Mathematical Analysis of Confirmed Cases of Coronavirus in South Korea to Assess Difference in Impact on Genders and Assessing Trends for Adequacy of Measures Employed

Jitendra Shreemali

Abstract: Coronavirus or COVID-19 has shaken the world like nothing in the last 70 years or more. The approach so far appears to be focused more on preventing spread of cases rather than eradicating the pandemic. While South Korea's efforts at preventing spread have produced commendable results, a mathematical analysis using polynomial regression indicates that more is needed than the presently employed measures to eradicate this pandemic or address the concern of fatalities on account of COVID-19. A vaccine could be that additional measure without which the pandemic may linger for a long time to come. Further, it is seen that though females appear to be more susceptible to falling prey to the pandemic, fatalities are greater among males indicating that gender may have a role in the manner in which the human body fights this pandemic.

Keywords: Coronavirus, COVID-19, Pandemic, Polynomial regression, Fatalities.

I. INTRODUCTION

As per Wikipedia [1], beginning with South Korea's announcement of the first confirmed case of COVID-19 on 20th January 2020, the number of confirmed cases began it's upward growth. This number grew by 20 on the 20th of February 2020 and by 58 or 70, the very next day i.e on 21st February 2020 bringing the cumulative tally to 346. This surge was attributed to a patient's participation in a church gathering. This was followed by measures aimed at reducing or eliminating opportunities through further spread through contact among humans. Such measures included cancellation of mass gathering, isolation of those infected and quarantining those in contact of the infected. The number of confirmed and suspected cases, as on date i.e. 30th March 2020 is 9661 and 13531 respectively while the number of recoveries and unfortunate deaths are 5228 and 158 respectively with the strain of virus being SARS-Cov-2. Wikipedia further reports that, as on 24th March 2020, South Korea witnessed a fatality rate of 1.33% , less than a third of the World Health Organization's global fatality rate of 4.34%. South Korea's approach to containing the pandemic has been praised

extensively and provides a model case for study. The objective of the study is to examine whether susceptibility to the virus is likely to be linked to the patient's gender and if one gender is more susceptible to the virus. Further, the study also aims to assess if the available data indicates adequate tapering of the trend of cases or added measures are likely to be needed in the days to come to eradicate the pandemic.

II. LITERAURE REVIEW

Data for the analysis carried out was taken from the Kaggle dataset [3] titled 'Data Science for Covid-19 (DS4C)', at the URL: https://www.kaggle.com/kimjihoo/coronavirusdataset. Given below is a graphical representation of total cases of COVID-19 as on 30th March 2020.
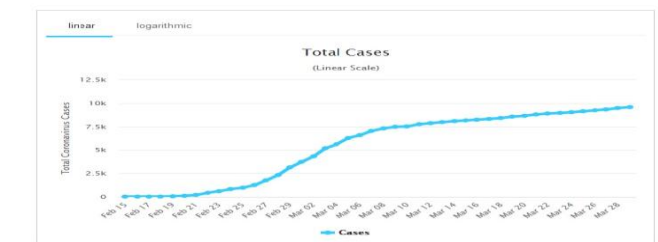
Fig.1: Covid-19 Cases (S.Korea) on linear scale
Source: https://www.worldometers.info/coronavirus/country/south-korea/

While the confirmed cases do appear to be plateauing after a rather steep rise in the last week of February and first week of March 2020, the slope on 30th March 2020 is, still, clearly not equal to zero suggesting that the incidence of cases have note stopped. This plateauing of confirmed cases is seen better when the data is viewed on a logarithmic scale as in Figure 2 below:
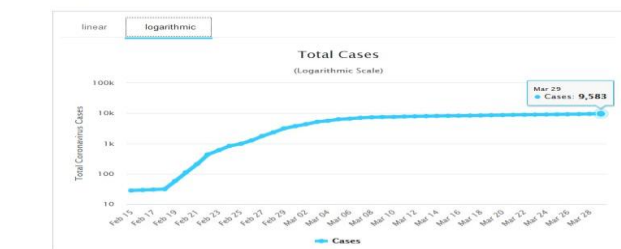

