

1 **Microbiota and Health Study: a prospective cohort of respiratory and diarrheal infections**  
2 **and associated risk factors in Bangladeshi infants under two years**

3

4 Karine Vidal <sup>1</sup>, Shamima Sultana <sup>2</sup>, Alberto Prieto Patron <sup>1</sup>, Aristeia Binia<sup>1</sup>, Mahbubur Rahman<sup>2</sup>,  
5 Iztiba Mallik Deeba<sup>2</sup>, Harald Brüßow <sup>1§</sup>, Olga Sakwinska <sup>1°</sup>, Shafiqul Alam Sarker <sup>2</sup>

6

7 °: Corresponding author

8

9 **Addresses**

10 <sup>1</sup> Nestlé Research, Nestlé Institute of Health Sciences, Lausanne, Switzerland

11 <sup>2</sup> International Center for Diarrheal Disease Research (icddr,b), Dhaka, Bangladesh

12 § Current address: University of Leuven, Leuven, Belgium

13 **Email addresses**

14 karine.vidal@rdls.nestle.com; shamima@icddrb.org; alberto.prietopatron@rdls.nestle.com;

15 aristeia.binia@rdls.nestle.com; haraldbruessow@yahoo.com; olga.sakwinska@rdls.nestle.com;

16 sasarker@icddrb.org

17

## 18 **Abstract**

### 19 **Background**

20 Early childhood respiratory and diarrheal infections are major causes of morbidity and mortality  
21 worldwide. There is a need to further assess the epidemiology through prospective and  
22 community-based studies to gain key insights that could inform preventative measures to reduce  
23 the risk of infectious disease in this vulnerable population. We aimed to analyze the burden and  
24 determinants of acute respiratory infection (ARI) and diarrhea episodes affecting infants during  
25 their first 2 years of life with state-of-the-art molecular technologies.

### 26 **Methods**

27 The ARI and diarrhea episodes were prospectively collected in a community-based, longitudinal  
28 cohort of infants (n=267) from birth to 2 years of life in Bangladesh. Women were recruited during  
29 the third trimester of pregnancy. Demographic, socioeconomic, and environmental information on  
30 the households was recorded. Nasopharyngeal and fecal samples were collected during regular  
31 scheduled visits from mother-infant pairs, and also from the infants during unscheduled visits for  
32 reported illnesses. New generation sequencing methods will be utilized to determine microbiota  
33 composition and function, supplemented by the state-of-the-art multiplex molecular detection  
34 technology for a wide range of bacterial and viral pathogens.

### 35 **Discussion**

36 This study sought to assess the epidemiology of both respiratory and gastrointestinal illnesses  
37 during the first 2 years of life in children from a peri-urban community of Dhaka, Bangladesh.  
38 Characteristics of the mothers, as well as birth characteristics of infants enrolled in the Microbiota  
39 and Health Study are presented here. We will determine any potential association between

40 microbiota composition and the abovementioned illnesses, and also examine the influence of  
41 known and hypothesized risk factors on the occurrence of infections. Such putative factors include  
42 environmental, socioeconomic, maternal, clinical, and selected genetic factors, namely the  
43 variation of the fucosyltransferase genes (*FUT2* and *FUT3*) of mothers and infants. This study will  
44 add to current knowledge about these early childhood infectious diseases, and will provide data to  
45 generate hypotheses for the development of nutritional approaches to be used as prevention  
46 strategies.

47 **Trial registration:** The study was retrospectively registered at [clinicaltrials.gov](https://clinicaltrials.gov) as NCT02361164  
48 (February 11, 2015).

49

## 50 **Keywords**

51 Respiratory tract infections; Diarrhea; Cohort study; Infants, Risk factors

## 52 **Background**

53 Despite notable progress made towards improving child health in the past several decades (1, 2)  
54 infectious diseases remain highly prevalent, with respiratory and diarrheal infections representing  
55 the highest burden (3). Worldwide in 2016, diarrhea and pneumonia caused over 1.4 million  
56 deaths, and approximately 80% of these deaths occurred in children under the age of two (2). Non-  
57 specific preventive measures, such as nutrition interventions and strategies to improve water,  
58 sanitation, and hygiene have resulted in a large decline in diarrhea in children under five years of  
59 age (4). However, this decrease was less pronounced for acute respiratory infections (ARI) (5)  
60 especially in low resource settings, leading to increased importance of ARI morbidity and  
61 mortality relative to diarrhea. The worldwide estimates of ARI prevalence vary widely. According  
62 to the Demographic Health Surveys (DHS) in low and middle-income countries, the percentage of  
63 reported cases of ARI symptoms in two weeks prior to the survey in children under age of five  
64 were as high as 10.0% in Haiti in 2015-2016; 5.4% in Bangladesh in 2014 and as low as 1.4% in  
65 Armenia in 2015-16 (6). It has to be noted that the comparisons among countries are challenging,  
66 mainly due to seasonality and the fact that DHS data are not available for the same years for all  
67 the locations.

68 Understandably, the majority of efforts are focused on lower respiratory tract infections (7),  
69 because of their high morbidity. Yet milder upper respiratory tract infections are approximately  
70 ten-fold more frequent (8, 9), thus representing substantial burden. Approximately two thirds of  
71 children below five years of age hospitalized in Bangladesh suffered from respiratory disorders.  
72 Economic costs of ARI episodes in lower income countries is high relative to the median per capita  
73 income. For example, a multi-site study in Northern India showed that among children aged under  
74 five years, the median direct cost of ARI represented over 5% or 10% of the family annual median

75 income depending if treated in public or private institutions respectively (10). Moreover, these  
76 estimates do not take into account the burden of increased risk of later health problems. For  
77 example, bronchiolitis in infancy was associated with increased prevalence of recurrent wheeze,  
78 asthma and allergy in early adulthood (11, 12).

79 Environmental risk factors for respiratory infections are well recognized, however the attributable  
80 fraction of variation is small (below 10%) [e.g. (9)], indicating that identification of other factors  
81 responsible for the differential susceptibility to respiratory infections is needed to optimize efforts  
82 aimed at prevention and treatment of ARI (2, 5). Emerging evidence indicates that commensal  
83 respiratory microbiota is an important modulator of susceptibility to respiratory infections (13, 14)  
84 as well as wheezing and asthma later in life (15). The commensal microbes could interact with  
85 respiratory pathogens in nasopharynx and modulate their infective potential (16, 17). It is also  
86 plausible that the respiratory microbiota play key roles in the immune maturation and maintenance  
87 of homeostasis as observed for gut commensals and thus impact respiratory health (16, 18).  
88 Interestingly, respiratory infections may negatively impact gut health indicating that so called gut-  
89 lung axis is bi-directional (19, 20). For example, in a model of acute influenza, airway disease  
90 affected the intestinal microbiota, resulting in higher susceptibility to bacterial pathogen invasion  
91 (21). Clinically, respiratory infections and diarrhea are often linked (22), suggesting the  
92 importance of investigating the two simultaneously (23).

