

THE LANCET

Global Health

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed.
We post it as supplied by the authors.

Supplement to: Verhulst A, Romero Prieto J, Alam N, et al. Divergent age patterns of under-5 mortality in south Asia and sub-Saharan Africa: a modelling study. *Lancet Glob Health* 2022; published online Sept 8. [https://doi.org/10.1016/S2214-109X\(22\)00337-0](https://doi.org/10.1016/S2214-109X(22)00337-0).

Contents

Webappendix 1: Age patterns of under-5 mortality in the vital registration (VR) model and potential divergent age patterns	2
Webappendix 2: Additional figures and tables	4
Webappendix 3: Data and methods for estimating mortality	16
Webappendix 4: Classification of the age pattern of under-5 mortality in each DHS	21

Webappendix 1: Age patterns of under-5 mortality in the vital registration (VR) model and potential divergent age patterns

Guillot et al. [1] compiled the Under-5 Mortality Database (U5MD) by collecting vital registration (VR) data from 25 high-income countries for the period 1841-2016. The database provides 1,741 annual distributions of under-5 deaths and corresponding death rates by detailed age. For 1,219 of these distributions of deaths (1920-2016), 22 age groups were harmonized with the following exact-age cut-off points: 0, 7, 14, 21, 28 days; 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 21 months; and 2, 3, 4, 5 y. The U5MD is publicly available at <https://web.sas.upenn.edu/global-age-patterns-under-five-mortality/>.

Guillot et al. summarized the age regularities of these harmonized 1,219 distributions of deaths into a log-quadratic model adapted from the work of Wilmoth et al. [2]. Let $q(x)$ denote the cumulative probability of dying from birth to age x . The model assumes $q(x)$ to be a log-quadratic function that depends on two entry values: (i) the under-5 mortality rate $q(5y)$, which determines the overall level of under-5 mortality; and (ii) parameter k that controls the shape of the age pattern of mortality between ages 0 and 5 y. Formally,

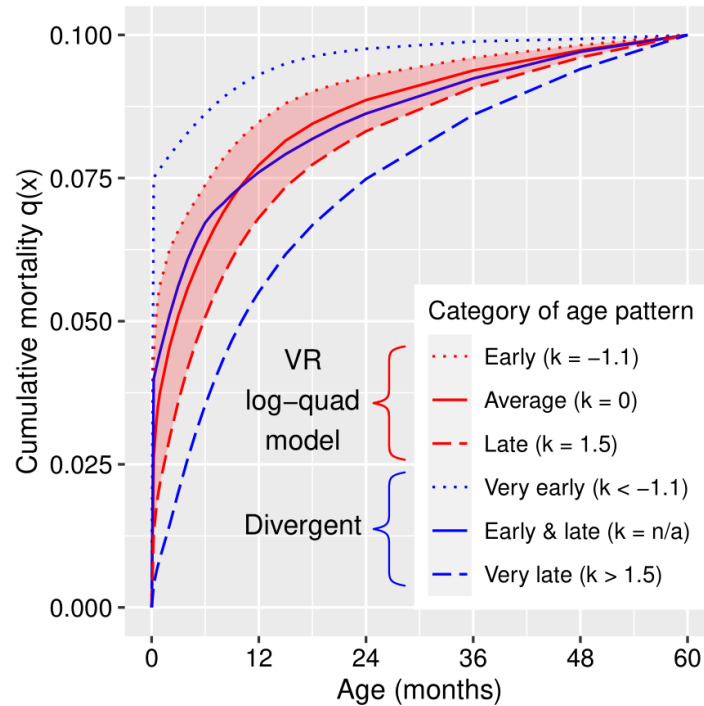
$$\ln[q(x)] = a_x + b_x \cdot \ln[q(5y)] + c_x \cdot \ln[q(5y)]^2 + v_x \cdot k \quad (1)$$

Age x takes on the 22 cut-off points reported above. Parameter k can vary continuously, with values ranging between -1.1 and +1.5 in the U5MD. Positive values of k correspond to later age patterns, that is with a higher-than-average concentration of mortality after 28 days of age at a given level of under-5 mortality. Negative values of k correspond to earlier age patterns, with a higher-than-average concentration of mortality before 28 days.

When $k = 0$, the VR model predicts a mortality schedule of 22 $q(x)$ values from birth to age 5 y corresponding to the model's average age pattern of mortality. Predictions can be obtained from any single rate or probability of dying within the age range 0 to 5 y, i.e., using any value of x as age boundaries. Parameter k can also be estimated by fitting multiple observed entry points (not used in the present study).

In Fig. S1, the red area corresponds to the VR model's range of variation in the predicted curve of cumulative mortality $q(x)$ for $q(5y) = .100$, with k ranging from -1.1 to +1.5. Examples of divergent age patterns—located outside the range of variation of the VR model—are stylized in blue. In some countries, $q(x)$ curves may follow a much earlier or much later age pattern than permitted in the model. The $q(x)$ curve could also cross over the average log-quadratic $q(x)$ curve as illustrated by the solid blue line. This crossover corresponds to an age pattern that is both earlier and later than the average observed in the U5MD at a given level of $q(5y)$, and thus not well represented by the model.

Fig. S1: Range of predictions of the VR model versus divergent age patterns for the curve of cumulative probabilities of dying $q(x)$ between ages 0 and 5 y.

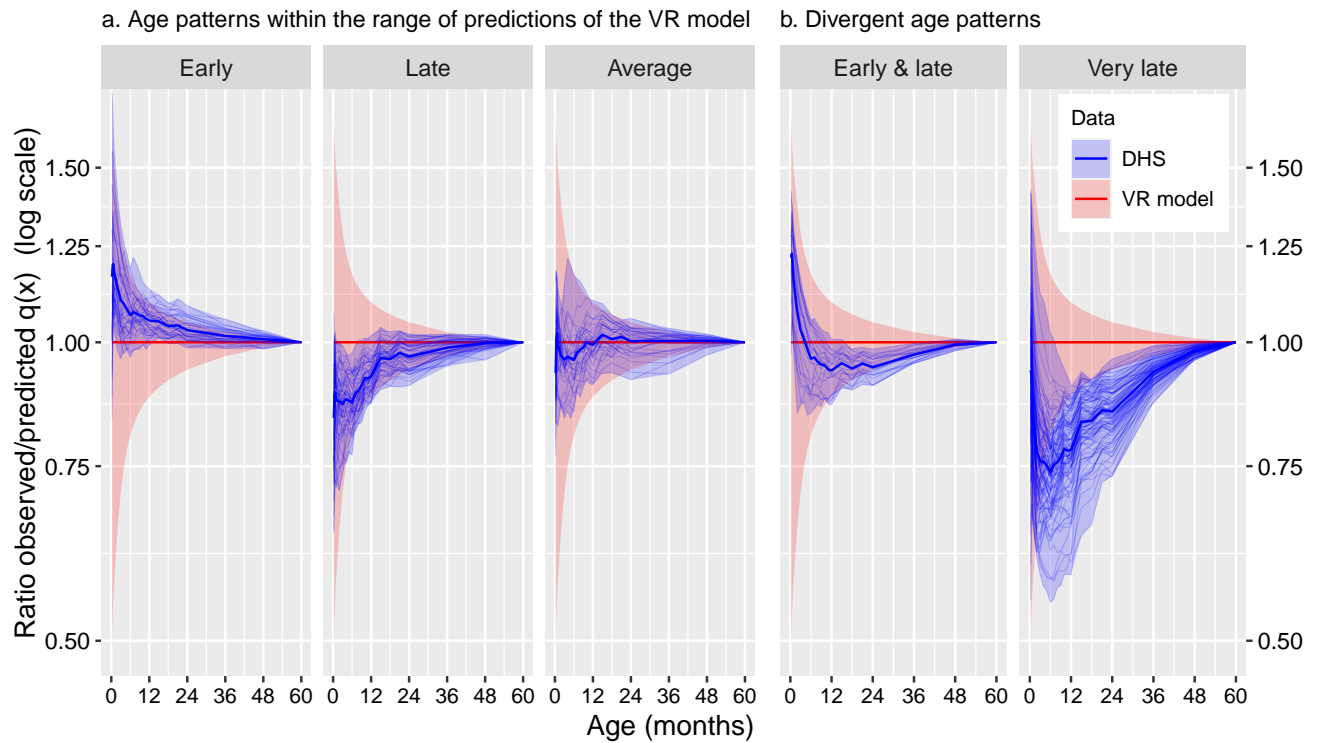


We generated all the results using the *logquad5q0* R package available at <https://github.com/verhulsta/logquad5q0>. This package implements the VR model used in this study

