

1 **TITLE:**

2 Cohort profile: The WHO Child Mortality Risk Stratification Multi-Country
3 Pooled Cohort (WHO-CMRS) to identify predictors of mortality through early
4 childhood

5
6 **Authors:**

7 Catherine Schwinger^{1*}, Siri Kaldenbach², James A. Berkley^{3, 4}, Judd L. Walson⁵, Alemayehu
8 Argaw⁶, Ranadip Chowdhury⁷, Tor A. Strand^{1,2}, Nigel Rollins⁸, on behalf of the WHO Risk
9 Stratification Working Group (WHO-RSWG)

10

11 **Affiliations:**

12 ¹ Centre for International Health, University of Bergen, Bergen, Norway

13 ² Department of Research, Innlandet Hospital Trust, Lillehammer, Norway

14 ³ Kenya Medical Research Institute (KEMRI)/ Wellcome Trust Research Programme, Kilifi,
15 Kenya

16 ⁴ Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine,
17 University of Oxford, Oxford, UK

18 ⁵ Department of International Health, Bloomberg School of Public Health, Johns Hopkins
19 University, Baltimore, USA

20 ⁶ Department of Food Technology, Safety and Health, Faculty of Bioscience Engineering,
21 Ghent University, Ghent, Belgium

22 ⁷ Centre for Health Research and Development, Society for Applied Studies, New Delhi, India

23 ⁸ Department of Maternal, Newborn, Child and Adolescent Health and Ageing, WHO,
24 Geneva, Switzerland.

25

26 *Corresponding author: Catherine Schwinger, Årstadveien 21, 5009 Bergen,
27 c.schwinger@uib.no

28

29

30

31 **Keywords:** Age-specific death rate, anthropometry, morbidity, low birthweight, preterm
32 birth, breastfeeding, LMIC

33

34 **ABSTRACT (284 words)**

35 **Purpose:** To provide details of a pooled dataset that will be used to estimate absolute and relative
36 mortality risks and other outcomes among children less than 59 months of age and the predictive
37 performance of common risk exposures, both individually and in combination.

38 **Participants:** Children from birth to five years of age recruited at health facilities or community
39 settings into 33 longitudinal observational or intervention studies in 18 low-and middle-income
40 countries.

41 **Findings to date:** The dataset includes 75,287 children with a median age of 3 months (IQR 1, 12) at
42 first measurement. In the pooled sample, 2,805 (3.7%) of the study children died. Data on
43 birthweight was recorded in 18 studies, and gestational age in 13 studies. Among these, 14% of the
44 included children were reported with low birthweight and 14% preterm birth. At first measurement,
45 33% of the children were stunted, 24% were wasted, and 35% underweight. 13% and 7% of
46 caregivers reported that their child had acute diarrhoea or acute lower respiratory tract infection
47 before the study visit, respectively. The proportion of children being breastfed at any study visit
48 decreased from 99% at age <6 months to 77% in the age group 12-23 months. Child characteristics
49 differed considerably between studies in the community and health care settings. The median study
50 period was 15 months (IQR 7.6 to 18.4 months).

51 **Future plans:** The WHO Child Mortality Risk Stratification Multi-Country Pooled Cohort (WHO-CMRS)
52 provides a large dataset including child, parental, and household characteristics from a diverse range
53 of geographical, community and health system settings; planned analyses will examine knowledge
54 gaps with the aim of informing global guidelines and their derivatives such as clinical management
55 tools and implementation guidance, and to inform future research agendas.

56

57 **STRENGTHS AND LIMITATIONS**

58 Strengths:

- 59 - A large dataset that integrates diverse geographical areas in LMICs, with children
60 recruited from community and health system settings.
- 61 - Prospective data collection and availability of key time-varying exposures, such as
62 nutritional status, common childhood morbidities, and feeding practices and non-
63 time varying exposures, such as birth characteristics and socio-demographics.
- 64 - Ability to estimate age-specific absolute mortality risks from different exposure
65 domains and combinations thereof.

66 Limitations:

- 67 - Non-systematic selection of included datasets.

68 - Heterogeneity of inclusion criteria, exposures collected and follow-up schedules
69 across studies.

70

71 **INTRODUCTION**

72 Sustainable Development Goal (SDG) 3 aims to end preventable deaths of newborns and children
73 under five years of age by 2030. Although the world has seen remarkable progress in reducing deaths
74 in children under 5 years, with a 59% reduction since 1990, preventable child mortality is still high.
75 The rate of decline has also decreased since 2015. In 2021, an estimated 5 million children under the
76 age of five died of largely preventable or treatable causes (1, 2). Furthermore, 53 countries must
77 double or triple their current pace in reducing the under-five mortality rates to meet this SDG. There
78 is an urgent need to identify and support improvements in strategies to protect child survival.

79

80 Communicable diseases, together with preterm birth, remain the leading (direct) causes of death in
81 children (1, 2), with pneumonia being the single most significant infectious cause of death in children
82 worldwide (3-5). Undernutrition is estimated to be associated with up to 50% of all deaths in under-
83 five children (6), and the risk of death increases with greater degrees of anthropometric deficit (7).
84 Nevertheless, there is longstanding debate within the nutrition research and practice communities
85 surrounding the most appropriate anthropometric measures to identify the risk of death (8, 9),
86 especially for infants under six months of age (10). Low birthweight (11, 12), preterm birth (13-15),
87 and sub-optimal breastfeeding practices (16) as well as maternal nutrition and vital status (17) are
88 also well-established risk factors. Focusing on single risk domains such as anthropometry or disease
89 severity may overly simplify the risks for an individual child who may present with more than one
90 disease and/or underlying conditions (18), and who may have additional environmental, social and
91 economic risk factors not captured by clinical risk assessments (19).

