Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2010;**39**:1103–1114
doi:10.1093/ije/dyq074

METHODOLOGY

Estimating the distribution of causes of death among children age 1–59 months in high-mortality countries with incomplete death certification

Hope L Johnson,* Li Liu, Christa Fischer-Walker and Robert E Black

Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

*Corresponding author: 615 N. Wolfe Street, Room E8537, Baltimore, MD 21205, USA. E-mail: hjohnson@jhsph.edu

Accepted 31 March 2010

Background Our objective was to develop a methodology to estimate causes of death among children age 1–59 months in high child mortality countries without adequate vital registration (VR) systems.

Methods We systematically reviewed community-based studies reporting at least two causes of death among children 1–59 months of age identified from published and unpublished sources. We included (i) studies conducted after 1979, (ii) for duration of 12 months or an exact multiple, (iii) with $\geqslant$25 deaths in children <5 years, (iv) each death represented once and (v) <25% of deaths due to unknown causes. A study-based multinomial logistic regression model was applied to country-level data to estimate causes of child death.

Results Of the 216 studies reviewed, 81 were included in the analysis comprising 79 067 under-5 deaths from 25 countries. After adjusting for risk factors and intervention coverage, the estimated distribution of causes of deaths in children 1–59 months of age in sub-Saharan Africa and Southeast Asia was: pneumonia (21 and 31%), diarrhoea (25 and 31%), malaria (26 and 2%), injury (3 and 4%), meningitis (3 and 4%), measles (3 and 2%) and other causes (20 and 27%), respectively.

Conclusion From studies reporting as few as two different causes of death, statistical modelling can be used to estimate the causes of child mortality for settings with incomplete VR. Pneumonia and diarrhoea remain the leading causes of death among children 1–59 months of age in sub-Saharan Africa and Southeast Asia.

Keywords Cause of death, child mortality, infant mortality, developing countries, Africa, Asia, statistical models

Introduction

International and national policymakers require credible disease burden estimates to establish public health priorities, develop disease prevention and control strategies, and evaluate progress towards achieving health objectives. Disease burden estimates are also important for identifying potential gaps in available information to prioritize areas of research.[1]

Mortality estimates ideally come from national vital registration (VR) systems that capture all deaths occurring in the country and use standardized coding systems to systematically categorize medically certified deaths by cause. Although 94% of the world's post-neonatal child deaths occur in sub-Saharan Africa and Southeast Asia, only 9 of 78 countries (12%) have reasonably complete cause of death certification (defined as >80%) (World Health Organization, unpublished results). Quantifying mortality in countries lacking sufficient VR systems is usually accomplished with national surveys asking women of reproductive age about the births and deaths of their children.

Surveys, community-based epidemiological studies and demographic surveillance sites (DSS) utilizing verbal autopsy methods are alternative sources of cause of death data in the absence of sufficient VR data.[2] Verbal autopsy methods attempt to assign cause of death based on signs, symptoms and events leading up to the death as reported by caretakers of the deceased and sometimes supported by medical documentation. Standardized tools and methods for verbal autopsy exist,[3] but inconsistency in implementation of the methods occur, thus interpretation of the findings from verbal autopsy studies must consider the quality of the methods utilized. Nevertheless, verbal autopsy studies have been validated and widely used.[4–6]

Previous methodologies deriving causes of child mortality in countries with incomplete VR data for the year 2000 modelled verbal autopsy data and could only include in their analysis only studies with a complete set of major causes of death, thus limiting the number of studies eligible for inclusion.[7,8] Furthermore, a new method that reflects the recent scale-up of child survival interventions is needed.

The objective of the present analysis was to develop a methodology to estimate causes of death among children 1–59 months of age adapting statistical methodology recently used by the Child Health Epidemiology Reference Group (CHERG) of the World Health Organization (WHO) and UNICEF to estimate causes of neonatal deaths[9] and adjusting for current coverage of child survival interventions.

Materials and methods

Sources of data

Post-neonatal infant and child cause of death data were identified from published DSS reports, existing reviews of the published literature, and by contacting researchers known to have recent mortality surveillance data. We worked collaboratively with members of the CHERG who conducted systematic literature reviews of the published literature for each of the main causes of child death (pneumonia, diarrhoea, malaria and measles). Literature searches were conducted in PubMed, CAB abstracts, SIGLE and the WHO Regional databases using disease-specific search terms as well as 'mortality', and 'cause of death'. Searches were conducted in all languages. The full text of each identified study was retrieved and two reviewers independently screened articles for eligibility for inclusion in the analysis.

Inclusion and exclusion criteria

We included community-based studies reporting at least two causes of death among children 1–59 months of age (i) conducted after 1979, (ii) for a duration of 12 months or an exact multiple due to the seasonal nature of many infectious causes of death, (iii) with $\geqslant$25 deaths in children <5 years of age, (iv) each death represented once and (v) <25% of deaths due to unknown causes. Studies conducted in sub-groups of the study population (e.g. intervention groups in clinical trials) and verbal autopsy studies that did not use a standard questionnaire or for which the methods could not be determined were excluded. We only included cause of death data from studies where the youngest child under mortality surveillance was age 0 or 1 months and where the oldest child was age 11–12, 35–36 or 59–60 months. Cause of death data from neonates was excluded where possible.

Data collection and management

For each article, two reviewers abstracted data characterizing the study design and setting, and the number or proportion of deaths due to each cause of death by age, year and location. Additional study data points were formed within articles presenting cause of death data by year and/or location. Information characterizing the study setting included: study location (city, urban/rural/mixed and WHO region), study years and duration in months, age of children under mortality surveillance, gross national income (GNI) per capita at purchasing power parity (PPP, international $), and under-5 mortality risk ($_5q_0$) per 1000 live births. A reported all-cause mortality rate ($_1m_0$ or $_5m_0$) was transformed into a probability of dying ($_1q_0$ or $_5q_0$) using standard life table equations[10] and $_1q_0$was transformed into $_5q_0$using the Coale–Demeny west model life tables[11] where necessary.