1. Guillot M, Romero Prieto J, Verhulst A, Gerland P. Modeling Age Patterns of Under-5 Mortality: Results From a Log-Quadratic Model Applied to High-Quality Vital Registration Data. *Demography*. 2022;59(1):321-47.
2. Wilmoth J, Zureick S, Canudas-Romo V, Inoue M, Sawyer C. A flexible two-dimensional mortality model for use in indirect estimation. *Population studies*. 2012;66(1):1-28.

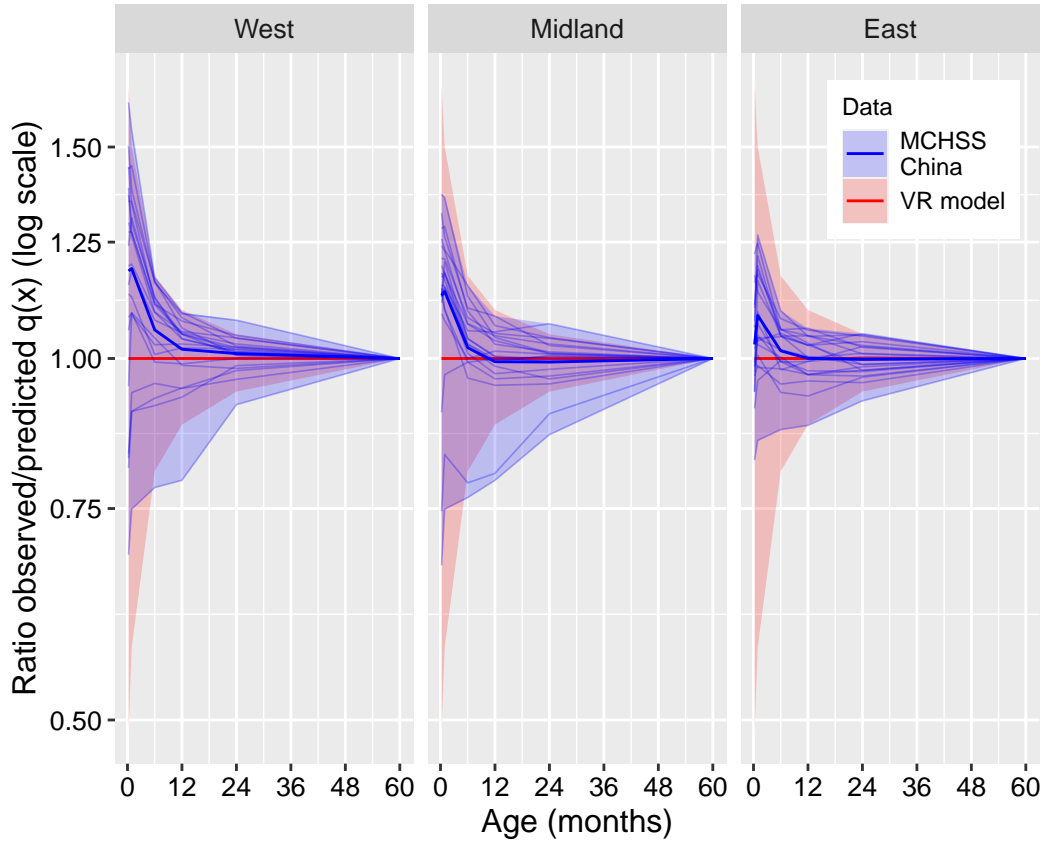
Webappendix 2: Additional figures and tables

Figure S2.1: Ratio of observed to predicted cumulative probabilities of dying from birth to age x ($q(x)$) controlling for $q(5y)$ across DHS.



Ratios were computed by dividing observed by predicted $q(x)$ values for 22 age groups between ages 0 and 5 y . The blue lines represent the ratios for each DHS. The thicker blue line represents the mean ratios across DHS. The red area represents the range of predictions of the VR model. Number of surveys (number of countries) per category: early 37 (20), late 59 (25), average 35 (20), early and late 22 (11), very late 120 (38).

Figure S2.2: Ratio of observed to predicted cumulative probabilities of dying from birth to age x ($q(x)$) controlling for $q(5y)$ for the data of the China National Maternal and Child Health Surveillance System (2001-2008, 2014-2015)

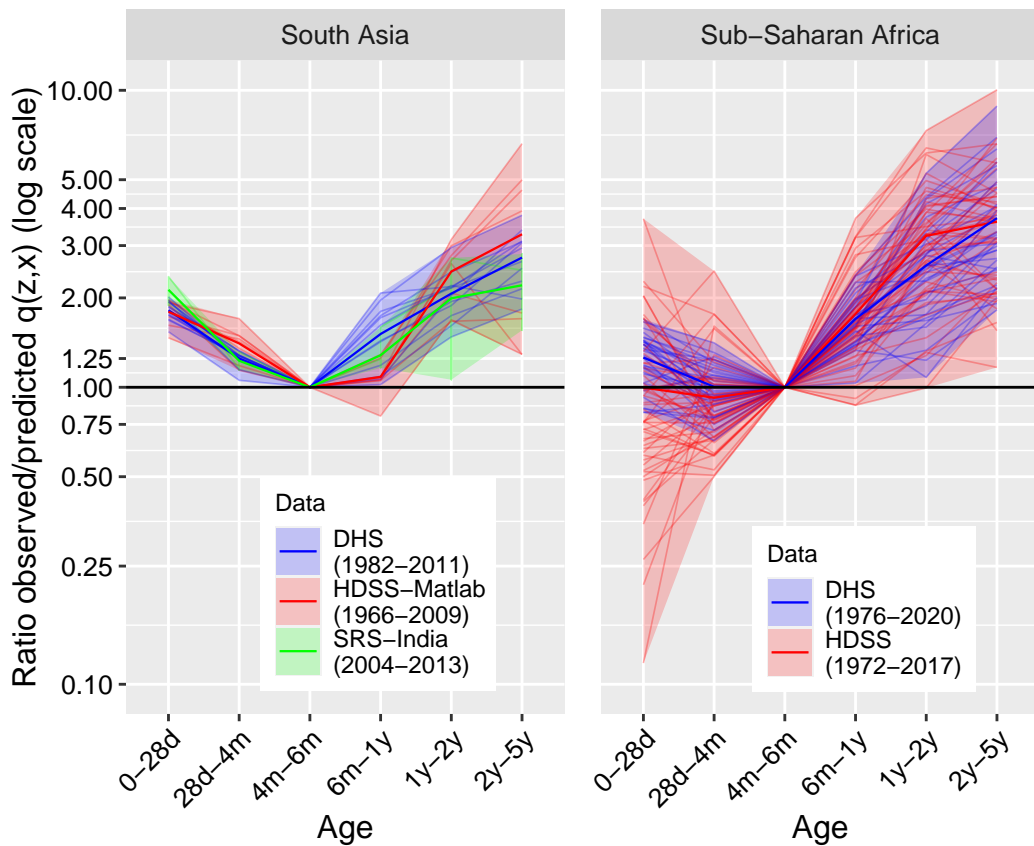


Ratios were computed by dividing observed by predicted $q(x)$ values for 6 age groups ($x = 7d, 28d, 6m, 12m, 2y, 3y, 5y$). The blue lines represent yearly ratios for each subnational area. The thicker blue line represents the mean ratios across areas. The red area represents the range of predictions of the vital registration model. Among the 60 lines, 50 correspond to an early or average age pattern of under-5 mortality (37 for early and 13 for average). We attribute the very late age patterns (4) to data quality issues.