92

93 The concept of risk stratification is integral to WHO care pathways, such as the Integrated
94 Management of Childhood Illness (IMCI) or Integrated Community Case Management. However,
95 these clinical tools largely focus on the severity of the presenting signs, such as dehydration or fast
96 breathing or chest indrawing. Although some evidence is emerging that more systematically
97 addresses the interaction of undernutrition and morbidity on mortality (20-23), there is limited
98 literature exploring the general concept of risk stratification at the individual level for vulnerable
99 children (24, 25). Identifying children at the highest risk of death, especially in resource-constrained
100 settings, may enable formal health systems to target these groups more effectively and improve
101 resource allocation and programmatic focus. In addition, identifying children at much lower risk of
102 death may provide the opportunity to redirect limited resources to achieve maximal impact.

103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137

We pooled multiple datasets to establish a large database that integrates individual-level data from a range of geographical, community and health system settings in LMICs. Analyses are in progress to estimate absolute and relative mortality risks in early childhood and predictive performance of common risk exposures, both individually and in combination. This initiative may provide evidence to inform global guidelines, clinical management tools and future research agendas benefitting infants and children at the highest risk of adverse outcomes. Here we describe the process of bringing the pooled dataset together as well as the main characteristics of the studies and children whose data are included. These will inform future analysis approaches.

COHORT DESCRIPTION

Selection of studies

Between August and November 2021, WHO convened the Risk Stratification Working Group (RSWG) and invited research investigators to consider specific research questions and analytical approaches for a set of risk stratification analyses related to child mortality. Researchers were invited based on known interest and having published relevant data. Other researchers requested to participate and actively contributed to the technical discussions. A list of RSWG members can be found in Table S1. Collectively, it was agreed that an individual-level pooled analysis of prospective cohort data would provide the highest quality evidence, enabling harmonization of an analytical approach, inclusion of smaller datasets, as well as flexibility for interaction and sensitivity analyses.

Datasets that could contribute to the analyses were identified following active outreach by all RSWG members within their various networks, emphasizing an inclusive approach.

In February 2022, all identified eligible study teams interested in sharing their data with the RSWG were asked to complete a Data Pooling Agreement. The agreement outlined conditions by which the Principal Investigators or representatives of host institutions of specific studies would share anonymized data with WHO for inclusion in the joint analyses, with the WHO being responsible for the overall coordination and additional technical support being provided by the University of Bergen. In total, 27 agreements were signed granting permission to pool 35 separate datasets. The WHO Ethics Review Committee granted an exemption in March 2022 with protocol ID: ERC.0003745.

The following inclusion criteria were set for datasets to be included in the pooled database:

- The dataset included a minimum set of variables defined as “required”, i.e., mortality, age, sex, and weight. Data on other anthropometric measures, and exposures within other risk

138 domains including morbidity, maternal health and nutrition status, pregnancy outcomes, and
139 clinical signs and symptoms were requested, but not required for inclusion in the dataset.

- 140 • The study included data on children <60 months of age.
- 141 • The study was conducted in a low-or low-middle income country (World Bank definition
142 (26).
- 143 • The study followed individual children longitudinally.
- 144 • An adequate description of the study population including the sampling and recruitment
145 strategy (e.g., random vs. convenience sampling, self-presentation vs. active community
146 screening), inclusion and exclusion criteria, and assessment procedures were available.
- 147 • Date of enrolment of participants from January 2000 onward.

148
149 Datasets were excluded if they were cross-sectional or did not include all essential data i.e. mortality,
150 age, sex, and weight.

151

152 **Data sources in the pooled dataset**

153 Up until August 2023 and following the exclusion of two datasets, 33 datasets had been included in
154 the pooled database (Figure 1). Individual studies contributing to the pooled database were
155 conducted in 18 different countries in five of the six WHO Regions: six in South-East Asia, one in the
156 Eastern Mediterranean region, 24 in sub-Saharan Africa, one in the Western Pacific region, and one
157 in Latin America. One study was conducted in more than one world region (Study 30, Table 1). Of the
158 studies included, 21 were randomized controlled trials (RCTs), and 12 were observational cohort
159 studies. Most RCTs (16 studies) assessed nutritional interventions such as multivitamins or specially
160 formulated and nutrient dense foods with or without additional educational or biomedical elements.
161 Included studies are described in Table S2.

162

163 **Figure 1:** Overview of the data included in the WHO Child Mortality Risk Stratification Multi-Country Pooled
164 Cohort (WHO-CMRS)

165

166 Study populations were categorized into three types: General Population (GP) when enrolment to
167 the original study was not based on an anthropometric deficit or the presence of an illness even if
168 some additional inclusion criteria may have been implemented; and Anthropometry-Selected (A-S) or
169 Illness-Selected (I-S) if the original study enrolled children based on an anthropometric deficit or
170 presence of signs or symptoms of an illness, respectively. For the categorization of geographical
171 regions of study implementation, we used WHO world regions (<https://www.who.int/countries>).

