Model-based retrospective estimates for COVID-19 or coronavirus in India: continued efforts required to contain the virus spread

India reported having 498 cases COVID-19 (including 40 foreign tourists) as of 23 March 2020 (ref. 1) and out of which 32 cases were recovered and 9 deaths occurred. The number of COVID-19 cases worldwide as of 23 March 2020, was 378,000 with 16,500 deaths[2]. Overall more than 1.5 million passengers were screened at airports in India, 17,237 people were clinically tested for COVID-19 as of 22 March 2020 (ref. 1). Although the first case of COVID-19 was reported in India on 30 January 2020, for a student who has returned from Wuhan, China[3], the confirmed cases started arising from the beginning of March. The government of India was quick to launch various levels of travel advisories beginning from 26 February 2020, with restrictions on travel to China and non-essential travel restrictions to Singapore, South Korea, Iran and Italy[3]. The latest travel advisory issued by the Government on 19 March temporarily bans all the commercial passenger international aircrafts arrival at any Indian airport for the period 22–29 March 2020 (ref. 4).

Practically to understand the community spread in the country there need to be a large sample testing to be conducted among those who have no travel history. Given the highly contagious nature of this virus COVID-19 and the experience of other countries which have peaked during January and March, people cannot wait for this confirmation to start preventive measures. The efforts to control by the Hon'ble Prime Minister Narendra Modi and the government through *Janata Curfew* (public curfew) on 22 March 2020, can be seen as the beginning of wide-scale public preventive measures. There were already various measures to prevent as described above. So we have provided two likely scenarios through our modelling approach by considering 10 and 20 COVID-19 infected individuals as of 1 March 2020, who were not quarantined and one situation which might sound unlikely with 50 infected who were not quarantined as of 1 March 2020. If the community level transmission had already started in late February or early March but the conclusion has arrived only in later in April by then several cases would have transmitted. Hence, this communication was aimed to assist the policy at the right time. Better to take extra precaution even if no community-level sample tests were not done than to neglect the epidemic spread before enough community level tests were conducted. When we studied case reporting patterns, then COVID-19 cases during the first month after the first case on 30 January, followed by 107 cases during

Appendix 1. Data, methods and models

The model specifically considered the proportion of populations in the year 2020 in the age groups (0–14 (27.1%), 15–64 (66.8%), 65+ (6.2%)), population density (464 km^2), level of urbanization in India (35%) from the sources[7,8]. In the first phase of the community spread the urban population will be affected. Uncontrolled transmission between urban to rural might lead to the second phase of the epidemic. We have assumed our results reflect the first phase of the spread. The two coupled differential equations $s(t) = -\beta s(t)k(t)$ and $k(t) = \beta s(t)k(t)$, where $s(t)$ and $k(t)$ represent susceptible and infected at time t, β is the average transmission rate from an infected to susceptible and β = 1.5E-09 for 15–64 age group and β = 7.5E-09 for 65+ age group. We further assumed that β is constant during 1–15 March 2020. We have divided model predicted numbers with the reported numbers during the same period to obtain degree of reporting, say, d_R time dependent degree of reporting, $d_R(t)$ is given by $d_R(t) = k(t)/R(t)$, where $R(t)$ represent reported COVID-19 cases at time t. We plotted $d_R(t)$ values for t = (1–15 March) shown in Figure 1. The above procedure also gives us under-reported COVID-19 cases (that includes under-diagnosed). Using reported and total cases constructed, the Meyer wavelets were constructed. The Meyer wavelets $\psi(\omega)$ are infinitely differentiable functions within a certain domain. See details in refs 5, 6, 12. These wavelets are accompanied by a function u given below

$$\psi(\omega) = \begin{cases} \frac{1}{\sqrt{2\pi}} \sin\left(\frac{\pi}{2} u\left(\frac{3|\omega|}{2\pi} - 1\right)\right) e^{\frac{i\omega}{2}} & \text{if } 2\pi/3 < |\omega| < 4\pi/3 \\ \frac{1}{\sqrt{2\pi}} \cos\left(\frac{\pi}{2} u\left(\frac{3|\omega|}{2\pi} - 1\right)\right) e^{\frac{i\omega}{2}} & \text{if } 4\pi/3 < |\omega| < 8\pi/3. \\ 0 & \text{otherwise} \end{cases}$$

Here $u(x) = 0$ for $x < 0$, and

$$u(x) = \begin{cases} x & \text{for } x \in (0,1) \\ 1 & \text{for } x > 1 \end{cases}.$$