Table S2.1: Frequency of the categories of age patterns of under-5 mortality observed in DHS by world region

Region	Early	Late	Average	Early & late	Very late	Discarded	Total
Asia	25	17	13	10	4	2	71
Central Asia	3	0	3	0	0	1	7
South Asia	9	1	1	10	2	0	23
Southeast Asia	2	13	6	0	2	0	23
West Asia	11	3	3	0	0	1	18
Europe	1	0	1	0	0	1	3
Latin America & Caribbean	7	28	10	0	2	1	48
North Africa	2	1	9	0	0	0	12
Oceania	0	1	0	0	0	0	1
Sub-Saharan Africa	2	12	2	12	114	0	142
Eastern Africa	0	7	1	4	46	0	58
Middle Africa	0	0	0	0	17	0	17
Southern Africa	2	4	1	1	1	0	9
Western Africa	0	1	0	7	50	0	58
All	37	59	35	22	120	4	277

Figure S2.3: Ratio of observed to predicted $q(z, x)$ controlling for $q(4m, 6m)$ across sources of data in South Asia and sub-Saharan Africa



$q(z, x)$ is used to denote the probability of dying between ages z and x . Ratios were computed by dividing observed by predicted $q(z, x)$ values for 6 age groups between ages 0 and 5 y . The lines represent the ratios for each dataset. The thicker line represents the mean ratios across datasets. The DHS include those with divergent age patterns of under-5 mortality.

Table S2.2: Mean ratio of observed to predicted $q(z, x)$ controlling for $q(4m, 6m)$ across sources of data in South Asia

Data	Country	Period	Resi.	n	Age group							
					0-28d	28d-4m	4m-6m	6m-1y	1y-2y	2y-5y		
DHS	Bangladesh	1983-1994	Total	3	1.77	1.29	1	1.20	1.85	2.99		
		1995-2011	Total	3	1.76	1.19	1	1.21	1.72	2.09		
	India	1982-1994	Total	2	1.91	1.30	1	1.91	2.31	3.02		
		1995-2006	Total	1	2.02	1.24	1	1.72	2.16	2.65		
	Nepal	1986-1994	Total	1	1.97	1.15	1	1.97	2.96	3.80		
		1995-2006	Total	2	1.72	1.28	1	1.71	2.18	2.55		
	All	1982-2011	Total	12	1.82	1.25	1	1.51	2.07	2.73		
			Rural	12	1.87	1.25	1	1.53	2.11	2.86		
	HDSS	Bangladesh	Matlab	1965-1979	Rural	3	1.95	1.52	1	1.14	2.87	5.41
				1980-1994	Rural	3	1.72	1.46	1	1.12	2.25	2.47
1995-2009				Rural	3	2.03	1.40	1	1.08	3.07	1.64	
Chakaria		2007-2011	Rural	1	2.31	1.40	1	0.89	2.10	1.74		
SRS		India	Central	2004-2013	Total	2	2.27	1.23	1	1.24	2.08	2.37
	East		2004-2013	Total	2	2.14	1.22	1	1.22	2.72	2.97	
	North		2004-2013	Total	2	2.01	1.16	1	1.24	1.38	1.49	
	South		2004-2013	Total	2	2.23	1.35	1	1.15	0.88	1.15	
	West		2004-2013	Total	2	2.15	1.26	1	1.43	2.11	1.97	
	CS		India	Tamil Nadu	1998-2000	Rural	1	2.68	2.83	1	—	—
Nepal		Sarlahi		1999-2000	Rural	1	1.82	1.46	1	0.76	1.52	1.75

$q(z, x)$ is used to denote the probability of dying between ages $z-x$. Ratios were computed by dividing observed by predicted $q(z, x)$ values for 6 age groups between ages 0 and 5 y . The DHS include those with divergent age patterns of under-5 mortality. n is the number of averaged estimates for each area and time period.

Table S2.3: Mean ratio of observed to predicted $q(z, x)$ controlling for $q(4m, 6m)$ across sources of data in sub-Saharan Africa

Country	Data	Period	Resi.	n	Age group					
					0-28d	28d-4m	4m-6m	6m-1y	1y-2y	2y-5y
Burkina Faso	DHS	1993-2010	Total	2	1.03	1.08	1	1.98	3.66	5.12
	HDSS	2001-2015	Rur./Urb.	5	0.50	0.81	1	2.35	4.53	4.31
Côte d'Ivoire	DHS	2002-2012	Total	1	1.47	1.24	1	1.76	2.50	3.26
	HDSS	2012-2016	Rural	1	0.50	0.72	1	1.95	3.39	3.18
Ethiopia	DHS	2001-2019	Total	3	1.52	1.13	1	1.37	1.39	2.42
	HDSS	2006-2017	Rur./Urb.	7	1.32	0.83	1	1.68	2.19	2.54
Gambia	DHS	2010-2020	Total	1	1.53	1.02	1	1.70	1.75	2.21
	HDSS	2009-2015	Rural	3	0.78	0.73	1	1.38	2.46	2.15
Ghana	DHS	1994-2015	Total	3	1.37	0.79	1	1.45	2.19	3.56
	HDSS	1995-2014	Rural	6	0.82	0.95	1	1.69	3.22	3.25
Malawi	DHS	2000-2016	Total	2	1.12	0.94	1	2.03	2.79	3.70
	HDSS	2003-2017	Rural	3	1.24	0.81	1	1.56	2.52	2.90
Mozambique	DHS	2001-2011	Total	1	0.97	1.10	1	1.49	1.60	2.54
	HDSS	2012-2016	Rural	1	0.49	0.59	1	2.00	5.26	3.93
Nigeria	DHS	2003-2019	Total	2	1.36	0.95	1	2.16	4.05	5.89
	HDSS	2013-2017	Rur./Urb.	1	1.71	1.00	1	1.60	3.70	2.13
Senegal	DHS	1976-2019	Total	12	1.35	1.10	1	1.73	2.93	4.46
	HDSS	1972-2016	Rural	18	1.25	0.99	1	1.74	3.65	4.94
Tanzania	DHS	1986-2016	Total	4	0.91	0.82	1	1.55	1.85	2.10
	HDSS	1988-2014	Rural	9	0.69	1.13	1	1.83	2.77	2.83
Uganda	DHS	2001-2016	Total	2	1.05	0.80	1	1.69	2.32	2.60
	HDSS	2007-2016	Rural	2	0.97	0.64	1	1.72	3.09	2.62
All	DHS	1976-2020	Total	33	1.26	1.00	1	1.70	2.57	3.71
			Rural	33	1.27	1.04	1	1.65	2.68	4.03
	HDSS	1972-2017	Rur./Urb.	56	1.00	0.92	1	1.77	3.24	3.62

$q(z, x)$ is used to denote the probability of dying between ages $z-x$. Ratios were computed by dividing observed by predicted $q(z, x)$ values for 6 age groups between ages 0 and 5 y . The DHSs include those with divergent age patterns of under-5 mortality. n is the number of averaged estimates for each area and time period.

