

April 26, 2022

Making Sense of Covid-19 Mortality Estimates for India

By: Murad Banaji

The government should not issue bad-faith rebuttals of international Covid-19 mortality estimates that challenge official figures. It should instead release data that could better resolve uncertainties.

The World Health Organization (WHO) estimates that India could have seen four million excess deaths during the Covid-19 pandemic, according to news reports. This would give India over 25% of the estimated worldwide total of 15 million pandemic excess deaths and place India as the country with—by far—the largest number of pandemic deaths in the world.

However, Indian government objections (MoHFW 2022d) have stalled the release of the WHO’s estimates, potentially delaying an important effort to understand the global impact of the pandemic. Several news outlets in India have carried the government response verbatim, or quoted it sympathetically, framing the story as “India” standing up to a global body.

While the WHO estimate is consistent with this previous scholarship, government objections also follow a familiar pattern.

We do not need to look very far back to see the dangers in nationalist discourse of this kind. This time last year, India was in the midst of a brutal second wave of Covid-19 infections. The second wave was preceded by official complacency founded on narratives of successful control of the epidemic and—crucially—low mortality. These narratives, which hinted strongly that Indians were less vulnerable to the pandemic than others, had devastating consequences.

Since then, a growing body of work has examined pandemic mortality in India and found it to be much higher than official figures. While the WHO estimate is consistent with this previous scholarship, government objections also follow a familiar pattern. Papers on pandemic mortality in India have elicited crude rebuttals from the health ministry, ironically titled “COVID-19: Myth vs. Facts” (MoHFW 2021a, 2021b, 2022a-d). These have labelled the academic work variously “ill-informed”, “speculative”, “fallacious” and “completely inaccurate”.

Putting the numbers in context

Until the WHO report is available, it is hard to comment on the details of their approach to estimating excess mortality in India. But, in brief: the WHO estimate is neither surprising nor implausible.

Firstly, an estimate of four million excess deaths in India is close to estimates in a number of published studies and preprints. These come from several research groups and use data from the civil registration system (CRS), various surveys, and other government systems such as the health management information system (HMIS). Most of this work has resulted in estimates of between 3 and 5 million excess deaths.

A toll of around four million would be consistent with international data on Covid-19 fatality rates and the wide spread of disease before the vaccination drive took off.

For example, a paper in *Science* by Prabhat Jha and co-authors (2022) estimates 3.2 million excess deaths; my own preprint with Aashish Gupta (2021b), finds a central estimate of 3.8 million excess deaths up to June 2021; a working paper by Anand, Sandefur and Subramanian (2021) gives estimates of 3.4 to 4.9 million; a paper by Guilmoto (2022) gives estimates of 3.2 to 3.7 million; a recent paper in *The Lancet* from the Institute for Health Metrics and Evaluation (IHME) (COVID-19 Excess Mortality Collaborators 2022) gives an estimate of 4.07 million. Some estimates go a little lower (2.7 million from Leffler et al. 2022), or considerably higher (6.3 million from Malani and Ramachandran 2022). Since the WHO estimate lies very much in this range, anyone familiar with this literature will not find it particularly surprising.

Secondly, the WHO estimate is in keeping with the epidemiology of Covid-19. A toll of around four million would be consistent with international data on Covid-19 fatality rates and the wide spread of disease before the vaccination drive took off (Banaji and Gupta 2021b). To dismiss the WHO estimate as much too high is not just to dismiss the data and previous literature, but also, effectively, continue to claim that Indians are somehow special and less vulnerable than others to the severe effects of Covid-19.

Thirdly, the WHO estimate suggests that India has spectacularly failed to count its pandemic deaths, with around eight times as many pandemic excess deaths as recorded Covid-19 deaths. But in this too, India is not alone. Several countries, mostly in Asia and Africa, have comparable or even higher ratios of excess deaths to official Covid-19 deaths (Karlinski and Kobak 2021, COVID-19 Excess Mortality Collaborators 2022). On the other hand, some Asian and African countries, such as Iran and South Africa, have done better in counting pandemic deaths, with excess deaths being two to three times official Covid-19 deaths.

|| The majority of deaths can be missed in official figures, especially when disease sweeps through marginalised populations.

Finally, thanks to the efforts of journalists, we have an increasing understanding of *how* deaths could have been undercounted so badly in India. [Ground-level reports](#) have demonstrated how easily the majority of deaths can be missed in official figures, especially when disease sweeps through marginalised populations with little access to testing and healthcare (Banaji, Gupta, and Kumarappan 2021). Journalistic work has also helped us understand the huge variability in the ratio of excess deaths to pandemic deaths between states: poverty, access to healthcare, policy, and institutionalised dishonesty, have all played a part.

