India: notes on pre-pandemic estimates of crude death rate and registration coverage[1]

The 2019 CRS report provides both national and subnational data on registered deaths, along with estimates of registration coverage. Estimated deaths are calculated using the Sample Registration Survey (SRS) which gives national and subnational (i.e., state- or Union Territory-) estimates of crude death rates (CDRs). The 2018 SRS annual statistical report (henceforth "the 2018 SRS report") is available here. The SRS report for 2019 is now available here, and an SRS bulletin for 2020 is available here.

National or subnational CDR values from the SRS reports can be multiplied by national or subnational population estimates for the year, based on population projections, to estimate total deaths in any given year. The level of death registration nationally or sub-nationally can then be estimated by comparing registered deaths and estimated total deaths.

However, a number of different calculations (below), including calculations based on SRS age-wise death rates, and calculations based on data from the National Family Health Survey (NFHS), suggest that SRS-CRS estimates may underestimate the crude death rate in India at national, and also subnational, levels.

The national 2018-2019 SRS-CRS estimates of CDR (6.0-6.2) and registration coverage (92-96%)

According to the 2019 CRS report, 7,641,076 deaths were registered nationally in 2019. The 2018 SRS report estimates India's CDR as **6.2** per 1K population, while the estimated 2019 national population was 1,332,900K. Assuming no change in the CDR between 2018 and 2019 we would expect around 6.2*1,332,900 = 8,263,980 deaths during 2019. This estimate would, indeed, imply registration coverage during 2019 to be **92%** (to the nearest full percent), as given in the 2019 CRS report.

If, instead of the 2018 value, we use the estimated national CDR of **6.0** per 1K population from the 2019 SRS report, we would expect around 6.0*1,332,900 = 7,997,400 deaths during 2019. This estimate would imply registration completion during 2019 to be around **96%.**

To summarise these calculations: the national CDR of 6.0-6.2 given in the SRS reports for 2018-19 implies a total of 8.0-8.3 million deaths in 2019, and registration coverage of 92-96%.

The United Nations' estimates

The United Nations estimated CDR in India to be **7.2** per 1K for the period 2015-20 (link to excel file). This is considerably higher than the 2018-2019 SRS estimates of 6.0-6.2%. At a CDR of 7.2, based on the population projections, we would expect 9.6 million deaths nationally in 2019, giving estimated registration completion of **80%**.

Estimates from subnational data in the SRS reports

Registration estimates for each state and Union Territory (UT) are also given in the 2019 CRS report, based on 2018 SRS estimates of crude death rates along with registered deaths in that state or UT in 2019. These subnational estimates are capped at 100%, i.e., if more deaths were registered

1 Updated on 14th June 2022, by Murad Banaji and Aashish Gupta.

than expected from the estimated CDR, then registration coverage is assumed to be 100%. For example, Andhra Pradesh had an estimated CDR of 6.7 in the 2018 CRS report (this dropped to 6.4 in the 2019 SRS report), and an estimated 2019 population of 52,221K, giving expected deaths in 2019 at around 350K (334K based on the 2019 CDR). However, 401,472 deaths were actually registered in the state in 2019, 15-20% more than expected. Formally speaking, this gives registration coverage in the state of 115-120% during 2019.

Like Andhra Pradesh, several other states, including Gujarat, Karnataka, Maharashtra and West Bengal, saw considerably more death registrations in 2019 than estimated from SRS-based CDR estimates. At the same time, 2018-19 SRS-based estimates for Uttar and Pradesh and Bihar (together) are of only 61-62% completion, equivalent to around 800,000 unregistered deaths in these states, at odds with national completion being above 90%.

One interpretation, perhaps unlikely, is that the additional deaths registered in states such as Andhra Pradesh during 2019 were actually of residents of other states. The most probable explanation is that the SRS underestimated true mortality in these states, and that their true CDR in 2019 was considerably higher than that estimated from the SRS.

We can use subnational data to estimate national CDR and registration coverage as follows. We assume that in each state or UT:

- whenever fewer deaths were registered than expected from the SRS subnational CDR, the true total deaths in the state was the SRS-estimated value; and
- whenever more deaths were registered than expected from the SRS subnational CDR, then the true total deaths in the state was the number of registered deaths.

Following this procedure across all states and UTs, we obtain an estimated total of 8,837,847 deaths nationally in 2019 (based on the 2018 SRS report); or 8,706,418 (based on the 2019 SRS report). The discrepancy of over 500,000 between these numbers and the corresponding SRS-CRS-based estimates of total deaths (around 8.0-8.3 million depending on whether we use values of CDR from the 2018 or 2019 SRS reports) arises for the reasons discussed above.

Overall, using subnational CDR estimates from the 2018 SRS report we find an estimated national CDR in 2019 of 8,837,847/1,332,900 = **6.6** per 1K, somewhat higher than the 2018 SRS-based estimate of 6.2. This higher CDR estimate gives estimated national registration coverage of 7,641,076/8,837,847, namely around **86%**, rather than 92%. If we use subnational CDR estimates from the 2019 SRS reports, the numbers change a little. We find estimated national CDR to be 8,706,418/1,332,900 = **6.5** per 1K as compared to the SRS-based estimate of 6.0. This gives estimated national registration coverage of 7,641,076/8,837,847, namely around **88%,** rather than 96%.