We also abstracted data on possible explanatory variables for the distribution of causes of death, including intervention coverage, from multiple sources in hierarchical order: (i) present article, (ii) other published reports from the study setting, (iii) sub-national surveys conducted in areas covering or representative of the study setting (e.g. DSS), (iv) national data from surveys (e.g. Demographic Health Surveys, Multiple Indicator Cluster Surveys) or public databases/reports (e.g. WHO, UNICEF,

UNAIDS, World Bank). Possible explanatory variables were explored as either continuous or dummy variables including: meningitis epidemics during the surveillance period, malaria risk index (based on MARA malaria endemicity[12] and Guerra's population at risk[13]), proportion of births with skilled attendant, proportion of females who are literate, proportion of children <5 years of age who are stunted, proportion of children <5 years of age who are underweight, antenatal human immunodeficiency virus (HIV) prevalence, proportion of children <5 years of age with diarrhoea receiving oral rehydration solution (ORS), proportion of children <5 years of age with fever receiving anti-malaria medication, proportion of children <5 years of age sleeping under insecticide-treated nets, measles vaccination coverage rate, *Haemophilus influenzae* type B vaccination coverage rate, proportion of population using improved sanitation facilities, and proportion of population using improved water source.

Abstracted data were directly entered into a customized Microsoft Access™ (Microsoft Corporation, Seattle, WA, USA) database, with built-in quality control checks to ensure completeness and accuracy of the data entered. Reviewer data was compared using Epi Info™ version 3.5.1 (US Centers for Disease Control and Prevention, Atlanta, GA, USA) and any discrepancies were re-abstracted by a third reviewer whose abstraction was considered as final.

Cause of death data

Deaths were grouped into one of the following seven causes of death: pneumonia, diarrhoea, malaria, injury, meningitis/encephalitis, measles or other causes. Deaths due to undetermined causes and HIV/acquired immune deficiency syndrome (AIDS) were removed from the analysis ($n = 3382$). Neonatal deaths and deaths due to neonatal causes were excluded ($n = 35\,564$) where possible and a dummy variable used in the statistical models to indicate confidence in whether a study includes neonatal data (yes, no, unsure). Deaths attributed to a category that included more than one of the seven causes were re-allocated to the respective causes of death based on relative importance of single causes in the same studies.[7] Despite accounting for an estimated one-fifth of all under-5 deaths, stunting, wasting and micronutrient deficiencies are infrequently reported as the underlying cause of child mortality in verbal autopsies or even in VR of deaths.[14] Because many malnutrition cases are precipitated by infections such as measles or diarrhoea, we re-allocated malnutrition causes to selected infection categories (diarrhoea, malaria, measles, meningitis, pneumonia) based on relative importance of cause categories in the same study and all remaining unallocated malnutrition deaths were added to 'other' cause of death category.[7]

Modelling of study data

As was done in the previous analyses by Morris *et al*. and Lawn *et al*.,[7,9] the selection of the 'base' cause of death category was limited to those present in every study included in the analysis; in our study these were pneumonia, diarrhoea or other. Next, we calculated the natural logarithm of the ratio of the proportion of each of the other six causes of death relative to the proportion of the 'base' cause using ordinary least squares (OLS) regression, adjusting for possible explanatory variables both identified *a priori* and with a significance level of $\leqslant 0.20$ in forward stepwise regression.[7,9] The OLS was performed using pneumonia and diarrhoea each as the 'base' cause and pneumonia was selected as the final 'base' cause for the present analysis due to higher R^2 values in most log ratio models compared to diarrhoea as the 'base'.

After selecting pneumonia as the final 'base' cause and covariates for the six log ratio models, studies were given a weight proportional to the inverse of the square root of total number of deaths and then included in a multinomial logistic regression (MLR) model with robust standard errors to obtain parameter estimates.[9] Standard errors were adjusted for within country clustering. R^2 values were obtained when fitting the log ratios using OLS with each study having equal weight. We assumed that in studies where any of the specified causes of death were not present, they had been categorized as 'other'. Therefore some of the deaths in the 'other' category in these studies were re-allocated to the missing causes of death using probability patterns from studies with the specified causes of death. The procedure was done stepwise, where the missing causes of death were first imputed for studies with only one cause not available, and then two causes, until missing values in all causes were imputed.

National and regional estimates

To obtain the country-level predicted distribution of causes of death for 2008 for countries with incomplete VR data in sub-Saharan Africa and Southeast Asia, we applied year 2008 country-level input data available from surveys and public databases (e.g. WHO, UNICEF, World Bank) to the MLR parameter estimates from study-level data. Modelled estimates were restricted to countries with similar under-5 mortality risk and GNI as represented in the studies included in the analysis. Next, we multiplied the predicted country-level proportion of causes of death by the draft estimate for year 2008 total number of non-HIV deaths among children age 1–59 months to estimate the number of child deaths due to each of the seven causes for each country (WHO, unpublished tables based on the 2008 Report on the global AIDS epidemic).[15] Deaths due to measles and malaria were set equal to zero in measles elimination and *Plasmodium falciparum* (PF)-free countries, and the remaining deaths were re-allocated to the other cause

of death categories by relative importance using the average of three permutations changing the order in which the adjustments were made (measles-free, PF-free, and both measles- and PF-free).

