BMJ Global Health

Community-level impacts of sanitation coverage on maternal and neonatal health: a retrospective cohort of survey data

Brittany L Kmush ,[1] Bhavneet Walia,[1] Anushruta Neupane,[1] Carolina Frances,[1] Idris Ahmed Mohamed,[1] Mahwish Iqbal,[2] David A Larsen[1]

To cite: Kmush BL, Walia B, Neupane A, *et al*. Community-level impacts of sanitation coverage on maternal and neonatal health: a retrospective cohort of survey data. *BMJ Global Health* 2021;**6**:e005674. doi:10.1136/bmjgh-2021-005674

Handling editor Soumitra S Bhuyan

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjgh-2021-005674).

Received 10 March 2021
Accepted 27 August 2021

Check for updates

[1]Department of Public Health, Syracuse University, Syracuse, New York, USA
[2]Department of Public Health and Preventive Medicine, SUNY Upstate Medical University, Syracuse, New York, USA

Correspondence to
Dr Brittany L Kmush; blkmush@syr.edu

ABSTRACT

Introduction Access to sanitation facilities (toilets or latrines) greatly improves human health. Low community sanitation coverage may lead to increased exposure to pathogens for households both with and without a sanitation facility.

Methods We created a retrospective cohort using Demographic and Health Surveys from 1990 through 2018. Using regression with matched women as a random intercept, we assessed the association between community-level sanitation coverage and neonatal mortality (Poisson model, *n*=1 254 862 live births, 187 datasets), small birth size (logit model, *n*=1 058 843 live births, 187 datasets) and anaemia (logit model, *n*=1 304 626 women, 75 datasets).

Results Among women with household sanitation, the incidence of neonatal death (incidence rate ratio: 0.85, 95% CI 0.77 to 0.93), the odds of small birth size (OR: 0.81, 95% CI 0.76 to 0.87) and anaemia (OR: 0.82, 95% CI 0.79 to 0.85) were lower for women in communities with 100% sanitation coverage compared with 1%–30% (p≤0.001 for all). There was no difference in neonatal deaths between women in communities with 31%–99% sanitation coverage compared with 1%–30% (p≥0.05). Among women without household sanitation, there were no differences in neonatal mortality by community sanitation (p≥0.05). The odds of small birth size were decreased (OR: 0.91, 95% CI 0.87 to 0.97, p=0.003) for women in communities with 61%–99% sanitation coverage compared with 1%–30%; there was no association with the other community sanitation categories (p≥0.05). The odds of anaemia were increased (OR: 1.08, 95% CI 1.06 to 1.11, p<0.001) for women living in communities with 0% sanitation coverage compared with 1%–30%, but no association with the other community sanitation categories (p≥0.05).

Conclusion Community sanitation coverage is associated with improved maternal and neonatal outcomes, particularly among women with household sanitation. This suggests that the impact of sanitation coverage on maternal and neonatal health is underestimated unless the community-level effects are considered.

WHAT IS ALREADY KNOWN?

⇒ Improved individual and household level coverage of adequate sanitation and faeces disposal facilities has been shown to reduce the risk of pathogens transmitted via human faeces.
⇒ Coverage of household sanitation in the community has been suggested to act via a 'herd protection' effect where low prevalence of household toilets leads to increased exposure to pathogens for all in the community, independently of sanitation at the household level.
⇒ The cumulative coverage of household sanitation in the community affects child survival and child health outcomes but has not yet been examined for maternal and neonatal health outcomes.

WHAT ARE THE NEW FINDINGS?

⇒ We found that among women with household sanitation, the incidence of neonatal death, the odds of small birth size and anaemia were lower for women who live in a community with 100% sanitation coverage compared with communities with 1%–30% sanitation coverage (p≤0.002 for all).
⇒ Among women without household sanitation, the odds of anaemia were increased (p<0.001) for women living in a community with 0% sanitation coverage compared with 1%–30% sanitation coverage but there was no association with small birth size (p=0.87) or neonatal death (p=0.51).

WHAT DO THE NEW FINDINGS IMPLY?

⇒ Lower community level sanitation coverage is associated with adverse maternal and neonatal outcomes, after accounting for household sanitation.
⇒ The greatest reductions in neonatal mortality, small birth size and maternal anaemia were seen as communities approached 100% sanitation coverage.
⇒ Interventions to improve household sanitation influence maternal and neonatal health at both the individual-level and community-levels, which need to be simultaneously accounted for when evaluating programme costs.

INTRODUCTION

Worldwide, access to sanitation is very low, with 2.5 billion people lacking access to an improved sanitation facility and over one billion people without access to any sanitation facility.[1] Despite international attention to increase access to improved sanitation, little progress was made to reach the Millennium Development Goal of reducing, by half, the population without access to basic sanitation.[1] Sustainable Development Goal six highlights improving access to adequate sanitation facilities, with an emphasis on ending open defecation. The risk of encountering human faeces is greatly increased in areas with low coverage of household toilets (community sanitation) or where households are lacking adequate sanitation and disposal facilities. Human faeces contain enormous amounts of pathogens; a single gram of faeces can contain over 10 million viral and bacterial pathogens, 10 000 protozoan cysts and up to 10 000 soil-transmitted helminth (STH) eggs.[2] Exposure to human faeces, and thus these pathogens, leads to a variety of human diseases, notably diarrheal diseases, soil-transmitted helminthiasis, schistosomiasis, undernutrition, iron-deficiency anaemia, trachoma and acute respiratory infections.[2]

Pregnant women are more likely to experience the severe consequences of many infectious diseases, particularly viruses and bacteria, primarily due to the shifting of the immune system during pregnancy to tolerate the developing fetus.[3] However, the connection between sanitation and maternal health is not always obvious as the risk factors for adverse health outcomes during pregnancy are diverse and may be removed, in time, from the maternal health outcomes.[4 5] Most evidence suggests that improvements in sanitation access would improve maternal health.[6–8] Recurrent infections, especially with multiple intestinal infections and diarrheal diseases, often lead to acute and chronic malnutrition in children and pregnant women. Chronic malnutrition can cause stunted growth and developmental delays as well as increasing the risk of future infections and chronic diseases later in life.[9]

Improved individual and household level access to adequate sanitation and faeces disposal facilities has been shown to reduce the risk of infections transmitted via human faeces.[10–13] However, household sanitation is only part of the solution. Coverage of household sanitation in the community has been suggested to act via a 'herd protection' effect where low prevalence of household toilets leads to increased exposure to pathogens for all in the community, independently of sanitation facilities at the household level.[14–16] Environmental enteropathy from continuous faecal-oral contamination is a major risk factor for disease and malnutrition in children.[16 17] Experts estimate that helminth infections would be drastically reduced with universal sanitation coverage; however, campaigns to control STH have largely focused on deworming efforts.[2] Several research teams, including our group, have demonstrated that the cumulative coverage of household sanitation facilities in the community affects child survival and child health outcomes, including stunting, diarrhoea and anaemia.[17–20] Additionally, universal, or near universal, coverage of household sanitation in the community leads to the greatest benefits for child health. Several observational and community-randomised controlled trials of total sanitation as well