Estimating excess mortality

Since the WHO estimates seem in keeping with the available data and what we know of India's epidemic, what are the objections? Before examining these, it may be helpful to step back and think about how we estimate excess deaths.

For this, we need two things: (i) knowledge of how many deaths to *expect* in some period; and (ii) knowledge of how many deaths *actually occurred* in that period. Excess deaths are precisely the difference between these two quantities.

|| The uncertainties go in both directions: they can equally lead to overestimation or underestimation of mortality.

Expected deaths: Estimating expected deaths means asking: *If the pandemic had not occurred then how many people would have died during the pandemic period?* To answer this counterfactual question, we necessarily need some element of modelling.

Even in countries with good historical data, there may be trends in mortality prior to the arrival of Covid-19—for example if the population is ageing, or healthcare provision is improving. There may also be year-on-year fluctuations in mortality—either random, or associated with external events like heat-waves or diseases.

In countries such as India, aside from trends in mortality, which are quite slight, there are trends—and uncertainties—in death registration (Banaji and Gupta 2021a) which add uncertainty to expected deaths.

Actual deaths: In countries with complete vital registration, we can rely on death registration data to tell us how many people actually died during some period.

In India, however, there are gaps. For some parts of the country, there is, so far, no CRS data for the pandemic period. Elsewhere, there are open questions about what fraction of deaths is captured in the available data.

Complicating matters further, there is evidence the pandemic disrupted death registration. It is likely that communities with poor death registration are also those with weak access to healthcare and, perhaps, higher pandemic mortality.

Where does this leave us? Estimating excess deaths in India is not straightforward: both halves of the calculation present complications. There are uncertainties and there is space for disagreements.

But, crucially, if we work carefully with the available data, then we find that:

1. The uncertainties go in both directions: they can equally lead to overestimation or underestimation of mortality.

2. The uncertainties are bounded: we can put a range on them. For example, as we might guess from the literature review above, it is extremely difficult, in any scenario, to infer fewer than 2.5 million pandemic excess deaths from the available data.

What are the Indian government’s objections?

Let us begin with objections that have at least a grain of truth or plausibility.

The data used is “unofficial”.

For example, much of the available CRS data comes from journalists, obtained either through Right to Information requests or by other means. If this is a problem, the central government could surely accelerate the release of official CRS data to solve this problem. But, in any case, does this really change the picture?

As a case study, consider Andhra Pradesh, a state with freely available CRS data on its online portal. (It is, however, unclear whether the central government views such data as official or not.) The data shows an alarming 250,000 to 300,000 excess death registrations during the pandemic, compared with 15,000 official Covid-19 deaths in the state (Banaji and Gupta 2021c). Is the government implying that a significant fraction of these excess registrations in Andhra Pradesh are a consequence of errors?

Ultimately, government rebuttals have not argued why unofficial or incomplete data should lead to systematic bias or change estimates in any significant way.

CRS data is available for some regions and not others. The size and diversity of India make extrapolation to the whole country risky.

There is some truth to this, but the objection is overstated. CRS data of reasonable quality is available for more than half of the national population, with more limited data covering much of the remainder.

|| There is some evidence that states with the poorest data availability might have seen the highest mortality.

Extrapolation could lead to estimates that are too high if the missing data is mostly from regions which were spared the worst of the epidemic. But there is no evidence for this. Quite to the contrary, there is some evidence that states with the poorest data availability might have seen the highest mortality. For example, Gujarat and Uttar Pradesh are omitted from some studies because adequate CRS data is unavailable. But limited data from these two states suggests that pandemic mortality has been higher than the national average (Acosta et al. 2021, Citizens for Justice and Peace and The Wire 2022).

Rather than dismissing extrapolation, the key is to try and take into account the risks. ‘Sensitivity analyses’ tell us how the results would change if assumptions were varied. The authors of the WHO analysis stress that they carried out “extensive sensitivity... and cross-validation analyses” to address such uncertainties (Knutson et al. 2022). Their report will no doubt detail the outcomes.

Surveys are prone to biases.

There are always such risks with survey data. But a lot of our understanding of mortality in India is founded on government survey data, such as from the Sample Registration Survey (ORGI 2020), and the National Family Health Survey (Government of India 2021). Our understanding of the evolution of the Indian Covid-19 epidemic is also founded on (sero)survey data.

Survey data can fail to give valid conclusions if there are systematic biases that are not recognised by the authors. The goal in analysing survey data is precisely to examine these risks and correct for possible biases. The government rebuttals have never demonstrated that the papers using survey data fail to do this.