93 To gain insights into the risk factors of ARI and diarrhea, we have conducted a prospective  
94 longitudinal study in Nandipara, a peri-urban community in Dhaka, Bangladesh. We have  
95 evaluated the role of both gastrointestinal and respiratory microbiota, including facultative  
96 pathogens, in susceptibility to ARI and diarrhea during the first two years of life. In addition, we

97 monitored known and hypothesized risk factors, including environmental, maternal,  
98 socioeconomic, clinical and selected genetic factors.

99

100

## 101 **Methods**

### 102 **Study objectives**

103 We conducted a prospective, population-based, longitudinal birth cohort study to evaluate the  
104 association between the microbiota (nasopharyngeal and stool) and the occurrence of respiratory  
105 and gastrointestinal infections in infants during their first two years of life. To this end, we  
106 surveyed the occurrence of respiratory infections and investigated the nasopharyngeal microbiota,  
107 as well as presence of respiratory pathogens. To address the same question in relation to diarrhea,  
108 we set out to determine the association between stool microbiota, intestinal inflammation  
109 biomarkers, presence of specific intestinal pathogens, and occurrence of diarrhea. A closely linked  
110 objective was to determine the influence of other factors on the risk of infections, such as  
111 environmental, maternal and specific genetic factors, namely the variation of the  
112 fucosyltransferase genes (*FUT2* and *FUT3*) of mothers and infants.

113

### 114 **Study location**

115 The study was conducted in Nandipara, a peri-urban community of Dhaka, the capital and largest  
116 city in Bangladesh (**Figure 1**). The region experiences a subtropical monsoon climate  
117 characterized by moderately warm temperatures (monthly average within 23 to 34°C during our  
118 trial) (24), high humidity (38 to 87%) and wide seasonal variations in rainfall, in three distinct  
119 seasons: a cool dry winter (November-February), a pre-monsoon hot season (March-May), and a

120 rainy monsoon (June–October) (25). The crude birth rate in urban area is 20.8 per 1000 population  
121 (26). The community of Nandipara was chosen for its close proximity (~8 km) to the International  
122 Center for Diarrheal Disease Research (icddr,b) campus, and presence of an icddr,b-run health  
123 outpost with experience in supporting research activities [e.g. (27)]. Nandipara covers an area of  
124 around 4.5 km<sup>2</sup> and with population density of 11500 per km<sup>2</sup> (Dr Sarker, unpublished 2016  
125 survey). All study participants lived within 4 km distance from the health outpost.

126

### 127 **Study participants**

128 Recruitment occurred from April 2013 to October 2014; the newborn infants were followed until  
129 two years of age (*i.e.* until October 2016). Antenatal enrolment of women and follow-up of the  
130 study infants are shown in **Figure 2**. Pregnant women (between 16 to 28 weeks of gestation) were  
131 identified by house-to-house visits and at the icddr,b community clinic by trained field health  
132 workers and 300 women were screened. Women who fulfilled the eligibility criteria and signed  
133 the informed written consent (n=285) were enrolled at 32 ± 1 weeks of gestation. Gestational age  
134 and parity of current pregnancy was confirmed by ultrasonography. For enrolment, pregnant  
135 women needed to be inhabitants of Nandipara, between 18–35 years, with any parity but with  
136 history of healthy and uncomplicated pregnancy, and having current singleton pregnancy, healthy  
137 and uneventful progression of pregnancy and expected to have a normal pregnancy related  
138 outcome. Exclusion criteria included Rhesus negative blood group; body mass index (BMI) of less  
139 than 18.5 or more than 35, history of diabetes, hypertension; and antibiotic treatment within 3  
140 weeks prior to this study.

141

142

143 **Data collection**

144 Data and sample collection were performed by the team consisting of health workers, research  
145 assistants and study nurse supervised by a physician. Demographic, socioeconomic, environmental  
146 and mother's health characteristics were documented via personal interview using a pre-designed  
147 epidemiological questionnaire comparable to the Bangladesh Demographic and Health Survey  
148 (BDHS). Household and mother characteristics of the study population are shown in **Table 1 and**  
149 **Table 2**, and birth characteristics of the infants in **Table 3**. Infant anthropometric data and feeding  
150 practice (exclusivity of breastfeeding and introduction of complementary foods) were recorded at  
151 scheduled visits, which were done at one month of age and subsequently every two months during  
152 the first year of life, and then quarterly, at 15, 18 and 24 months of age. Weight was recorded using  
153 digital weighting scale nearest 10 g (Seca, United Kingdom) and the length was measured with a  
154 manual length board to the nearest 1 cm (icddr,b-designed and validated). Mid upper arm  
155 circumference (MUAC) was recorded with a standard tape measure to the nearest 0.1 cm. In  
156 addition to the regular scheduled visits, weekly home visits were done by the study nurse  
157 supervised by a physician for the active surveillance of respiratory and gastrointestinal infections.  
158 Furthermore, mothers were asked to contact the nurse whenever her infant experienced symptoms  
159 like cough, running/blocked nose, fever, ear discharge, rapid breathing or refusal of feeding, or  
160 passage of loose stool. In such events, extra home visit was performed at the home, or mother was  
161 asked to present herself and the child to the health outpost. Each ARI symptoms were recorded as  
162 adverse events (AE) with the precise starting and ending dates. Concomitant medications  
163 (including antibiotic treatment), AE and serious adverse events (SAE) were documented in a  
164 standard electronic case report form by the medical officer.

165



## 166 **Case definitions**

167 An episode of ARI was defined as a period with one or more of the following ARI symptoms:  
168 cough, runny nose, nasal congestion, ear discharge, rapid breathing or refusal of feeding. In line  
169 with other studies (28, 29), a new episode was defined as an episode starting after 7 symptom-free  
170 days from a previous episode. Each ARI episode was further categorized as either upper or lower  
171 respiratory tract infections (URTI and LRTI, respectively). An URTI was defined as any or  
172 combination of the following respiratory symptoms in the presence or the absence of fever: cough,  
173 runny nose, nasal congestion, ear discharge, and breathing difficulty; and an LRTI was defined as  
174 the presence of cough or respiratory difficulty plus any of the following: fever, fast breathing,  
175 lower chest in-drawing, lung findings in auscultation (rales or rhonchi). A diarrhea episode was  
176 defined according to the guidelines outlined by the World Health Organization (WHO), as the  
177 passage of three or more loose or liquid stools per day. A new diarrhea episode was defined when  
178 a period of symptom-free of more than 48 hours was observed.