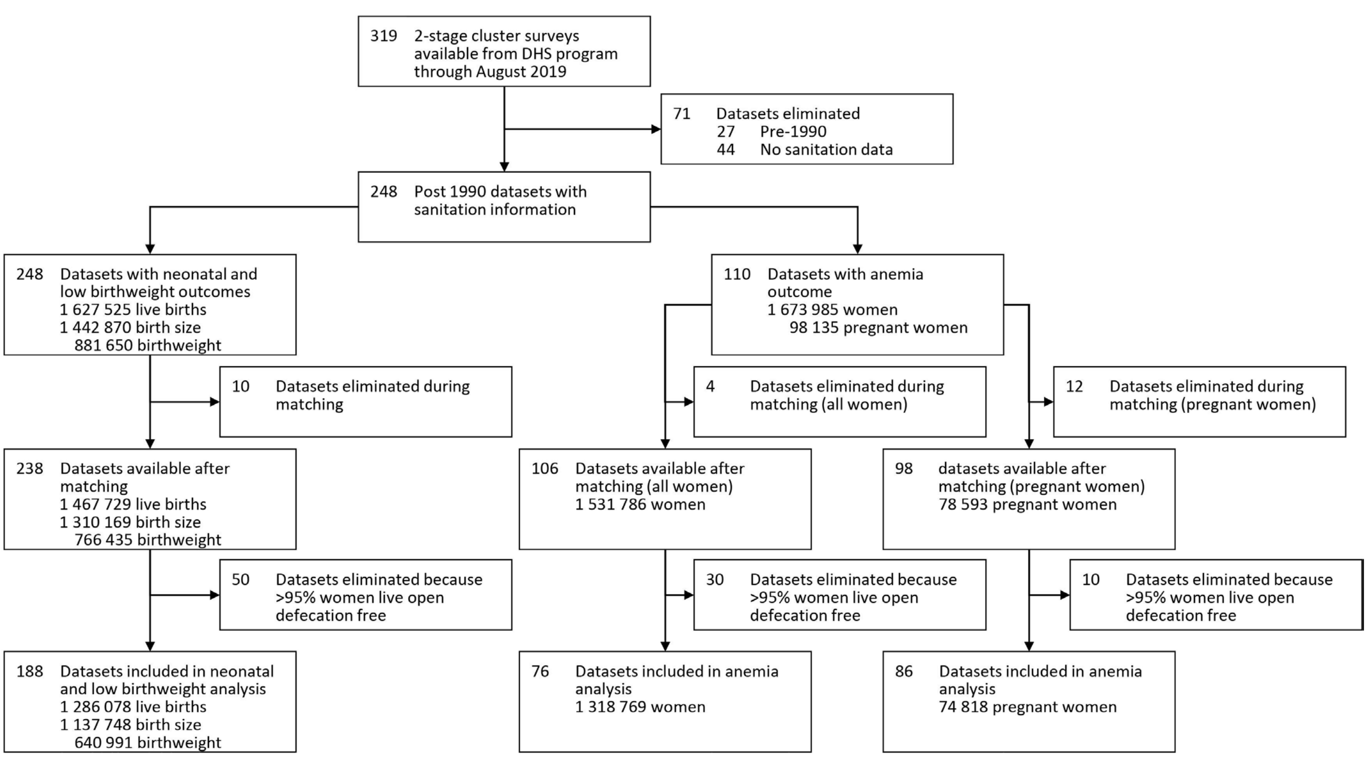

Figure 1 Flow chart of included demographic health surveys. DHS, Demographic and Health Surveys.

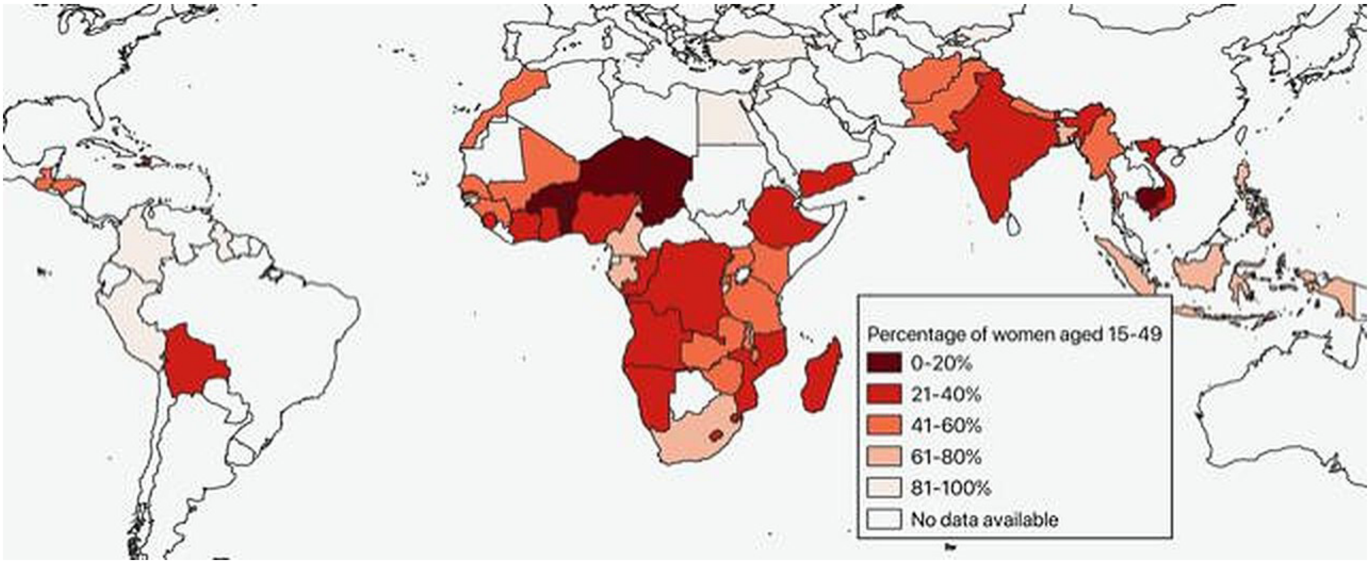

Figure 2 Per cent of women living in communities with 100% sanitation coverage as measured by the most recent DHS. DHS, Demographic and Health Surveys.

as a modelling study have found increasing sanitation to improve child health,[14 17 19 20] although this association is not always seen.[21 22]

Here, we present a global analysis of 248 Demographic and Health Surveys (DHS) examining the association between household and community sanitation coverage and maternal and neonatal health outcomes. This is the first study to examine this association on a global level. Most estimates of the impact of improved sanitation only account for benefits of household-level access. However, we argue that both the public health and individual well-being impacts are underestimated due to the community-level effects of increasing coverage of improved sanitation facilities. Measuring the magnitude of the relationship between sanitation coverage and maternal-child health creates a rationale for stronger public policy related to sanitation access.

METHODS

Study design, setting, participants and data sources

We created a retrospective cohort using the birth histories of more than one million women from 248 two-stage cluster sampled DHS conducted between 1990 and 2018 that measured sanitation access and maternal and neonatal health outcomes.[23] The DHS programme was established by United States Agency for International Development (USAID) to collect comparable data on health indicators across countries, with a focus on countries that receive USAID assistance.[24 25] The surveys are nationally representative two-stage samples, with first clusters being selected with probability proportionate to estimated population and second households being selected randomly. Women from age 15 to 49 years old in selected households are administered a questionnaire about the household and her health experience. The DHS is carried out in 5 years, overlapping intervals starting in 1984. Since its inception, there have been over 300 surveys completed in over 90 countries with over two million biomarker tests completed.[24 25] DHS programme staff provide technical assistance at all stages of the data collection process to ensure that data is collected accurately, reliably and comparably across countries and overtime.[24] Methodology and questionnaires are published online (available at https://dhsprogram.com) and can be compared over time.

Variables

The primary exposures of interest are household and community-level sanitation coverage, which are assessed reliably in the DHS methodologies.[26] We classified women as living in households that report using any type of sanitation facility, including both improved and unimproved facilities, or not having access to any sanitation facility. We defined community as the survey sampling area or cluster. We calculated the proportion of households in each area or cluster that have a sanitation facility to serve as the measure of community-level sanitation coverage. Since we are interested in the incremental effects of community level sanitation, we eliminated datasets where greater than 95% of the households are in communities (sampling area or cluster) with 100% sanitation coverage from further analysis. Community level sanitation was categorised at 0%, 1%–30%, 31%–60%, 61%–99% and 100%.

