do with prevention. Identifying
152 the appropriate factors related to getting infected and mortality would help plan in advance.

153 Till now the only valid and detailed analysis and correlation data relating to COVID 19 outcomes has
154 been related to age and sex variables. [11]

155 This is the first analysis to ascertain certain additional factors related to adverse outcomes
156 associated with COVID-19 infection. What is the correlation and its strength as far as getting infected
157 and mortality were concerned? Can we build a predictive model on the same using a linear
158 regression analysis? What is the recovery rate country-wise? Is there a difference depending on the
159 geographical region? How does GDP share of the country in question and the number of physicians
160 & hospital bed per 1000 patient available impacts survival? These are crucial questions which could
161 help us identify important logistics related factors contributing to the direction the infected
162 population would go.

163 We found a statistically significant 89% association with death once an individual gets infected with
164 COVID-19. There were 25 expected deaths per 1000 patient infected. This is in keeping with the case
165 fatality rates that are evolving worldwide, [12] with a significant percentage of these deaths
166 effecting the elderly.[13] With the help of regression analysis we could build a mathematical model
167 to predict death depending on the number infected.

168 Germany has a very low death rate in spite of being on the top 6 countries being affected. This could
169 be attributed to the highest physician and hospital bed per 1000 patient ratio among the top 6
170 countries. It should also be noted here that recovery rates in China improved dramatically after an
171 increase in the number of beds made available, along with strong quarantine measures that were
172 applied. This analysis bolsters the notion that higher number of hospital beds per 1000 population
173 correlates with a lower death rate. This might provide the basis to make an appropriate calculation
174 and get a proportional amount of beds prepared in advance.

175 It is interesting to note that GDP share has no correlation whatsoever with both the infection rates
176 as well as mortality. Wealth and economic resources may therefore not be adequate in tackling the
177 scourge of this novel coronavirus.

178 What is most revealing about this data, however, is that higher testing rates improved recovery
179 rates. Increased number of tests were significantly correlated with 1) a reduction in the number of
180 new cases, 2) a reduction in the number of active cases and 3) most importantly with a reduction in
181 the number of critically ill patients. This would indicate that the best option to deter this pandemic is
182 higher number of tests being conducted in all populations. Intuitively one could surmise that
183 increasing tests would identify the infected individuals early in the disease process providing more
184 robust and aggressive treatment to this population, improving morbidity and probably mortality.in
185 addition it would warn the exposed to take strict steps to quarantine themselves, reducing rates of
186 new and active cases.

187 **5.1 Study limitations:**

188 Firstly, correlation studies do not mean causality. The associations mentioned above are indicative of
189 certain trends only. However, at an early stage of a novel disease trends can definitely be a good
190 approximation for a large data-related analysis.

191 Secondly, due to extreme paucity of data from many countries, they were excluded in certain
192 analysis for example the one on testing frequency and outcomes. This weakness was overcome by
193 including the top 15 countries wherefrom the majority of the data were more homogenous and
194 devoid of significant outliers.

195 Thirdly, the correlation values are not indicative strong associations- both positive and negative.
196 However, this is due to the fluidity of the situation. We do not desire the numbers to go up to give us
197 opportunity for a robust analysis. This is precisely why we attempted to identify trends from the
198 updated data and guide our health care system to gear up for a better assessment of the situation.

199 Fourthly, the use of GDP as a measure of disease outcomes cuts both ways. If it is an indicator of
200 surplus wealth, then it is expected that a large proportion of the same could be channelized into the
201 health care facility and hence the outcomes. Hence, we choose to include both GDP as well as share
202 of the World's GDP in this analysis to overcoming this confounding.

203 **5.2 Strength of the study:**

204 In contrast to the association data related to mortality rates and risk factors for the same like sex
205 age, other co-morbidities, this analysis is probably the first in its kind to look at the correlation
206 between infection rate/mortality and GDP, physician & hospital bed per 1000 patient population,
207 testing frequency and its relation with reduced new cases being detected as well as new deaths
208 being prevented. Although these are only trends, we believe these data will open up newer avenues
209 for health-care sector planning leading to effective management strategies.

210

211

212 **6. Conclusion:**

213 This is the first study to indicate that increasing resources in the form of increasing number of beds
214 per 1000 population reduces death from COVID 19 and improves recovery thereof. More
215 importantly this is the first dataset analysis which reveals that increasing number of tests for the
216 novel coronavirus would drastically reduce incidence and possibly death rates and morbidity from
217 this dreaded epidemic.

218

219 **Conflict of interest:** None to declare

220 **Funding:** None

221 **Acknowledgements:**

222 1. Mr. Milan Majumder: Independent statistician, Pune, India

223 2. Mr. Kingshuk Bhattacharjee: Independent Biostatistician, Kolkata, India.