In order to ensure cause of death estimates reflected the effect of intervention scale-up in 2008, we attempted to take into account the impact of these interventions on the distribution of causes of death in the MLR model. Several of the intervention coverage variables did not meet the significance threshold to be retained in the model possibly due to a limited number of contemporary studies reflecting a time period of recent scale-up of child survival interventions. Therefore, we made post-hoc adjustments to the country-level cause of death estimates for malaria, pneumonia and meningitis to account for intervention coverage and effectiveness using the average effect of intervention based on six different permutations changing the order of performing the adjustments, then deaths 'saved' by the intervention were reallocated to the other cause of death categories by relative importance. Malaria deaths were adjusted for the proportion of children <5 years of age using insecticide treated nets (ITNs) and the effectiveness of ITNs in preventing malaria mortality.[16] To account for the effect of *H. influenzae* type B (Hib) pediatric vaccination on: (i) pneumonia deaths—we multiplied the vaccine efficacy and immunization coverage for three doses of Hib (Hib3) by the vaccine efficacy of Hib3 against chest X-ray confirmed pneumonia[17]; (ii) for meningitis deaths in meningitis belt countries—we performed a fixed effects meta-analysis estimating the proportion of meningitis deaths due to Hib and multiplied this value by the Hib3 vaccine efficacy[17] and Hib3 immunization coverage; (iii) for meningitis deaths in non-meningitis-belt countries—we multiplied the vaccine efficacy of Hib3 against all-cause meningitis mortality[17] and Hib3 coverage. The final country-level cause-specific estimates were then aggregated to provide year 2008 estimates for the respective regions.

Methods for obtaining uncertainty estimates are described elsewhere,[9] briefly, we performed a jackknife analysis restricted to studies that converged (94% of included studies) to estimate the standard error and entered this standard error into Monte-Carlo simulations (1000 iterations) to provide a range of plausible estimates (the 2.5th and 97.5th centiles of the corresponding distributions were used).

Results

Of the 216 studies reviewed, 49 did not meet inclusion criteria most often due to lack of cause of death data or data only available for a single cause of death ($n=16$) or inability verify use of a standardized method for ascertainment of cause of death ($n=16$). Studies without cause of death data abstractable for study purposes (e.g. age range of children under mortality surveillance excluded infants or included children >60 months of age; only reported deaths due to a single cause) ($n=66$) and those with identical cause of death data reported in multiple studies ($n=20$) were also excluded from the analysis.

The remaining 81 study data points were included in the analysis and comprised 79 067 under-5 deaths from 25 countries, predominantly from sub-Saharan Africa and Southeast Asia (see Supplementary Table 1 available as Supplementary data at *IJE* online). No studies were identified from Western Asia, Central America, North America, or Europe, and only one study from Oceania is represented in the analysis. Characteristics of the included studies by region are presented in Table 1. Most studies were from low-income, rural, high-mortality settings. Only one-third of studies included any child mortality surveillance since year 2000 and most contemporary studies were from India. A disproportionate number of Asian studies were from settings with very low skilled attendance at birth, typically India and Bangladesh.

Each geographic region had a single country with a greater number of child deaths in the analysis compared with other countries (i.e. sub-Saharan Africa: Senegal; Southeast Asia: Bangladesh). Verbal autopsy data from nine DSS, most from sub-Saharan Africa are included in the analysis. The median number of child deaths in the included studies was 164 (range: 26–25 753). Final assignment of cause of death was ascertained using a computer algorithm ($n=2$), other standardized algorithm ($n=15$), independent expert(s) ($n=16$), a panel of experts ($n=17$), or was not further specified by authors ($n=31$).

All included study data points reported child deaths due to pneumonia, diarrhoea and other causes. Only one-quarter of study data points reported child deaths due to all seven categories, and most of these were from India. Deaths due to causes in the 'other' category were typically other communicable diseases (7.4%), other causes not further specified (53.4%), non-communicable diseases (11.3%) and congenital disorders or other conditions arising from the perinatal period (27.9%). As seen in Figure 1, there was substantial variability in the proportion of child deaths due to each cause observed across study data points.

Table 2 shows the MLR parameter estimates from included studies. The proportion of diarrhoea deaths was reduced compared with pneumonia with increased access to skilled attendant and proportion of children <5 years of age with diarrhoea receiving ORS. Compared with pneumonia, malaria death in children was associated with higher malaria risk index; and the proportion of child deaths due to measles was reduced in more contemporary mortality surveillance (indicated by later mid-study year). Overall, the explanatory variables in the malaria:pneumonia log ratio model account for most of the variation

Table 1. Characteristics of study data points included in the analysis, by region ($N = 81$)