Maternal health and birth outcomes available in these datasets include neonatal mortality, low birth weight and anaemia. Women are asked to provide a full birth history, including the date of birth of each child and if the child is still alive. If the child is not alive, the woman is asked for the age at death in days, weeks, months and/or years.[27] We defined a neonatal death as a child who died during the first 28 days or 4 weeks of

Table 1 Descriptive frequencies of demographics and socioeconomic characteristics for women in the datasets before matching

Characteristic	0% community sanitation coverage (n, %)	1%–30% community sanitation coverage (n, %)	31%–60% community sanitation coverage (n, %)	61%–99% community sanitation coverage (n, %)	100% community sanitation coverage (n, %)	χ^2 p value
	N=96 933	N=217 438	N=167 602	N=499 345	N=647 100	
Parity						<0.001
1	18 971 (19.6%)	46 704 (21.5%)	38 527 (23.0%)	119 832 (24.0%)	182 198 (28.2%)	
2	13 576 (14.0%)	33 189 (15.3%)	27 076 (16.2%)	80 894 (16.2%)	124 087 (19.2%)	
3	4747 (4.9%)	12 688 (5.8%)	9865 (5.9%)	27 007 (5.4%)	38 734 (6.0%)	
4	46 631 (48.1%)	97 567 (44.9%)	71 409 (42.6%)	213 382 (42.7%)	241 848 (37.4%)	
≥5	13 008 (13.4%)	27 290 (12.6%)	20 725 (12.4%)	58 230 (11.7%)	60 233 (9.3%)	
Wealth quintile						<0.001
1	52 434 (54.2%)	92 018 (42.4%)	55 677 (33.3%)	123 505 (24.8%)	98 640 (15.3%)	
2	20 336 (21.0%)	53 527 (24.7%)	40 784 (24.4%)	107 096 (21.5%)	113 400 (17.5%)	
3	14 749 (15.2%)	41 709 (19.2%)	35 807 (21.4%)	110 310 (22.1%)	130 368 (20.2%)	
4	7341 (7.6%)	22 579 (10.4%)	23 971 (14.3%)	89 576 (18.0%)	139 652 (21.6%)	
5	1966 (2.0%)	7311 (3.4%)	11 163 (6.7%)	68 123 (13.7%)	164 287 (25.4%)	
Mother's education						<0.001
None	67 388 (69.6%)	117 996 (54.6%)	71 694 (43.2%)	157 957 (31.9%)	107 664 (16.8%)	
Some primary	15 969 (16.5%)	40 198 (18.6%)	35 375 (21.3%)	126 935 (25.6%)	103 577 (16.2%)	
Completed primary	4041 (4.2%)	15 978 (7.4%)	15 583 (9.4%)	58 992 (11.9%)	81 527 (12.7%)	
Higher than primary	9415 (9.7%)	42 018 (19.4%)	43 280 (26.1%)	151 496 (30.6%)	346 911 (54.2%)	
Antenatal care quintile*						<0.001
1	14 317 (35.1%)	37 596 (31.6%)	27 457 (27.1%)	77 975 (23.5%)	94 209 (20.5%)	
2	10 077 (24.7%)	28 651 (24.1%)	25 528 (25.2%)	82 742 (25.0%)	132 447 (28.8%)	
3	9185 (22.5%)	29 232 (24.6%)	26 636 (26.3%)	85 478 (25.8%)	115 403 (25.1%)	
4	4182 (10.3%)	12 132 (10.2%)	11 466 (11.3%)	50 163 (15.1%)	69 923 (15.2%)	
5	3025 (7.4%)	11 229 (9.4%)	10 365 (10.2%)	35 194 (10.6%)	47 733 (10.4%)	
Urban residence	91 314 (94.2%)	201 884 (92.8%)	142 957 (85.3%)	355 060 (71.1%)	269 283 (41.6%)	<0.001
Mother <150 cm	14 320 (21.9%)	41 456 (26.3%)	31 417 (26.3%)	72 625 (22.8%)	85 088 (19.7%)	<0.001
Male child	49 562 (51.1%)	112 769 (51.9%)	86 824 (51.8%)	256 967 (51.5%)	333 468 (51.5%)	<0.001
Delivery in a facility	29 943 (30.9%)	97 533 (44.9%)	87 464 (52.2%)	295 282 (59.2%)	482 432 (74.6%)	<0.001

life. A sensitivity analysis was completed, including children who died at 1 month of age as a neonatal death. In order to minimise recall errors, the analyses of neonatal mortality were limited to births and neonatal deaths within 2 years previous to the survey.

Women are also asked to report both on the relative size of their most recently born baby at birth (very small, smaller than average, average, larger than average, very large) and if their baby was weighed at birth. If the baby was weighed, the weight is recorded directly from a health card (if it exists) or the women are asked to recall the birth weight. We classified a low birth weight baby as those born to women who recalled their children to be 'very small' at birth. We also performed a sensitivity analysis where a reported or recorded birth weight of <2500 g was considered low birth weight. We excluded women who did not report a birth in the previous 2–5 years and women who could not recall a birth weight from this outcome analysis.

In many DHS surveys, a blood sample is taken from the women to assess circulating haemoglobin as a marker of anaemia status, usually with the HemoCue system. Moderate and severe anaemia was defined as any woman (pregnant or not) with smoking and altitude adjusted haemoglobin <100 g/L. Altitude and smoking adjustments are standard in the DHS and made

Table 2 Descriptive frequencies of outcomes and exposures for women in the datasets before matching

	Household sanitation	0% community Sanitation Coverage (n, %)	1%–30% community Sanitation Coverage (n, %)	31%–60% community Sanitation Coverage (n, %)	61%–99% community Sanitation Coverage (n, %)	100% community Sanitation Coverage (n, %)	Total (n, %)
N women reporting a neonatal death (%)	Yes	N/A	577 (1.9%)	1421 (1.9%)	6699 (1.8%)	5608 (1.5%)	14 305 (1.7%)
	No	2139 (2.2%)	4105 (2.2%)	1865 (2.1%)	1367 (2.0%)	N/A	9476 (2.2%)
N women reporting a death at 30 days or less (%)	Yes	N/A	668 (2.2%)	1618 (2.2%)	7576 (2.0%)	6306 (1.7%)	16 168 (1.9%)
	No	2466 (2.6%)	4623 (2.5%)	2117 (2.3%)	1552 (2.3%)	N/A	10 758 (2.5%)
N women reporting a very small perceived birth size (%)	Yes	N/A	1473 (5.3%)	3576 (5.4%)	16 341 (4.9%)	12 345 (4.0%)	33 735 (4.6%)
	No	6114 (7.0%)	10 103 (6.0%)	4997 (6.2%)	3548 (6.0%)	N/A	24 762 (6.2%)
N women reporting a low-birth-weight baby (%)	Yes	N/A	1807 (12.8%)	4375 (12.5%)	22 633 (11.3%)	25 253 (10.1%)	54 068 (10.8%)
	No	3124 (14.1%)	9504 (14.4%)	5454 (14.5%)	4032 (13.2%)	N/A	22 114 (14.2%)
N women with moderate or severe anaemia (%)	Yes	N/A	5017 (12.5%)	10 932 (11.9%)	40 372 (11.1%)	38 273 (9.9%)	94 594 (10.7%)
	No	11 353 (15.9%)	30 171 (14.5%)	14 383 (13.9%)	8 510 (13.6%)	N/A	64 417 (14.5%)
N pregnant women with moderate or severe anaemia (%)	Yes	N/A	641 (26.5%)	1351 (25.5%)	5285 (24.3%)	4120 (21.2%)	11 397 (23.3%)
	No	1916 (32.6%)	4281 (30.1%)	1934 (29.5%)	1123 (28.4%)	N/A	9254 (30.2%)

N/A-not applicable.

Table 3 Association between neonatal mortality (death at ≤28 days) and level of sanitation coverage in the community for women with and without household level sanitation

Household sanitation access	Community-level sanitation access	Unadjusted incidence rate ratio (95% CI)	P value	Adjusted incidence rate ratio (95% CI)	P value
		n=832 663		*n*=823 179	
Yes	1%–30%	Reference	N/A	Reference	N/A
Yes	31%–60%	1.00 (0.91 to 1.10)	0.99	0.99 (0.90 to 1.09)	0.84
Yes	61%–99%	0.92 (0.85 to 1.01)	0.09	0.92 (0.84 to 1.01)	0.07
Yes	100%	0.84 (0.77 to 0.92)	<0.001	0.85 (0.77 to 0.93)	<0.001
		n=434 965		*n*=431 683	
No	0%	0.99 (0.93 to 1.04)	0.61	0.99 (0.94 to 1.05)	0.79
No	1%–30%	Reference	N/A	Reference	N/A
No	31%–60%	0.93 (0.87 to 0.99)	0.03	0.93 (0.87 to 1.00)	0.05
No	61%–99%	0.95 (0.88 to 1.02)	0.18	0.96 (0.89 to 1.04)	0.34

Poisson regression.
Unadjusted models included the dataset.
Adjusted models included mother's age, mother's age squared, parity, wealth quintile, mother's education, antenatal care quality score, urban or rural, mother's stunting, child's gender, facility delivery or not and dataset.
N/A, not applicable.

using WHO recommendations.[28] A subset analysis was completed using only women who were pregnant at the time of the survey. Women without measured haemoglobin were eliminated from the analysis.