179

## 180 **Sample collection and planned analysis**

181 At enrollment, a low vaginal swab sample was collected from each mother to assess vaginal  
182 microbiota composition. As well, a blood sample (5ml) was collected for routine biochemistry  
183 (such as serum protein and hemoglobin levels), blood group including Rhesus factor and to enable  
184 future analyses (e.g. antibody titers against specific pathogens). At delivery, a cord blood sample  
185 (5 ml) was collected for future determination of the neonate nutritional status (e.g. retinol binding  
186 protein and pre-albumin) and passive transfer of immunity from mother to infant (e.g. antibodies  
187 to *S. pneumoniae* and *H. influenzae* pathogens). Nasopharyngeal and stool samples were collected  
188 from the mother and her infant at birth or within 1-2 days post-delivery, and during scheduled

189 visits. Whenever possible, a nasopharyngeal or stool sample was collected when infant  
190 experienced ARI or diarrhea, respectively. Nasopharyngeal samples were collected using flocked  
191 swabs for adult or pediatric swabs as per manufacturer's instruction (Copan Diagnostics, Italy).  
192 The determination of four common facultative respiratory pathogens was performed at iccdr,b by  
193 bacterial culture using standard clinical diagnostics methods. The remaining part of the sample  
194 was frozen and stored for future analysis of viral and bacterial respiratory pathogens colonization  
195 and nasopharyngeal microbiota composition. Stool were collected for future determination of  
196 pathogen colonization, gut microbiota composition and metabolome, as well as analysis of  
197 intestinal inflammatory biomarkers. Saliva samples were collected from both the mother and the  
198 infants using Oragene DNA kits (OG-500 and OG-575, respectively) to allow for future DNA  
199 extraction and determination of *FUT2* and *FUT3* polymorphisms. All biological samples were  
200 initially stored at -20°C and subsequently transferred to -80°C for shipment on dry ice to Nestlé  
201 Research where they were stored at -80°C until analysis.

202

### 203 **Participant management**

204 The enrolled mothers benefited from the antenatal medical checkups (including uro-genital health  
205 assessment, blood pressure, and blood group and hemoglobin measurements), medical advice, and  
206 basic medication such as iron and vitamin supplement, if required. Mothers with severe anemia  
207 and negative Rh factor were referred to public hospitals, including Dhaka Medical College  
208 Hospital (DMCH), Dhaka or Bangabaondhu Sheikh Mujib Medical University (BSMMU), for  
209 management and advice. Enrolled children benefited from regular follow-up of growth and  
210 development, as well as prompt, free and appropriate treatment of respiratory and diarrheal  
211 infections, such as antalgic, antihistamine, oral rehydration solution (ORS) with zinc, and

212 antibiotics. If necessary, referral to other hospital were made. The children were vaccinated  
213 according to the UNICEF guided Expanded Program for Immunization run by the Government of  
214 Bangladesh.

215

## 216 **Statistical considerations**

217 The present study is an exploratory study with a broad range of variables of interest with little or  
218 no pre-existing data available at the time the study was designed to guide a priori statistical  
219 analysis. Nevertheless, based on the clinical expertise in the target community and published  
220 information, we have attempted sample size calculation based on the incidence of ARI in early  
221 childhood and tentatively linked it to the estimates of colonization with facultative respiratory  
222 pathogens. Prevalence of colonization in the population was estimated to approximately be 30%  
223 and the average number of ARI episodes per infant was five (7, 30-32). No study disaggregated  
224 the number of ARI episodes by colonization with respiratory pathogens, we thus assumed that  
225 colonized infants suffer 25-33% more episodes than non-colonized infants. While the variation of  
226 ARI incidence was unknown, we assumed it as high as 80% of the group mean. We estimated the  
227 minimum sample size of each group to detect statistically significant difference in ARI episodes  
228 by 30% between colonized and non-colonized groups. Assuming 5% significance level and a  
229 statistical power of 90%, 225 subjects were required. With expected 10% dropout rate, 250 infants  
230 had to be included, therefore 300 pregnant mothers were recruited in this study. Data analysis was  
231 conducted using full analysis set (FAS) without imputation of missing data. The FAS includes  
232 pairs of mother and infant who satisfied criteria of target population, *i.e.* mothers' eligibility  
233 criteria and infants who were born alive and have at least one post-baseline visit available. When  
234 the date of completion/discontinuation was equal to the date of last visit, the subject was classified

235 as a dropout at this visit; if the date of completion/discontinuation was later than the date of last  
236 visit, the subject was classified as a dropout at next visit. Weight-for-age, height-for-age and  
237 weight-for-length z-scores will be calculated according to WHO child growth standards (33).

238

239

## 240 **Discussion**

241 Here we present the study protocol of a prospective, community-based, longitudinal cohort study  
242 with detailed follow up of 267 infants during the first two years of life. This study was designed  
243 to gather data on infant health with focus on respiratory and gastrointestinal infections, collect  
244 samples from infants as well as their mothers, and assemble information about numerous potential  
245 risk factors and confounders.

246

### 247 **Study population**

248 The participants of the present study inhabited peri-urban community of Nandipara. The majority  
249 of the households included in the study had monthly income higher than 7000 taka, corresponding  
250 to approximately 90 dollars or 390 international purchasing power parity (PPP) dollars between  
251 2013 and 2016. The use of a socio-demographic questionnaire similar to the BDHS 2014 allows a  
252 direct comparison of key socioeconomic, sanitation and health indicators recorded in the present  
253 cohort with those reported at national level. It revealed that the population included in the present  
254 cohort is similar to the third and fourth wealth quintile of Bangladeshi urban households. Housing  
255 corresponds to median income families in Bangladesh, with houses built in the majority with  
256 cement floor (86%) and a single room (70%). Appropriate drinking water treatment method (57%),  
257 handwashing inside the household (16%), washing hands with soap after defecating (73%), before  
258 eating (22%) and sharing toilet facility with other households (73%) were in line with BDHS 2014

259 findings (26). However, some of the factors considered as important risk contributors to childhood  
260 infections, such as indoor pollution from cooking were less frequent in the current population. For  
261 example, 95% of households used gas for cooking compared to 82% using solid fuel with 50%  
262 using wood in the BDHS 2014 (26).

263 In this cohort study, 26% of the children were delivered by Caesarean section, which is very close  
264 to the 23% national average, and reflecting the very rapid increase from 4% in 2004 (26).  
265 According to BDHS 2014 (26), 62% of births were delivered at home, a figure similar to 55%  
266 observed in the present study.

