**Supplementary appendix**

Supplement to: Ram U, Jha P, Gerland P, et al. **Age-specific and sex-specific adult mortality risk in India in 2014: analysis of 0∙27 million nationally surveyed deaths and demographic estimates from 597 districts**

The attached tables are available in excel format by emailing the lead authors ([cghr@smh.ca](mailto:cghr@smh.ca))

**Contents**

|  |  |  |
| --- | --- | --- |
|  | **Description** | **Page nos.** |
| Web table 1 | UN 2014 live births and age-specific deaths for India | 2 |
|  | Determinants of child and adult mortality | 3 |
| Web table 2 | Descriptions of the variables, source and year of data, percentage of missing cases, overall mean and standard deviation (before and after imputation for missing cases) of the known correlated in explaining premature mortality used in the regression analysis for district mortality risk and age standardised mortality for children below 15 years and men and women aged 15-69 , India 2014 | 4-6 |
| Web table 3 | Population, death, age-standardised mortality rates (ASMR) and district level conditional probability of dying (CpoD\*) at an age of 15-69 years for men, India, 2014 | 7 |
| Web table 4 | Population, death, age-standardised mortality rates (ASMR) and district level conditional probability of dying (CpoD\*) at an age of 15-69 years for women, India, 2014 | 8 |
| Web table 5 | Population (15-69 years), deaths (15-69 years), death rate (15-69 years), age-standardized death rate (15-69 years), district level conditional probability of dying at an age of 15-69 years for men, India, 2014 | 9-18 |
| Web table 6 | Population (15-69 years), deaths (15-69 years), death rate (15-69 years), age-standardized death rate (15-69 years), district level conditional probability of dying at an age of 15-69 years for women, India, 2014 | 19-28 |
| Web table 7 | Total number of districts, districts having higher \*/lower † conditional probability of dying (CpoD\*) at an age of 15-69 years for men and women, India, 2014 | 29 |
| Web table 8 | Results of the regression analysis for district mortality risk and age standardised mortality for children below 15 years and men and women aged 15-69 years | 30 |
|  |  |  |

**Inputs to methods:**

The main text provides the key methods used for the calculations of the age and gender-specific death rates. These are additional details.

**UN national deaths for India – 2014:** We use UN live births and deaths for India for the year 2014 for males and females.

|  |  |  |  |
| --- | --- | --- | --- |
| **Web table 1: UN 2014 live births and age-specific deaths for India** | | | |
| **Births/Deaths** | **Male** | **Female** | **Both Sex** |
|  |  |  |  |
| **Live Births** | 14,586,171 | 13,169,226 | 27,755,397 |
|  |  |  |  |
| **Deaths** |  |  |  |
| 0 | *566,983* | *521,917* | *1,088,900* |
| 1-4 | *155,321* | *181,257* | *336,720* |
| 5-9 | *55,323* | *55,262* | *110,598* |
| 10-14 | *53,105* | *51,470* | *104,585* |
| 15-19 | *77,641* | *77,833* | *155,492* |
| 20-24 | *115,343* | *100,566* | *215,915* |
| 25-29 | *137,354* | *97,522* | *234,851* |
| 30-34 | *161,585* | *94,635* | *256,161* |
| 35-39 | *194,597* | *101,389* | *295,898* |
| 40-44 | *229,007* | *117,242* | *346,143* |
| 45-49 | *275,785* | *147,279* | *422,944* |
| 50-54 | *340,678* | *193,273* | *533,819* |
| 55-59 | *420,621* | *271,122* | *691,628* |
| 60-64 | *494,210* | *372,306* | *866,457* |
| 65-69 | *540,328* | *458,241* | *998,578* |
| 70-74 | *553,783* | *514,207* | *1,068,063* |
| 75-79 | *500,472* | *502,762* | *1,003,355* |
| 80-84 | *375,260* | *402,473* | *777,860* |
| 85+ | 344,006 | 419,349 | 763,544 |
| **All Ages** | **5,591,404** | **4,680,106** | **10,271,509** |

**State Death Envelopes:**

1. For the 20 larger states and for India nationally, the age-specific death rates (ASDR) for males and females were averaged from the SRS for the years 2011, 2012 and 2013 (Source: Registrar General of India. Sample registration system statistical report 2011, 2012 and 2013. New Delhi: Government of India, 2014; Table – 8 in each). We excluded observations where the rates were 0.0.
2. The ASDR for Uttarakhand state were available for the year 2010 only.
3. For the remaining 14 smaller states and union territories (Chandigarh, Sikkim, Arunachal Pradesh, Nagaland, Manipur, Mizoram, Tripura, Meghalaya, Daman and Diu, Dadra Nagar and Haveli, Goa, Lakshadweep, Pondicherry and Andaman and Nicobar Islands), the ASDR rates were available only up to 2006 and for the year 2010.
4. For Lakshadweep, the average ASDR for females at ages 15-19 was zero for all the years and thus we replaced it with the average of Kerala for both sexes.

**Determinants of child and adult mortality**

An initial list of 47 public health, geographic and demographic variables were collected from various sources for 597 districts in India (Web table 2 for the complete description of the variables, sources and percentage of missing entries). Of the 47 variables, 22 were used for further analyses – a greater number was compiled to assist with missing data analyses. While only 5.2% of the complete 47 variable dataset was missing, a list wise deletion of missing data for the 22 final variable dataset would result in the excessive exclusion of 362 districts from analysis. Multiple imputations for all 47 variables of the larger dataset were therefore conducted using AMELIA II – Version 1.7.2 1. R statistical software (www.r-project.org) was used for the logistic regression analysis with the lme4 package 2. The models were fit by maximum likelihood (Laplace Approximation) and the betas were converted to odds ratios via an exponential function. The average of five imputed datasets was used for further analyses. Descriptive statistics for all of the study variables and imputed values are found in Web table 2; the imputed means and standard deviations for all of the variables were within the expected range of the initial dataset.

For the multivariable analysis, a principal component analysis was applied to 17 variables with Varimax rotation. All KMO values for the individual variables were above 0.5 and the overall MSA was 0.74, indicating the data were sufficient for the analysis. The Bartlett’s test of sphericity 3 (χ²(136)= 9285.866, p<0.001) showed that there were patterned relationships between the items. Three factors were retained that explained a cumulative variance of 56%. Web table 2 shows the factor loading after rotation using a significant factor criterion of 0.35.

Regressions on the age–standardized mortality rates at ages 15–69 for males and females and for children ages 0-14 were conducted to identify key correlates of mortality. Logistic regressions were used to assist with interpretation by generating odds-ratios and a random intercept model was used to account for state-level variation. Tertiles were created for the variable water and sanitation, sex ratio (all ages M/F), state language spoken, smoking prevalence in males, and malaria exposure, as well as the three factors for “general public health indicators”, “population density” and “access to primary care units”. The original explanatory variables were all continuous. For the final logistic regression presented in the Web table 7, page 29, we converted the explanatory variables into tertiles where the lowest tertile was coded to 1, the second to 2 and the highest tertile to 3. A simpler mixed-effect linear regression was repeated with similar explanatory power (not shown)

**References**

1. Honaker J, Gary K. What To Do About Missing Values In Time Series Cross-Section Data*. American Journal of Political Science* 2010; 54(3): 561-81.
2. Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 2013; 4: 133–42.
3. Dziuban C D, Shirkey E C. When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin* 1974; 81(6): 358-61.