Table S2.4: Mean ratio of observed to predicted $q(z, x)$ controlling for $q(4m, 6m)$ across DHS

	Age Pattern	Period	n	Age group					
				0-28d	28d-4m	4m-6m	6m-1y	1y-2y	2y-5y
All DHS	Early	1977-2018	37	1.42	1.18	1	1.60	1.22	1.06
	Without South Asia	1977-2016	30	1.31	1.20	1	1.76	1.22	0.96
	Average	1977-2018	35	1.12	1.05	1	1.50	1.34	1.20
	Late	1977-2018	59	1.04	1.15	1	1.74	1.91	1.64
	Early/Late	1976-2020	22	1.56	1.13	1	1.50	1.86	2.40
	Very Late	1976-2019	120	1.13	0.95	1	1.71	2.50	3.74
<hr/>									
	Region								
DHS with divergent age patterns	South Asia	1981-1999	9	1.79	1.25	1	1.54	2.11	2.86
		2000-2012	3	1.92	1.26	1	1.44	1.95	2.36
		1981-2012	12	1.82	1.25	1	1.51	2.07	2.73
	Sub-Saharan Africa	1976-1999	62	1.12	0.94	1	1.55	2.41	3.68
		2000-2020	64	1.17	0.96	1	1.81	2.44	3.58
		1976-2020	126	1.15	0.95	1	1.68	2.43	3.63
	Western Africa	1976-1999	25	1.28	0.95	1	1.57	2.86	4.88
		2000-2020	32	1.28	0.99	1	1.81	2.71	4.11
		1976-2020	57	1.28	0.97	1	1.71	2.77	4.45
	Rest of sub-Saharan Africa	1977-1999	37	1.01	0.93	1	1.53	2.11	2.86
		2000-2019	32	1.07	0.93	1	1.8	2.18	3.05
		1977-2020	69	1.04	0.93	1	1.66	2.14	2.95

$q(z, x)$ is used to denote the probability of dying between ages $z-x$. Ratios were computed by dividing observed by predicted $q(z, x)$ values for 6 age groups between ages 0 and 5 y . n is the number of averaged estimates for each area and time period.

Table S2.5: Ratio of observed to predicted $q(z, x)$ controlling for $q(4m, 6m)$ for the India 2005-06 DHS at subnational levels

Survey	Region (State)	Age group		
		0-28d	4m-6m	2y-5y
India 2005-06	Central	2.03 [1.75,2.30]	1 [1,1]	2.59 [2.02,3.16]
	Chhatisgarh	2.32 [1.20,3.44]	1 [1,1]	2.99 [0.75,5.23]
	Delhi	1.50 [0.25,2.74]	1 [1,1]	1.20 [-0.27,2.68]
	Madhya Pradesh	1.63 [1.19,2.07]	1 [1,1]	2.02 [1.06,2.98]
	Rajasthan	2.15 [1.26,3.04]	1 [1,1]	2.12 [0.83,3.40]
	Uttar Pradesh	2.14 [1.72,2.55]	1 [1,1]	2.93 [2.02,3.84]
	East	2.22 [1.68,2.75]	1 [1,1]	4.17 [2.77,5.57]
	Bihar	1.86 [1.20,2.52]	1 [1,1]	3.78 [1.86,5.71]
	Jharkhand	2.45 [1.37,3.54]	1 [1,1]	5.30 [1.66,8.93]
	Orissa	2.28 [1.22,3.33]	1 [1,1]	3.89 [1.20,6.58]
	West Bengal	2.83 [0.68,4.98]	1 [1,1]	3.93 [-0.09,7.95]
	North	1.52 [1.18,1.85]	1 [1,1]	1.92 [1.18,2.67]
	Arunachal Pradesh	1.00 [0.66,1.35]	1 [1,1]	2.19 [0.90,3.47]
	Assam	1.89 [0.97,2.81]	1 [1,1]	2.81 [0.33,5.28]
	Himachal Pradesh	1.76 [0.04,3.49]	1 [1,1]	0.82 [-0.45,2.08]
	Haryana	1.01 [0.47,1.55]	1 [1,1]	1.44 [0.03,2.86]
	Jammu and Kashmir	1.57 [0.46,2.68]	1 [1,1]	0.81 [-0.12,1.75]
	Meghalaya	1.19 [0.15,2.23]	1 [1,1]	4.93 [-0.11,9.96]
	Manipur	1.36 [0.60,2.12]	1 [1,1]	2.44 [0.89,4.00]
	Mizoram	1.01 [0.21,1.82]	1 [1,1]	1.78 [-0.19,3.75]
	Nagaland	0.99 [0.57,1.42]	1 [1,1]	2.70 [1.16,4.24]
	Punjab	1.55 [0.28,2.82]	1 [1,1]	1.16 [-0.07,2.40]
	Sikkim	1.17 [-0.10,2.45]	1 [1,1]	1.02 [-0.69,2.72]
	Tripura	1.36 [0.64,2.08]	1 [1,1]	1.19 [0.09,2.28]
	Uttaranchal	2.48 [-0.85,5.82]	1 [1,1]	2.42 [-1.24,6.08]
	Andhra Pradesh	2.64 [0.96,4.32]	1 [1,1]	1.79 [-0.05,3.62]
	South	2.09 [1.45,2.73]	1 [1,1]	1.83 [0.96,2.69]
	Karnataka	1.72 [1.06,2.39]	1 [1,1]	1.97 [0.79,3.16]
	Kerala	2.70 [1.49,3.90]	1 [1,1]	0.98 [-0.37,2.32]
	Tamil Nadu	3.18 [-2.18,8.55]	1 [1,1]	3.44 [-2.45,9.34]
	West	2.17 [1.26,3.07]	1 [1,1]	1.81 [0.72,2.90]
	Gujarat	2.70 [1.00,4.41]	1 [1,1]	2.58 [0.26,4.91]
	Goa	1.59 [-2.34,5.52]	1 [1,1]	1.65 [-2.65,5.95]
	Maharashtra	1.91 [0.80,3.02]	1 [1,1]	1.46 [0.23,2.69]

$q(z, x)$ is used to denote the probability of dying between ages $z-x$. Ratios significantly higher than 1 are highlighted.

Table S2.6: Ratio of observed to predicted $q(z, x)$ controlling for $q(4m, 6m)$ across DHS with divergent age pattern of under-5 mortality