267

## 268 **Study strengths**

269 Studies of both respiratory and diarrheal infections during the first two years of life conducted in  
270 community settings are rare (23), and to our knowledge our study is the first one performed in a  
271 low income country, particularly in combination with commensal microbiota determination.  
272 Active surveillance of infections conducted by community-based team supervised by a physician  
273 is a major strength of our study, resulting in comprehensive reporting of infections. Restricted  
274 geographic area of the surveyed community facilitated communication between families and the  
275 study team, and quality medical care received by the study participants provided further incentive  
276 to report cases of infections. In addition, the recruitment period extending over 18 months will  
277 help to disentangle the effect of the season and of infant age on the susceptibility to infections. The  
278 antenatal recruitment of pregnant women and collection of antenatal samples also constitute study  
279 strength, as it permits to assess pre-natal maternal factors known or hypothesized to contribute to  
280 the risk of childhood infections. The drop-out rate from the study was relatively low, with 90%  
281 and 82% retention at one year and two years of life, respectively. The main cause of drop-out was

282 migration out from Nandipara, and thus unlikely to introduce major bias. Another important  
283 strength of the study is the comprehensive analysis of biological samples, including state of the art  
284 molecular detection of infectious agents, covering virus and bacteria, coupled with description of  
285 commensal microbiome, metabolome and gut health biomarkers.

286

## 287 **Study limitations**

288 The recruitment was limited to one site and peri-urban settings potentially limiting the conclusions  
289 of our study, although it has to be noted that this feature allowed high quality of active surveillance  
290 of the participants. To focus on mothers and infants who were healthy at baseline and presented  
291 no specific risk factors, we excluded women younger than 18 years and with BMI below 18  
292 recorded in the third trimester of pregnancy. However, this has likely led to exclusion of  
293 individuals with most severe health problems. Thus, when interpreting the future findings it has to  
294 be considered that the conclusions should be applied to this group and not to the general  
295 population. In the absence of chest radiography to diagnose pneumonia in this study, the distinction  
296 between upper and lower respiratory tract infections will be based on clinical assessment,  
297 potentially leading to some misclassification. Likewise, the precision of the estimate of gestational  
298 age in the present study, and resulting prematurity rate is limited. Our estimates were based on the  
299 ultrasound performed in the third trimester, as the recall of Last Menstrual Period was unreliable  
300 in this population according to the clinical practice. Nevertheless, the records of gestational age  
301 conformed to expectations (min: 32.9 weeks, max: 44.7 weeks; average: 39.5), with the exception  
302 of low prematurity rate of 8 compared to 14% WHO 2012 estimate for Bangladesh (34). However,  
303 late pregnancy enrolment precluded the observation of early preterm births, and the exclusion  
304 criteria eliminated some of the most risky pregnancies potentially explaining the observed lower

305 rate of prematurity. In the present study, 16 of infants were classified as low birth weight (*i.e.*  
306 <2500 g), less than Bangladesh average of 23% reported in the National Low Birth Weight Survey  
307 conducted in 2015 (35). However, we have noted that for 14% infants, the birth weight was  
308 recorded as exactly 2500 g, suggesting inadvertent bias that commonly occurs in surveys,  
309 especially of populations with high prevalence of low birth weight (36). Future analyses will need  
310 to address this bias, as outlined by Blanc and colleagues (36).

311 **List of abbreviations**

312 AE: adverse event

313 ARI: acute respiratory infections

314 BDHS: Bangladesh Demographic and Health Survey

315 BMI: body mass index

316 FAS: full analysis set

317 FUT: fucosyltransferase

318 Icdhr, b: International Centre for Diarrheal Disease Research, Bangladesh

319 LRTI: lower respiratory tract infections

320 URTI: upper respiratory tract infections

321 WHO: World Health Organization



## 322 **Declarations**

### 323 *Ethics approval and consent to participate*

324 The local independent Institutional Review Board (IRC) comprising Research Review Committee  
325 (RRC) and Ethical Review Committee (ERC) approved the study protocol (Protocol ID: PR-  
326 12051) (September 2012), its amendments (December 2013, September 2014 and March 2017),  
327 and the related study protocol (Protocol ID: PR-15098) (December 2015). The study was designed  
328 and conducted in accordance with the declaration of Helsinki and in full conformity with the  
329 principles of the Belmont Report (April, 1979). The study was performed in compliance with the  
330 guidelines on Good Clinical Practice provided by International Conference on  
331 Harmonization/Commission on European Communities (Brussels 1991). Trained staff informed  
332 all the potential participants about the purposes, duration, procedures/methods, risks and benefits  
333 of the study, as well as the terms of confidentiality. Prior to the start of the study, all participants  
334 signed an informed written consent.

335

### 336 *Availability of data and materials*

337 The datasets used and/or analyzed during the current study are available from the corresponding  
338 author on reasonable request.

339

### 340 *Conflict of interest statement*

341 KV, PPA, AB, HB and OS were employees of Nestec Ltd. (now the Société des Produits Nestlé  
342 S.A.) when this study was conducted; SS and SAS received research funding from Nestec Ltd.  
343 The other authors report no conflict of interest. The study was internally funded by Nestec Ltd.  
344 (Vevey, Switzerland).

345

346

347 *Author's contributions*

348 HB, SAS and OS designed the study. SS and SAS were responsible for participant recruitment and  
349 follow up and the collection of samples. KV, PPA, AB, HB, MR, IMD and OS conducted data  
350 analysis and interpretation. KV and OS wrote the manuscript. SS, PPA, HB, and SAS contributed  
351 to writing of the manuscript. All authors read and approved the final manuscript.

352

353 *Acknowledgments*

354 We are grateful to all the families of the infants who participated in the study, as well as the local  
355 staff, namely the field health workers, the community nurses, the research assistants (Nur Jahan  
356 Akther, Ishrat Akanda, Nasreen Sultana, and Sakina Begum) and the medical officer (Zennat  
357 Arfin) who made this study possible. We acknowledge Amelie Goyer for the clinical project  
358 management as well as Maya Shevlyakova and Pavla Kadlecova for their expert statistical support.  
359 We are also thankful to Christèle Close for providing medical writing services. We would like to  
360 thank Nashmil Emami, Elisabeth Forbes-Blom and Armin Alaedini for helpful comments on the  
361 manuscript.