172 **Table 1:** Brief description of studies included in the WHO Child Mortality Risk Stratification Multi-Country Pooled Cohort (WHO-CMRS)

Nr	Short title	Main publication(s) (PMID)	Country of implementation	Years of implementation	Study design	Sample size ¹	Study category ²	Age at enrolment (months)	Median Follow-up time (months, IQR)
1	AMANHI Bangladesh	29163938; 34999881	Bangladesh	2014-2016	Cohort	938	GP	<1 mo	23.9 (23.8, 24)
2	AMANHI Pakistan	29163938; 34999881	Pakistan	2014-2016	Cohort	1,138	GP	<1 mo	23.7 (23.6, 23.8)
3	CARING	25886587; 28911749	India	2013-2015	RCT	2,993	GP	<1 mo	17.7 (17.5, 17.7)
4	CHILD 2	32518499; 30718805; 29472265	Tanzania	2007-2009	RCT	2,398	GP	<1 mo	17.2 (9.3, 17.6)
5	CRI study	21793745; 20028725; 17916587	India	2002-2006	Cohort	373	GP	<1 mo	43.3 (39.6, 47.1)
6	IECDZ	29689045; 28588962	Zambia	2014-2015	RCT	526	GP	6-12 mo	27.1 (13.2, 27.6)
7	ILiNS-ZINC	25816354; 25521188; 6362661	Burkina Faso	2010-2012	RCT	3,220	GP	9 mo	8.8 (8.6, 9)
8	LAO zinc	32153900; 30580974; 32612816	Lao	2015-2017	RCT	3,406	GP	6-23 mo	7.7 (7.6, 7.9)
9	Low BW	18297930	Burkina Faso	2004-2005	Cohort	1,103	GP	<1 mo	12 (12, 12)
10	MAL-ED Bangladesh	25305298	Bangladesh	2010-2012	Cohort	265	GP	<1 mo	35 (32, 35)
11	MAL-ED India	25305300	India	2010-2012	Cohort	247	GP	<1 mo	33.1 (26.1, 35.1)
12	MISAME I	18996870	Burkina Faso	2004-2006	RCT	1,221	GP	<1 mo	11.4 (3.3, 12.2)
13	MISAME II	19812173	Burkina Faso	2006-2008	RCT	1,084	GP	<1 mo	11.8 (8.2, 12.3)
14	MISAME III	36745684	Burkina Faso	2019-2020	RCT	1,608	GP	<1 mo	10.7 (8.0, 11.0)
15	MRCG Keneba	26559544; 30753251	Gambia	2005-2016	Cohort	560	GP	<1 mo	12.3 (3.8, 19.2)

16	Pemba	29163938; 34999881	Tanzania	2014-2018	Cohort	4,109	GP	<1 mo	36 (32.5, 37.1)
17	PM2A Guatemala	30184223	Guatemala	2011-2015	RCT	4,212	GP	<1 mo	22.9 (22.7, 23.0)
18	PNS	17409323; 34151973	Tanzania	2001-2004	RCT	7,956	GP	<1 mo	11.6 (10.7, 16.9)
19	PROMIS Burkina Faso	31454347	Burkina Faso	2015-2017	RCT	2,113	GP	0-17 mo	17.7 (17.2, 18.0)
20	PROMIS Mali	PMC5343313; PMC6711497	Mali	2015-2017	RCT	1,132	GP	6-23.9 mo	18 (18, 18.1)
21	ROSE	34375336	Niger	2014-2015	RCT	2,507	GP	1.4-1.8 mo	23.8 (23.3, 24.3)
22	South Africa Study	37058529	South Africa	2018-2021	RCT	386	GP	<1 mo	25.8 (8.8, 30.2)
23	ZINC 720	27489011	Burkina Faso	2010-2012	RCT	6,245	GP	6-27 mo	3.6 (0, 7.1)
24	3 foods for MAM	22170366; 23256140; 25419681	Malawi	2009-2011	RCT	2,707	A-S	6-59 mo	0.5 (0.5, 1.5)
25	10% milk RUTF	20980648	Malawi	2008-2009	RCT	1,873	A-S	6-59 mo	1 (0.5, 1.4)
26	Abx for SAM	23363496	Malawi	2009-2011	RCT	2,742	A-S	6-59 mo	1 (0.5, 1.5)
27	ComPAS	29690916; 32645109; 33534818	Kenya, South Sudan	2017-2018	RCT	4,072	A-S	6-59 mo	2.1 (1.2, 3.2)
28	DIVIDS	21628364	India	2007-2010	RCT	2,079	A-S	<1 mo	6 (5.1, 6)
29	HI MAM	33963734	Sierra Leone	2018-2020	RCT	1,286	A-S	6-59 mo	5.5 (5.5, 5.5)
30	CHAIN	35427524	Kenya, Uganda, Malawi, Burkina Faso, Bangladesh, Pakistan	2016-2019	Cohort	3,101	I-S	2-23 mo	6.1 (5.8, 6.4)
31	SMART 0-6 Phase1	37182535; 26608641	Uganda	2018-	Cohort	2,647	I-S	0-6 mo	6 (6, 6)
32	SMART 6-60 Phase0	37182535; 26608641	Uganda	2017-	Cohort	1,273	I-S	6-60 mo	6 (6, 6)
33	SMART 6-60 Phase1	37182535; 26608641	Uganda	2018-	Cohort	3,767	I-S	6-60 mo	6 (6, 6)

173 ¹ Number of children in the datasets received for the WHO Child Mortality Risk Stratification Multi-Country Pooled Cohort (WHO-CMRS)

174 ² Categorized based on inclusion and exclusion criteria of individual studies. Anthropometry selected (A-S): enrolment based on an anthropometric deficit; Illness selected (I-S): enrolment
175 based on the presence of an illness; General Population (GP): study population enrolled from a general population without consideration of an anthropometric deficit or the presence of
176 illness; other inclusion criteria were however, implemented.