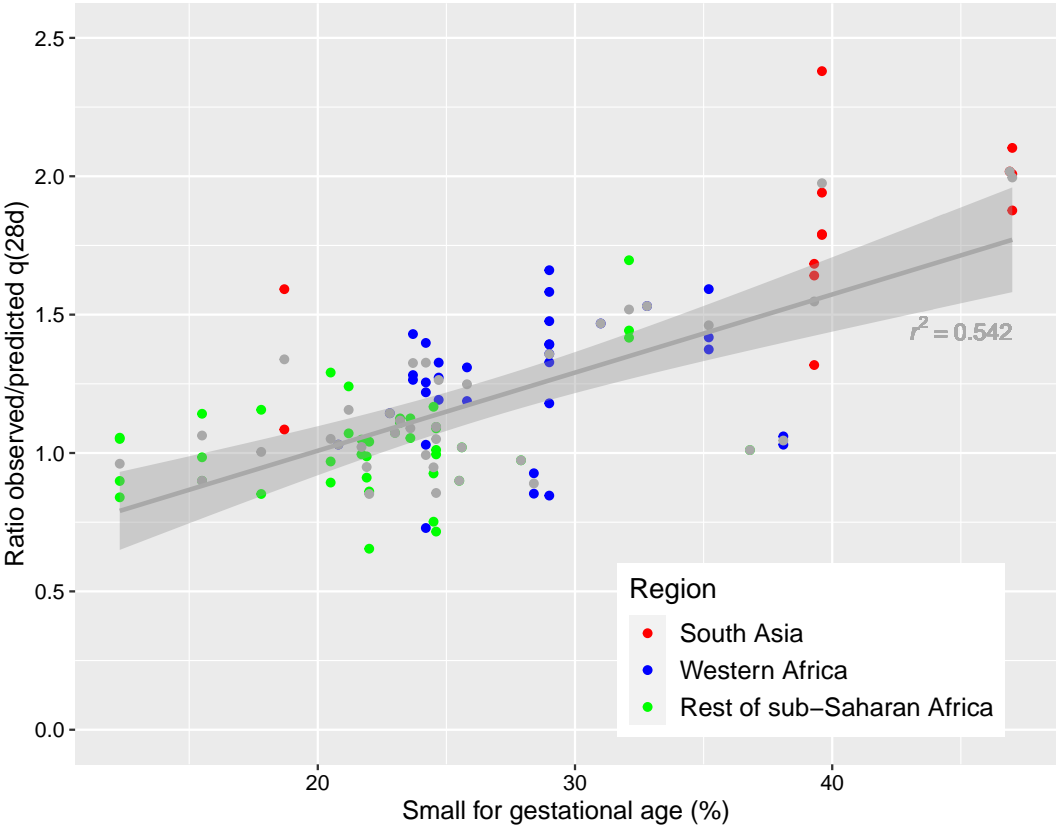
Country	Year	Age group		
		0-28d	4m-6m	2y-5y
South Asia				
Bangladesh	1993-94	1.86 [1.62,2.11]	1 [1,1]	3.39 [2.61,4.16]
	1996-97	1.75 [1.43,2.08]	1 [1,1]	3.07 [2.11,4.02]
	1999-00	1.70 [1.41,1.98]	1 [1,1]	2.52 [1.83,3.20]
	2004	1.54 [1.26,1.82]	1 [1,1]	1.83 [1.23,2.43]
	2007	1.79 [1.31,2.27]	1 [1,1]	2.15 [1.31,2.98]
	2011	1.94 [1.45,2.44]	1 [1,1]	2.28 [1.44,3.11]
India	1992-93	1.87 [1.74,2.00]	1 [1,1]	2.95 [2.60,3.29]
	1998-99	1.94 [1.80,2.08]	1 [1,1]	3.09 [2.69,3.49]
	2005-06	2.02 [1.82,2.21]	1 [1,1]	2.65 [2.25,3.05]
Nepal	1996	1.97 [1.65,2.28]	1 [1,1]	3.80 [2.81,4.78]
	2001	1.76 [1.43,2.09]	1 [1,1]	3.11 [2.10,4.12]
	2006	1.68 [1.32,2.05]	1 [1,1]	1.98 [1.22,2.75]
Western Africa				
Benin	1996	1.11 [0.91,1.30]	1 [1,1]	4.02 [2.93,5.12]
	2001	0.99 [0.84,1.14]	1 [1,1]	3.15 [2.33,3.98]
	2006	1.19 [1.06,1.32]	1 [1,1]	4.58 [3.83,5.33]
	2017-18	1.31 [1.11,1.51]	1 [1,1]	4.85 [3.76,5.94]
Burkina Faso	1993	1.15 [0.99,1.32]	1 [1,1]	3.71 [2.83,4.58]
	1998-99	0.94 [0.80,1.07]	1 [1,1]	4.33 [3.24,5.41]
	2003	1.03 [0.88,1.18]	1 [1,1]	5.58 [4.45,6.71]
	2010	1.03 [0.91,1.15]	1 [1,1]	4.66 [3.79,5.54]
Côte d'Ivoire	1994	1.44 [1.21,1.67]	1 [1,1]	4.03 [3.00,5.06]
	1998-99	1.18 [0.80,1.56]	1 [1,1]	2.06 [0.47,3.66]
	2011-12	1.47 [1.19,1.75]	1 [1,1]	3.26 [2.17,4.36]
Gambia	2019-20	1.53 [1.05,2.01]	1 [1,1]	2.21 [1.16,3.26]
Ghana	1988	1.69 [1.32,2.06]	1 [1,1]	6.99 [4.46,9.53]
	1993	1.47 [1.19,1.76]	1 [1,1]	4.69 [3.01,6.37]
	1998	1.18 [0.85,1.52]	1 [1,1]	4.65 [2.79,6.51]
	2003	1.44 [1.05,1.84]	1 [1,1]	4.13 [2.57,5.70]
	2008	1.26 [0.86,1.65]	1 [1,1]	3.35 [1.68,5.02]
	2014	1.40 [0.94,1.86]	1 [1,1]	3.21 [1.54,4.87]
Guinea	1999	1.14 [0.99,1.29]	1 [1,1]	3.36 [2.66,4.06]
	2005	1.19 [1.04,1.35]	1 [1,1]	2.85 [2.23,3.47]
	2012	1.33 [1.10,1.55]	1 [1,1]	4.33 [3.25,5.40]
	2018	1.27 [1.00,1.54]	1 [1,1]	4.61 [3.22,6.00]
Liberia	1986	1.30 [1.14,1.45]	1 [1,1]	1.60 [1.19,2.02]
	2013	1.03 [0.82,1.24]	1 [1,1]	2.59 [1.81,3.36]
	2019-20	1.22 [0.83,1.60]	1 [1,1]	2.15 [1.32,2.98]
Mali	1987	1.58 [1.31,1.85]	1 [1,1]	6.52 [4.43,8.62]
	1995-96	1.54 [1.39,1.69]	1 [1,1]	5.10 [4.20,6.00]
	2001	1.59 [1.43,1.76]	1 [1,1]	5.05 [4.15,5.96]
	2006	1.42 [1.26,1.57]	1 [1,1]	5.01 [3.90,6.12]
	2012-13	1.59 [1.23,1.95]	1 [1,1]	6.35 [4.40,8.30]
	2018	1.37 [1.14,1.61]	1 [1,1]	6.40 [4.88,7.91]

