Review of: "What is the impact of COVID-19 era on annual tuberculosis notifications in India? A comparative study (2017-2022)"

Kandaswamy Paramasivan[1]

1 Government of Tamil Nadu

Potential competing interests: No potential competing interests to declare.

The researchers have investigated the impact of COVID-19 on annual tuberculosis notifications in India. I have several concerns about the study.

1. The authors used two different types of data. The first type is shown in Tables 1 and 2. Tables 3 and 4 show the second type. The first category includes all TB (tuberculosis) new case notifications from India's 36 states and union territories that are electronically transmitted, reported, and communicated via the Government of India's web-based NIKSHAY platform. The second category examines the average number of newly detected tuberculosis cases reported by public and private health care facilities. It appears that the results are based on a sample. For the readers to understand, this requires a proper explanation.
2. The paper is merely a retelling of the data obtained from these two sources, with no rigour in the analysis. It is understandable that there would be fewer tuberculosis notifications due to the severe restrictions imposed on people and the closure of many institutions during pandemic-induced lockdown periods. This has caused the drop.
3. A simple linear regression is used to perform the counterfactual analysis. The prediction during and after the pandemic is based on the pre-pandemic period. Figure 1 depicts a counterfactual analysis of newly discovered tuberculosis case notifications per lakh population.This is oversimplified and naive. No statistical or any learning model has been built and such model built should have been validated with high prediction accuracy by using appropriate error metric.

4. Traditional time series forecasting techniques include the Auto-Regressive Integrated Moving Average method, the Holt-Winters Model, and the Generalized Additive Method. The AI community employs Bayesian Structural Time Series Methods and Deep Learning tools such as convolutional neural networks (CNN), long-short term memory (LSTM), recurrent neural networks (RNN), and temporal convolutional networks (TCN). Both traditional statistical techniques and ML and DL methods require some historical data. The training data from 2017 to 2019 used by the researchers may be sufficient for the predictions to be accurate.

5. The authors have commented on and questioned the validity of the press release's claim (refer to item 23 in the reference list). The study lacks both the data and the rigorous methodology to make such a claim.
6. Overall, the study should be backed up by more and different data, as well as the right use of counterfactual analysis.