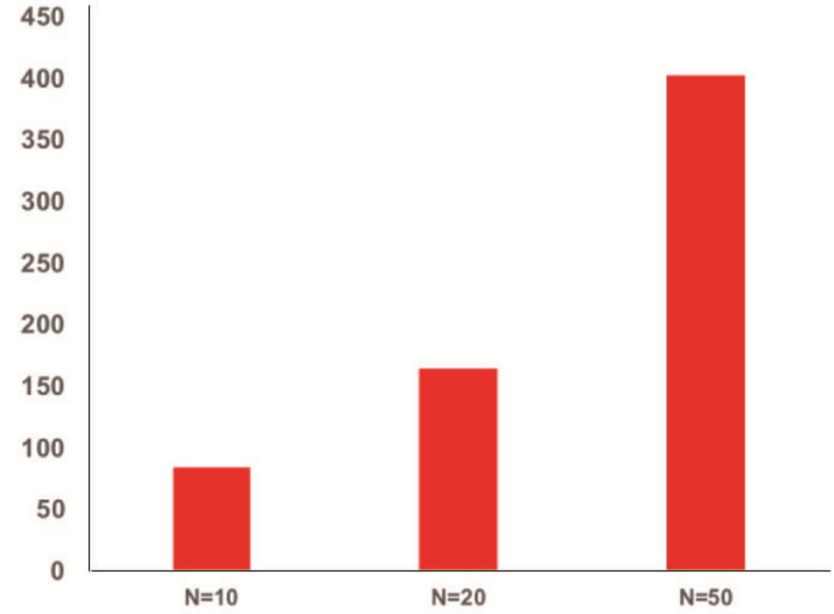

Figure 1. Undetected COVID-19 per each detected during 1–15 March 2020 under the community spread assumption when N number are infected by 1 March 2020. We also believe $N = 50$ could be unrealistic beginning because the epidemic has not picked to have these many number as of 1 March 2020.

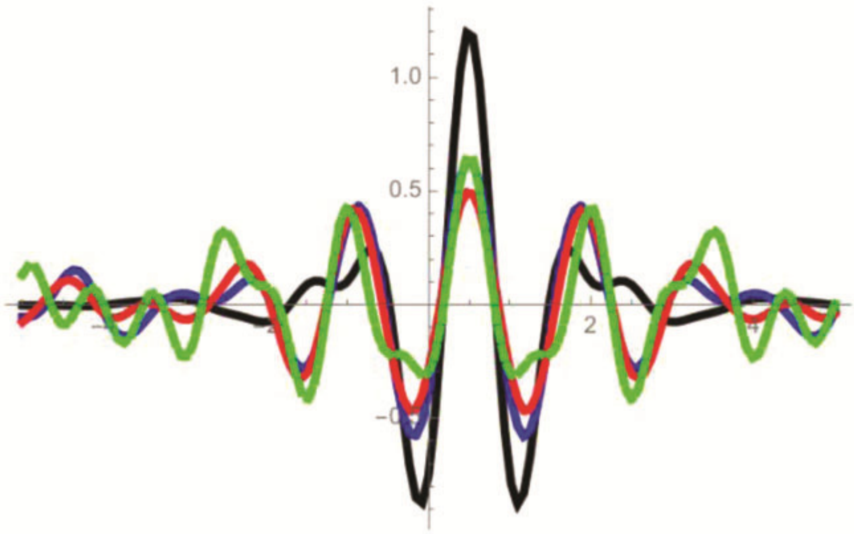

Figure 2. Meyer wavelets for the reported numbers (black), undetected at a rate of 1 per 84 (blue), undetected at a rate of 1 per 164 (red), undetected at a rate of 1 per 402 (green).

1–15 March. During the next eight days of 16–23 March, additional 361 cases have been reported. This could indicate that the spread mechanism had initiated a couple of weeks ago.

To understand the current status of COVID-19 spread during the first two weeks of March, model-based estimates were developed for India based on India-centric data. We applied them on the harmonic analysis methods and dynamic models developed earlier for emerging epidemics[5] and modified for COVID-19 epidemic[6]. See Appendix 1 for the details on data, models and methods. The model estimates show that the number of COVID-19 infected would be 9225 (if there were 10 infected individuals as of 1 March 2020, who were not taking any precautions to spread), 17,986 (if there were 20) and 44,265 (if there were 50). We have used China-level parameters that were used in the earlier paper[6]. We have plotted the Meyer wavelets that were plotted for the reported numbers and for the model-based adjusted numbers to give us an idea of the degree of variation in reporting. As the testing for COVID-19 was not done widely for general population but done for travelers from abroad or returned travelers, so for the general population we are not sure whether $N = 50$ on 1 March is a real beginning. Even if medium range values were happened during 1–15 March then that would lead to very large infections in the future because controlled social distances and other precautions would be not practically possible for a longer period.