Country	Year	Age group		
		0-28d	4m-6m	2y-5y
Niger	1992	0.99 [0.88,1.11]	1 [1,1]	5.48 [4.36,6.59]
	1998	1.02 [0.89,1.14]	1 [1,1]	5.29 [4.41,6.16]
	2006	1.03 [0.87,1.19]	1 [1,1]	5.44 [4.15,6.73]
	2012	1.06 [0.90,1.22]	1 [1,1]	7.26 [5.67,8.85]
Nigeria	1990	1.38 [1.14,1.62]	1 [1,1]	7.05 [5.11,8.99]
	2003	1.20 [1.01,1.39]	1 [1,1]	4.27 [3.06,5.48]
	2008	1.26 [1.16,1.36]	1 [1,1]	4.80 [4.15,5.44]
	2013	1.28 [1.17,1.39]	1 [1,1]	4.83 [4.09,5.56]
	2018	1.43 [1.27,1.59]	1 [1,1]	6.95 [5.80,8.10]
Senegal	1986	1.42 [1.14,1.70]	1 [1,1]	8.85 [6.10,11.60]
	1992-93	1.19 [0.99,1.39]	1 [1,1]	5.74 [4.30,7.18]
	1997	1.34 [1.10,1.59]	1 [1,1]	6.35 [4.71,7.98]
	2005	1.36 [1.12,1.60]	1 [1,1]	5.42 [4.02,6.81]
	2010-11	1.48 [1.18,1.77]	1 [1,1]	4.93 [3.52,6.34]
	2012-13	1.39 [0.93,1.86]	1 [1,1]	3.95 [2.08,5.81]
	2014	0.85 [0.57,1.13]	1 [1,1]	2.68 [1.55,3.81]
	2015	1.33 [1.00,1.65]	1 [1,1]	3.80 [2.35,5.25]
	2016	1.18 [0.80,1.56]	1 [1,1]	2.90 [1.71,4.09]
	2017	1.58 [1.19,1.97]	1 [1,1]	3.04 [1.96,4.13]
Sierra Leone	2018	1.39 [0.75,2.04]	1 [1,1]	2.89 [0.89,4.90]
	2019	1.66 [0.93,2.39]	1 [1,1]	2.99 [1.36,4.62]
	2013	0.93 [0.83,1.02]	1 [1,1]	2.04 [1.66,2.42]
Togo	2019	0.85 [0.72,0.99]	1 [1,1]	2.92 [2.18,3.66]
	1988	1.27 [0.97,1.57]	1 [1,1]	4.81 [2.92,6.70]
Togo	1998	1.32 [1.07,1.58]	1 [1,1]	5.21 [3.82,6.59]
	2013-14	1.14 [0.91,1.38]	1 [1,1]	4.32 [3.07,5.57]
Rest of sub-Saharan Africa				
Angola	2015-16	1.02 [0.80,1.24]	1 [1,1]	3.72 [2.60,4.84]
Burundi	1987	0.89 [0.69,1.09]	1 [1,1]	4.22 [2.61,5.84]
	2010	1.05 [0.87,1.23]	1 [1,1]	3.25 [2.37,4.14]
	2016-17	1.13 [0.90,1.35]	1 [1,1]	4.12 [3.03,5.21]
Cameroon	1991	1.65 [1.16,2.14]	1 [1,1]	6.24 [3.48,9.01]
	1998	1.32 [0.98,1.66]	1 [1,1]	5.07 [3.03,7.10]
	2004	0.99 [0.85,1.14]	1 [1,1]	3.77 [2.89,4.64]
	2011	1.07 [0.90,1.24]	1 [1,1]	4.29 [3.22,5.35]
	2018	1.24 [0.98,1.50]	1 [1,1]	4.63 [3.33,5.93]
Central African Republic	1994-95	0.95 [0.81,1.10]	1 [1,1]	1.88 [1.42,2.34]
Chad	1996-97	1.09 [0.94,1.24]	1 [1,1]	3.80 [2.96,4.64]
	2004	1.02 [0.85,1.20]	1 [1,1]	3.24 [2.20,4.28]
Congo	2014-15	1.01 [0.90,1.13]	1 [1,1]	3.85 [3.04,4.66]
	2005	1.01 [0.77,1.26]	1 [1,1]	2.71 [1.75,3.67]
Congo Democratic Republic	2011-12	1.09 [0.77,1.41]	1 [1,1]	4.37 [2.74,6.00]
	2007	0.99 [0.78,1.20]	1 [1,1]	2.98 [2.03,3.94]
Ethiopia	2013-14	0.91 [0.77,1.05]	1 [1,1]	3.39 [2.58,4.20]
	2000	1.46 [1.25,1.67]	1 [1,1]	3.94 [3.07,4.80]
	2005	1.23 [1.03,1.43]	1 [1,1]	3.70 [2.72,4.68]
	2011	1.42 [1.14,1.69]	1 [1,1]	3.08 [2.06,4.09]
Ethiopia	2016	1.44 [1.09,1.80]	1 [1,1]	2.00 [1.16,2.84]
	2019	1.70 [0.87,2.52]	1 [1,1]	2.17 [0.83,3.51]

Country	Year	Age group		
		0-28d	4m-6m	2y-5y
Gabon	2000	1.18 [0.83,1.53]	1 [1,1]	2.53 [1.50,3.57]
	2012	1.09 [0.60,1.59]	1 [1,1]	2.25 [0.86,3.63]
Kenya	1989	0.72 [0.58,0.86]	1 [1,1]	1.76 [1.17,2.35]
	1993	0.78 [0.63,0.93]	1 [1,1]	1.70 [1.18,2.23]
	1998	0.59 [0.48,0.70]	1 [1,1]	1.19 [0.80,1.57]
	2008-09	1.14 [0.87,1.41]	1 [1,1]	1.56 [0.97,2.15]
Madagascar	1992	1.02 [0.84,1.21]	1 [1,1]	3.05 [2.14,3.95]
	1997	0.99 [0.81,1.17]	1 [1,1]	2.78 [2.08,3.49]
	2003-04	1.23 [0.87,1.59]	1 [1,1]	3.64 [2.00,5.27]
	2008-09	0.90 [0.74,1.06]	1 [1,1]	2.74 [1.98,3.49]
Malawi	1992	0.87 [0.73,1.02]	1 [1,1]	2.18 [1.54,2.81]
	2000	0.90 [0.82,0.99]	1 [1,1]	2.47 [2.08,2.85]
	2004	0.84 [0.70,0.98]	1 [1,1]	2.44 [1.94,2.93]
	2010	1.11 [0.97,1.25]	1 [1,1]	4.07 [3.32,4.83]
	2015-16	1.12 [0.92,1.33]	1 [1,1]	3.32 [2.49,4.15]
Mozambique	1997	1.07 [0.89,1.24]	1 [1,1]	2.01 [1.39,2.63]
	2011	0.97 [0.83,1.12]	1 [1,1]	2.54 [1.93,3.15]
Namibia	1992	1.15 [0.87,1.43]	1 [1,1]	2.04 [1.17,2.91]
	2000	0.84 [0.54,1.15]	1 [1,1]	1.87 [0.97,2.78]
Rwanda	1992	1.34 [1.13,1.54]	1 [1,1]	4.94 [3.67,6.20]
	2000	1.11 [0.96,1.27]	1 [1,1]	3.40 [2.68,4.13]
	2005	1.06 [0.92,1.19]	1 [1,1]	3.54 [2.76,4.33]
	2007-08	0.84 [0.65,1.03]	1 [1,1]	3.15 [2.19,4.12]
	2010	1.05 [0.88,1.22]	1 [1,1]	3.53 [2.66,4.4]
	2014-15	0.90 [0.69,1.11]	1 [1,1]	2.67 [1.81,3.53]
Sao Tome and Principe	2008-09	0.90 [0.40,1.40]	1 [1,1]	2.95 [0.61,5.30]
Sudan	1989-90	1.23 [1.04,1.43]	1 [1,1]	3.36 [2.43,4.30]
Tanzania	1991-92	0.95 [0.78,1.11]	1 [1,1]	2.39 [1.77,3.02]
	1996	0.82 [0.70,0.95]	1 [1,1]	1.82 [1.41,2.23]
	2004-05	0.82 [0.69,0.95]	1 [1,1]	2.14 [1.62,2.66]
	2010	0.85 [0.69,1.01]	1 [1,1]	1.92 [1.34,2.49]
	2015-16	1.16 [0.90,1.41]	1 [1,1]	2.50 [1.70,3.30]
Uganda	1988-89	1.10 [0.92,1.27]	1 [1,1]	3.07 [2.20,3.94]
	1995	0.81 [0.68,0.93]	1 [1,1]	3.14 [2.33,3.96]
	2000-01	0.85 [0.71,0.98]	1 [1,1]	2.68 [1.97,3.40]
	2006	0.75 [0.65,0.86]	1 [1,1]	2.20 [1.74,2.67]
	2011	0.93 [0.74,1.11]	1 [1,1]	2.48 [1.73,3.24]
	2016	1.17 [0.97,1.36]	1 [1,1]	2.71 [2.01,3.41]
Zambia	1992	0.84 [0.73,0.95]	1 [1,1]	2.03 [1.54,2.51]
	1996	0.73 [0.63,0.83]	1 [1,1]	2.16 [1.67,2.64]
	2001-02	0.73 [0.62,0.83]	1 [1,1]	2.24 [1.73,2.74]
	2007	0.89 [0.75,1.03]	1 [1,1]	2.43 [1.78,3.09]
	2013-14	0.97 [0.80,1.14]	1 [1,1]	3.07 [2.28,3.87]
	2018	1.29 [0.94,1.64]	1 [1,1]	3.40 [2.21,4.59]
Zimbabwe	1988	1.36 [0.89,1.82]	1 [1,1]	2.86 [1.36,4.36]
	1994	1.12 [0.76,1.49]	1 [1,1]	2.27 [1.20,3.33]
	1999	0.85 [0.60,1.11]	1 [1,1]	1.88 [1.14,2.62]

$q(z,x)$ is used to denote the probability of dying between ages $z-x$. Ratios significantly higher than 1 are highlighted.

