

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/341398579>

Undetected! -- Estimating reporting lag for COVID-19 in Sri Lanka

Article · May 2020

CITATIONS
0

READS
11

1 author:



[Vanodya Perera](#)

Linear Squared (pvt) Ltd, Sri Lanka

1 PUBLICATION 0 CITATIONS

SEE PROFILE

Undetected!

Estimating reporting lag for COVID-19 in Sri Lanka

Beyond visualizations

Since 11th March 2020 (first COVID-19 patient in Sri Lanka), we have seen several visualizations about the present situation of the disease. Most of those efforts are focused on reporting the detected cases on a daily basis.

However, without scientifically assessing the spread of the disease, and forecasting the future, it would be extremely difficult to make decisions with respect to;

- Lockdown enforcement and relaxation (suppress-lift) strategy
- Adequacy of medical facilities such as hospital beds, critical care facilities, etc.

But,

Forecasting the future by analysing the historical '**reported cases**' is misleading because of the reporting delay. Reporting delay of cases is a common drawback for real-time risk assessment of epidemics. Also, the reporting delay is highly dependent on the amount of daily testing conducted. (Especially in Sri Lanka, due to our limited testing capacity)

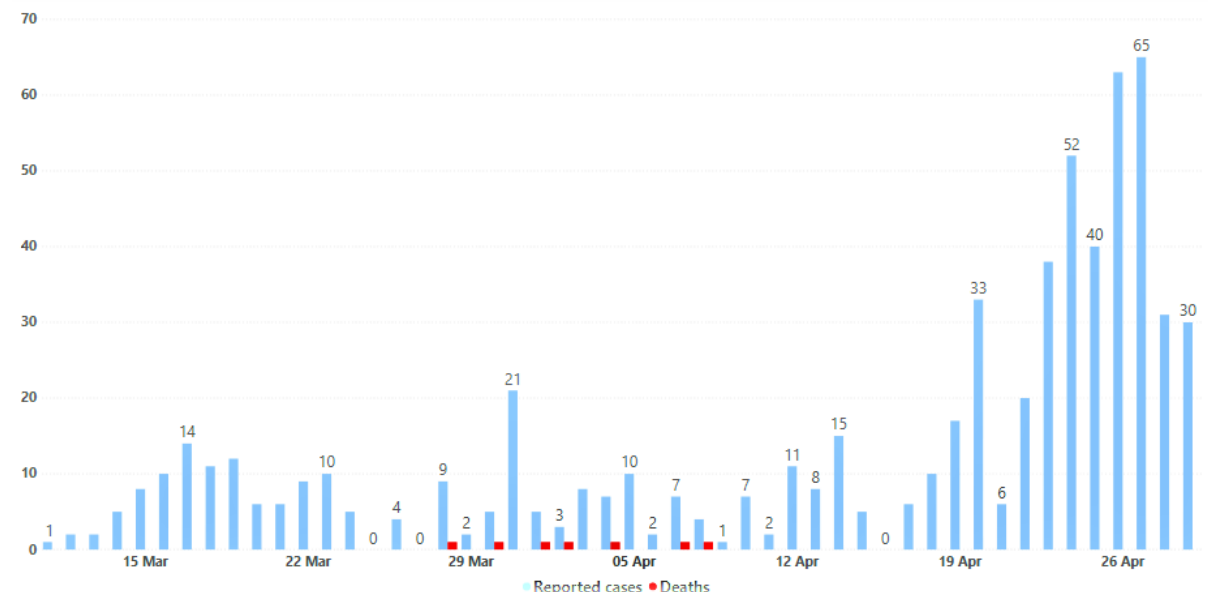
The statistical technique used for delay-adjusting is called 'Now-casting'. The aim of nowcasting, or "predicting the present," is to estimate the number of 'occurred-but-not-yet-reported events' at any given time based on the reported cases by considering the impact of reporting delays.

Now-casting is regularly used in other countries to accurately track disease spread, even though it does not seem like a common practice in Sri Lanka. [Here](#) is an example of the H1N1 pandemic in 2009.

This analysis is collaborative work with [Ishanga Udatiyawala](#), [Ayesha Kasturiarachchi](#), [Ishan Buddhika](#), and Dr. [Sankha Muthu Poruthotage](#). We have attempted at nowcasting the number of daily infections with publicly available data in Sri Lanka. It is obviously subjected to certain assumptions, which we will explicitly mention as we go.

How many reported cases so far?

The daily updates of reported COVID-19 cases can be obtained from the Epidemiology Unit at the Ministry of Health of Sri Lanka through their [website](#) The data includes the number of reported cases and the deaths from 11 March 2020 to 29 April 2020.



Daily reported COVID-19 cases and deaths for Sri Lanka

Do we know the reporting delay in these cases?