Fig.2: Covid-19 Cases (S.Korea) on logarithmic scale

Source: https://www.worldometers.info/coronavirus/country/south-korea/

A bar-plot of new cases each day further supports the contention that efforts to bring the pandemic under control are showing results. Figure 3 below shows this bar-plot:
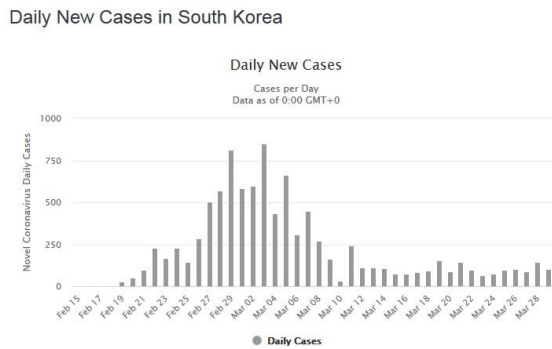

Figure 3: Bar-plot of new cases everyday (S.Korea)

Source: https://www.worldometers.info/coronavirus/country/south-korea/

III. ANALYSIS AND DISCUSSIONS

A plot of the number of male and female cases during the 3 week period from 2-22nd March 2020 is presented below.
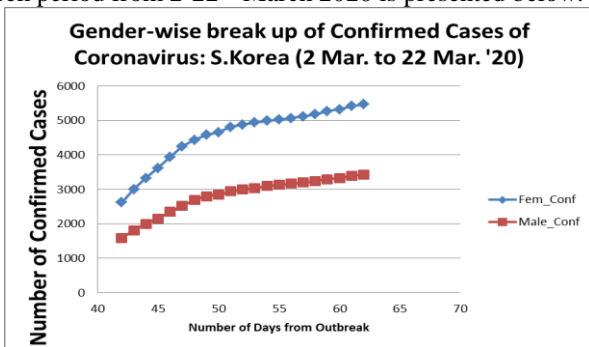

Fig.4: Genderwise breakup: Confirmed cases in S.Korea
The graph above indicates that female patients are consistently above 60% of confirmed cases of Covid-19 suggesting the possibility that the females could be more susceptible to attack by the virus than males. This is further examined later in the study but the shape of the two curves above also indicates a difference in the way confirmed cases are growing among females and males with the plateauing being more distinct in the case of males as compared to females.

The trend lines for quadratic, cubic, bi-quadratic and 5th order polynomial are shown below along with the respective R^2 value. In these figures below, the mathematical equation representing the trend of confirmed cases for female and male cases are indicated by 'Y' and x represents the number of days from the onset of the pandemic.
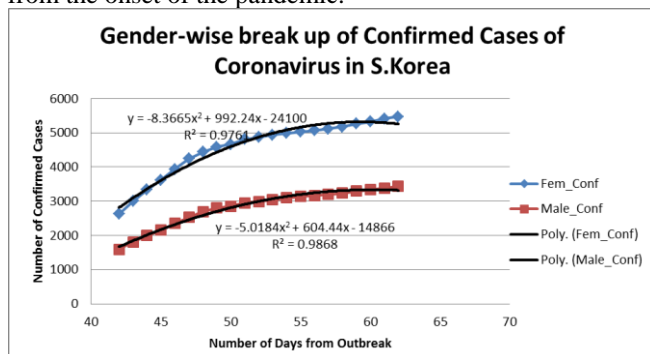

Fig.5: Fitting a quadratic polynomial to confirmed cases

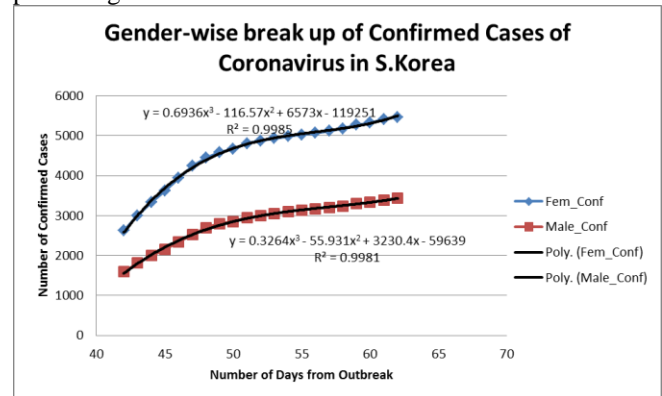The falling shape of the trend line is at variance with the prevailing trend.