Figure S2.4: Relationship between the prevalence of small for gestational age in 2010 and the ratio of observed to predicted neonatal mortality ($q(28d)$) controlling for $q(4m,6m)$ across DHS from South Asia and sub-Saharan Africa collected between 2000-2020



The gray points are average estimates for a same country. The regression line is based on these points. Source: Lee et al., 2013 for the estimates of small for gestational age.

Webappendix 3: Data and methods for estimating mortality

We downloaded the Demographic and Health Surveys (DHS) from <https://dhsprogram.com> on April 5, 2022. We used all the publicly-available standard surveys with few exceptions. We discarded 12 surveys¹ whose data quality was shown to be insufficient through comparison with other sources of information. These surveys are those that were not retained by the UN Inter-agency Group for Child Mortality Estimation (UN IGME) (<https://childmortality.org/>) for monitoring the U5MR. These surveys deviated strongly from the other sources of data potentially because of omission of deaths or sampling errors.

We obtained the Health and Demographic Surveillance System (HDSS) data from two sources: (1) the INDEPTH Data Repository (50 datasets downloaded on August 25, 2021 from <https://www.indepth-ishare.org/>), and (2) from bilateral agreements (7 datasets marked by an asterisk in Table S3, appendix p 19). We selected datasets that overlapped at least one year with the period covered by DHS estimates in a given country (appendix p 8-9).

We selected only two cohort studies. 20 cohort studies were initially selected based on a systematic search for another study [1]. Among these 20, we selected 7 studies from South Asia and sub-Saharan Africa chosen for their sample sizes (> 1000 births) as well as for the duration and thoroughness of their mortality follow-up (minimum 6 months for all births). However, since these cohort studies were designed for another purpose than estimating mortality, they incorporated exclusion criteria (such as the survival of the child after 3 days) that bias the mortality estimates at early ages. Therefore, we only selected two studies that we considered adequate for the purpose of mortality estimation. Issues associated with cohort studies and their promising use for the purpose of estimating mortality is analyzed in a forthcoming paper (unpublished data; Erchick DJ et al.).

In these three sources of information (DHS, HDSS, and CS), the data were collected at the individual level and included the dates of birth and death (or age at death). This allowed us to use the same direct procedure to estimate mortality before 5 y. In the three cases, we estimated age-specific death rates using a standard event/exposure procedure [2]: The age-specific death rates were obtained dividing the number of deaths by the number of person-years within each age interval for a given cohort or period for the 22 age groups used the study. We then computed probabilities of dying under the assumption of a constant force of mortality within each age group. For DHS, we computed 10-y period estimates in order to address potential transference of births. Such type of error is generally attributed to interviewers. Typically, to avoid a supplementary module of questions about recent births, interviewers might displace (or omit) births beyond the cut-off date defined for that purpose (5 y prior to survey) [3]. For HDSS, we computed 5-y

¹ Afghanistan 2015, Albania 2017-18, Armenia 2015-16, Benin 2011-12, Dominican Republic 1999, El Salvador 1985, Eswatini 2006-07, Gambia 2013, India 2015-16, Mexico 1987, Sierra Leone 2008, Timor-Leste 2016.

estimates as provided by INDEPTH. CS estimates are computed for birth cohorts born over a 2-y period.

For DHS, we also used a jackknife variance estimation method to produce uncertainty estimates [4]. The standard error of the ratio of the observed/predicted probabilities of dying for each age group was estimated across subsamples generated by deleting one cluster (primary sampling unit) from the main sample. We computed 95% uncertainty intervals as ± 1.96 * standard error.

The Sampling Registration Systems (SRS) are routine demographic surveys collecting vital events at the national scale. In 2018, we obtained data of the SRS of India from The Million Death Study (MDS) based at the University of Toronto (<https://www.cghr.org/projects/million-death-study-project/>). The MDS provided a list of deaths with the date of occurrence for the period 2004-2013 by state and place of residence. However, we did not have access to direct measurements of the exposure to the risk of dying (births or populations). We circumvented this obstacle by using the death rates by state and sex published in the annual SRS reports (downloaded on November 1, 2019 from <https://censusindia.gov.in>). In these reports, the death rates were available for 0-1 and 1-4 age groups and for 20 states out of 28. By dividing the death counts by the corresponding death rates, we estimated the mid-populations aged 0-1 and 1-4. In order to obtain these populations by single year of age, we interpolated the 0-4 age group using a cubic interpolation method: We estimated the mid-population for the 5-9 age group and we interpolated the cumulative populations using 0, 1, 5 and 10 y of age as knots. We then estimated the exposure for the 22 age groups, assuming that the mid-population was proportional to the length of each age interval. Finally, to produce more stable mortality estimates, we summed the deaths counts and mid-populations into 5 regions: Central, East, North, South, and West.²

Based on this information, we calculated age-specific mortality rates by dividing the number of deaths in each age interval by the corresponding mid-population. We computed these rates for two 5-y periods (2004-08 and 2009-2013). Finally, we calculated cumulative probabilities of dying under the assumption that the force of mortality was constant within each age interval as for the previous sources of data.

Lastly, we obtained aggregated annual counts of births and deaths from the China National Maternal and Child Health Surveillance System (1996-2015) [5]. As an exception among the data used in our study, the aggregated counts of deaths had more limited age breakdowns: 7, 28 days; 6, 12 months; 2, 3, 5 y. In order to estimate the exposures for each age group, we reconstructed birth cohorts from the available data. Since we estimated under-5 mortality, data from the previous 5 years were necessary to approximate the exposures and build period life tables for a given year. However, because the surveillance was expanded between 2008 and 2009, the data from 2008 or

² Central Region: Chhattisgarh, Madhya Pradesh, Rajasthan, Uttar Pradesh, Delhi. East: Bihar, Orissa, Jharkhand, West Bengal. North: Himachal Pradesh, Jammu & Kashmir, Haryana, Punjab. Northeast: Assam. South: Andhra Pradesh, Karnataka, Kerala, Tamil Nadu. West: Gujarat, Maharashtra. Assam being the only state representing Northeast India, we included it into the East region

earlier could not be combined with data from 2009 onwards. Hence, we produced mortality estimates only for the periods 2001-2008 and 2014-2015. After computing deaths and persons-years for each age groups, we used the same method to derive age-specific mortality rates and cumulative probabilities of dying than for the previous data sources. For each annual cohorts, we produced estimates for six strata: three regions (East, Midland, and West) divided into rural and urban areas.