India has launched several social distancing measures and personal hygiene measures during the second week of March[9]. As soon as more data becomes available in the coming weeks on the impact of these measures in controlling the spread, that will help to generate more valid model-based predictions stratified on state level epidemic, level of health infrastructure and magnitude of preventive strategies employed. Identification of individuals at risk is more important when the virus is highly contagious and there were proposals to capture the community level severity through AI models[10], and now casting techniques are also very helpful[11]. There are speculations that the virus spread will be restricted in hot temperature countries like India, and thus factors like treatment can be introduced in the models as soon as credible data is available. Natural positives factors like very hot summers and the urbanization level which is low in comparison with other COVID-19 affected countries such as China (61%) and Italy (83%) might play better roles in controlling the spread in India in the future.

The economic impact of COVID-19 will be widespread and significant but the government has already initiated steps to intervene. To remain prepared we need data based decisions which help us to shape our interventions. The projections produced by the model and after their validation can be used to determine the scope and scale of measures that governments need to initiate.

In conclusion, if the current mathematical model results can be validated within the range provided here, then the social distancing and other prevention, treatment policies that the central and various state governments and people are currently implementing should continue until new cases are not seen. Importantly spread to urban to rural populations should be controlled.

Authors' contributions: All the authors contributed in writing. A. S. R. S. Rao designed the study, developed the methods, collected data, performed analysis, computing, wrote the first draft. S. G. Krantz designed the study, contributed in writing, editing the draft. T. Kurien, R. Bhat and K. Sudhakar have contributed in data, editing and commenting. All the authors approved the final manuscript.

1. https://www.mohfw.gov.in/ and ICMR (accessed on 22 March 2020).
2. 2020_coronavirus_pandemic_in_India, Wikipedia (accessed on 21 March 2020).
3. Consolidated Travel advisory in view of COVID-19 (26 February 2020), Government of India, Ministry of Health & FW (accessed on 21 March 2020).
4. Additional travel advisory for novel coronavirus disease (COVID-19) (19 March 2020, Government of India, Ministry of Health & Family Welfare (accessed on 21 March 2020).
5. Krantz S. G., Polyakov P. G. and Rao, A. S. R. S., *J. Theoret. Biol.*, 2020, **494**, 110243.
6. Krantz S. G. and Rao, A. S. R. S., Level of under-reporting including under-diagnosis before the first peak of COVID-19 in various countries: Preliminary Retrospective Results Based on Wavelets and Deterministic Modeling. *Infection Control and Hospital Epidemiology* (*to appear*). Also available at IMAG Wiki NIBIB/NIH (accessed on 27 March 2020).
7. Worldometer.info for India (accessed on 19 March 2020).
8. The World Bank Open Data. https://data.worldbank.org/ (accessed on 19 March 2020).
9. Indian Council of Medical Research, Department Of Health Research, Revised Strategy of COVID19 testing in India (version 3, dated 20 March 2020).
10. Rao, A. S. R. S. and Vazquez, J. A., *Infect. Control Hospital Epidemiol.*, 2020, 1–18; doi:10.1017/ice.2020.61.
11. Wu, J. T., Leung, K. and Leung, G. M., *Lancet*, 2020, **395**(10225), 689–697; https://doi.org/10.1016/S0140-6736(20)-30260-9.
12. Krantz, S. G., *A Panorama of Harmonic Analysis*, The Carus Mathematical Monographs, Number 27, Mathematical Association of America, 1999.

ACKNOWLEDGEMENT. We thank the reviewer editorial board member for constructive comments which helped us in revising the article.

Received 20 March 2020; revised accepted 24 March 2020

ARNI S. R. SRINIVASA RAO[1,*]
STEVEN G. KRANTZ[2]
THOMAS KURIEN[3]
RAMESH BHAT[4]
SUDHAKAR KURAPATI[5]

[1]*Division of Health Economics and Modeling,*
Department of Population Health Sciences,
Medical College of Georgia,
Augusta University, Augusta, USA and
Laboratory for Theory and Mathematical Modeling,
Division of Infectious Diseases,
Department of Medicine,
Medical College of Georgia,
Augusta, GA 30912, USA
[2]*Department of Mathematics,*
Washington University in St Louis,
Campus Box 1146,
One Brookings Drive,
St Louis, Missouri 63130, USA
[3]*Department of Medicine,*
Pondicherry institute of Medical Sciences,
Puducherry 605 014, India
[4]*NMIMS University,*
Mumbai 400 056, India
[5]*Formerly with CDC,*
World Bank and USAID
**For correspondence.*
e-mail: arrao@augusta.edu