Fig.6: Fitting a cubic polynomial to confirmed cases
The cubic equation provides a higher R^2 value and appears to fit more snugly with the data available. The R^2 value is observed to be higher than 0.998 for confirmed cases among the female as well as male population. Further, the coefficients are not high thereby eliminating the possibility of the trend line over-fitting the data at the expense of possibly not fitting satisfactorily at future data points.
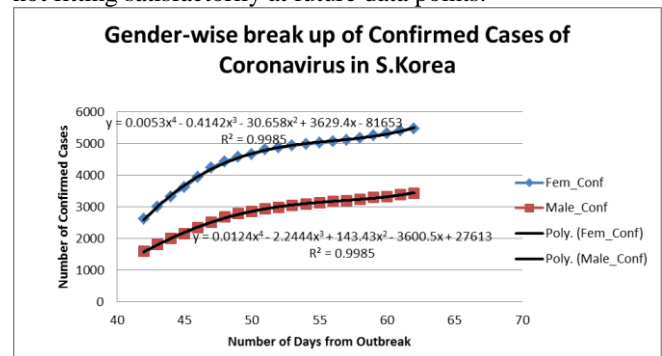

Fig. 7: Fitting bi-quadratic polynomial: Confirmed cases
As expected, adding higher order polynomials increases the R^2 value but a analysis of the slope of the curve indicates a very rapidly increasing slope for both, females and males that is at variance with the prevailing trends. The figure next depicts the 5th order polynomial for confirmed cases. The trend line for 5th order ^{P}olynomial was also not found to conform to the observed behavior as it produced a very high value of negative slope when the actual data does not show such behavior.
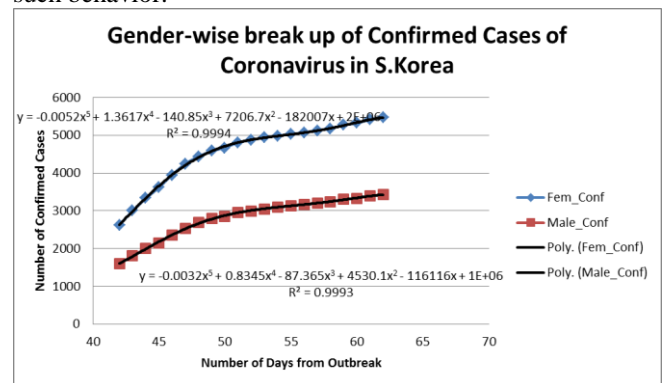

Fig. 8: Fitting a 5th order polynomial to confirmed cases
The analysis above suggests that a cubic polynomial is best suited to explain the trends, mathematically.

However, analysis of the slope (dY/dx) of the cubic equation leads to the following outcome in this case:

$(dY/dx)_{Female} = 2.0808\ x^2 – 233.14\ x + 6573$

$(dY/dx)_{Male} = 0.9792\ x^2 – 111.862\ x + 3230.4$

Solving the quadratic equations above for roots to ascertain the maxima throws up complex roots raising a question if the existing trend would ever taper-off in the foreseeable future. In both cases it is seen that:

$B^2 – 4\ A.C. < 0$ (-354.13 for Females and -139.72 for Males) It is, therefore, possible that the pandemic may require something more than the prevailing measures and quite possible, a vaccine may be the answer to what more is required to eradicate the pandemic and may well be the real need of the hour with the present measures contributing to reduced incidence and reduced fatalities.