1. Katz J, Lee ACC, Kozuki N, et al. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *The Lancet*. 2013;382(9890):417-25.
2. Hill K. Direct estimation of child mortality from birth histories. In: Moultrie T, Dorrington R, Hill A, Hill K, Timæus I, Zaba B, editors. *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population; 2013.
3. Silva R. Child Mortality Estimation: Consistency of Under-Five Mortality Rate Estimates Using Full Birth Histories and Summary Birth Histories. *PLOS Medicine*. 2012;9(8):e1001296.
4. Pedersen J, Liu J. Child Mortality Estimation: Appropriate Time Periods for Child Mortality Estimates from Full Birth Histories. *PLOS Medicine*. 2012;9(8):e1001289.
5. He C, Liu L, Chu Y, et al. National and subnational all-cause and cause-specific child mortality in China, 1996-2015: a systematic analysis with implications for the Sustainable Development Goals. *Lancet Glob Health*. 2017;5(2):e186-e97.

Table S3: Used datasets

Country	DHS Datasets	Country	DHS Datasets
Albania	2008-09	Jordan	1990, 1997, 2002, 2007, 2009, 2012, 2017-18
Angola	2015-16	Kazakhstan	1995, 1999
Armenia	2000, 2005, 2010	Kenya	1989, 1993, 1998, 2003, 2008-09, 2014
Azerbaijan	2006	Kyrgyz Republic	1997, 2012
Bangladesh	1993-94, 1996-97, 1999-00, 2004, 2007, 2011, 2014, 2017-18	Lesotho	2004, 2009, 2014
Benin	1996, 2001, 2006, 2017-18	Liberia	1986, 2007, 2013, 2019-20
Bolivia	1989, 1994, 1998, 2003, 2008	Madagascar	1992, 1997, 2003-04, 2008-09
Brazil	1986, 1996	Malawi	1992, 2000, 2004, 2010, 2015-16
Burkina Faso	1993, 1998-99, 2003, 2010	Maldives	2009, 2016-17
Burundi	2010, 2016-17	Mali	1987, 1995-96, 2001, 2006, 2012-13, 2018
Cambodia	2000, 2005, 2010, 2014	Mexico	1987
Cameroon	1991, 1998, 2004, 2011, 2018	Moldava	2005
Central African Rep.	1994-95	Morocco	1987, 1992, 2003-04
Chad	1996-97, 2004, 2014-15	Mozambique	1997, 2003, 2011, 2015-16
Colombia	1986, 1990, 1995, 2000, 2005, 2010, 2015	Myanmar	1992, 2000, 2006-07, 2011,
Comoros	1996	Namibia	1996, 2001, 2006, 2011, 2016
Congo (DRC)	2007, 2013-14,	Nicaragua	1998, 2001
Congo	2005, 2011-12	Niger	1992, 1998, 2006, 2012
Côte d'Ivoire	1994, 1998-99, 2011-12	Nigeria	1990, 2003, 2008, 2013, 2018
Dominican Republic	1986, 1991, 1996, 2002, 2007, 2013	Pakistan	1990-01, 2006-07, 2012-13, 2017-18
Ecuador	1987	Papua New Guinea	2016-17
Egypt	1988, 1992, 1995, 2000, 2003, 2005, 2008, 2014	Paraguay	1990
Ethiopia	2000, 2005, 2011, 2016, 2019	Peru	1986, 1991-92, 1996, 2000, 2004-06, 2007-08, 2009, 2010, 2011, 2012
Gabon	2000, 2012	Philippines	1993, 1998, 2003, 2008, 2013, 2017
Gambia	2019-20	Rwanda	1992, 2000, 2005, 2007-08, 2010, 2014-15
Ghana	1988, 1993, 1998, 2003, 2008, 2014	Sao Tome and Principe	2008-09
Guatemala	1987, 1995, 1998-99, 2014-15	Senegal	1986, 1992-93, 1997, 2005, 2010-11, 2012-13, 2014, 2015, 2016, 2017, 2018, 2019
Guinea	1999, 2005, 2012, 2018	Sierra Leone	2013, 2019
Guyana	2009	South Africa	1998, 2016
Haiti	1994-95, 2000, 2005-06, 2012, 2016-17	Sri Lanka	1987
Honduras	2005-06, 2011-12	Sudan	1989-90
India	1992-93, 1998-99, 2005-06	Tajikistan	2012, 2017
Indonesia	1987, 1991, 1994, 1997, 2002-03, 2007, 2012, 2017		

Country	DHS Datasets
Tanzania	1991-92, 1996, 1999, 2004-05, 2010, 2015-16
Thailand	1987
Timor-Leste	2009-10
Togo	1988, 1998, 2013-14
Trinidad and Tobago	1987
Tunisia	1988
Turkey	1993, 1998, 2003, 2008, 2013
Uganda	1988-99, 1995, 2000-01, 2006, 2011, 2016
Ukraine	2007
Uzbekistan	1996
Vietnam	1997, 2002
Yemen	1991-92, 2013
Zambia	1992, 1996, 2001-02, 2007, 2013-14, 2018
Zimbabwe	1988, 1994, 1999, 2005-06, 2010-11, 2015

Country (Site)	HDSS Datasets
Bangladesh	
Chakaria	2007-11
Matlab	1966-2009*
Burkina Faso	
Nanoro	2011-15
Nouna	2001-05, 2006-10, 2011-15
Ouagadougou	2011-15
Côte d'Ivoire	
Taabo	2012-16
Ethiopia	
Gilgel-Gibe	2006-10, 2011-15
Kilite-Awlaelo	2010-14
Kersa	2012-16
Harar	2012-16
Dabat	2013-17
Arba-Minch	2011-15
Gambia	
Farafenni	2009-2013*, 2011-15,
Basse	2011-2015*
Ghana	
Navrongo	1995-99, 2000-04, 2005-09, 2010-14
Kintampo	2010-14
Dodowa	2007-11
Malawi	
Karonga	2003-07, 2008-12, 2013-17

Country (Site)	HDSS Datasets
Mozambique	
Chokwe	2012-16
Nigeria	
Cross-River	2013-17
Senegal	
Bandafassi	1972-76, 1977-81, 1982-86, 1987-91, 1992-96, 1997-01, 2002-06, 2007-11, 2012-16
Mlomp	1987-91, 1992-96, 1997-01, 2002-06, 2007-11,
Niakhar	1990-96*, 1997-2002*, 2003-08*, 2009-2015*
Tanzania	
Ifakara	2000-04, 2005-09, 2010-14
Rufiji	2000-04, 2005-09, 2010-14
Magu	1998-02 2003-07, 2008-12
Uganda	
Iganga-Mayuge	2007-11, 2012-16

Country (Site)	CS Datasets
India	
Tamil Nadu	1998-2000
Nepal	
Sarlahi	1999-2000

Country	SRS datasets
China	
MCHSS	1996-2015*
India	
MDS	2004-2013*

* Data obtained by bilateral agreement

Webappendix 4: Classification of the age pattern of under-5 mortality in each DHS

The plots below display 22 ratios of observed to predicted cumulative probabilities of dying ($q(x)$) between ages 0 and 5 y for each DHS. These ratios represent the deviation of the observed mortality from the average prediction of the vital registration (VR) model based on the value of $q(5y)$. The central blue line represents the ratios at each age. The red area represents the range of predictions of the VR model ($k = [-1.1, 1.5]$) (appendix p 2).

The surveys are classified relative to the boundaries of the red area as described in the main text. We allowed the deviation to go across these boundaries when the deviation was fewer than 2.5% or not statistically significant. We also allowed one of the 22 ratios to be out of these boundaries.

We discarded 4 surveys from this classification (Albania 2008-9, Armenia 2010, Brazil 1996, and Kazakhstan 1999) for they display an unusual late and early age pattern. No other survey collected in the same country or region showed a similar age pattern. Consequently, we attributed this unusual pattern to data quality issues such as strong omissions of deaths and/or random variation due to few deaths/low mortality.