177 **Data cleaning**

178 The pooled dataset underwent a data cleaning process before analysis (Figure 1). Datasets with one
179 row per child (wide format) were re-structured into a long format with one row for each observation
180 at different time-points of a child (and thus multiple rows per child). All duplicate rows were omitted
181 from the dataset (949 observations). Some duplicates occurred because continuous variables in
182 some datasets were recorded with lower precision, e.g., age was recorded as age in months with one
183 decimal place such that follow-up visits more frequent than 0.1 months could not be differentiated.
184 All observations for individual visits without information on age (22,342 observations) or sex (1,047
185 observations) were omitted from the analysis. We created a missing category for variables with
186 missing values in individual participants but that were generally captured in the study.
187 Measurements of exposures that occurred after age 59 months were omitted from the analyses (n =
188 160); age at death after age 59 months was retained. For 16 children, an age <0 was recorded; these
189 observations were omitted from analyses. Anthropometric measurements were converted to z-
190 scores for length/ height-for-age (HAZ), weight-for-age (WAZ) and weight-for-length/-height (WHZ)
191 according to the WHO Child Growth Standards (27) using the zscore06 command in Stata. Mid-upper
192 arm circumference (MUAC) was included as an absolute measure (in mm). Cut-offs for outliers were
193 decided before commencing the analyses (MUAC >300mm, WAZ >10 and <-10, length <35cm, HAZ
194 >10 and <-12 as well as WHZ >20 and <-10) and individual records were examined; consequently, 54
195 values (i.e., weight, length, or MUAC) and the corresponding Z-score were re-assigned as 'Missing'.
196 For dichotomous variables, observations with values other than 0 and 1 were recoded after
197 confirmation with study PI or deleted (2 observations re-assigned as 'Missing').

198

199 **Variable availability and definition**

200 For each variable (Table 2), a defined format was requested from each study team. In case of
201 deviation, a detailed description of the discrepancy to the requested format, definition and/or
202 assessment method was provided by the study team. The requested format and the availability of
203 data in each study is shown in Supplemental Tables S3 and S4.

204

205 *Anthropometry*

206 Moderate anthropometric deficit is defined as a Z-score <-2 and \geq -3, and a severe deficit as a Z-score
207 <-3. MUAC is included as a simple measurement (mm) and categorized as \geq 125 mm (no deficit), 115-
208 124 mm (moderate deficit) and <115 mm (severe deficit) (28). For children 0-1 month of age, we also
209 applied cut-offs at 115 mm and 105 mm, and for children 2-5 months, 120mm and 110mm,
210 respectively. Other anthropometric measures such as height-for-age difference (HAD in cm) as well

211 as changes in HAD (29), growth velocity (30), and indicators suggested in NICE guidelines (31) were
212 estimated.

213

214 *Morbidity*

215 All morbidity variables were dichotomous (yes/no). Haemoglobin concentration (g/dL), peripheral
216 arterial oxygen saturation (%), and blood glucose concentration (mmol/L) were included as
217 continuous variables. The main morbidity variables used in the primary analyses were diarrhoea,
218 acute lower respiratory tract infection (ALRTI) and severe illness. Twenty-nine studies had
219 information on diarrhoea. In 3 studies, diarrhoea was assessed by a physician or trained field worker;
220 in the other 24 studies information was from parental report of the child having any episode of
221 diarrhoea within the last 14 days. Study definitions of LRTI were used: in 5 studies, LRTI referred to
222 children diagnosed with pneumonia by a physician; in 10 studies, this referred to children reported
223 by caretakers to have pneumonia or acute lower respiratory tract infection according to WHO
224 definitions based on symptoms in the last 14 days (WHO classification of signs indicating pneumonia
225 have been revised between 2000 and 2023; studies applied classifications recommended at the time
226 of study implementation). The WHO definitions of pneumonia are sensitive and not specific and
227 include fast breathing or lower chest indrawing and a history of difficulty breathing (32-34). Severe
228 illness was defined as being admitted to hospital or having any of the WHO IMCI danger signs, i.e.,
229 the child is not able to drink or breastfeed, the child has had convulsions, and the child is lethargic or
230 unconscious (34). The IMCI danger sign of ‘vomiting everything’ was not available in any of the
231 studies. Fifteen datasets had data to generate this dichotomous severe illness variable. In addition,
232 we created a variable that tallied the total number of severe illness characteristics that were present
233 (range 0 to 4). Other morbidity variables were available (see Table 2) but not used in primary
234 analyses as they were setting-specific, only available in a minority of children in the pooled cohort
235 (<50%), anticipated to have a low sensitivity and/or specificity for mortality, or of secondary interest.

236

237 *Other risk exposures*

238 Data on “any breastfeeding” practices was available from 26 studies; in 16 studies was defined as any
239 breastfeeding within the past 24 hours reported by the caregiver. Two studies reported
240 breastfeeding practices over longer intervals, i.e., the past 3-4 days (study no. 5) and the past month
241 (study no. 4). Descriptions of this variable were not detailed in the remaining eight studies. The
242 information if the child was exclusively breastfed (i.e. no other foods or fluids than breastmilk was
243 given) was available in 16 studies. The following non-time-varying variables were also available for
244 inclusion in the primary analyses: low birthweight (LBW; birthweight <2500g; 17 studies), pre-term
245 birth (PTB; <37 completed weeks of gestation at birth; 13 studies), mother as primary caregiver (14

246 studies), maternal age (25 studies), maternal educational level at enrolment (20 studies,
247 respectively), and rural vs. urban residence (33 studies). Methods used to establish gestational age
248 varied by study and included date of last menstrual period, early ultrasound, and study-specific
249 algorithms. We categorized maternal educational level as no, primary, and secondary education and
250 above. In case of continuous variables, we used the cut-offs of 0, 1-6 and >6 years of education to
251 categorize the level accordingly.