Bias

Women in households with sanitation or living in communities with high sanitation coverage are likely to be predisposed to less risk of anaemia and poor birth outcomes, independent of sanitation access. We used two separate strategies to account for this selection bias and potential confounding. First, we stratified our analyses by women in households with any sanitation and women in households without any sanitation. We considered modelling household access jointly in the models. However, we felt that stratification was a better method for examining our research question. Second, we used exact matching on community-level measures to circumvent the inherent selection bias of living in communities with sanitation coverage. Women were matched on the following parameters using the MatchIt

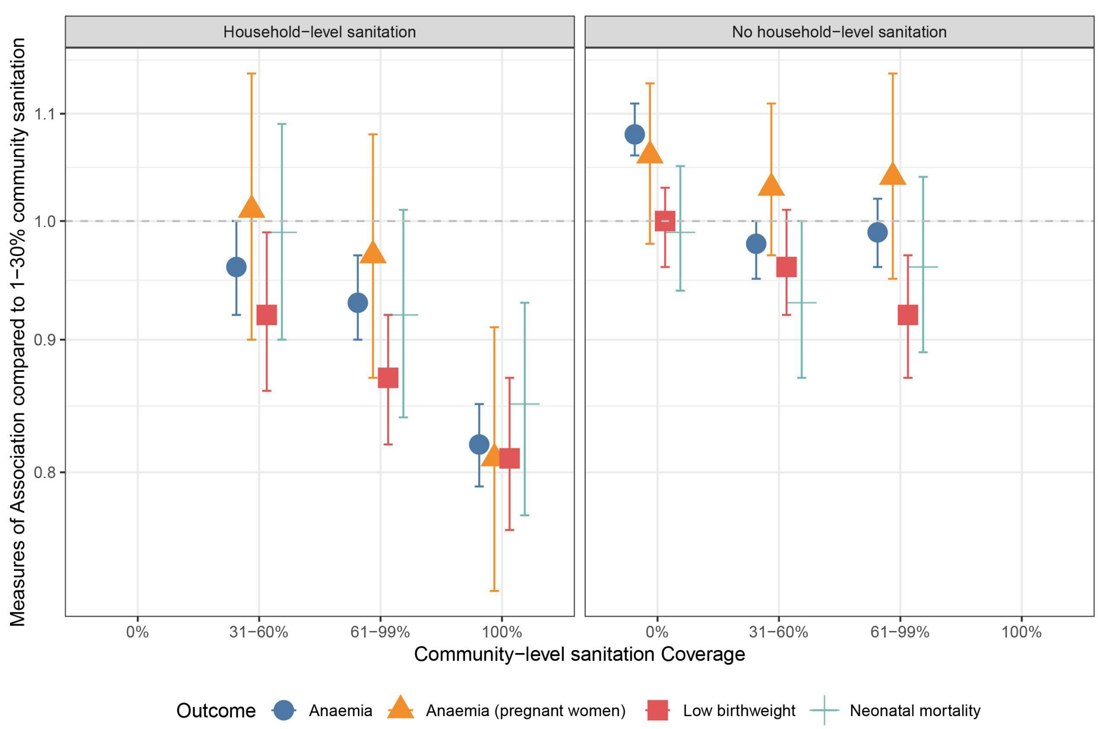

Figure 3 Adjusted relationship between sanitation and various outcomes.

Table 4 Association between very small reported birth size and level of sanitation coverage in the community for women with and without household level sanitation

Household sanitation	Community-level sanitation coverage	Unadjusted OR (95% CI)	P value	Adjusted OR (95% CI)	P value
		n=700 088		*n*=698 529	
Yes	1%–30%	Reference	N/A	Reference	N/A
Yes	31%–60%	0.92 (0.86 to 0.98)	0.01	0.92 (0.86 to 0.99)	0.02
Yes	61%–99%	0.85 (0.80 to 0.90)	<0.001	0.87 (0.82 to 0.92)	<0.001
Yes	100%	0.78 (0.73 to 0.83)	<0.001	0.81 (0.76 to 0.87)	<0.001
		n=360 930		*n*=360 314	
No	0%	0.96 (0.93 to 1.00)	0.03	1.00 (0.96 to 1.03)	0.82
No	1%–30%	Reference	N/A	Reference	N/A
No	31%–60%	0.91 (0.87 to 0.95)	<0.001	0.96 (0.92 to 1.01)	0.11
No	61%–99%	0.86 (0.81 to 0.90)	<0.001	0.91 (0.87 to 0.97)	0.003

Logistic regression.
Unadjusted models included the dataset.
Adjusted models included mother's age, mother's age squared, parity, wealth quintile, mother's education, antenatal care quality score, urban or rural, mother's stunting, child's gender, facility delivery or not and dataset.
N/A, not applicable.

package[29] in R V.3.2.3:[30] the tertile household wealth within the woman's primary sampling unit, the tertile of access to improved water source within the woman's primary sampling unit, whether the woman completed primary education or not, whether the woman's household was above or below the median for household-level wealth and the survey dataset. Women are matched across clusters, not just within clusters. We used tertiles rather than quintiles to reduce the number of covariate patterns, which leads to fewer dropped observations when using exact matching.

Statistical analysis

Descriptive statistics were compared using Student's t-test for continuous variables and χ^2 test for categorical variables. We used a generalised linear model with the matched group as a random intercept to assess an adjusted association between the exposures and outcomes of interest, stratified by household sanitation. Community sanitation coverage from 1% to 30% was chosen as the reference group, so there could be a common reference across all analyses performed. The high or low categories (0% or 100% community sanitation coverage) are only available

Table 5 Association between moderate and severe anaemia (smoking and altitude adjusted haemoglobin <10 g/dL) and level of sanitation coverage in the community for women with and without household level sanitation

Household sanitation	Community-level sanitation coverage	Unadjusted OR (95% CI)	P value	Adjusted OR (95% CI)	P value
		n=871 979		*n*=865 864	
Yes	1%–30%	Reference	N/A	Reference	N/A
Yes	31%–60%	0.95 (0.92 to 0.99)	0.01	0.96 (0.92 to 0.995)	0.03
Yes	61%–99%	0.93 (0.90 to 0.96)	<0.001	0.93 (0.90 to 0.97)	<0.001
Yes	100%	0.81 (0.78 to 0.84)	<0.001	0.82 (0.79 to 0.85)	<0.001
		n=4 40 822		*n*=4 38 762	
No	0%	1.10 (1.07 to 1.13)	<0.001	1.08 (1.06 to 1.11)	<0.001
No	1%–30%	Reference	N/A	Reference	N/A
No	31%–60%	0.97 (0.95 to 0.997)	0.03	0.98 (0.95 to 1.00)	0.05
No	61%–99%	0.99 (0.96 to 1.02)	0.39	0.99 (0.96 to 1.02)	0.52

Logistic regression.
Unadjusted models included the dataset.
Adjusted models included mother's age, mother's age squared, parity, wealth quintile, mother's education, urban or rural and dataset.
N/A, not applicable.