Characteristics	Sub-Saharan Africa ($N = 27$) *n* (%)	Southeast Asia ($N = 41$) *n* (%)	All others[a] ($N = 13$) *n* (%)	Total ($N = 81$) *n* (%)
Characteristics of study setting				
Mid-study year				
1980–89	10 (37.0)	10 (24.4)	7 (53.9)	27 (33.3)
1990–99	10 (37.0)	13 (31.7)	5 (38.5)	28 (34.6)
2000–06	7 (25.9)	18 (43.9)	1 (7.7)	26 (32.1)
Study setting				
Urban	1 (3.7)	6 (14.6)	4 (30.8)	11 (13.6)
Rural	20 (74.1)	28 (68.3)	5 (38.5)	52 (64.2)
Mixed	6 (22.2)	7 (17.1)	4 (30.8)	18 (22.2)
Quintiles of under-5 mortality risk				
7.8–80.3	7 (25.9)	5 (11.9)	5 (38.5)	17 (21.0)
80.8–103.1	1 (3.7)	10 (23.8)	5 (38.5)	16 (19.8)
105.0–123.2	6 (22.2)	11 (26.2)	0 (0.0)	16 (19.8)
127.5–156.7	6 (22.2)	9 (21.4)	1 (7.7)	16 (19.8)
160.0–347.1	7 (25.9)	7 (16.7)	2 (15.4)	16 (19.8)
GNI per capita				
<750	6 (22.2)	16 (39.0)	0 (0.0)	22 (27.2)
750–999	13 (48.2)	4 (9.8)	2 (15.4)	19 (23.5)
1000–1999	1 (3.7)	5 (12.2)	2 (15.4)	8 (9.9)
2000–2999	1 (3.7)	16 (39.0)	2 (15.4)	19 (23.5)
⩾3000	6 (22.2)	0 (0.0)	7 (53.9)	13 (16.1)
Malaria risk index				
0–0.99	8 (29.6)	41 (100.0)	13 (100.0)	62 (76.5)
1.00–1.99	13 (48.2)	0 (0.0)	0 (0.0)	13 (16.1)
2	6 (22.2)	0 (0.0)	0 (0.0)	6 (7.4)
Proportion of births with skilled attendant (%)				
1–24	3 (11.1)	19 (46.3)	1 (7.7)	23 (28.4)
25–49	11 (40.7)	13 (31.7)	4 (30.8)	28 (34.6)
50–74	8 (29.6)	6 (14.6)	6 (46.2)	20 (24.7)
75–100	5 (18.5)	3 (7.3)	2 (15.4)	10 (12.4)
Children <5 years of age who are stunted (%)				
0–20	0 (0.0)	0 (0.0)	5 (38.5)	5 (6.2)
21–40	18 (66.7)	1 (2.4)	8 (61.5)	27 (33.3)
41–60	9 (33.3)	40 (97.6)	0 (0.0)	49 (60.5)
Children <5 years of age with diarrhoea receiving ORS (%)				
6–24	7 (25.9)	5 (12.2)	4 (30.8)	16 (19.8)
25–49	15 (55.6)	23 (56.1)	6 (46.2)	44 (54.3)
50–76	5 (18.5)	13 (31.7)	3 (23.1)	21 (25.9)
Population using improved sanitation facilities (%)				
0–19	6 (22.2)	12 (29.3)	1 (7.7)	19 (23.5)
20–39	11 (40.7)	19 (46.3)	5 (38.5)	35 (43.2)
40–59	9 (33.3)	10 (24.4)	6 (46.2)	25 (30.9)
60–84	1 (3.7)	0 (0.0)	1 (7.7)	2 (2.5)
Measles vaccine coverage rate (%)				
<50	3 (11.1)	9 (22.0)	1 (7.7)	13 (16.1)
50–69	13 (48.2)	15 (36.6)	1 (7.7)	29 (35.8)
70–89	9 (33.3)	12 (29.3)	10 (76.9)	31 (38.3)
90–100	2 (7.4)	5 (12.2)	1 (7.7)	8 (9.9)

(Continued)

Table 1. Continued

Characteristics	Sub-Saharan Africa ($N=27$) n (%)	Southeast Asia ($N=41$) n (%)	All others[a] ($N=13$) n (%)	Total ($N=81$) n (%)
Cause of death data				
Age range in months				
0–<59	1 (3.7)	2 (4.9)	3 (23.1)	6 (7.4)
0–<59[b]	0 (0.0)	0 (0.0)	1 (7.7)	1 (1.2)
1–<59	1 (3.7)	4 (9.8)	2 (15.4)	7 (8.6)
0–59	13 (48.1)	14 (34.1)	5 (38.5)	32 (39.5)
0–59[b]	2 (7.4)	1 (2.4)	1 (7.7)	4 (4.9)
1–59	10 (37.0)	20 (48.8)	1 (7.7)	31 (38.3)
Total no. of deaths in children <5 years of age				
<50	1 (3.7)	14 (34.2)	1 (7.7)	16 (19.8)
50–199	7 (25.9)	15 (36.6)	7 (53.9)	29 (35.8)
200–499	8 (29.6)	7 (17.1)	2 (15.4)	17 (21.0)
500–999	6 (22.2)	1 (2.4)	1 (7.7)	8 (9.9)
>1000	5 (18.5)	4 (9.8)	2 (15.4)	11 (13.6)
Availability of data by cause of death				
Pneumonia	27 (100.0)	41 (100.0)	13 (100.0)	81 (100.0)
Diarrhoea	27 (100.0)	41 (100.0)	13 (100.0)	81 (100.0)
Malaria	19 (70.4)	17 (41.5)	2 (15.4)	38 (46.9)
Injury	13 (48.2)	31 (75.6)	6 (46.2)	50 (61.7)
Meningitis	13 (48.2)	22 (53.7)	2 (15.4)	37 (45.7)
Measles	13 (48.2)	31 (75.6)	3 (23.1)	47 (58.0)
Other	27 (100.0)	41 (100.0)	13 (100.0)	81 (100.0)
Number of study data points reporting all causes of death	5 (18.5)	15 (36.6)	0 (0.0)	20 (24.7)

[a]Includes studies from the Latin America and Caribbean, Oceania and North African geographic regions.
[b]Inclusion of neonates in the study was unknown.

observed; whereas little variability was explained by the covariates in the diarrhoea:pneumonia and other:pneumonia ratios.

The observed and estimated distribution of causes of death from the study data is presented in Table 3 for all study data points ($n=81$) and those reporting all causes of death ($n=20$) predominantly from India (15/20). The leading causes of death estimated from the study-level MLR models were pneumonia, diarrhoea and malaria, with overlapping uncertainty ranges. As seen in Figure 2, leading causes of death for children 1–59 months old in sub-Saharan Africa were malaria, diarrhoea and pneumonia with each accounting for 21–26% of all child deaths. For this age group in Southeast Asia, pneumonia and diarrhoea comprised >60% of all deaths. Deaths due to AIDS are not included in these estimates but are available from UNAIDS.[15] Only 9 of the 81 study data points reported AIDS as a cause of death ($n=353$ deaths) although some of the AIDS deaths may have remained in 'other' causes.