To summarise these calculations: subnational data from the 2018-2019 SRS reports implies a CDR of 6.5-6.6, equivalent to a total of 8.7-8.8 million deaths in 2019, and national registration coverage of 86-88%. But the data also strongly suggests that estimates of subnational CDR from the SRS may be systematically too low; and hence even these estimates of coverage may be too high.

Estimates using SRS age-wise death rates

The 2018 and 2019 SRS reports give estimated percentages of the population in different age brackets (Table 1 in the reports) and also estimated death rates in different age brackets (Table 8 in the reports). This data can also be used to estimate the national CDR.

The reports give estimated percentages of the population in different age brackets to only one decimal place. This level of precision, especially in the older age groups where the percentages are small, is very unsatisfactory for the purposes of estimating population and mortality in different age-brackets. Nevertheless, we can use these estimates along with the projected 2019 population to estimate total 2019 populations in each age-bracket. We can then use age-wise death rates to estimate deaths in each age group, and hence total deaths in 2019. This process gives an estimate of 8.38 million deaths in 2019 (based on death rates in the 2018 SRS report), and 8.18 million deaths (based on death rates in the 2019 SRS report).

If we dig a little deeper, however, we find that the SRS estimated population fractions in the older age-groups could be too low. According to Table 1 in the 2018 and 2019 SRS reports, about 5.2% of the population was 65 or above. But already by the 2011 census, 5.5% of the population was 65 or above and, according to the population projections, this should have risen to 6.2% by 2016 and 6.9% by 2021. We would thus expect around 6.4%-6.7% of the population to be over 65 during 2018-2019, considerably higher than the 5.2% estimated in the SRS reports. This discrepancy makes a fairly major difference to estimated CDR.

The population projections give projected population pyramids for 2016 and 2021, but not intervening years. We can use either estimated 2016 or 2021 population pyramids to estimate 2019 CDR. Estimates depend on the fraction of the over-80s who are assumed to be over 85, which is not given in the projected population pyramids, so we set this to be the fraction estimated in the 2018 and 2019 SRS, namely, 0.375.

Using the 2018 SRS age-wise death rates and the projected 2016 age pyramid. We expect 8.78 million deaths in 2016 which, over an estimated population of 1,290 million gives a CDR of **6.8** per 1K. This CDR estimate, applied to the estimated 2019 population of 1,333 million gives 9.1 million deaths in 2019, corresponding to registration coverage in 2019 of **84%**.

Using the 2018 SRS age-wise death rates and the projected 2021 age pyramid. We expect 10.23 million deaths in 2021 which, over an estimated population of 1361 million gives a CDR of **7.5** per 1K. This CDR estimate, applied to the estimated 2019 population of 1,333 million gives 10.0 million deaths in 2019, corresponding to registration coverage in 2019 of **76%**.

Using the 2019 SRS age-wise death rates and the projected 2016 age pyramid. We expect 8.52 million deaths in 2016 which, over an estimated population of 1,290 million gives a CDR of **6.6** per 1K. This CDR estimate, applied to the estimated 2019 population of 1,333 million gives 8.8 million deaths in 2019, corresponding to registration coverage in 2019 of **87%**.

Using the 2019 SRS age-wise death rates and the projected 2021 age pyramid. We expect 9.93 million deaths in 2021 which, over an estimated population of 1361 million gives a CDR of **7.3** per 1K. This CDR estimate, applied to the estimated 2019 population of 1,333 million gives 9.7 million deaths in 2019, corresponding to registration coverage in 2019 of **79%**.

To summarise these calculations: age-wise CDRs from the 2018-2019 SRS reports, combined with population projections, imply a CDR of 6.6-7.5, equivalent to a total of 8.8-10.0 million deaths in 2019, and national registration coverage of 76-87%.

Estimates based on the National Family Health Survey (NFHS-5)

Recently released NFHS summary reports also suggest that official death registration levels in the CRS reports, including the subnational estimates, may be overestimated. The NFHS asks

respondents about deaths of any usual family member in the last three years, and whether the death was registered. Estimated completion from the NFHS, along with the number of registered deaths in a region can be used to estimate total deaths and hence a CDR in each state or UT.

If we compare observed registration completion during 2018 in the NFHS data to estimated completion using the SRS and CRS data, we find that the SRS-CRS consistently overestimates completion relative to the NFHS in the larger states, with only a few exceptions (see Figure 1). This suggests that the SRS consistently underestimates the crude death rate.