252

253 **Principles and plans for statistical analysis**

254 The RSWG agreed that future analyses will use all-cause mortality as the main outcome, accounting
255 for repeated measurements in the individual child, clustering within studies, differences in age at
256 inclusion, as well as duration and frequency of follow-up in the individual studies. It was also agreed
257 that all analyses should be presented by age group and sex. RCTs are considered as single parallel
258 cohorts with no adjustment for effect/ no effect of intervention. As per definition, the interventions
259 in each study were randomly assigned independent of the exposures in our analyses. Any
260 intervention was therefore considered equally to any co-intervention.

261

262 The aim is to report absolute risk and risk differences (as percentage or percentage points) with 95%
263 confidence intervals in addition to relative measures, such as relative risk or odds ratio. Widely used
264 tools to improve the quality of reporting, such as the “Strengthening the Reporting of Observational
265 Studies in Epidemiology” (STROBE) statement, recommend that both relative and absolute measures
266 of association together with their level of precision are reported. The rationale for this is that relative
267 measures alone can be misleading due to differences in baseline risk (risk in the reference group)
268 obscuring the clinical significance of the effect; absolute risks can therefore be highly informative for
269 decision-making (35).

270

271 Age- and sex-specific mortality risk for each selected exposure will be estimated individually. An
272 interaction term between the exposure and age will be included in the statistical models because we
273 assumed that the association between the exposures and death are modified by age. Results will be
274 presented by the following age groups: 0-6, 6-11, 12-24, and 24-59 months. When data allows, and
275 especially for specific analyses focusing on children under 6 months of age, age groups will be further
276 refined. Further, the interaction of each anthropometric predictor with other predictors will be
277 measured on an additive scale. We will also examine the extent to which study-level characteristics,
278 such as study population type and geography modify the association between selected predictors
279 and the risk of death.

280

281 Not all studies have data on all possible risk exposures. To understand how individual studies
282 potentially drive estimates, forest plots will be used to visually represent the data characteristics of
283 studies. Sensitivity analyses will be performed to understand the effect of outliers or missingness. For
284 each analysis, a detailed statistical analysis plan will be developed and reported with findings.

285

286 **COHORT CHARACTERISTICS**

287 The pooled dataset includes 75,287 children (49.7% male) with 546,459 observations between 0 and
288 59 months of age. Twenty-nine percent of all observations were between birth and 5 months of age
289 (156,451 observations), 29% at ages 6-11 months (157,197 observations), 33% at age 12-23 months
290 (180,998 observations), and 9% at age 24-59 months (51,813 observations). Children from sub-
291 Saharan Africa make up the largest part of the pooled cohort data (n = 58,581; 78%). Studies in
292 South-East Asia included 8,756 children (12%), and 7,952 children were from other world regions
293 (11%).

294

295 Two thirds (66%, n= 49,740) of all children were enrolled into 23 studies that did not use a health
296 condition or anthropometry as an inclusion criterion (GP), 20% (n= 14,759) were enrolled in six
297 studies based on an anthropometric deficit (A-S), and 14% (n= 10,788) in four studies based on the
298 presence of an illness (I-S). Sixteen (48%) of the included studies were birth cohorts (15 of the GP,
299 one of the A-S and none of the I-S studies) with a total of 32,284 children. Of the 16 studies defined
300 as birth cohorts, 12 studies enrolled pregnant women, one study enrolled neonates within 12 hours
301 and 1 study within 48 hours after birth (study nos. 4 and 28; Table 1). For two studies, the study
302 protocol allowed for enrolment up to 17 days after birth (nos. 10 and 11; Table 1). The median length
303 of follow-up of the individual children was 15 months (IQR 7.6 to 18.4 months); follow-up time was
304 longer in the GP studies (median 17 months) compared to the A-S and I-S populations (3 and 6
305 months, respectively). The number of scheduled visits varied substantially between studies, ranging
306 from 2 to 81, with a median of 12 visits (IQR 7-19) in the pooled cohort. The median number of visits
307 differed between study population types: 14 in the GP, 6 in the A-S and 4 in the I-S studies.

308

309 Table 2 summarizes characteristics of the children and their mothers included in the pooled cohort.
310 The median age at enrolment was 3 months; 23% (n = 17,383) children were enrolled within 17 days
311 of birth, and 49% (n =36,648) children before the age of two months. GP studies enrolled children at
312 younger ages; the median age at enrolment was 1 month in the GP studies, 13 months in the A-S
313 studies, and 11 months in the I-S studies.

314

315 **Table 2:** Maternal and child characteristics in the WHO Child Mortality Risk Stratification Multi-Country Pooled
316 Cohort (WHO-CMRS)