Discussion

Our analysis demonstrated that statistical modelling can be used to estimate the causes of child mortality for settings without adequate VR systems using studies with verbal autopsy data reporting as few as two causes of death. This analysis builds upon the statistical methods used in the initial analysis of the proportional distribution of under-5 deaths by cause for countries in year 2000.[7,8] The prior analysis applied OLS and Seemingly Unrelated Regression methods to studies reporting under-5 deaths due to all five cause of death categories (pneumonia, diarrhoea, malaria, measles, other) ascertained by verbal autopsy methods to estimate the leading causes of child mortality.[7] The use of models that estimate multiple causes of death simultaneously has a possible advantage over single-cause models in that the estimated number of deaths due to proportionate causes cannot exceed the total number of deaths.

We adapted MLR methods used to estimate country-level distribution of causes of neonatal

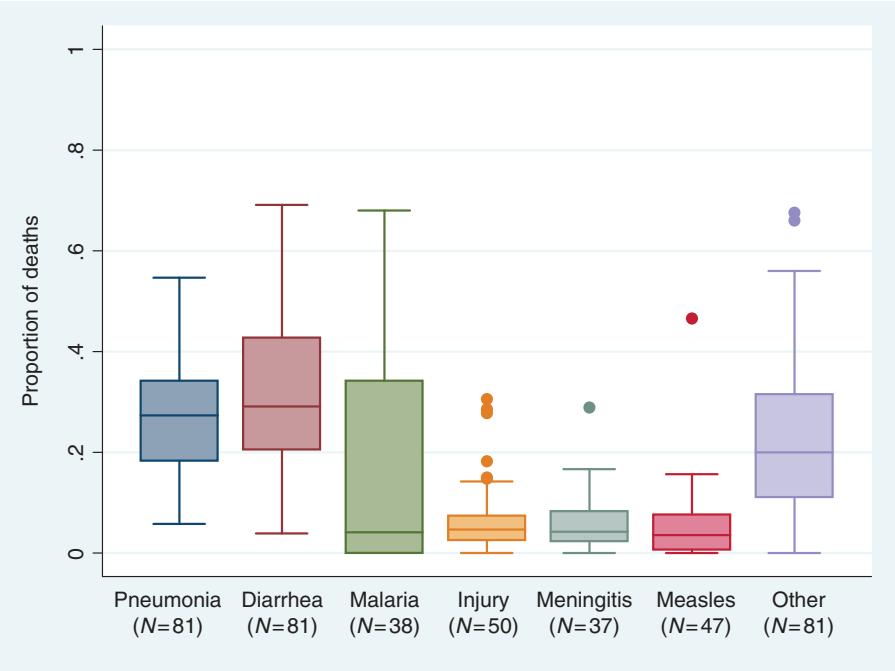

Figure 1. Distribution of causes of child death across study data points included in the analysis ($N=81$). Measures of spread [range; standard deviation (SD)] of study-level proportional mortality for causes with outlier values are: (i) injury (0–0.305; 0.070), (ii) meningitis (0–0.289; 0.058), (iii) measles (0–0.466; 0.076), (iv) other (0–0.676; 0.157)

deaths for 2000 to estimate the distribution of causes of child mortality in 2008 for the sub-Saharan African and Southeast Asian countries with incomplete death certification and adjusted estimates for the potential effect of child survival interventions during this time period.[9] A detailed discussion of the differences in the application of OLS and MLR statistical methods to study-level data and potential impact on the modelled cause of death estimates is provided elsewhere.[9] Use of MLR statistical methods allowing for simultaneous estimation of multiple causes of death from studies reporting a limited number of causes of death categories permitted inclusion of a larger number of studies and allowed for estimation of seven causes of death. Probability patterns of observed and missing causes of death across studies were used to predict and redistribute deaths from 'other' to the missing causes of death, thus using more of the available cause of death data. Additionally, we adjusted country-level cause of death estimates for intervention coverage and effectiveness to reflect the recent scale-up of child survival interventions in many countries.

Our analysis relies on existing studies that collected and reported cause of death data for purposes other than the analysis presented here, and the limitations of using these study-level data and applying statistical methods to estimate the distribution of causes of child death have been discussed in related analyses.[7,9] Briefly, the ability of the model to estimate parameters for each of the six MLR equations and country-level cause of death distribution is dependent on the quality and representativeness of the study data. To avoid extrapolating study data to dissimilar settings, we limited our prediction to countries meeting specific criteria based on characteristics of the included studies. Because our analysis included very few studies from lower child mortality risk settings, we only estimated the proportional cause of death distribution for countries within the range of the input countries i.e. an under-5 mortality risk of >25 per 1000 live births and a GNI $\leqslant$$7510.