Reporting delay can occur due to several reasons;

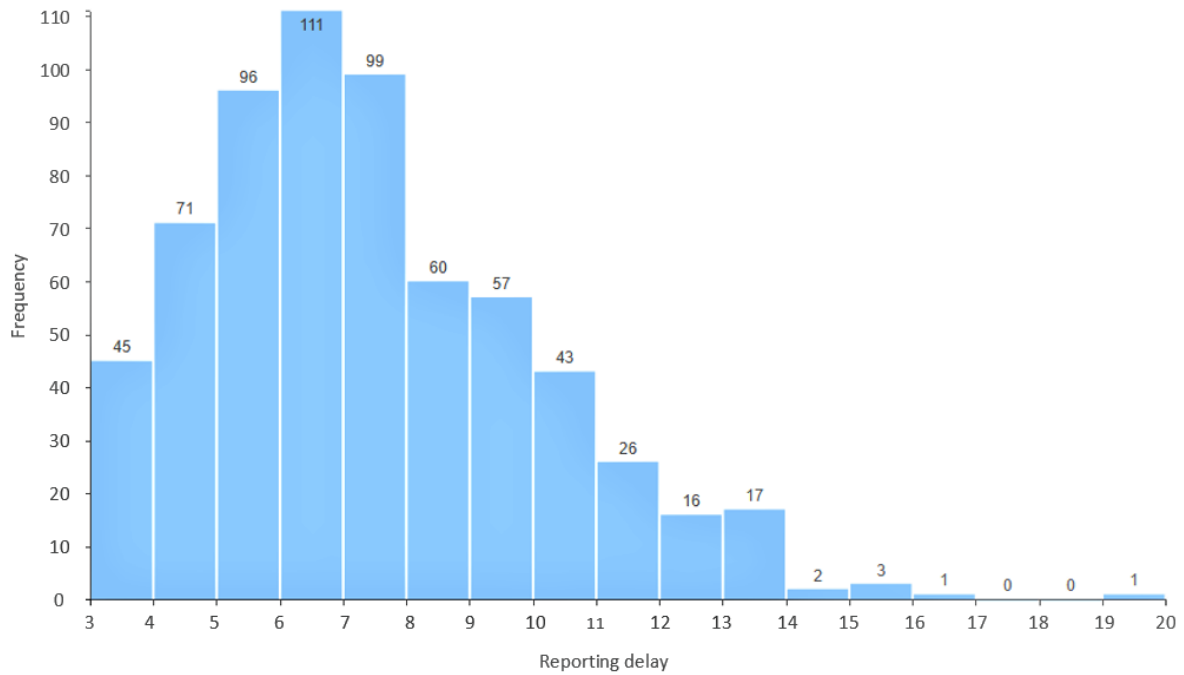
- (i) Time from infection until symptom onset
- (ii) Time from symptom onset until the patient seeks medical attention
- (iii) Time from medical service until initiation of a laboratory test
- (iv) Time from test initiation until its result and subsequent reporting

(i) is the [incubation period](#) which is not easy to measure (the period between exposure to an infection and the appearance of the first symptom). (ii) depends on factors such as public awareness about the disease and accessibility to medical services. Even though part (iii) and (iv) are controllable, they depend on the medical resources in a country that is impractical to increase at short notice.

For this analysis, we define the reporting delay to be the time between infection of a case and the case is registration [(i) + (ii) + (iii) + (iv)]. Due to the lack of data about the exposure dates for all reported cases in Sri Lanka, let's assume a delay distribution based on literature.

In China, the median reporting delay is reported [here](#) as 12 days (range 8–18 days) in early January and 3 days (range 1–7 days) in early February. Singapore, the mean reporting delay is reported [here](#) as 6.4 days (95% CI: 5.8, 6.9).

By considering the above, we have **assumed** a reporting delay for COVID-19 in Sri Lanka follows a Gamma distribution with a mean of 8 days and a standard deviation of 2.6 days.



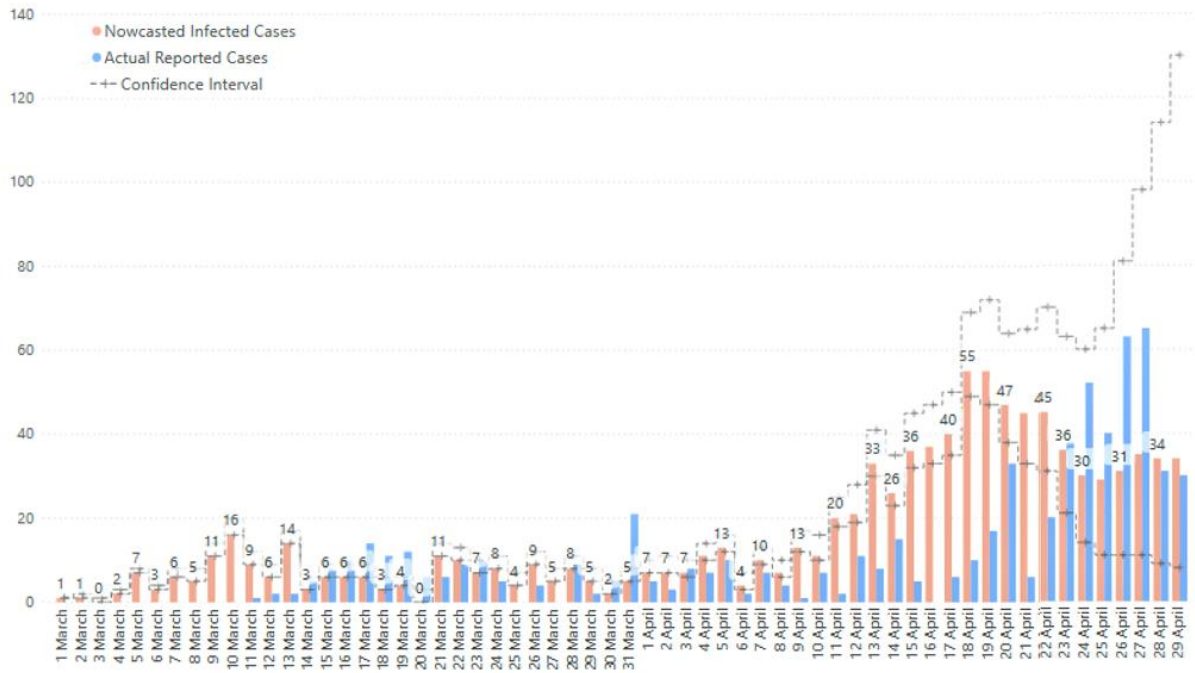
The assumed reporting delay distribution