Characteristic	N ¹	% (n)	median (IQR)
Maternal characteristics at enrolment			
Mother alive ²	42,890	99.1% (42,519) ²	
Mother primary caretaker	37,362	95.8% (35,809)	
Maternal age, years	55,416		25 (21, 30)
Maternal education	49,051		
no education		32% (15,734)	
any primary education		23% (11,306)	
any secondary or higher education		45% (21,011)	
Mother HIV positive	16,807	6% (941)	
Child's characteristics at enrolment			
Child's sex, male	75,287	49.7% (37,447)	
Child's age, months	75,287		3 (0.1, 12)
0		23.1% (17,383)	
0-5		29.8% (22,410)	
6-11		21.3% (16,056)	
12-23		18.4% (13,879)	
24-59		7.4% (5,559)	
Low birthweight, <2500 grams	29,924	13.5% (4,043)	
Median birthweight, grams	29,924		3000 (2700, 3400)
Preterm birth (reported), before 37 weeks of gestation	34,067	13.6% (4,627)	
Median (IQR) WAZ	68,157		-1.4 (-2.5, -0.4)
Median (IQR) HAZ	67,655		-1.4 (-2.4, -0.4)
Median (IQR) WHZ	65,533		-0.9 (-2.0, 0.2)
Median (IQR) MUAC mm	58,902		120 (106, 134)
Proportion underweight (WAZ <-2)	68,157	35.5% (24,188)	
Proportion stunted (HAZ <-2)	67,655	33.3% (22,517)	
Proportion wasted (WHZ <-2)	65,533	24.1% (15,765)	
Proportion low MUAC (MUAC <125mm)	58,902	60% (35,354)	
Child characteristics at any visit			
Child's age, months	546,459 ¹		10.5 (5.1, 17.0)
0-1		9.7% (52,871)	
2-5		19.0% (103,580)	
6-11		28.8% (157,197)	
12-17		21.3% (116,539)	
18-23		11.8% (64,459)	
24-59		9.5% (51,813)	
Acute diarrhea ⁴	385,332	12.5% (47,997)	
Acute lower respiratory tract infection ⁵	169,545	7.3% (12,326)	
Severe illness ⁶	315,414	12.6% (39,720)	
Any breastfeeding ⁷ at any visit			
at age 0-5 months	103,593	98.6% (102,188)	
at age 6-23 months	240,631	85.5% (205,831)	
at age 24-59 months	31,250	18.0% (5,634)	

317 ¹ Number of children.

318 ²In one study, maternal vital status was assessed at each follow-up time; of 386 mothers, 2 were dead at enrolment and 11
319 mothers died during the study.

320 ³Number of observations.

321 ⁴Any diarrheal episode reported by a physician or caregiver within the last 14 days.

322 ⁵Pneumonia or acute lower respiratory tract infection reported by a clinician or caregiver within the last 14 days.

323 ⁶Any of the following: being admitted to hospital, not being able to drink or breastfeed, having had convulsions, being
324 lethargic or unconscious.

325 ⁷Any breastfeeding within the past 14 days reported by caregiver at the study visit.

326

327 At enrolment, 33% of children were stunted (HAZ <-2), 24% were wasted by WHZ (WHZ <-2), 60%
328 were wasted by MUAC (MUAC <125mm), and 36% underweight (WAZ <-2). The anthropometric
329 status differed considerably between included studies and age-groups both at enrolment (Tables S5,
330 S6a, S6b) and throughout the study period (Figure S1). However, due to the heterogeneity in
331 inclusion criteria as well as ages at enrolment of the individual studies, it should be noted that the
332 means in each of these study population or age-groups depend on the varying contributions from
333 individual studies.

334

335 Twelve percent of all children were reported to have an episode of diarrhoea within the prior 14 days
336 (at any visit). When groups were analysed based on study population type, the prevalence of
337 diarrhoea was somewhat lower for the youngest age group (<6 months) and then relatively stable at
338 around 7% (GP children), 32% (A-S children) and 50% (I-S children), respectively. For LRTI, the
339 prevalence in GP children varied between 5% for children younger than 12 months and 1% for older
340 children. Among studies categorized as I-S, the prevalence was relatively stable at around 40%. LRTI
341 was not captured in any of the A-S studies.

342

343 The proportion of children with low birthweight was 13% in GP, 5% in A-S and 21% in I-S studies.
344 However, only one study in each of the A-S (1,112 children) and I-S cohorts (2,145 children) had
345 information on birthweight. None of the A-S studies reported data on gestational age at birth. In GP
346 studies, the prevalence of reported preterm birth was 16%, and in the I-S studies it was 4%. The
347 proportion of children currently receiving any breastfeeding reduced from almost universal practice
348 in the age group <6 months (91-99%) to less than 20% in the oldest age group (24-59 months).

349 Children enrolled in studies with inclusion criteria based on having an illness (I-S) were less likely to
350 be breastfed at any age (Figure 2). Overall, mothers had a median age of 24 years (IQR 20-28.5
351 years); 33% had no formal education, 20% had primary education and 47% had an educational level
352 of secondary school or above.

353

354 **Figure 2:** Proportion of children in the WHO Child Mortality Risk Stratification Multi-Country Pooled Cohort
355 (WRS-CMRS) being breastfed at time of follow-up according to age group and study population type; GP (not
356 enrolled by anthropometric deficit or illness), A-S (anthropometry-selected), and I-S (illness-selected)

357

358 There were 2,805 children who died within the study period. The proportion of children who died
359 differed according to how studies enrolled their population: 2.5% (n = 1,264) in the GP studies, 2.5%
360 (n = 366) in the A-S studies, and 11.1% (n = 1,201) in the I-S studies. There were 145 children (all from
361 GP studies) for whom the age at death was not documented. For these children, we estimated the
362 age at death by adding half of the median follow-up time for the original study to the age at last
363 measurement. The distribution of deaths by age (frequency) and by study population type is
364 presented in Figure 3. (Note: Age distribution of deaths also reflects age at enrolment). The median
365 observed follow-up time from enrolment to death was 1 month (IQR 0.2, 4.7 months) for the pooled
366 cohort; with 4.1 (0.8, 9.5) months in the GP studies, 1 (0.5, 2.3) month in the A-S studies and 0.3 (0.1,
367 1.4) months in the I-S studies. Among children with available information on HIV status (16,190
368 children), 653 were reported to be HIV infected (4%); of these 653 children, 140 (21%) died during
369 the study period.