The cause of death distribution for sub-Saharan Africa and Southeast Asia estimated here cannot be validated by national VR data because no countries in these regions with similar characteristics have adequate VR systems, but we can compare our estimates with other modelled estimates. Although we included almost double the number of studies in the analysis and used more contemporary data, the uncertainty ranges around the estimates presented here include the point estimates for causes in the year 2000 analysis.[7] Pneumonia and diarrhoea were the leading causes of death in each region followed by malaria in sub-Saharan Africa for 2008 and are consistent with year 2000 estimates. Estimates of 'other' causes for 2000 are included in the uncertainty range for 2008 but the point estimates are considerably greater in the estimates presented here. We are not aware of existing estimates of child injury or meningitis deaths with

Table 2. OLS and MLR model parameter estimates for data points included in the analysis ($N = 81$)

Explanatory variable	OLS estimates				MLR estimates		
	R^2	Parameter estimate	95% CI	*P*-value	Parameter estimate	95% CI	*P*-value
Diarrhoea:pneumonia ratio							
Country in WHO Western Pacific region	0.24	−1.745	−2.703 to −0.787	0.001	−1.491	−2.076 to −0.907	0.000
Age of oldest child under mortality surveillance was 12 or 36 months		0.499	0.084 to 0.914	0.019	0.424	−0.085 to 0.933	0.102
Mid-study year		−0.015	−0.036 to 0.005	0.132	−0.029	−0.059 to 0.001	0.058
Proportion of children <5 years of age with diarrhoea receiving ORS		−0.010	−0.019 to 0.001	0.040	−0.008	−0.015 to −0.000	0.037
Proportion of births with skilled attendant		−0.004	−0.012 to 0.002	0.174	−0.007	−0.018 to 0.003	0.178
Malaria:pneumonia ratio							
Malaria risk index	0.74	3.156	2.730 to 3.582	0.000	2.099	1.546 to 2.653	0.000
Country in WHO Western Pacific region		−1.161	−3.113 to 0.791	0.240	−3.494	−145.150 to 138.162	0.961
Injury:pneumonia ratio							
Percent of births with skilled attendant	0.33	−0.035	−0.051 to −0.018	0.000	−0.008	−0.022 to 0.006	0.254
Mid-study year		0.036	−0.019 to 0.090	0.195	0.013	−0.034 to 0.061	0.577
Under-5 mortality risk, per 1000 live births		−0.008	−0.014 to −0.002	0.010	0.001	−0.004 to 0.006	0.708
Age of oldest child under mortality surveillance was 12 or 36 months		−1.293	−2.395 to −0.192	0.022	−0.828	−1.434 to −0.222	0.007
Measles vaccination coverage rate		0.015	−0.005 to 0.034	0.137	0.018	0.005 to 0.030	0.006
Meningitis:pneumonia ratio							
Mid-study year	0.45	0.122	0.064 to 0.180	0.000	0.000	−0.164 to 0.165	0.997
Age of youngest child under mortality surveillance was 1 month		1.115	0.331 to 1.900	0.006	−0.108	–1.030 to 0.813	0.818
GNI per capita, (Purchasing Power Parity, international dollars)		−0.001	−0.001 to −0.000	0.001	–0.000	−0.001 to 0.001	0.815
Proportion of children <5 years of age who are underweight		−0.072	–0.112 to −0.031	0.001	−0.032	−0.115 to 0.051	0.456
Under-5 mortality risk, per 1000 live births		−0.009	−0.015 to –0.003	0.006	−0.004	−0.040 to 0.033	0.839
Proportion of children <5 years of age who are stunted		0.057	−0.001 to 0.114	0.053	0.039	−0.031 to 0.110	0.275
Measles:pneumonia ratio							
Proportion of children <5 years of age who are stunted	0.27	0.069	0.024 to 0.114	0.003	0.017	−0.037 to 0.071	0.537
Mid-study year		−0.081	−0.135 to −0.028	0.003	−0.049	−0.116 to 0.018	0.153
Country in WHO region of the Americas		−1.414	−3.168 to 0.340	0.112	0.841	−1.470 to 3.151	0.476
Proportion of the population living in an urban area		0.009	−0.004 to 0.021	0.175	0.007	−0.002 to 0.015	0.152

(Continued)

Table 2. Continued

Explanatory variable	OLS estimates				MLR estimates		
	R^2	Parameter estimate	95% CI	*P*-value	Parameter estimate	95% CI	*P*-value
Country in WHO Eastern Mediterranean or South-East Asia regions		−0.960	−2.037 to 0.118	0.080	−1.077	−2.012 to −0.142	0.024
Other:pneumonia ratio							
Age of oldest child under mortality surveillance was 12 or 36 months	0.08	1.383	0.211 to 2.554	0.021	0.306	−0.836 to 1.448	0.600
Mid-study year		0.014	−0.044 to 0.073	0.627	0.021	−0.043 to 0.084	0.519
Inclusion of neonates in the mortality surveillance was unknown		1.076	−0.722 to 2.874	0.237	0.650	−0.927 to 2.227	0.419

CI: confidence interval. R^2 values obtained when fitting the log ratio using OLS regression with each data point having equal weight.

which we can compare. The uncertainty around our estimates of child deaths by cause are wide but do not incorporate all uncertainty in the model input data or post-hoc adjustments to cause-specific estimates, thus caution should be used when interpreting the results presented here. The quality of cause of death ascertainment methods and model covariate data contribute to the uncertainty in our estimates. Verbal autopsy methods have been validated and described elsewhere,[4–6] but use of a standardized method, preferably a computer-based algorithm, could improve the observed variability in our model parameter estimates. We only included cause of death data from DSS sites known to use standard verbal autopsy methods,[3] but it was difficult to properly evaluate the methods used in many of the epidemiological studies based on the limited information provided in the publication. Development of methods to evaluate the quality of verbal autopsy methods in studies reporting cause of child death data and use of data quality score or measure to exclude studies not meeting quality standards or adjustment for data quality through statistical models could reduce the uncertainty in our estimates. Measuring the between study heterogeneity in the proportional cause of death distribution and identifying differences in study setting or design that might explain these differences and incorporating this into our study methods could also improve our ability to estimate causes of child mortality.

Quality of covariate data is another potential source of residual variability. Studies frequently did not report the covariate data of interest, thus we relied on other sources of covariate data frequently representative of the sub-national or national level, thus covariate data may not necessarily represent the study setting. The paucity of study and national-level data for potential explanatory variables such as breastfeeding, use of anti-malaria medications, and interventions to reduce micronutrient deficiency, reduced our ability to account for variability or make adjustments to our estimates for the impact of these interventions.