## 362 **References**

- 363 1. Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national  
364 causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated  
365 systematic analysis. *Lancet* (London, England). 2015;385(9966):430-40.
- 366 2. UNICEF. United Nations Inter-agency Group for Child Mortality Estimation (UN IGME).  
367 'Levels and Trends in Child Mortality: Report 2017, Estimates Developed by the UN Inter-agency  
368 Group for Child Mortality Estimation', United Nations Children's Fund, New York. 2017.
- 369 3. Walker CLF, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, et al. Global burden of  
370 childhood pneumonia and diarrhoea. *Lancet* (London, England). 2013;381(9875):1405-16.
- 371 4. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM, Jr. Water, sanitation,  
372 and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and  
373 meta-analysis. *The Lancet Infectious diseases*. 2005;5(1):42-52.
- 374 5. Leung DT, Chisti MJ, Pavia AT. Prevention and Control of Childhood Pneumonia and  
375 Diarrhea. *Pediatric clinics of North America*. 2016;63(1):67-79.
- 376 6. ICF 2012. The DHS Program STATcompiler. Funded by USAID [Available from:  
377 <http://www.statcompiler.com>.
- 378 7. Rudan I, Boschi-Pinto C, Biloglav Z, Mulholland K, Campbell H. Epidemiology and  
379 etiology of childhood pneumonia. *Bulletin of the World Health Organization*. 2008;86(5):408-16.
- 380 8. Kusel MM, de Klerk NH, Holt PG, Keadze T, Johnston SL, Sly PD. Role of respiratory  
381 viruses in acute upper and lower respiratory tract illness in the first year of life: a birth cohort  
382 study. *The Pediatric infectious disease journal*. 2006;25(8):680-6.
- 383 9. Vissing NH, Chawes BL, Rasmussen MA, Bisgaard H. Epidemiology and Risk Factors of  
384 Infection in Early Childhood. *Pediatrics*. 2018;141(6).

- 385 10. Peasah SK, Purakayastha DR, Koul PA, Dawood FS, Saha S, Amarchand R, et al. The cost  
386 of acute respiratory infections in Northern India: a multi-site study. *BMC public health*.  
387 2015;15:330.
- 388 11. Fauroux B, Simoes EAF, Checchia PA, Paes B, Figueras-Aloy J, Manzoni P, et al. The  
389 Burden and Long-term Respiratory Morbidity Associated with Respiratory Syncytial Virus  
390 Infection in Early Childhood. *Infectious diseases and therapy*. 2017;6(2):173-97.
- 391 12. Sigurs N, Aljassim F, Kjellman B, Robinson PD, Sigurbergsson F, Bjarnason R, et al.  
392 Asthma and allergy patterns over 18 years after severe RSV bronchiolitis in the first year of life.  
393 *Thorax*. 2010;65(12):1045-52.
- 394 13. Biesbroek G, Tsvitvadze E, Sanders EA, Montijn R, Veenhoven RH, Keijser BJ, et al.  
395 Early respiratory microbiota composition determines bacterial succession patterns and respiratory  
396 health in children. *American journal of respiratory and critical care medicine*. 2014;190(11):1283-  
397 92.
- 398 14. Bosch A, de Steenhuijsen Piters WAA, van Houten MA, Chu M, Biesbroek G, Kool J, et  
399 al. Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health  
400 Consequences. A Prospective Cohort Study. *American journal of respiratory and critical care*  
401 *medicine*. 2017;196(12):1582-90.
- 402 15. Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N, et al. The infant nasopharyngeal  
403 microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell*  
404 *host & microbe*. 2015;17(5):704-15.
- 405 16. Man WH, de Steenhuijsen Piters WA, Bogaert D. The microbiota of the respiratory tract:  
406 gatekeeper to respiratory health. *Nature reviews Microbiology*. 2017;15(5):259-70.

- 407 17. Vissers M, de Groot R, Ferwerda G. Severe viral respiratory infections: are bugs bugging?  
408 Mucosal immunology. 2014;7(2):227-38.
- 409 18. de Steenhuijsen Piters WA, Sanders EA, Bogaert D. The role of the local microbial  
410 ecosystem in respiratory health and disease. Philosophical transactions of the Royal Society of  
411 London Series B, Biological sciences. 2015;370(1675).
- 412 19. Marsland BJ, Trompette A, Gollwitzer ES. The Gut-Lung Axis in Respiratory Disease.  
413 Annals of the American Thoracic Society. 2015;12 Suppl 2:S150-6.
- 414 20. Taylor SL, Wesselingh S, Rogers GB. Host-microbiome interactions in acute and chronic  
415 respiratory infections. Cellular microbiology. 2016;18(5):652-62.
- 416 21. Yildiz S, Mazel-Sanchez B, Kandasamy M, Manicassamy B, Schmolke M. Influenza A  
417 virus infection impacts systemic microbiota dynamics and causes quantitative enteric dysbiosis.  
418 Microbiome. 2018;6(1):9.
- 419 22. Dilantika C, Sedyaningsih ER, Kasper MR, Agtini M, Listiyaningsih E, Uyeki TM, et al.  
420 Influenza virus infection among pediatric patients reporting diarrhea and influenza-like illness.  
421 BMC infectious diseases. 2010;10:3.
- 422 23. Lambert SB, Ware RS, Cook AL, Maguire FA, Whiley DM, Bialasiewicz S, et al.  
423 Observational Research in Childhood Infectious Diseases (ORChID): a dynamic birth cohort  
424 study. BMJ open. 2012;2(6).
- 425 24. World Weather Online. Nandipara Historical Weather [Available from:  
426 <https://www.worldweatheronline.com/lang/en-au/nandipara-weather-history/bd.aspx>.
- 427 25. Banglapedia National Encyclopedia of Bangladesh. Season [Available from:  
428 <http://en.banglapedia.org/index.php?title=Season>.

- 429 26. National Institute of Population Research and Training (NIPORT). Mitra and Associates,  
430 and ICF International 2016. *Bangladesh Demographic and Health Survey 2014*. Dhaka,  
431 Bangladesh, and Rockville, Maryland, USA:  
432 NIPORT, Mitra and Associates, and ICF International.
- 433 27. Sarker SA, Mahmud H, Davidsson L, Alam NH, Ahmed T, Alam N, et al. Causal  
434 relationship of *Helicobacter pylori* with iron-deficiency anemia or failure of iron supplementation  
435 in children. *Gastroenterology*. 2008;135(5):1534-42.
- 436 28. Budge PJ, Griffin MR, Edwards KM, Williams JV, Verastegui H, Hartinger SM, et al. A  
437 household-based study of acute viral respiratory illnesses in Andean children. *The Pediatric*  
438 *infectious disease journal*. 2014;33(5):443-7.
- 439 29. Koch A, Sorensen P, Homoe P, Molbak K, Pedersen FK, Mortensen T, et al. Population-  
440 based study of acute respiratory infections in children, Greenland. *Emerging infectious diseases*.  
441 2002;8(6):586-93.
- 442 30. Kamal N, Joarder AH, Chowdhury AA, Khan AW. Prevalence of chronic suppurative otitis  
443 media among the children living in two selected slums of Dhaka City. *Bangladesh Medical*  
444 *Research Council bulletin*. 2004;30(3):95-104.
- 445 31. Roy E, Hasan Kh Z, Haque F, Siddique AK, Sack RB. Acute otitis media during the first  
446 two years of life in a rural community in Bangladesh: a prospective cohort study. *Journal of health,*  
447 *population, and nutrition*. 2007;25(4):414-21.
- 448 32. Zaman K, Baqui AH, Yunus M, Sack RB, Bateman OM, Chowdhury HR, et al. Acute  
449 respiratory infections in children: a community-based longitudinal study in rural Bangladesh.  
450 *Journal of tropical pediatrics*. 1997;43(3):133-7.