In the assumed delay distribution,

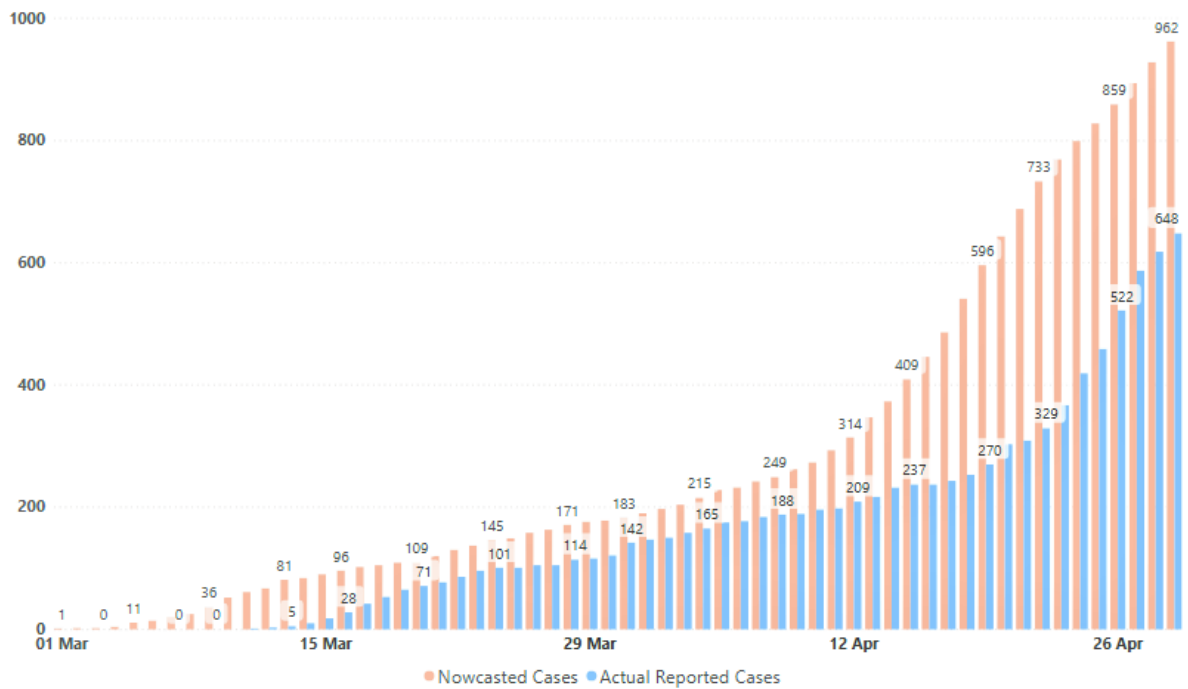
- 18% of the cases have been reported with a delay of less than 5 days from exposure
- 56% of the cases have been reported with a delay of 6–9 days from exposure
- 26% of the cases have been reported with a delay of more than 10 days from exposure

Let's see what we get...

We used the nowcast modeling R package '[NobBS](#)' to obtain the daily nowcasts from 01st March 2020 to 29th April 2020.



- The total nowcasted number of infected cases till 29th April = 962
- The total number of reported cases till 29th April = 648



Is all this necessary?

Analysing reported cases alone doesn't really help in preparation for new COVID-19 cases as it is a lagging indicator. Despite limitations in data availability, nowcasting is a powerful tool to aid public health decision making.

What's next?

Given that we have the nowcasted infection counts (which is probably a better estimate of the present situation), the next step would be to forecast the future behaviour of the disease spread. The reproductive number (R_0) - which represents the number of secondary infections from a single infection) is one of the key parameters in such forecasts. R_0 can be calculated using the nowcasted infection counts we obtained here.