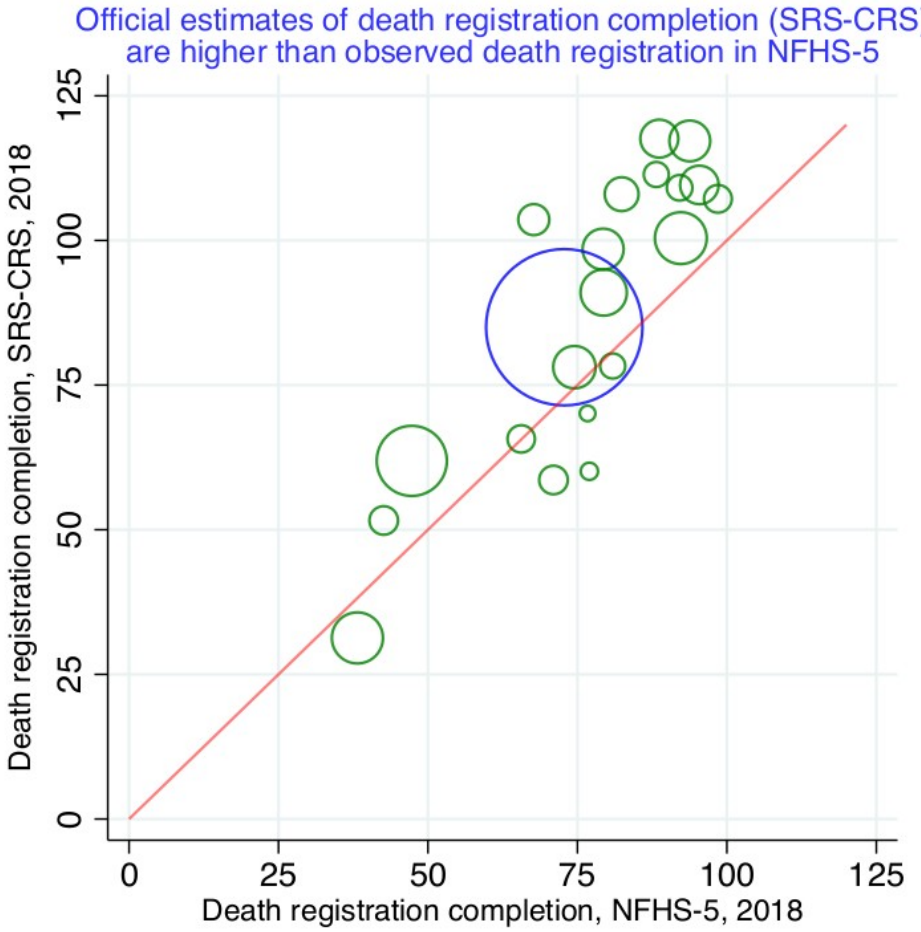

Figure 1: *Estimated death registration completion based on NFHS data and SRS-CRS data. Registration estimates from NFHS-5 for 2018 are calculated from NFHS micro-data. Green circles show major states, as classified by the SRS, and the blue circle shows India. Circle sizes represent 2018 population, as estimated by the National Committee on Population.*

Using NFHS data for 2018, this approach leads to an estimated national CDR of **7.2**. The corresponding estimated registration coverage in 2018 was 73%. This compares to an SRS-CRS based estimate of 85%. If the national CDR remained unchanged at the same level in 2019, this would imply a total of 9.65 million deaths in 2019, corresponding to registration completion in 2019 of **79%**.

Summary of estimates, and discussion

In the table below we summarise the various estimates of the 2019 crude death rate discussed above and the consequent estimates of registration completion during 2019.

	2019 CRS and 2018-19 SRS (national data)	**2019 CRS and 2018-19 SRS (subnational data)**	**United Nations estimate**	**2018-19 SRS age-wise death rates using projected 2016 age pyramid**	**2018-19 SRS age-wise death rates using projected 2021 age pyramid**	**NFHS-5 based estimate**
CDR	6.0-6.2	6.5-6.6	7.2	6.6-6.8	7.3-7.5	7.2
registration coverage	92-96%	86-88%	80%	84-87%	76-79%	79%

We see that SRS-CRS estimates using aggregate national data give the lowest estimates of CDR. Even the SRS-CRS estimates using subnational data are somewhat lower than estimates from other sources. Notably, there is a cluster of estimates of CDR at around 7.0-7.2, coming from both the NFHS, and SRS age-wise data coupled with population projections. These estimates correspond to registration completion of around 80-82% nationally during 2019.

Overestimating registration coverage is equivalent to underestimating mortality. Since excess mortality calculations rely on estimates of baseline mortality, we also risk underestimating excess mortality during the pandemic period and underestimating other quantities such as COVID-19 infection fatality rate based on excess deaths.

The data available from a number of sources suggests that taking at face value the SRS-CRS estimated pre-pandemic national registration completion of 90% or higher is problematic. Such high estimates of completion are inconsistent with subnational SRS data, with age-stratified mortality rates from the SRS, and with estimates of completion based on the NFHS. Several calculations presented above suggest that it is more likely that around 9.5 million deaths occurred nationally in 2019, corresponding to a CDR of around 7.2, and registration completion of around 80%.