451 33. World Health Organization. Child growth standards [Available from:  
452 <http://www.who.int/childgrowth/standards/en/>.

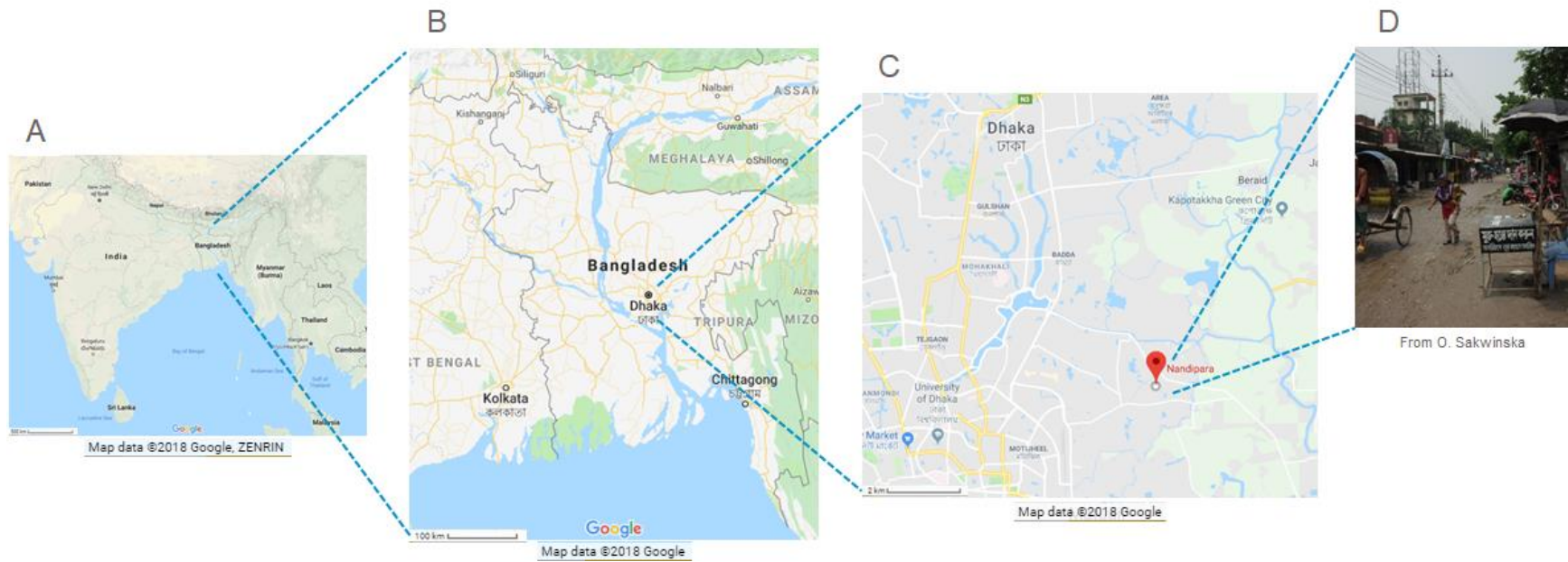
453 34. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller AB, Narwal R, et al. National,  
454 regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since  
455 1990 for selected countries: a systematic analysis and implications. Lancet (London, England).  
456 2012;379(9832):2162-72.

457 35. National Nutrition Services. National Low Birth Weight Survey (NLBWS) Bangladesh,  
458 2015. Institute of public health nutrition, directorate general of health services, Ministry of Health  
459 & family welfare. Government of the People's Republic of Bangladesh; 2017.

460 36. Blanc AK, Wardlaw T. Monitoring low birth weight: an evaluation of international  
461 estimates and an updated estimation procedure. Bulletin of the World Health Organization.  
462 2005;83(3):178-85.

463

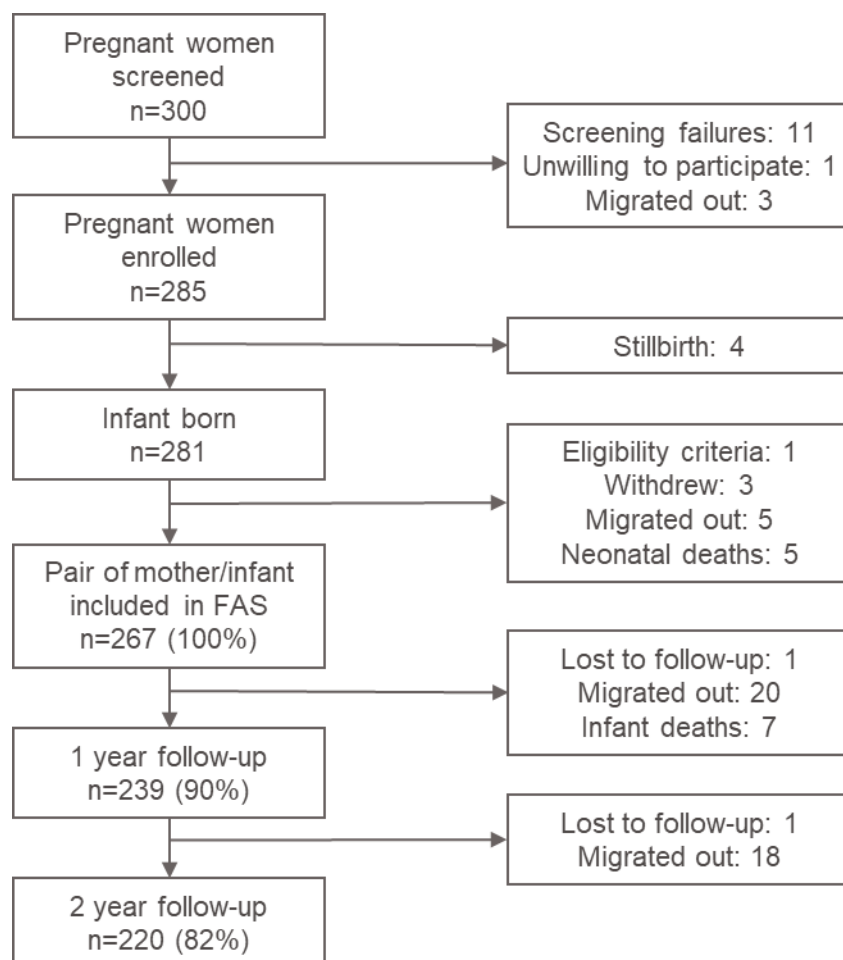
464 **Figure 1.** Study area. (A) Location of Bangladesh in South Asia. (B) Map of Bangladesh, with the capital Dhaka. (C) Road map showing  
465 the location of Nandipara, a peri-urban area of Dhaka. (D) Photography of Nandipara.



