SCIENTISTS SEE FLAWS IN SUTRA'S APPROACH TO MODELLING PANDEMIC

Relevant for: Developmental Issues | Topic: Health & Sanitation and related issues

Emergency care:An isolation facility set up in a banquet hall in New Delhi.Shiv Kumar Pushpakar

With close to 4,00,000 cases being added every day, questions are being raised by some scientists on whether a government-backed model, called SUTRA, to forecast the rise and ebb of the COVID-19 pandemic, may have had an outsize role in creating the perception that a catastrophic second wave of the pandemic was unlikely in India.

An official connected with the COVID-19 management exercise said, on condition of anonymity, that the SUTRA model input was "an important one, but not unique or determining".

The SUTRA group had presented their views to V.K. Paul, who chaired a committee that got inputs from several modellers and sources. "The worst-case predictions from this ensemble were used by the National Empowered Group on Vaccines and the groups headed by Dr. Paul to take measures. However, the surge was several times what any of the modellers had predicted," the official said.

On May 2, the SUTRA group put out a statement, carried by the Press Information Bureau, that the government had solicited its inputs where they said a "second wave" would peak by the third week of April and stay at around 1 lakh cases. "Clearly the model predictions in this instance were incorrect," it noted.

Past its peak

SUTRA (Susceptible, Undetected, Tested (positive), and Removed Approach) first came into public attention when one of its expert members announced in October that India was "past its peak".

After new cases reached 97,000 a day in September, there was a steady decline and one of the scientists associated with the model development, M. Vidyasagar, said at a press conference then that the model showed the COVID burden was expected to be capped at 10.6 million symptomatic infections by early 2021, with fewer than 50,000 active cases from December. In October, at that time, there were 7.4 million confirmed cases of which about 7,80,000 were active infections.

Computational biologist Mukund Thattai, of the National Centre for Biological Sciences, Bengaluru, in a Twitter thread on May 1 summarised instances of the SUTRA forecasts being far out of bounds of the actual case load.

"The so-called Covid 'supermodel' commissioned by the Govt of India is fundamentally flawed," he tweeted. "Based on Prof. Agrawal's [Manindra Agrawal of IIT-Kanpur] own posts, it was quite clear that the predictions of the SUTRA model were too variable to guide government policy. Many models got things wrong but the question is why the government continued to rely on this model, than consult epidemiologists and public health experts," Mr. Thattai told *The Hindu* .

Mr. Agrawal was among the mathematicians involved in developing the model. In an email to

The Hindu , Mr. Agrawal said that the model, which had multiple purposes, didn't work well on a metric of "predicting the future under different scenarios".

He said unlike many epidemiological models that extrapolated cases based on the existing number of cases, the behaviour of the virus and manner of spread, the SUTRA model chose a "data centric approach". The equation that gave out estimates of what the number of future infections might be and the likelihood of when a peak might occur, needed certain 'constants'. These numbers kept changing and their values relied on the number of infections being reported at various intervals. However, the equation couldn't tell when a constant changed. A rapid acceleration of cases couldn't be predicted in advance.

'Danger of overfitting'

Rahul Siddharthan, a computational biologist at the Institute of Mathematical Sciences, in an email said no model, without external input from real-world data, could have predicted the second wave. However, the SUTRA model was problematic as it relied on too many parameters, and recalibrated those parameters whenever its predictions "broke down".

"The more parameters you have, the more you are in danger of 'overfitting'. You can fit any curve over a short time window with 3 or 4 parameters. If you keep resetting those parameters, you can literally fit anything," he said.

According to Mr. Agrawal, one of the main reasons for the model not gauging an impending, exponential rise was that a constant indicating contact between people and populations went wrong. "We assumed it can at best go up to pre-lockdown value. However, it went well above that due to new strains of virus."

Further the model was 'calibrated' incorrectly. The model relied on a serosurvey conducted by the ICMR in May that said 0.73% of India's population may have been infected at that time. " I have strong reasons to believe now that the results of the first survey were not correct (actual infected population was much lower than reported). This calibration led our model to the conclusion that more than 50% population was immune by January. In addition, there is also the possibility that a good percentage of immune population lost immunity with time," Mr. Agrawal said.

In the SUTRA approach, the factor by which reported cases differ from actual ones is a parameter in the model that could be estimated from just reported government data (covid19india.org), according to Mr. Agrawal. "I understand it may appear a bit mysterious, but the math shows how. This, in fact, is one of our central contributions," he told *The Hindu* . This has been described in a preprint research paper that has been available online since January.

(With inputs from Shubashree Desikan)