Table 6 Association between moderate and severe anaemia (smoking and altitude adjusted haemoglobin <10 g/dL) and level of sanitation coverage in the community for pregnant women with and without household level sanitation

Household sanitation	Community-level sanitation coverage	Unadjusted OR (95% CI)	P value	Adjusted OR (95% CI)	P value
		n=44 800		n=44 357	
Yes	1%–30%	Reference	N/A	Reference	N/A
Yes	31%–60%	1.02 (0.90 to 1.14)	0.80	1.01 (0.90 to 1.14)	0.81
Yes	61%–99%	0.95 (0.86 to 1.06)	0.40	0.97 (0.87 to 1.08)	0.56
Yes	100%	0.78 (0.70 to 0.88)	<0.001	0.81 (0.72 to 0.91)	<0.001
		n=29 742		*n*=29 583	
No	0%	1.09 (1.01 to 1.17)	0.02	1.06 (0.98 to 1.13)	0.13
No	1%–30%	Reference	N/A	Reference	N/A
No	31%–60%	1.01 (0.94 to 1.09)	0.72	1.03 (0.97 to 1.11)	0.30
No	61%–99%	1.00 (0.91 to 1.09)	0.93	1.04 (0.95 to 1.14)	0.41

Logistic regression.
Unadjusted models included the dataset.
Adjusted models included mother's age, mother's age squared, parity, wealth quintile, mother's education, urban or rural and dataset.
N/A, not applicable.

in half of the analyses performed. For neonatal mortality, we used a log link with a Poisson distribution. For low birth weight and anaemia, we used a logit link. Neonatal mortality and low birth weight were adjusted for the age of the mother, the age of the mother squared, parity (1, 2, 3, 4, 5 or more), wealth quintile, education (none, some primary, completed primary, higher than primary), antenatal care quality quintile (principle components analysis of the binary variables that occurred during the pregnancy: weighed, height measured, urine sample given, blood pressure measured, blood sample given, told about pregnancy complications, took antimalarial drugs, took antiparasitic drugs and iron supplementation), urban versus rural, mother's stunting (at least 150 cm or not), sex of the child, place of delivery (home, facility, not recorded) and the dataset. For anaemia, we adjusted for age of the mother, age of the mother squared, parity (0, 1, 2, 3, 4, 5 or more), wealth quintile, education (none, some primary, completed primary, higher than primary), urban versus rural and dataset. The age of the mother squared term accounts for the U-shaped curve seen for adverse maternal and neonatal outcomes; both younger and older age are associated with an increased risk, a linear term only cannot account for that type of relationship. We included these covariates to decrease the potential for confounding, with variable selection determined a priori based on the literature surrounding maternal and child health outcomes.[4 18] Models were checked for collinearity using a variance inflation factor check. The only covariates found to be collinear were age and age squared of the mother, as to be expected. Anyone with missing variables were eliminated from the analysis. Generalised linear model analysis was completed in Stata V.13. Two-sided p values less than 0.05 were considered significant. These are individual, although related, research questions chosen a priori, therefore 0.05 was an appropriate cut-off. Statistical analysis scripts can be found in the online supplemental appendix.

Patient and public involvement

Patients and public were not involved in this analysis of previously collected data.

RESULTS

We identified 319 two stage DHS cluster surveys available for analysis as of August 2019. Of those, 71 were eliminated from further analysis because they were conducted pre-1990 or did not collect information on household sanitation (figure 1). The remaining 248 datasets were conducted in 73 different countries from 1990 to 2018. Despite gains in sanitation access, a large proportion of women still live in communities with open defecation (figure 2). The demographic and socioeconomic characteristics of the women varied significantly by community sanitation coverage (table 1). The average age of the women was 26.9 years (SD 7.6 years) and was statistically different across community sanitation coverage (Student's t-test p<0.001). However, due to the large number of women in these datasets, even small differences can appear to be significant. Descriptive frequencies suggest a trend of a lower prevalence of adverse maternal and neonatal outcomes in communities with increased community sanitation (table 2).

Before matching, 248 datasets contained vital information from 1 627 525 live births, birth size information from 1 442 870 live births and birth weight information from 881 650 live births. After matching and removing datasets where >95% of women live in communities with 100% sanitation, 188 datasets contained birth history information from 1 286 078 live births, birth size information from 1 137 748 women and birth weight information from 640 991 women. Before matching, 110 datasets measured haemoglobin in 1 673 985 women, including 98 135 pregnant women. After matching and eliminating datasets where >95% of women live in communities with 100% sanitation, 76 datasets measured haemoglobin in 1 318 769 women and 85 datasets measured anaemia in 74 818 pregnant women.

Among women with household sanitation, the incidence of neonatal mortality was decreased in women living in communities with 100% sanitation coverage (incidence rate ratio (IRR): 0.85, 95% CI 0.77 to 0.93) compared with women living in communities with 1%–30% sanitation coverage. There was no difference between women in communities with 31%–99% sanitation coverage and 1%–30% sanitation coverage (table 3, figure 3). Among women without household sanitation, there were no differences in neonatal mortality by community sanitation coverage (table 3, figure 3). Similar trends were seen in the sensitivity analysis when all deaths in the first 30 days of life were considered a 'neonatal' death. However, some associations that were not statistically significant in the primary analysis were statistically significant in this analysis, possibly due to the larger number of deaths (online supplemental table 1).

Among women with household sanitation, the odds of a very small baby at birth decreased, incrementally, in women in communities with 31%–60% sanitation coverage (OR: 0.92, 95% CI 0.86 to 0.99), 61%–90% sanitation coverage (OR: 0.87, 95% CI 0.82 to 0.92), and 100% sanitation coverage (OR: 0.81, 95% CI 0.76 to 0.87) compared with women in communities with 1%–30% sanitation coverage (table 4, figure 3). Among women without household sanitation, the odds of a very small baby at birth were only decreased in those with 61%–99% sanitation coverage (OR: 0.91, 95% CI 0.87 to 0.97) compared with women in communities with 1%–30% sanitation coverage (table 4, figure 3). However, in the sensitivity analysis when this analysis was restricted to those babies weighed at birth, the odds of low birth weight were only lower among women with household sanitation in communities with 100% sanitation coverage (online supplemental table 2).

Among women with household sanitation, the odds of moderate or severe anaemia decreased, incrementally, among women in communities with 31%–60% sanitation coverage (OR: 0.96, 95% CI 0.92 to 0.995), 61%–99% sanitation coverage (OR: 0.93, 95% CI 0.90 to 0.97) and 100% sanitation coverage (OR: 0.82, 95% CI 0.79 to 0.85) compared with women in communities with 1%–31% sanitation coverage (table 5, figure 3). Among women without household sanitation, the odds of moderate and severe anaemia increased in women in communities with 0% sanitation coverage (OR: 1.08, 95% CI 1.06 to 1.11) compared with women in communities with 1%–30% sanitation coverage. No differences were seen in women in communities with >31% sanitation coverage (table 5, figure 3). However, in pregnant women, the only differences were seen among women with household sanitation in communities with 100% sanitation (OR: 0.81, 95% CI 0.72 to 0.91) compared with women in communities with 1%–30% sanitation coverage (table 6, figure 3).

DISCUSSION

We find that lower community level sanitation coverage is generally associated with increased risk of adverse maternal and neonatal outcomes, after accounting for household sanitation. The greatest gains in reducing neonatal mortality, small birth size and maternal anaemia were seen among households with sanitation as communities approached 100% sanitation coverage. These results strengthen growing evidence of the impact of environmental enteropathy and STH on maternal and neonatal health. Both neonatal mortality and low birth weight were reduced with increasing community-level sanitation, likely at least in part due to reduced impacts of environmental enteropathy. A recent report from Uganda linked environmental enteropathy with shorter gestational age and smaller birth size.[31] The risk of anaemia was also greatly reduced with increasing community-level sanitation coverage. Infections with hookworm, and possibly other STH species, are a major risk factor for anaemia in women and children,[32–34] and community-level sanitation coverage is a strong indicator of STH infection.[35] Furthermore, deworming during pregnancy also reduces the risk of neonatal mortality and low birth weight outcomes.[36]

These results suggest that the greatest health benefits from sanitation access will only be realised with universal coverage as environmental enteropathy and STH transmission are minimised. Sanitation access is largely a problem of inequity, with improved access to sanitation in lower income countries being limited to wealthy neighbourhoods in urban areas.[37] Sanitation systems are expensive, with recent estimates suggesting \$3–5 per person per year for septic-based sanitation in urban areas.[38] Our results strengthen the need to improve sanitation for all. Expanding sanitation does improve the lives of those without household-level access and decreases the risk of adverse maternal and neonatal outcomes for those with household-level access. Unfortunately, a recent report on progress towards Sustainable Development Goal 6 (ensuring universal sanitation access) finds that the world is 'alarmingly off-track to deliver sanitation for all by 2030'.[39] Furthermore, the situation is even more dire when considering sanitation coverage as a community-level indicator (figure 2).