466



467 **Figure 2.** Flowchart of cohort study participants screened, enrolled and followed between April  
468 2013 and October 2016.



469

470 **Table 1.** Characteristics of the households of the study participants.

| Characteristics<br>(n=267 unless stated)         | Category                        | Frequency<br>(%) |
|--|---------------------------------|------------------|
| <b>Household socioeconomic characteristics</b>   |                                 |                  |
| <b>Household income (taka per month)</b>         | <7'000                          | 78 (29.2)        |
|  | ≥7'000 <10'000                  | 104 (39.0)       |
|  | >10'000                         | 85 (31.8)        |
| <b>House ownership<sup>1</sup></b>               | Own                             | 54 (24.5)        |
|  | Rent                            | 166 (75.5)       |
| <b>Mother's education</b>                        | Illiterate                      | 19 (7.1)         |
|  | Primary                         | 167 (62.5)       |
|  | Secondary                       | 66 (25.7)        |
|  | High Secondary                  | 14 (5.2)         |
|  | University                      | 1 (0.4)          |
| <b>Mother's occupation</b>                       | Housewife                       | 245 (91.8)       |
|  | Student                         | 2 (0.7)          |
|  | Working                         | 20 (7.5)         |
| <b>Father's occupation</b>                       | Service                         | 66 (24.7)        |
|  | Business                        | 42 (15.7)        |
|  | Grocery shopkeeper              | 5 (1.9)          |
|  | Day labor                       | 69 (25.8)        |
|  | Rickshaw puller                 | 46 (17.2)        |
|  | Auto Rickshaw driver            | 30 (11.2)        |
|  | Other occupations               | 8 (3.0)          |
|  | Jobless                         | 1 (0.4)          |
| <b>Food insecurity<sup>1a</sup></b>              | Never                           | 194 (88.2)       |
|  | Rarely                          | 22 (10.0)        |
|  | Sometimes                       | 2 (0.9)          |
|  | Often                           | 2 (0.9)          |
| <b>Household possession of goods<sup>1</sup></b> | Almirah/wardrobe                | 94 (42.7)        |
|  | Bed                             | 209 (95.0)       |
|  | Refrigerator                    | 79 (35.9)        |
|  | Television                      | 159 (72.3)       |
|  | Radio                           | 3 (1.4)          |
|  | Mobile phone/phone land line    | 206 (93.6)       |
|  | Access to internet <sup>b</sup> | 29 (13.2)        |
|  | Bicycle                         | 26 (11.8)        |
| <b>Household size</b>                            |                                 |                  |
| <b>Adults present during the day<sup>1</sup></b> | 1                               | 129 (58.6)       |
|  | 2                               | 71 (32.3)        |
|  | 3                               | 9 (4.1)          |
|  | ≥4                              | 11 (5.0)         |

| <b>Characteristics<br/>(n=267 unless stated)</b>                        | <b>Category</b>             | <b>Frequency<br/>(%)</b> |
|---|-----------------------------|--------------------------|
| <b>Household size</b>   |                             |                          |
| <b>Adults present at night<sup>1</sup></b>                              | 1                           | 4 (1.8)                  |
|   | 2                           | 154 (70.0)               |
|   | 3                           | 18 (8.2)                 |
|   | ≥4                          | 44 (20.0)                |
| <b>Children between 5- and 16-year old during the day<sup>1</sup></b>   | 0                           | 98 (44.5)                |
|   | 1                           | 77 (35.0)                |
|   | ≥2                          | 45 (20.5)                |
| <b>Children between 5- and 16-year old present at night<sup>1</sup></b> | 0                           | 96 (43.6)                |
|   | 1                           | 77 (35.0)                |
|   | ≥2                          | 47 (21.4)                |
| <b>Children &lt;5-year old present during the day<sup>2</sup></b>       | 1                           | 186 (84.9)               |
|   | 2                           | 32 (14.5)                |
|   | 3                           | 1 (0.5)                  |
| <b>Children &lt;5-year old present at night<sup>2</sup></b>             | 1                           | 186 (84.9)               |
|   | 2                           | 31 (14.2)                |
|   | 3                           | 1 (0.5)                  |
| <b>Housing characteristics</b>  |                             |                          |
| <b>Number of rooms</b>  | 1                           | 186 (69.7)               |
|   | 2                           | 63 (23.6)                |
|   | 3                           | 13 (4.9)                 |
|   | ≥4                          | 5 (1.9)                  |
| <b>Flooring material<sup>1</sup></b>                                    | Cement                      | 188 (85.5)               |
|   | Wood                        | 11 (5.0)                 |
|   | Earth                       | 21 (9.5)                 |
| <b>Wall material<sup>1</sup></b>  | Tin                         | 123 (55.9)               |
|   | Solid                       | 96 (43.6)                |
|   | Mud                         | 1 (0.5)                  |
| <b>Roof material<sup>1</sup></b>  | Tin                         | 178 (80.9)               |
|   | Concrete                    | 39 (17.7)                |
|   | Plastic/mud                 | 3 (1.4)                  |
| <b>Household sanitation facilities</b>                                  |                             |                          |
| <b>Type of toilet facility<sup>1</sup></b>                              | Flush to piped sewer system | 1 (0.5)                  |
|   | Flush to septic tank        | 90 (40.9)                |
|   | Flush to pit latrine        | 76 (34.5)                |
|   | Pit latrine with slab       | 50 (22.7)                |
|   | Flush to somewhere else     | 2 (0.9)                  |
|   | Hanging toilet              | 1 (0.5)                  |
| <b>Toilet facility shared with other households<sup>1</sup></b>         | Yes                         | 162 (73.6)               |
|   | No                          | 58 (26.4)                |