370

371 **Figure 3:** Distribution deaths in the WHO Child Mortality Risk Stratification Multi-Country Pooled Cohort (WHO-
372 CMRS) by age (months) and study population type

373

374 **CONCLUDING REMARKS**

375 The aim of this cohort profile is to describe the process of assembling the WHO-CMRS cohort as well
376 as the characteristics of included studies and enrolled children. As expected, children enrolled to
377 studies based on anthropometric criteria or the presence of illness differed from those resembling
378 the general population. In addition to anthropometry and morbidity status, there were differences in
379 individual characteristics such as age at enrolment and study design features such as number and
380 frequency of scheduled study visits or assessment methods such as determining gestational age.
381 These differences will need to be taken into account in future analyses. Although the heterogeneity
382 of included studies posed some challenges in the pooling of datasets and interpretation of findings, it
383 also provides opportunities to explore specific differences these characteristics through sensitivity
384 analyses.

385

386 **FURTHER DETAILS**

387 **Data sharing statement:** All contributing research teams have acknowledged that the pooled data
388 can only be used for collaborative activities within the WHO RSWG, with no transfer of ownership.

389 Respective Principal Investigators (see Supplementary Table S1) can be contacted for access to their
390 data.

391

392 **Funding declaration:**

393 WHO gratefully acknowledges funding support from USAID to establish the pooled database and
394 conduct the main analyses.

395

396 **Contributors statement:**

397 - Conceptualization: Rajiv Bahl, Nigel Rollins

398 - Study design: Alemayehu Argaw, Rajiv Bahl, James A. Berkley, Ranadip Chowdhury, Siri
399 Kaldenbach, Nigel Rollins, Catherine Schwinger, Tor Strand, Judd Walson with input from all
400 members of the WHO-RSWG

401 - Data cleaning and preparation: Siri Kaldenbach, Catherine Schwinger, Tor Strand

402 - Statistical analysis: Catherine Schwinger, Tor Strand

403 - Data acquisition and interpretation: all

404 - Drafting the manuscript: Catherine Schwinger, Tor Strand, Siri Kaldenbach, Nigel Rollins, Jay
405 Berkley

406 - Reviewing the manuscript: all

407

408 **Disclaimer:**

409 The authors alone are responsible for the views expressed in this paper and they do not necessarily
410 represent the views, decisions, or policies of the institutions with which they are affiliated.

411

412 **References:**

413 1. Perin J, Mulick A, Yeung D, Villavicencio F, Lopez G, Strong KL, et al. Global, regional, and national
414 causes of under-5 mortality in 2000-19: an updated systematic analysis with implications for the
415 Sustainable Development Goals. *Lancet Child Adolesc Health*. 2021.

416 2. UN IGME. Levels & Trends in Child Mortality - Report 2022. New York: UNICEF, WHO, World Bank
417 Group, United Nations; 2022.

418 3. IHME. Pushing the Pace: Progress and Challenges in Fighting Childhood Pneumonia. Seattle, WA:
419 Institute for Health Metrics and Evaluation; 2014.

420 4. Collaborators GDaI. Global burden of 369 diseases and injuries in 204 countries and territories,
421 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*.
422 2020;396(10258):1204-22.

423 5. McAllister DA, Liu L, Shi T, Chu Y, Reed C, Burrows J, et al. Global, regional, and national estimates
424 of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015:
425 a systematic analysis. *Lancet Glob Health*. 2019;7(1):e47-e57.

426 6. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, de Onis M, et al. Maternal and child
427 undernutrition and overweight in low-income and middle-income countries. *Lancet*. 2013.