Data of confirmed cases shows that female patients make over 60% of the total confirmed cases consistently on all days of the three week period. Countrymeters [4] site (https://countrymeters.info/en/Republic_of_Korea) indicates that the population of males and females is very close to each other with females making up 50.2% of the population. For a country with this split of female-male population, the p value for confirmed female patients being consistently above 60% is close to zero suggesting greater susceptibility of females to the virus.
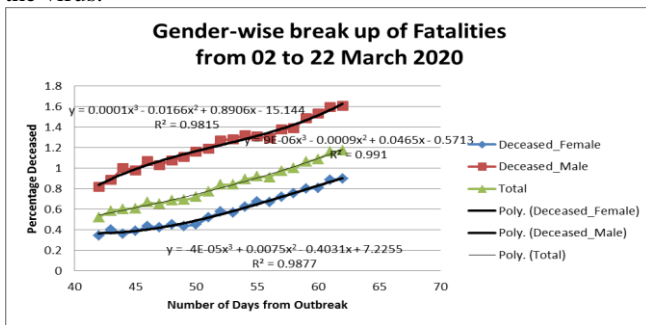

Fig. 9: Gender-wise breakup of Fatalities

This shows a significant change from earlier trends with fatalities among male population being consistently higher. While the confirmed cases show that incidence among females is consistent above 60% of total conformed cases, the cases of fatalities are consistently below 50% of the total suggesting the possibility of female patients coping with the virus better. As far trends of fatalities goes, the trend here are again not very reassuring with the percentage of fatalities (fatalities/confirmed cases expressed as percentage) showing a constantly rise trend during the period in question. Further computing the maxima for total deaths again throws up non-real (i.e. complex) roots. The continuously rising trend here is in keeping with actual data as shown in the figure below:
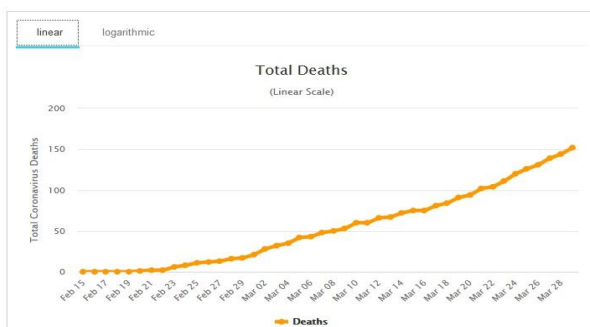

Fig. 10: Actual trend of Fatalities in S.Korea

IV. CONCLUSIONS

The study above indicates that even though South Korea has done commendably as far as effort to contain the pandemic is concerned, the efforts may not be adequate either to eradicate it or address the problem of fatalities. Added solutions are urgently needed to eradicate the pandemic. A vaccine could be the added measure but that was not the central subject of this study. Further, it is seen that while females are observed to be consistently above 60% of the confirmed cases, fatalities are consistently higher among the male patients suggesting the female members of the population cope better with the after infection virus despite being more susceptible initially.

LIMITATIONS OF THE STUDY

The mathematical analysis assumes similar behavior of the dependent and independent variables within the three week period of study (02-22 March 2020). Since preventive measures have evolved in all countries, this assumption may have introduced some error.

ACKNOWLEDGMENT

REFERENCES

1. Wikipedia page on "2020 coronavirus pandemic in South Korea" at https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_in_South_Korea.
2. Kaggle dataset titled 'Data Science for Covid-19 (DS4C)', at the URL: https://www.kaggle.com/kimjihoo/coronavirusdataset
3. Worldmeter site on Coronavirus cases at the URL: https://www.worldometers.info/coronavirus/country/south-korea/
4. Countrymeters Site (https://countrymeters.info/en/Republic_of_Korea)

AUTHORS PROFILE

Prof (Dr.) Jitendra Shreemali is a graduate from IIT Madras with post graduate from IIM Bangalore. He is working as Professor of the Department of Computer Science and Engineering of Techno India NJR Institute of Technology Udaipur. He has worked in reputed companies in India & abroad for about a decade and half followed by about two decades of academic/ research/ training experience. He has taught students upto doctoral level students, his range of subjects/courses including operations management, research methodology, and data science besides others. His areas of work include data science, optimization, mathematical modeling and machine learning. Email: jitendrapshreemali@gmail.com