After reporting the distribution of the most common causes of child death, a large proportion of deaths in most studies were aggregated into a category of 'other causes' not further specified, thus it is possible that deaths due to less common causes such as injuries, meningitis and measles either did not occur or were included in this 'other causes' category and it is unclear if the studies where all cause of death categories were specified can be generalized in the model. As a result, the proportions of the less common causes could be under- or over-estimated. Approaching authors directly for the observed causes of death and study-level covariate data could improve the model parameter estimates and ability to explain the variability in our MLR models.

Two-thirds of included studies reported on mortality during 1980–99, thus there were few studies with

Table 3. Observed and estimated distribution of causes of death among children age 1–59 months with uncertainty estimates from (i) studies included in the analysis and (ii) country-level data

	(i) Study-level data					(ii) Country-level data				
	Mean proportion of deaths among all studies in analysis ($N = 81$)	Multinomial model		Mean proportion of deaths among studies with all causes reported ($N = 20$)		Sub-Saharan Africa		South/East Asia	
Cause	Observed (%)	Estimated (%)	Observed (%)	Observed (%)	Uncertainty range[a]	Estimated (%)	Uncertainty range[a]	Estimated (%)	Uncertainty range[a]
Pneumonia	29.5	24.4	25.7	25.7	17.5–25.6	20.6	13.9–22.1	30.5	19.4–37.2
Diarrhoea	19.7	27.2	26.7	26.7	19.8–34.0	24.8	16.6–32.8	31.3	17.2–45.2
Malaria	6.8	16.2	18.6	18.6	10.5–25.9	25.5	15.7–40.2	1.5	0.6–6.9
Injury	6.8	3.4	2.9	2.9	2.7–7.4	3.1	2.2–6.4	3.9	2.1–12.6
Meningitis	1.0	3.1	4.7	4.7	2.8–11.0	2.9	2.3–10.3	3.5	1.9–18.6
Measles	1.0	2.6	5.4	5.4	1.9–5.0	3.0	2.0–6.3	1.9	1.1–4.6
Other	35.2	23.0	16.1	16.1	16.6–28.4	20.2	14.0–27.0	27.4	15.9–39.0
Total	100.0	100.0	100.0	100.0		100.0		100.0	

[a]Estimated using Monte Carlo simulation based on standard errors from jackknife estimation.

cause of death data representative of the last decade when there has been recent scale-up of child survival interventions, including widespread introduction of Hib paediatric vaccination, use of insecticide-treated bed nets and measles elimination efforts. The few studies with post-1999 verbal autopsy data available were from mortality surveys conducted predominantly in India and may not represent the distribution of causes of child mortality in other settings with varying coverage levels of child survival interventions. With a limited number of studies from settings in which these interventions have been introduced at scale representative of the current context, rather than including intervention coverage as a covariate in the model, we made post-hoc adjustments to country-level estimates to account for current levels of use of Hib paediatric vaccination and insecticide-treated nets and the estimated effectiveness of these interventions that might have led to over/under-estimation of deaths due to related causes. Additional contemporary cause of death data could improve the ability to adjust mortality estimates for intervention coverage.

Rates of child mortality are highest in the sub-Saharan African and Southeast Asian geographic regions where complete death registration is rare. Although adequate VR systems with complete cause of death reporting is preferable, modelled estimates can provide information essential for monitoring and evaluating the health status of the population over time and an evidence-base for defining national health priorities. Despite the limitations of available community-based mortality surveillance data, we demonstrated the application of improved statistical methodology to estimate the distribution of causes of child mortality in low-income, high-mortality countries lacking adequate VR systems. The causes of neonatal deaths have been estimated with a similar model to that described in this article[9] and child deaths due to AIDS have also been modelled[8,15]; these estimates would need to be combined with our results to derive the complete distribution of the causes of death in children <5 years of age. Previous estimates suggested that AIDS accounted for 6% of child deaths in sub-Saharan Africa, which would make it the fourth most common cause after malaria, diarrhoea and pneumonia.[8] The advancement of methodology for disease burden estimation is ongoing through several collaborative initiatives and will remain an important mechanism for establishing national and international health priorities; measuring the impact of increased access to life-saving child survival interventions; and progress towards achieving the year 2015 Millennium Development Goal 4 target to reduce child mortality by two-thirds.[18]

Supplementary Data

Supplementary Data are available at *IJE* online.

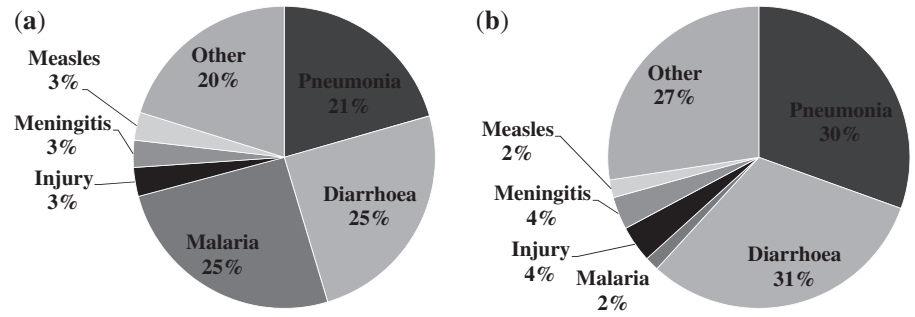

Figure 2. Distribution of causes of death among children age 1–59 months in 2008 for (a) Sub-Saharan African and (b) Southeast Asian high-mortality countries with incomplete death certification. Deaths due to AIDS are not included in the estimates. Country-level proportional mortality by cause for 2008 was multiplied by total number of deaths among children age 1–59 months (excluding HIV deaths) for each country, then deaths by cause were aggregated to the regional level

Funding

The U.S. Fund for UNICEF and The Bill & Melinda Gates Foundation.