- 428 7. Olofin I, McDonald CM, Ezzati M, Flaxman S, Black RE, Fawzi WW, et al. Associations of suboptimal
429 growth with all-cause and cause-specific mortality in children under five years: a pooled analysis
430 of ten prospective studies. *PLoS One*. 2013;8(5):e64636.
- 431 8. Aneja S, Kumar P, Choudhary TS, Srivastava A, Chowdhury R, Taneja S, et al. Growth faltering in
432 early infancy: highlights from a two-day scientific consultation. *BMC Proc*. 2020;14(Suppl 12):12.
- 433 9. WaST TIG. Wasting and Stunting - overcoming the divide: Emergency Nutrition Network (ENN),
434 Wasting-Stunting (WaSt) Technical Interest Group (TIG); 2020 [2022/01/30]. Available from:
435 <https://www.enonline.net/ourwork/reviews/wastingstunting>.
- 436 10. Angood C, McGrath M, Mehta S, Mwangome M, Lung'aho M, Roberfroid D, et al. Research
437 priorities to improve the management of acute malnutrition in infants aged less than six months
438 (MAMI). *PLoS Med*. 2015;12(4):e1001812.
- 439 11. Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al. Maternal and child
440 undernutrition: global and regional exposures and health consequences. *Lancet*.
441 2008;371(9608):243-60.
- 442 12. Upadhyay RP, Martines JC, Taneja S, Mazumder S, Bahl R, Bhandari N, et al. Risk of postneonatal
443 mortality, hospitalisation and suboptimal breast feeding practices in low birthweight infants from
444 rural Haryana, India: findings from a secondary data analysis. *BMJ Open*. 2018;8(6):e020384.
- 445 13. Katz J, Lee AC, Kozuki N, Lawn JE, Cousens S, Blencowe H, et al. Mortality risk in preterm and
446 small-for-gestational-age infants in low-income and middle-income countries: a pooled country
447 analysis. *Lancet*. 2013;382(9890):417-25.
- 448 14. Hazel EA, Erchick DJ, Katz J, Lee ACC, Diaz M, Wu LSF, et al. Neonatal mortality risk of vulnerable
449 newborns: A descriptive analysis of subnational, population-based birth cohorts for 238 203 live
450 births in low- and middle-income settings from 2000 to 2017. *BJOG*. 2023.
- 451 15. Lawn JE, Ohuma EO, Bradley E, Idueta LS, Hazel E, Okwaraji YB, et al. Small babies, big risks: global
452 estimates of prevalence and mortality for vulnerable newborns to accelerate change and improve
453 counting. *Lancet*. 2023;401(10389):1707-19.
- 454 16. Pretorius CE, Asare H, Genuneit J, Kruger HS, Ricci C. Impact of breastfeeding on mortality in sub-
455 Saharan Africa: a systematic review, meta-analysis, and cost-evaluation. *Eur J Pediatr*.
456 2020;179(8):1213-25.
- 457 17. Victora CG, Christian P, Vidaletti LP, Gatica-Domínguez G, Menon P, Black RE. Revisiting maternal
458 and child undernutrition in low-income and middle-income countries: variable progress towards
459 an unfinished agenda. *Lancet*. 2021;397(10282):1388-99.
- 460 18. English M, Berkley J, Mwangi I, Mohammed S, Ahmed M, Osier F, et al. Hypothetical performance
461 of syndrome-based management of acute paediatric admissions of children aged more than 60
462 days in a Kenyan district hospital. *Bull World Health Organ*. 2003;81(3):166-73.
- 463 19. The Childhood Acute Illness and Nutrition Network -CHAIN. Childhood mortality during and after
464 acute illness in Africa and south Asia: a prospective cohort study. *Lancet Glob Health*.
465 2022;10(5):e673-e84.
- 466 20. Kirolos A, Blacow RM, Parajuli A, Welton NJ, Khanna A, Allen SJ, et al. The impact of childhood
467 malnutrition on mortality from pneumonia: a systematic review and network meta-analysis. *BMJ*
468 *Glob Health*. 2021;6(11).
- 469 21. GBD 2016 Diarrhoeal Disease Collaborators. Estimates of the global, regional, and national
470 morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the
471 Global Burden of Disease Study 2016. *Lancet Infect Dis*. 2018;18(11):1211-28.
- 472 22. Wen B, Brals D, Bourdon C, Erdman L, Ngari M, Chimwezi E, et al. Predicting the risk of mortality
473 during hospitalization in sick severely malnourished children using daily evaluation of key clinical
474 warning signs. *BMC Med*. 2021;19(1):222.
- 475 23. Karunaratne R, Sturgeon JP, Patel R, Prendergast AJ. Predictors of inpatient mortality among
476 children hospitalized for severe acute malnutrition: a systematic review and meta-analysis. *Am J*
477 *Clin Nutr*. 2020;112(4):1069-79.
- 478 24. Keezer MR, Sisodiya SM, Sander JW. Comorbidities of epilepsy: current concepts and future
479 perspectives. *Lancet Neurol*. 2016;15(1):106-15.

- 480 25.Valderas JM, Starfield B, Sibbald B, Salisbury C, Roland M. Defining comorbidity: implications for
481 understanding health and health services. *Ann Fam Med*. 2009;7(4):357-63.
- 482 26.World Bank. World Bank Country and Lending Groups 2022 [Available from:
483 [https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups)
484 [lending-groups](https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups).
- 485 27.WHO. WHO Child Growth Standards: Methods and development: Length/height-for-age, weight-
486 for-age, weight-for-length, weight-for-height and body mass index-for-age. Geneva: World Health
487 Organization Multicentre Growth Reference Study; 2006.
- 488 28.WHO. Guideline: updates on the management of severe acute malnutrition in infants and
489 children. Geneva: World Health Organization; 2013.
- 490 29.Leroy JL, Ruel M, Habicht JP, Frongillo EA. Using height-for-age differences (HAD) instead of
491 height-for-age z-scores (HAZ) for the meaningful measurement of population-level catch-up in
492 linear growth in children less than 5 years of age. *BMC Pediatr*. 2015;15:145.
- 493 30.WHO. WHO Child Growth Standards- Growth velocity based on weight, length and head
494 circumference: methods and development. Geneva: World Health Organization; 2009.
- 495 31.Gonzalez-Viana E, Dworzynski K, Murphy MS, Peek R, Committee G. Faltering growth in children:
496 summary of NICE guidance. *BMJ*. 2017;358:j4219.
- 497 32.WHO. Handbook: IMCI integrated management of childhood illness. Geneva: World Health
498 Organization; 2005.
- 499 33.WHO. Integrated Management of Childhood Illness (IMCI): Chart Booklet. Geneva, Switzerland:
500 WHO; 2014.
- 501 34.WHO. Integrated Management of Childhood Illness: management of the sick young infant aged up
502 to 2 months. IMCI chart booklet. Geneva: WHO; 2019.
- 503 35.Rothman KJ. *Epidemiology - An Introduction*. 2nd edition ed. New York: Oxford University Press;
504 2012.
- 505

Contacted for data sharing

31 teams

2 teams: no response

2 teams: no signed agreement

Signed data pooling agreement

27 teams

35 datasets

Excluded using inclusion criteria

- 1 study: no weight data

- 1 study: cross-sectional study

Received data for pooling

25 teams

33 datasets

Data-pooling

Pooled dataset572 491 datapoints /
76 375 children**Data cleaning**

- Rows dropped from dataset

- Re-structuring: 1 518

- Duplicates: 949

- Missing sex: 1 047

- Missing age: 22 342

- Age >60 months: 160

- Age <0 months: 16 (data error)

- Single values set to missing

- Anthropometry: 54

- Diarrhea: 1

- Fever: 1

Analysis pooled dataset546 459 datapoints /
75 287 children