The WHO report uses the WHO’s Global Health Estimates to set expected levels of mortality.

This objection is to the WHO estimates (MoHFW 2022d), but similar objections (MoHFW 2022a) were earlier raised about the use of United Nations estimates of expected mortality in India.

It is unclear why the government has objections to using data curated by global bodies like the WHO and the UN. But the objections do highlight that there is uncertainty around baseline mortality in India.

Do such uncertainties about baseline mortality make a huge difference to estimates of pandemic excess deaths? Not really.

To understand the scale of the uncertainty, we can ask: how many people died in India in 2019? Government data leads to estimates ranging from 8.3 million to over 9 million. Meanwhile, the UN estimates, also based on Indian government data (Gerland 2014), would put the figure at over 9.5 million (Banaji and Gupta 2021a).

Do such uncertainties about baseline mortality make a huge difference to estimates of pandemic excess deaths?

Not really. Excess death estimates could go up or down by around 15%. Four million excess deaths could plausibly become 3.4 million or 4.6 million; but it could not drop to 0.5 million.

Importantly, assuming higher baseline mortality does not necessarily lead to higher estimates of excess deaths. For example, suppose UN estimates are used to set expectations of baseline mortality for comparison with survey data, as in the paper by Jha and co-authors (Jha et al. 2022). In that case we get *lower* estimates of excess mortality than we would obtain by using Indian government estimates of baseline mortality.

Excess deaths estimates rely on modelling.

This ‘criticism’ hints that the estimates are based on complex and questionable assumptions.

We need *some* modelling to estimate excess deaths, even in countries with high-quality mortality data. But the process can be transparent and fairly simple. None of the rebuttals have demonstrated—or even tried to demonstrate—why any underlying assumptions of the models used should lead to systematic overestimation of excess mortality.

A particular objection, in this vein, is that the WHO used a global model to make predictions for India (MoHFW, 2022d). A global model would not, of course, invalidate the results, unless we had reason to believe there were some special factors at play in India. But, in any case, the authors of the WHO estimates have explicitly denied this claim, stating, “We stress that for India the global predictive covariate model is not used and so the estimates of excess mortality are based on data from India only.” (Knutson et al. 2022).

Further objections: incoherent and confused

There are other objections from the government, which appear to be contradictory, incoherent, or meant merely to distract and raise doubts.

CRS data.

One health ministry statement claims that “all the births and deaths in the country get registered” (MoHFW 2021b). This is, of course, false, as government data plainly shows (Banaji and Gupta 2021a). But, while overstating the completeness of death registration, the same statement dismisses the conclusion — from CRS data — of a huge mortality surge. It is unclear if the ministry wants us to trust or distrust registration data as a basis for estimating mortality.

Mitigation and vaccination.

Another statement, from March, claims that the estimates are faulty because they do not take into account “multiple pandemic management efforts” (MoHFW 2022c) This appears to be a [category error](#), implying that if efforts such as mitigation and vaccination are taking place, then comparing pandemic deaths to pre-pandemic deaths is no longer a valid approach to computing excess mortality! Hopefully mitigation and vaccination reduce excess mortality; but they do not change how we measure it.

Test positivity.

We see similar confusion when the government claims that “variation in Covid-19 positivity rate within India was not considered for modelling purposes.” (MoHFW 2022d)

It is unclear how test positivity is supposed to relate to levels of mortality, or why it should be considered for mortality estimation. If test positivity is being taken as a proxy for disease spread, then there is ample serosurvey data from national and state-level surveys to inform us instead.

Conclusions

When all is said and done, the objections raised by the Indian government do not hold water. Many are incoherent and absurd. Others begin with a grain of truth—gaps in the data, risks in extrapolation and surveying, and so forth—but go on to conclude, without justification, that mortality must have been overestimated without demonstrating any systematic bias in the work.

|| There has been plenty of time for government to strengthen mortality surveillance and to gather data.

Reconstructing pandemic mortality in India is a challenge. There are a range of estimates in the literature, and we will never be able to put a single number on the scale of the catastrophe. But a growing body of evidence using *data from many different sources* points in the same direction: the pandemic toll has been many times higher than the 500,000 or so official Covid-19 deaths.

It has been almost a year since data started to emerge showing that most pandemic deaths were going uncounted in India. Even before this data emerged I [wrote in this magazine](#) (Banaji 2021) that there were strong indications of major Covid-19 death undercounting. Since then, there has been plenty of time for the government to strengthen mortality surveillance and to gather data which would shed more light on the scale of pandemic mortality in India.