224

225 **References:**

226 1. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, Shi Z, Hu Z, Zhong W, Xiao G. Remdesivir and
227 chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res.
228 2020 Mar;30(3):269-271.

229 2. Yao X, Ye F, Zhang M, Cui C, Huang B, Niu P, Liu X, Zhao L, Dong E, Song C, Zhan S, Lu R, Li H, Tan
230 W, Liu D. In Vitro Antiviral Activity and Projection of Optimized Dosing Design of Hydroxychloroquine
231 for the Treatment of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Clin Infect Dis.
232 2020 Mar 9. pii: ciaa237. doi: 10.1093/cid/ciaa237. [Epub ahead of print]

- 233 3. Coronavirus disease (COVID-2019) situation reports. World Health Organization. [Online] Available
234 at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>
235 [Accessed on: 22nd March 2020]
- 236 4. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in
237 Relation to COVID-19 in Italy. JAMA 2020. [Online] Available at:
238 <https://jamanetwork.com/journals/jama/fullarticle/2763667>
239 [Accessed on: 24nd March 2020]
- 240
- 241 5. GPP. The World Bank. [Online] Available at:
242 <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD>
243 [Accessed on: 22nd March 2020]
- 244 6. Physicians (per 1,000 people). World Health Organization's Global Health Workforce Statistics,
245 OECD, supplemented by country data. [Online] Available at:
246 <https://data.worldbank.org/indicator/SH.MED.PHYS.ZS>
247 [Accessed on: 22nd March 2020]
- 248 7. Hospital beds (per 1,000 people) Data are from the World Health Organization, supplemented by
249 country data. [Online] Available at: <https://data.worldbank.org/indicator/SH.MED.BEDS.ZS>
250 [Accessed on: 22nd March 2020]
- 251 8. Novel Coronavirus (2019-nCoV) SITUATION REPORT - 1 21 JANUARY 2020. World Health
252 Organization. [Online] Available at: [https://www.who.int/docs/default-
253 source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf)
254 [Accessed on: 22nd March 2020]
- 255 9. Anderson KG, Rambaut A, Lipkin WI, Holmes EC, Gary RF. The proximal origin of SARS-CoV-2.
256 [Online] Available at: <https://www.nature.com/articles/s41591-020-0820-9.pdf?origin=ppub>
257 [Accessed on: 22nd March 2020]
- 258 10. Subbaraman N. Coronavirus tests: researchers chase new diagnostics to fight the pandemic.
259 Nature 2020. [Online] Available at: <https://www.nature.com/articles/d41586-020-00827-6>
260 [Accessed on: 24nd March 2020]
- 261 11. Zhou F, Du R, Fan G, Liu Y, Xiang J, Wang Y, et al. Clinical course and risk factors for mortality of
262 adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet 2020.
263 [Online] Available at: [https://www.thelancet.com/action/showPdf?pii=S0140-6736%2820%2930566-
264 3](https://www.thelancet.com/action/showPdf?pii=S0140-6736%2820%2930566-3)
265 [Accessed on: 24nd March 2020]
- 266 12. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States,
267 February 12–March 16, 2020. MMWR Morb Mortal Wkly Rep. ePub: 18 March 2020. DOI:
268 <http://dx.doi.org/10.15585/mmwr.mm6912e2>

269 13. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus
270 (2019-nCoV) in Wuhan, China. *J Med Virol.* 2020 Apr;92(4):441-447. doi: 10.1002/jmv.25689. Epub
271 2020 Feb 12.

272

273 **Figure & Table Legends:**

274 Fig 1: Scatter plot comparing Death and Cases identified (COVID-19) with correlation data: Linear
275 regression analysis with line of best fit.