The major limitations of this analysis come from using, large, cross-sectional datasets. Information bias and recall bias are inherent in these surveys due to the nature of self-reported responses. However, we do not expect the presence of these biases to vary significantly by exposure or outcome status. Additionally, household sanitation access in these surveys has been shown to be subject to social desirability bias. Additionally, birth histories in the past 2 years were used, although only current sanitation access was analysed, possibly leading to misclassification.

However, improvement in sanitation access in a short time would likely only have occurred in a small proportion of the women surveyed. Sanitation access is oversimplified into two categories, which is necessary in order to quantify community-level sanitation. Unmeasured confounders are also a limitation, notably handwashing behaviours. Handwashing has been suggested to be a confounder in the relationship between health outcomes and sanitation.[40 41] While handwashing facilities are recorded in a small portion of DHS, individual handwashing behaviour is not well captured. Additionally, the community is poorly defined as the survey unit or cluster, as the primary sampling unit could comprise several villages. However, if we were able to capture the variation in community sanitation coverage within each cluster with a more precise estimate, we would likely see stronger associations between community sanitation coverage and maternal and neonatal health outcomes. Missing data are another issue as not all surveys capture our outcomes of interest.

The impact of access to sanitation on maternal and neonatal health is underestimated unless the community-level effects are adequately considered. Our results demonstrate that interventions to improve household sanitation influence maternal and neonatal health at both the individual-level and community-levels, which need to be simultaneously accounted for when evaluating programme costs.

Present affiliations The present affiliation of Idris Ahmed Mohamed is: University of Minnesota School of Public Health, MN, Minneapolis, USA and Mahwish Iqbal is: Department of Family Medicine, SUNY Upstate Medical University, NY, Syracuse, USA.

Twitter Brittany L Kmush @brittlk17

Contributors BLK designed the study, analysed and interpreted the data and drafted the manuscript. BW conceptualised the study and analysed the data. AN, CF, IAM and MI acquired and analysed the data. DAL conceptualised the study, analysed the data and contributed to drafting the manuscript. All authors contributed to critically revising the work for important intellectual content, approved the final version and agree to be accountable for all aspects of the work.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data is available to the public from the Demographic Health Surveys (DHS) established by United Stated States Agency for International Development (USAID) available at https://dhsprogram.com.

ORCID iD

Brittany L Kmush http://orcid.org/0000-0003-2143-7554

REFERENCES

1 UNICEF, WHO. *Progress on sanitation and drinking-water: 2015 update and mdg assessment*, 2015.
2 Mara D, Lane J, Scott B, *et al*. Sanitation and health. *PLoS Med* 2010;7:e1000363.
3 Klein SL, Roberts CW. *Sex hormones and immunity to infection*. New York: Springer, 2010.
4 Campbell OMR, Benova L, Gon G, *et al*. Getting the basic rights - the role of water, sanitation and hygiene in maternal and reproductive health: a conceptual framework. *Trop Med Int Health* 2015;20:252–67.
5 Benova L, Cumming O, Campbell OMR. Systematic review and meta-analysis: association between water and sanitation environment and maternal mortality. *Trop Med Int Health* 2014;19:368–87.
6 Alvarez JL, Gil R, Hernández V, *et al*. Factors associated with maternal mortality in Sub-Saharan Africa: an ecological study. *BMC Public Health* 2009;9:462.
7 Andoh SY, Umezaki M, Nakamura K, *et al*. Correlation between national income, HIV/AIDS and political status and mortalities in African countries. *Public Health* 2006;120:624–33.
8 Cheng JJ, Schuster-Wallace CJ, Watt S, *et al*. An ecological quantification of the relationships between water, sanitation and infant, child, and maternal mortality. *Environ Health* 2012;11:4.
9 Guerrant RL, DeBoer MD, Moore SR, *et al*. The impoverished gut-a triple burden of diarrhoea, stunting and chronic disease. *Nat Rev Gastroenterol Hepatol* 2013;10:220–9.
10 Speich B, Croll D, Fürst T, *et al*. Effect of sanitation and water treatment on intestinal protozoa infection: a systematic review and meta-analysis. *Lancet Infect Dis* 2016;16:87–99.
11 Wolf J, Prüss-Ustün A, Cumming O, *et al*. Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression. *Trop Med Int Health* 2014;19:928–42.
12 Strunz EC, Addiss DG, Stocks ME, *et al*. Water, sanitation, hygiene, and soil-transmitted helminth infection: a systematic review and meta-analysis. *PLoS Med* 2014;11:e1001620.
13 Ziegelbauer K, Speich B, Mäusezahl D, *et al*. Effect of sanitation on soil-transmitted helminth infection: systematic review and meta-analysis. *PLoS Med* 2012;9:e1001162.
14 Fuller JA, Eisenberg JNS. Herd protection from drinking water, sanitation, and hygiene interventions. *Am J Trop Med Hyg* 2016;95:1201–10.
15 Fuller JA, Villamor E, Cevallos W, *et al*. I get height with a little help from my friends: herd protection from sanitation on child growth in rural Ecuador. *Int J Epidemiol* 2016;45:460–9.
16 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;374:1032–5.
17 Hammer J, Spears D. Village sanitation and child health: effects and external validity in a randomized field experiment in rural India. *J Health Econ* 2016;48:135–48.
18 Larsen DA, Grisham T, Slawsky E, *et al*. An individual-level meta-analysis assessing the impact of community-level sanitation access on child stunting, anemia, and diarrhea: evidence from DHS and MICs surveys. *PLoS Negl Trop Dis* 2017;11:e0005591.
19 Pickering AJ, Djebbari H, Lopez C, *et al*. Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. *Lancet Glob Health* 2015;3:e701–11.
20 Vyas S, Kov P, Smets S, *et al*. Disease externalities and net nutrition: evidence from changes in sanitation and child height in Cambodia, 2005-2010. *Econ Hum Biol* 2016;23:235–45.
21 Clasen T, Boisson S, Routray P, *et al*. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth

infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *Lancet Glob Health* 2014;2:e645–53.
22 Patil SR, Arnold BF, Salvatore AL, *et al*. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Med* 2014;11:e1001709.
23 ICF. *Demographic and Health Surveys (various) [Datasets]. Funded by USAID*, 2017.
24 USAID. *The DHS Program - Quality information to plan. monitor and improve population. health. and nutrition programs*, 2017.
25 Rutstein SO, Rojas G. *Guide to DHS statistics: demographic and health surveys methodology*. USIAS, 2006.
26 World Health Organization. *Core questions on drinking-water and sanitation for household surveys*. Geneva: World Health Organization, 2006.
27 World Health Organization. *Global reference list of 100 core health indicators. 2015: metadata. 2017*, 2017.
28 Sullivan KM, Mei Z, Grummer-Strawn L, *et al*. Haemoglobin adjustments to define anaemia. *Trop Med Int Health* 2008;13:1267–71.
29 Ho DE, Imai K, King G, *et al*. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011;42.
30 R Development Core Team. *R: a language and environment for statistical computing*, 2008.
31 Lauer JM, Duggan CP, Ausman LM, *et al*. Biomarkers of maternal environmental enteric dysfunction are associated with shorter gestation and reduced length in newborn infants in Uganda. *Am J Clin Nutr* 2018;108:889–96.
32 Gyorkos TW, Gilbert NL. Blood drain: soil-transmitted helminths and anemia in pregnant women. *PLoS Negl Trop Dis* 2014;8:e2912.
33 Crompton DWT, Whitehead RR. Hookworm infections and human iron metabolism. *Parasitology* 1993;107:S137–45.
34 Scott SP, Chen-Edinboro LP, Caulfield LE, *et al*. The impact of anemia on child mortality: an updated review. *Nutrients* 2014;6:5915–32.
35 Oswald WE, Stewart AEP, Kramer MR, *et al*. Association of community sanitation usage with soil-transmitted helminth infections among school-aged children in Amhara region, Ethiopia. *Parasit Vectors* 2017;10:91.
36 Walia B, Kmush BL, Lane SD, *et al*. Routine deworming during antenatal care decreases risk of neonatal mortality and low birthweight: a retrospective cohort of survey data. *PLoS Negl Trop Dis* 2021;15:e0009282.
37 Local Burden of Disease WaSH Collaborators. Mapping geographical inequalities in access to drinking water and sanitation facilities in low-income and middle-income countries, 2000-17. *Lancet Glob Health* 2020;8:e1162–85.
38 Delaire C, Peletz R, Haji S, *et al*. How much will safe sanitation for all cost? Evidence from five cities. *Environ Sci Technol* 2021;55:767–77.
39 Anonymous *State of the World's Sanitation: An urgent call to transform sanitation for better health. environments. economies and societies*. New York: United Nations Children's Fund (UNICEF) and the World Health Organization, 2020.
40 Ejemot RI, Ehiri JE, Meremikwu MM, *et al*. Hand washing for preventing diarrhoea. *Cochrane Database Syst Rev* 2008:CD004265.
41 Ejemot-Nwadiaro RI, Ehiri JE, Arikpo D, *et al*. Hand washing promotion for preventing diarrhoea. *Cochrane Database Syst Rev* 2015:CD004265.