| <b>Characteristics<br/>(n=267 unless stated)</b>                    | <b>Category</b>   | <b>Frequency<br/>(%)</b>  |
|---|---|---|
| <b>Hand washing</b>   |   |   |
| <b>Place of hand washing<sup>1</sup></b>                            | Inside house<br>Outside house<br>< 3 steps<br>> 3 and < 10 steps<br>> 10 steps  | 35 (15.9)<br>185 (84.1)<br>26 (11.8)<br>130 (59.1)<br>29 (13.2)     |
| <b>Hand washing after defecating<sup>1</sup></b>                    | No<br>Yes<br>With water<br>With soap<br>With mud  | 0 (0.0)<br>220 (100.0)<br>21 (9.5)<br>197 (89.5)<br>2 (0.9)         |
| <b>Hand washing after cleaning infant who defecated<sup>1</sup></b> | No<br>Yes<br>With water<br>With soap<br>With mud  | 1 (0.5)<br>219 (99.5)<br>21 (9.6)<br>196 (89.5)<br>2 (0.9)          |
| <b>Hand washing before feeding infant<sup>1</sup></b>               | No<br>Yes<br>With water<br>With soap  | 4 (1.8)<br>216 (98.2)<br>154 (71.3)<br>62 (28.7)                    |
| <b>Hand washing before preparing food<sup>1</sup></b>               | No<br>Yes<br>With water<br>With soap  | 171 (77.7)<br>49 (22.3)<br>38 (77.6)<br>11 (22.4)                   |
| <b>Hand washing before eating<sup>1</sup></b>                       | No<br>Yes<br>With water<br>With soap  | 4 (1.8)<br>216 (98.2)<br>159 (73.6)<br>57 (26.4)                    |
| <b>Household drinking water</b>                                     |   |   |
| <b>Main source of drinking water<sup>1</sup></b>                    | Piped into dwelling<br>Piped to yard/plot<br>Tube well<br>Dug well protected<br>Dug well unprotected<br>Surface water | 4 (1.8)<br>33 (15.0)<br>8 (3.6)<br>168 (76.4)<br>5 (2.3)<br>2 (0.9) |
| <b>Location of drinking water<sup>1</sup></b>                       | Dwelling<br>Yard/plot or elsewhere  | 27 (12.3)<br>193 (87.7)   |
| <b>Water treatment prior to drinking<sup>1</sup></b>                | None<br>Boiled<br>Filtered  | 95 (43.2)<br>117 (53.2)<br>8 (3.6)                                  |

| <b>Characteristics<br/>(n=267 unless stated)</b>   | <b>Category</b>                     | <b>Frequency<br/>(%)</b> |
|--|-------------------------------------|--------------------------|
| <b>Cooking amenity</b>   |                                     |                          |
| <b>Place of cooking<sup>1</sup></b>  | In the house                        | 47 (21.4)                |
|  | In a separate building              | 80 (36.4)                |
|  | Outdoors                            | 93 (42.3)                |
| <b>Cooking fuel<sup>1</sup></b>  | Natural gas/liquefied petroleum gas | 209 (95.0)               |
|  | Wood                                | 11 (5.0)                 |
| <b>Family history of respiratory illness</b>   |                                     |                          |
| <b>Person in family with chronic cough<br/>(&gt;3 months/year)</b>   | No                                  | 244 (91.4)               |
|  | Yes                                 | 23 (8.6)                 |
| <b>Smoker in family</b>  | No                                  | 89 (33.3)                |
|  | Yes                                 | 178 (66.7)               |
| <sup>1</sup> Data missing for 47 participants; <sup>2</sup> Data missing for 48 participants.<br><sup>a</sup> In the past 4 weeks, how often did you worry that your household would not have enough food?<br><sup>b</sup> In this household do you have access to internet through smart phone or computer? |                                     |                          |

471

472 **Table 2.** Characteristics of the mothers at baseline.

| <b>Characteristics<br/>(n=267 unless stated)</b>  | <b>Category</b> | <b>Frequency (%) or mean <math>\pm</math> SD<br/>(median [min, max])</b> |
|---|-----------------|--|
| <b>Age (years)</b>  |                 | 23.7 $\pm$ 3.8 (23.0 [18, 35])   |
| <b>Height (cm)<sup>1</sup></b>  |                 | 149.0 $\pm$ 6.0 (148.5 [130.0, 175.5])                                   |
| <b>Weight (cm)</b>  |                 | 52.0 $\pm$ 8.1 (50.0 [37.5, 83.0])                                       |
| <b>Blood group</b>  | A+              | 66 (24.7)  |
|   | AB+             | 24 (9.0)   |
|   | B+              | 90 (33.7)  |
|   | O+              | 87 (32.6)  |
| <b>Gravida</b>  | 1               | 85 (31.8)  |
|   | 2               | 88 (33.0)  |
|   | 3               | 63 (23.6)  |
|   | 4               | 23 (8.6)   |
|   | 5               | 6 (2.2)  |
|   | 6               | 2 (0.7)  |
| <b>Age of last child (years)</b>  |                 | 5.3 $\pm$ 2.6 (5.0 [1, 15])  |
| <b>Place of previous delivery<sup>3</sup></b>   | Home            | 125 (77.6)   |
|   | Hospital/clinic | 36 (22.4)  |
| <b>Event at current labor</b>   | Normal          | 246 (92.1)   |
|   | Prolonged       | 21 (7.9)   |
| <sup>1</sup> Data missing for 1 participant; <sup>2</sup> Data missing for 3 participants; <sup>3</sup> Data missing for 21 participants. |                 |  |

473

474 **Table 3.** Characteristics at birth of the infants included in the birth cohort study.

| <b>Characteristics<br/>(n=267 unless stated)</b>  | <b>Category</b>           | <b>Frequency (%) or mean ±<br/>SD (median [min, max])</b> |
|---|---------------------------|---|
| <b>Sex</b>  | Female                    | 140 (52.4)  |
|   | Male                      | 127 (47.6)  |
| <b>Gestational age</b>  | >=37 weeks                | 246 (92.1)  |
|   | <37 weeks                 | 21 (7.9)  |
| <b>Place of delivery</b>  | Hospital/Clinic           | 120 (44.9)  |
|   | Home                      | 147 (55.1)  |
| <b>Mode of delivery</b>   | Vaginal birth             | 199 (74.5)  |
|   | Caesarean section         | 68 (25.5)   |
| <b>Season</b>   | Pre-monsoon (Mar-May)     | 89 (33.3)   |
|   | Rainy monsoon (Jun-Oct)   | 115 (43.1)  |
|   | Cool dry winter (Nov-Feb) | 63 (23.6)   |
| <b>Weight<sup>1</sup></b>   | >=2,500g                  | 220 (83.7)  |
|   | <2,500g                   | 43 (16.3)   |
| <b>Weight (kg)<sup>1</sup></b>  |                           | 2.8 ± 0.4 (2.8 [1.7, 4.2])                                |
| <b>Length (cm)<sup>1</sup></b>  |                           | 46.9 ± 2.0 (47.2 [41.0, 52.02])                           |
| <b>Number of siblings<sup>2</sup></b>   | 0                         | 111 (42.0)  |
|   | 1                         | 101 (38.3)  |
|   | >=2                       | 52 (19.7)   |
| <sup>1</sup> Data missing for 4 participants; <sup>2</sup> Data missing for 3 participants. |                           |   |

475