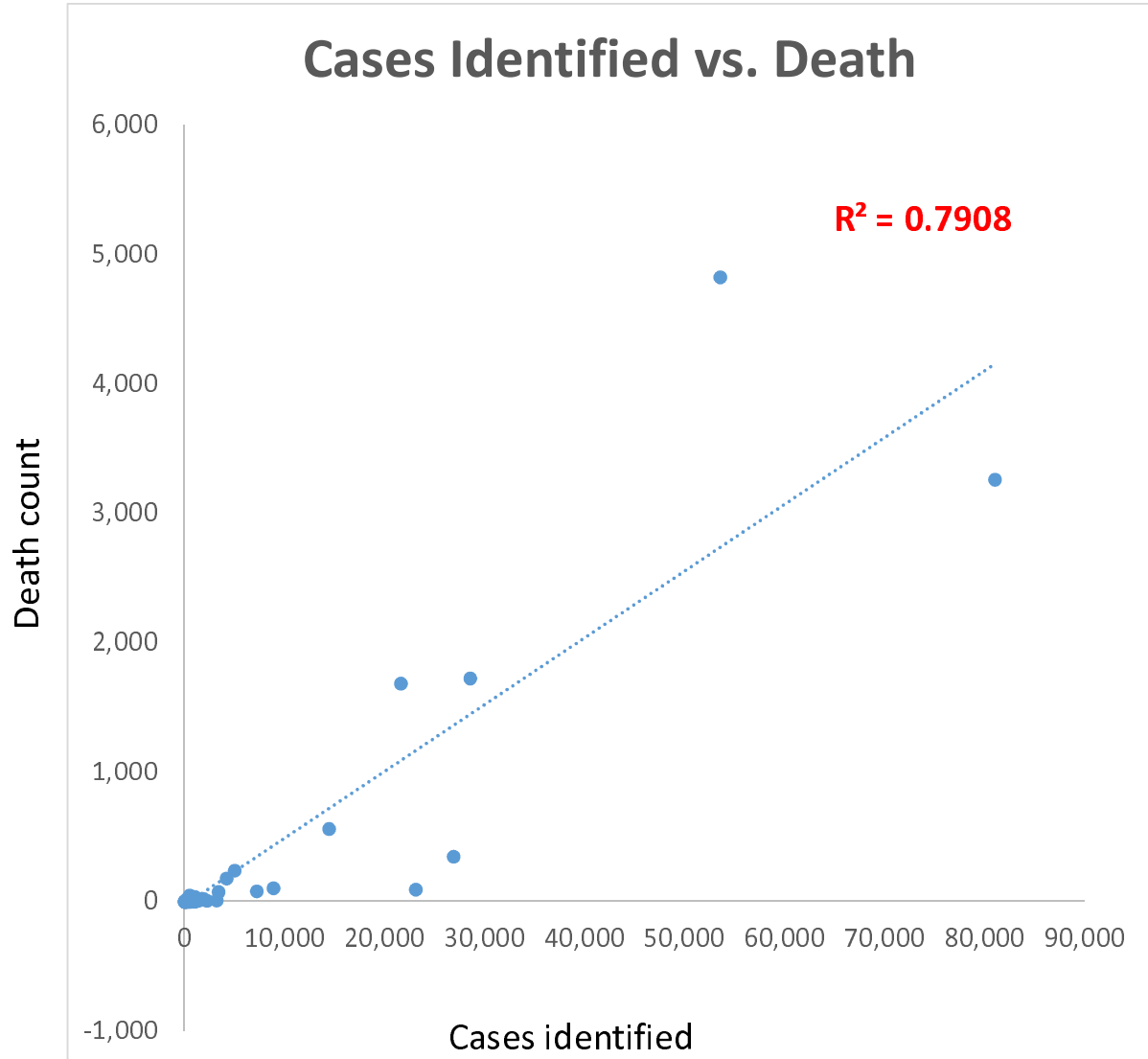
276 Fig 2: Regression analysis with regression model comparing total death and those infected with
277 COVID-19: Raw data.

278 Fig 3: Correlation statistics from the top 15 infected countries with testing frequency included.

279 Table 1: Comparison of total cases and GDP & Physician and hospital bed per 1000 patient
280 population from the top 6 infected countries.

281

282



SUMMARY
OUTPUT

Regression Statistics

Multiple R	0.889290998
R Square	0.79083848
Adjusted R Square	0.788704178
Standard Error	286.4700615
Observations	100

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	30408189.34	30408189.34	370.5374243	4.55661E-35
Residual	98	8042379.42	82065.09612		
Total	99	38450568.76			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	-26.7237135	29.86559959	-0.894799162	0.373086889	-85.9910239	32.5435969	-85.9910239	32.5435969
Total Cases	0.0515849660	0.002679829	19.24934867	4.55661E-35	0.046266933	0.056903	0.046266933	0.056903

SUMMARY OUTPUT

Regression Statistics

Multiple R	0.889290998
R Square	0.79083848
Adjusted R Square	0.788704178
Standard Error	286.4700615
Observations	100

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	30408189.34	30408189.34	370.5374243	4.55661E-35
Residual	98	8042379.42	82065.09612		
Total	99	38450568.76			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	-26.7237135	29.86559959	-0.894799162	0.373086889	-85.9910239	32.5435969	-85.9910239	32.5435969
Total Cases	0.051584966	0.002679829	19.24934867	4.55661E-35	0.046266933	0.056903	0.046266933	0.056903

Country	Total Cases	Share of World GDP	Hospital beds per 1000 patients	Physicians per 1000 patients
<i>China</i>	81,054	0.1512	3.56	1.7855
<i>Italy</i>	53,578	0.024	3.6	3.9174
<i>Spain</i>	28,572	0.0162	3.2	3.8493
<i>United States</i>	26,909	0.2408	3	2.5858
<i>Germany</i>	23,129	0.0456	8.2	4.1383