Supplement

Supplementary Table 1: Association between neonatal mortality (death at ≤30 days) and level of sanitation access in the community for women with and without household level access to sanitation.

Household Sanitation	**Community-level sanitation Coverage**	**Unadjusted Incidence Rate Ratio (95% CI)**	***P*-value**	**Adjusted Incidence Rate Ratio (95% CI)**	***P*-value**
		n=832 663		*n*=823 179	
Yes	1-30%	Reference	N/A	Reference	N/A
Yes	31-60%	0.98 (0.89-1.07)	0.64	0.98 (0.89-1.07)	0.60
Yes	61-99%	0.89 (0.82-0.97)	0.01	0.90 (0.83-0.98)	0.02
Yes	100%	0.82 (0.75-0.89)	<0.001	0.84 (0.77-0.92)	<0.001
		n=434 965		*n*=431 683	
No	0%	0.97 (0.92-1.02)	0.27	0.98 (0.93-1.04)	0.51
No	1-30%	Reference	N/A	Reference	N/A
No	31-60%	0.92 (0.86-0.98)	0.01	0.93 (0.87-1.00)	0.04
No	61-99%	0.93 (0.87-1.01)	0.07	0.96 (0.89-1.03)	0.25

Poisson Regression
Unadjusted models included the dataset
Adjusted models included mother's age, mother's age squared, parity, wealth quintile, mother's education, antenatal care quality score, urban or rural, mother's stunting, child's gender, facility delivery or not, and dataset.
N/A-not applicable

Supplementary Table 2: Association between low birth weight (<2500g at birth) and level of sanitation access in the community for women with and without household level access to sanitation.

Household Sanitation Access	**Community-level sanitation Access**	**Unadjusted Odds Ratio (95% CI)**	***P*-value**	**Adjusted Odds Ratio (95% CI)**	***P*-value**
		n=474 789		*n*=472 646	
Yes	1-30%	Reference	N/A	Reference	N/A
Yes	31-60%	0.98 (0.92-1.04)	0.53	0.98 (0.92-1.04)	0.56
Yes	61-99%	0.97 (0.92-1.03)	0.34	0.98 (0.93-1.04)	0.49
Yes	100%	0.89 (0.84-0.94)	<0.001	0.91 (0.86-0.97)	0.002
		n=151 687		*n*=151 321	
No	0%	0.95 (0.91-1.00)	0.04	0.98 (0.94-1.03)	0.43
No	1-30%	Reference	N/A	Reference	N/A
No	31-60%	0.98 (0.93-1.04)	0.49	1.03 (0.97-1.08)	0.37
No	61-99%	0.97 (0.92-1.03)	0.36	1.04 (0.97-1.10)	0.26

Logistic Regression
Unadjusted models included the dataset
Adjusted models included mother's age, mother's age squared, parity, wealth quintile, mother's education, antenatal care quality score, urban or rural, mother's stunting, child's gender, facility delivery or not, and dataset.
N/A-not applicable

Appendix

Statistical Analysis Script

NOTE: Files names and locations will need be changed to match where they are saved on your computer.