Even now, many uncertainties could be resolved by accelerating the release of “official” CRS data for the pandemic period and conducting a large-scale mortality survey to fill in the gaps. Instead, government efforts seem to be focussed on rebutting scientific work.

|| Government objections have been focussed on the inconvenient conclusions of these studies.

These rebuttals reek of bad faith. They claim that all the studies massively overestimate excess mortality, but never acknowledge that published estimates could equally be *too low*. For example, one problem with using the CRS is that marginalised communities, where death registration is weak, could have been hit hardest and yet this impact would not be reflected in the data. Incorrect estimates of pre-pandemic mortality could also lead to underestimation of the pandemic toll.

There are no doubt valid criticisms to be made of the data and methodologies in each paper on pandemic mortality in India. But government objections have been focussed on the inconvenient conclusions of these studies.

It is a sign of the times that some media coverage portrays the government’s refusal to acknowledge the pandemic toll as an act of patriotic defiance. It is an absurd version of patriotism that discards data and evidence, and disrespects the millions of families for whom Covid-19 brought tragedy.

Murad Banaji is a London-based mathematician. He has been tracking India's Covid-19 epidemic since its very early days.

References:

Acosta, R. J., et al. (2021). All-cause excess mortality in the State of Gujarat, India, during the COVID-19 pandemic (March 2020-April 2021). *medRxiv*.

Anand A., Sandefur J., and Subramanian A. (2021) Three new estimates of India’s all-cause excess mortality during the COVID-19 pandemic. *Center for Global Development*.

Banaji, M. (2021). *Estimating Covid-19 Fatalities in India*. *The India Forum*.

Banaji, M. and Gupta, A. (2021a). *India: Notes on Estimates of Crude Death Rate and Registration Coverage*.

Banaji M. and Gupta, A. (2021b). *Estimates of pandemic excess mortality in India based on civil registration data*. *medRxiv*.

Banaji, M. and Gupta, A. (2021c). *Excess mortality in Andhra Pradesh during the COVID-19 pandemic: A factsheet*.

Banaji, M., Gupta, A. and Kumarappan L. (2021) What local news reports tell us about Covid-19 mortality in rural areas of North and Central India. *Scroll*.

Citizens for Justice and Peace and The Wire (2022). In 2021, COVID Deaths Surged in Stunning Fashion in Eastern UP, Investigation Finds. *The Wire*.

COVID-19 Excess mortality collaborators (2022). Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. *Lancet* 399:1513-36.

Government of India (2021). *National Family Health Survey*.

Gerland, P. (2014). UN Population Division's methodology in preparing base population for projections: case study for India. *Asian Population Studies* 10 (3): 274–303.

Guilmoto, C. (2022). An alternative estimation of the death toll of the Covid-19 pandemic in India. *PLoS ONE* 17(2).

Jha P., Deshmukh Y., et al. (2022). COVID mortality in India: National survey data and health facility deaths. *Science* 375: 667-671.

Karlinsky, A. and Kobak, D. (2021). Tracking excess mortality across countries during the COVID- 19 pandemic with the World Mortality Dataset. *eLife* 2021;10:e69336

Knutson V., et al. (2022). Extract from “Estimating Country-Specific Excess Mortality During the COVID-19 Pandemic”.

Leffler, C. T., Lykins, V., Das, S., Yang, E. and Konda, S. (2022) Preliminary Analysis of Excess Mortality in India During the COVID-19 Pandemic. *The American Journal of Tropical Medicine and Hygiene*. (April 2022)

Malani, A. and Ramachandran, S. (2022). Using Household Rosters from Survey Data to Estimate All-cause Mortality during COVID in India. *Becker Friedman Institute Working Paper*.

MoHFW (Ministry of Health and Family Welfare) (2021a). COVID19 Mortality: Myth Vs. Facts. *Bulletin*, 22 July 2021.

MoHFW (2021b). COVID19 Mortality: Myth Vs. Facts. *Bulletin*, 27 July 2021.

MoHFW (2022a). COVID-19 Myth vs Facts. *Bulletin*, 14 January 2022.

MoHFW (2022b). COVID-19 Myth vs Facts. *Bulletin*, 17 February 2022.

MoHFW (2022c). COVID-19 Myth vs Facts. *Bulletin*, 11 March 2022.

MoHFW (2022d). COVID-19 Myth vs Facts *Bulletin*, 16 April 2022.

ORGI (Office of the Registrar General and Census Commissioner of India) (2020). *Sample Registration System Bulletin (Various Years)*.

Singh K. D. and Nolen, S. (2022). India Is Stalling the WHO's Efforts to Make Global Covid Death Toll Public. *New York Times*. (16 April 2022)