Acknowledgements

This work was conducted in collaboration with the CHERG of WHO/UNICEF to estimate the causes of child mortality. We thank all CHERG members for their comments and collaboration throughout this process. We also thank the following individuals for their support to this project: Sidharth Ramji (Maulana Azad Medical College), Colin Mathers (WHO), Richard Cibulskis (WHO), Neff Walker and Martin Aryee (Johns Hopkins University Bloomberg School of Public Health) and Simon Cousens (London School of Hygiene and Tropical Medicine) for their data contributions and statistical support; Thom Eisele (Tulane University School of Public Health & Tropical Medicine), Igor Rudan and Harry Campbell (University of Edinburgh Medical School), and Chris Sudfeld (Harvard School of Public Health) and Neal Halsey (Johns Hopkins University Bloomberg School of Public Health) for sharing literature reviews; and our data abstractors Laura Lamberti, Sabriya Linton, Cristina Garcia, Aimee Summers (Johns Hopkins University Bloomberg School of Public Health).

Conflict of interest: None declared.

KEY MESSAGES

- Publicly available verbal autopsy data reporting as few as two causes of child death can be used to model the proportional distribution of causes of child mortality in similar settings with incomplete death certification.
- Despite recent introduction and scale-up of child survival interventions and reductions in overall child mortality, the leading causes of post-neonatal infant and child mortality and respective proportionate mortality distribution for sub-Saharan Africa and Southeast Asia was consistent with previous estimates.
- Heterogeneity in observed study-level estimates of the proportion of deaths by cause may in part be due to inconsistent methods of verbal autopsy and this could be reduced by universal use of standardized methods and definitions in ascertainment of cause of death.
- Increased availability of contemporary, nationally representative data on the causes of child mortality, and explanatory variables such as intervention coverage, will improve our ability to model cause of death data in the context of scale-up of child survival interventions.

References

[1] Rudan I, Lawn J, Cousens S *et al*. Gaps in policy-relevant information on burden of disease in children: a systematic review. *Lancet* 2005;**365:**2031–40.

[2] INDEPTH Network. *An International Network of Field Sites with Continuous Demographic Evaluation of Populations and their Health in Developing Countries*. http://www.indepth-network.org/ (13 December 2009, date last accessed).

[3] World Health Organization. *Verbal Autopsy Standards: Ascertaining and Attributing Cause of Death*.Geneva: World

Health Organization; 2007. http://www.who.int/whosis/mort/verbalautopsystandards/en/index.html (13 December 2009, date last accessed).

[4] Kalter HD, Gray RH, Black RE, Gultiano SA. Validation of postmortem interviews to ascertain selected causes of death in children. *Int J Epidemiol* 1990;**19:**380–86.

[5] World Health Organization. Measurement of overall and cause-specific mortality in infants and children: memorandum from a WHO/UNICEF meeting. *Bull World Health Organ* 1994;**72:**707–13.

[6] Setel PW, Rao C, Hemed Y *et al*. Core verbal autopsy procedures with comparative validation results from two countries. *PLoS Med* 2006;**3:**1282–91.

[7] Morris SS, Black RE, Tomasković L. Predicting the distribution of under-five deaths by cause in countries without adequate vital registrations systems. *Int J Epidemiol* 2003;**32:**1041–51.

[8] Bryce J, Boschi-Pinto C, Shibuya K, Black RE. WHO Child Health Epidemiology Reference Group. WHO estimates of the causes of death in children. *Lancet* 2005;**365:**1147–52.

[9] Lawn JE, Wilczynska-Ketende K, Cousens SN. Estimating the causes of 4 million neonatal deaths in the year 2000. *Int J Epidemiol* 2006;**35:**706–18.

[10] Preston SH, Heuveline P, Guillot M. *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell Publishing, 2001.

[11] United Nations. *Manual X: Indirect Techniques for Demographic Estimation*. New York: United Nations, 1983. http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm (29 December 2009, date last accessed).

[12] Adjuik M, Bagayoko M, Binka F *et al*. *Towards an Atlas of Malaria Risk in Africa*. Durban: MARA/ARMA; 2008. http://www.mara.org.za/ (29 December 2009, date last accessed).

[13] Guerra CA, Gikandi PW, Tatem AJ *et al*. The limits and intensity of *Plasmodium falciparum* transmission: Implications for malaria control and elimination worldwide. *PLoS Med* 2008;**5:**300–11.

[14] Black RE, Allen LH, Bhutta ZA *et al*. Maternal and child undernutrition: Global and regional exposures and health consequences. *Lancet* 2008;**371:**243–60.

[15] United Nations. *2008 Report on the Global AIDS Epidemic*.http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/2008/2008_Global_report.asp (29 December 2009, date last accessed).

[16] Eisele TP, Larsen D, Steketee RW. Protective efficacy of interventions for preventing malaria mortality in children in Plasmodium falciparum endemic areas. *Int J Epidemiol* 2010;**39 (Suppl 1):**i88–101.

[17] Theodoratou E, Johnson S, Jhass A *et al*. The effect of Haemophilus influenzae type b and pneumococcal conjugate vaccines on childhood pneumonia incidence, severe morbidity and mortality. *Int J Epidemiol* 2010;**39 (Suppl 1):**i172–85.

[18] United Nations. General Assembly, 56th session. *Road Map Towards the Implementation of the United Nations Millennium Declaration: Report of the Secretary-General*. New York: United Nations, 2001.