Stata Script

```
***********************
***Creating dataset***
***********************
use "/YEAR DHS COUNTRY household file.dta" , clear
gen notoilet = 1 if hv205 == 31
replace notoilet = 0 if hv205!=31 & hv205!=.
egen opendefecation = mean(notoilet), by (hv001)
gen surfacewater = 1 if hv201 == 43
replace surfacewater = 0 if hv201!=43 & hv201!=.
egen communitysurfacewater = mean(surfacewater), by(hv001)
egen communitywealth = mean(hv270), by (hv001)
gen preg = 0 if v213==1
replace preg = 1 if v213==2
gen child_a_twin = b0_1
replace child_a_twin = 1 if b0_1>0
drop if child_a_twin==1
rename hv001 v001
rename hv002 v002
keep v001 v002 hv025 notoilet-communitywealth
merge 1:m v001 v002 using "/YEAR DHS COUNTRY mother file.dta"
drop if _merge==1
drop _merge
gen str countryname = "COUNTRY"
gen str dataset = "COUNTRY DHS YEAR"
keep v001-communitywealth communitywealthtertile-dataset
save "/COUNTRY DHS YEAR.dta", replace

*Repeat for all datasets included in analysis*

use "/COUNTRY DHS YEAR.dta", clear
append using "/COUNTRY DHS YEAR 2.dta"
*Repeat for all datasets included in analysis*
compress
drop if notoilet==.
save "/Compiled.dta", replace
keep if preg==1
save "/Compiled pregos.dta", replace

************************
***Creating variables***
************************

**********************
**Exposure variables**
**********************
```

```
**Household Sanitation**
gen notoilet = 1 if hv205 == 31
replace notoilet = 0 if hv205!=31 & hv205!=.
gen toilet=0
replace toilet=1 if notoilet==0
replace toilet=. if notoilet==.

**Community Sanitation Coverage (1-30% as the Reference Group)**
egen opendefecation = mean(notoilet), by (hv001)
gen commsan=1-opendefecation
gen commsancat=0
replace commsancat=1 if commsan==0
replace commsancat=2 if commsan>.30
replace commsancat=3 if commsan>.60
replace commsancat=4 if commsan==1

**********************
***Outcome variables**
**********************

*Index to birth history*
gen index_to_birth_history = midx_1

*DROPPING THE WOMEN WITHOUT A BIRTH IN LAST 2 YEARS*
drop if midx_1==.

**Neonatal Mortality**
*DHS codes this with a 3 digit number. The first digit (hundreds place) is
1=days, 2=weeks, 3=months. The second and third digit then are the number of
days, weeks or months
gen age_at_death = 999
replace age_at_death = b6_1-100 if b6_1<200
replace age_at_death = 30 if b6_1==201
gen death=0
replace death=1 if age_at_death<=30

**Death at <=28 days (neonatal death)**
gen death2=death
replace death2=0 if age_at_death>=29

**Size of Child at Birth**
*Missing = 99
gen birthsize = m18_1
gen vsmall=1
replace vsmall=0 if birthsize!=5
replace vsmall=. If birthsize>=8

**Birth Weight (kilos - 3dec.)**
**800-6,000, not weighed = 9996, don’t know = 9998, missing = 9999
*2: Not weighed
*99 :don’t know, and missing
gen birthweight = m19_1
gen lbw=1
replace lbw=0 if birthweight>=2500
replace lbw=. If birthweight>=9996

**Anemia**
```

```
*0 no anemia or mild anemia
*1 moderate or severe anemia
gen modsevanemia = 0 if v457>2 & v457!=.
replace modsevanemia = 1 if v457<3

**********************
**Matching Variables**
**********************

*Household wealth in the sampling unit*
egen communitywealth = mean(hv270), by (hv001)
xtile communitywealthtertile = communitywealth, nq(3)

*Access to improved water source in sampling unit*
gen surfacewater = 1 if hv201 == 43
replace surfacewater = 0 if hv201!=43 & hv201!=.
egen communitysurfacewater = mean(surfacewater), by(hv001)
xtile communitywatertertile = communitysurfacewater, nq(3)

*Mother's education - completed primary or not*
*0 coded as incomplete primary or no education
*1 coded as completed primary
gen primaryeduc = 0 if v149!=.
replace primaryeduc = 1 if v149>1 & v149!=.

*Household above or below median household wealth*
*1 coded as below median wealth
*2 coded as above median wealth
xtile richpoor = v191, nq(2)

***********************
*Regression Covariates*
***********************

*Age of the mother at birth (in years) and age squared*
gen age = b2_1-v010
gen age2=age*age

*Parity*
gen parity = 2
replace parity = 1 if bord_1==1
replace parity = 3 if b11_1<25
replace parity = 4 if bord_1>2
replace parity = 5 if bord_1>2 & b11_1<25

*Wealth Quintile*
*1 : 1st quintile
*2 : 2nd quintile
*3 : 3rd quintile
*4 : 4th quintile
*5 : 5th quintile
*99 : missing
gen wealth = v190

*Mother's education*
*0 : none
*1 : some primary
```

```
*2 : completed primary
*3 : higher than primary
*99 : missing
gen educ = 0 if v149==0
replace educ = 1 if v149==1
replace educ = 2 if v149==2
replace educ = 3 if v149>2 & v149!=.

*Antenatal Care Quality*
*During pregnancy: Weighed
*no = 0
*yes = 1
*missing = 99
gen weighed_during_pregnancy = m42a_1
replace weighed_during_pregnancy = . if m42a_1>1

*During pregnancy: Height Measured
*no = 0
*yes = 1
*missing = 99
gen measured_during_pregnancy = m42b_1
replace measured_during_pregnancy = . if m42b_1>1

*During pregnancy: Urine Sample Given
*no = 0
*yes = 1
*missing = 99
gen urine_sample_during_pregnancy = m42d_1
replace urine_sample_during_pregnancy = . if m42d_1>1

*During pregnancy: Blood Pressure Measured
*no = 0
*yes = 1
*missing = 99
gen bp_taken_during_pregnancy = m42c_1
replace bp_taken_during_pregnancy = . if m42c_1>1

*During pregnancy: Blood Sample Given
*no = 0
*yes = 1
*missing = 99
gen blood_sample_during_pregnancy = m42e_1
replace blood_sample_during_pregnancy = . if m42e_1>1

*During Pregnancy: Told About Pregnancy Complications
*no = 0
*yes = 1
*don't know = 99
*missing = 99
gen told_about_pregnancy_comp = m43_1
replace told_about_pregnancy_comp = . if m43_1>1

*During Pregnancy: Took fansidar for malaria
*no = 0
*yes = 1
*don't know = 99
*missing = 99
```

```
gen fansidar_during_pregnancy = m49a_1
replace fansidar_during_pregnancy = . if m49a_1>1

*During Pregnancy: Drugs for Intestinal Parasites
*no = 0
*yes = 1
*don't know = 99
*missing = 99
gen drugs_intestinal_parasites = m60_1
replace drugs_intestinal_parasites = . if m60_1>1

*During Pregnancy: Iron Supplementation
*0 : No
*1 : Yes
gen ironsupplementation = 0 if m45_1==0
replace ironsupplementation = 1 if m45_1==1

*Running principle components analysis of binary variables
pca weighed_during_pregnancy measured_during_pregnancy
bp_taken_during_pregnancy urine_sample_during_pregnancy
blood_sample_during_pregnancy told_about_pregnancy_comp
fansidar_during_pregnancy drugs_intestinal_parasites ironsupplementation
* Predicting the score
predict ancscore
*Creating quintiles
xtile ancquality = ancscore, nq(5)

*Urban/Rural*
*0 : Rural
*1 : Urban
*99 : missing
gen urban = 0 if v025==1
replace urban = 1 if v025==2
replace urban = 99 if urban==.

*Mother's Stunting*
*0 not stunted, at least 150 cm (4.9 ft)
*1 stunted, smaller than 150 cm (4.9 ft)
gen stunted = 0 if v438<2000
replace stunted = 1 if v438<1500

*Sex of Child*
*1 = Male
*0 = Female
gen male = 0
replace male = 1 if b4_1==1

*Place of delivery*
* 0 = in a home
* 1 = in a facility
gen facilitydelivery = 0 if m15_1<20 | m15_1==96
replace facilitydelivery = 1 if m15_1>19 & m15_1!=96 & m15_1!=.

*******************************
***Generalized Linear Models***
*******************************
```

```
import delimited "/Compiled and matched.csv", clear

***REMOVING THE DATASETS WITH <5% OF HOUSEHOLDS NOT HAVING ACCESS TO
SANITATION
by dataset, sort: egen notoiletdataset=mean(notoilet)
drop if notoiletdataset<=.05
egen datasetnumeric= group(dataset)

*Drop Twins
***Child is a Twin***
*0 : No
*1 : Yes
gen child_a_twin = b0_1
replace child_a_twin = 1 if b0_1>0
drop if child_a_twin==1

xtset subclass

**********************
**Neonatal Mortality**
**********************

***Unadjusted regression***

xtpoisson death2 i.commsancat i.datasetnumeric if toilet==1,
exposure(exposure) irr
xtpoisson death2 i.commsancat i.datasetnumeric if toilet==0,
exposure(exposure) irr

xtpoisson death i.commsancat i.datasetnumeric if toilet==1,
exposure(exposure) irr
xtpoisson death i.commsancat i.datasetnumeric if toilet==0,
exposure(exposure) irr

***Adjusted regression***

xtpoisson death2 i.commsancat age age2 i.parity i.wealth i.educ i.ancquality
urban stunted male i.facilitydelivery i.datasetnumeric if toilet==1,
exposure(exposure) irr
xtpoisson death2 i.commsancat age age2 i.parity i.wealth i.educ i.ancquality
urban stunted male i.facilitydelivery i.datasetnumeric if toilet==0,
exposure(exposure) irr

xtpoisson death i.commsancat age age2 i.parity i.wealth i.educ i.ancquality
urban stunted male i.facilitydelivery i.datasetnumeric if toilet==1,
exposure(exposure) irr
xtpoisson death i.commsancat age age2 i.parity i.wealth i.educ i.ancquality
urban stunted male i.facilitydelivery i.datasetnumeric if toilet==0,
exposure(exposure) irr

*******************
**Low Birthweight**
*******************

***Unadjusted regression***

xtlogit lbw i.commsancat i.datasetnumeric if toilet==1, or
```

```
xtlogit lbw i.commsancat i.datasetnumeric if toilet==0, or

xtlogit vsmall i.commsancat i.datasetnumeric if toilet==1, or
xtlogit vsmall i.commsancat i.datasetnumeric if toilet==0, or

***Adjusted regression***

xtlogit lbw i.commsancat age age2 i.parity i.wealth i.educ i.ancquality urban
stunted male facilitydelivery i.datasetnumeric if toilet==1, or
xtlogit lbw i.commsancat age age2 i.parity i.wealth i.educ i.ancquality urban
stunted male facilitydelivery i.datasetnumeric if toilet==0, or

xtlogit vsmall i.commsancat age age2 i.parity i.wealth i.educ i.ancquality
urban stunted male facilitydelivery i.datasetnumeric if toilet==1, or
xtlogit vsmall i.commsancat age age2 i.parity i.wealth i.educ i.ancquality
urban stunted male facilitydelivery i.datasetnumeric if toilet==0, or

**********
**Anemia**
**********

***Unadjusted regression***

xtlogit modsevanemia i.commsancat i.datasetnumeric if toilet==1, or
xtlogit modsevanemia i.commsancat i.datasetnumeric if toilet==0, or

***Adjusted regression***

xtlogit modsevanemia i.commsancat age age2 i.parity i.wealth i.educ hv025
i.datasetnumeric if toilet==1, or
xtlogit modsevanemia i.commsancat age age2 i.parity i.wealth i.educ hv025
i.datasetnumeric if toilet==0, or

**Anemia in pregnant women**

import delimited "/Compiled and matched anemia pregos.csv", clear

***REMOVING THE DATASETS WITH <5% OF HOUSEHOLDS NOT HAVING ACCESS TO
SANITATION
by dataset, sort: egen notoiletdataset=mean(notoilet)
drop if notoiletdataset<=.05
egen datasetnumeric= group(dataset)

xtset subclass

***Unadjusted regression***

xtlogit modsevanemia i.commsancat i.datasetnumeric if toilet==1, or
xtlogit modsevanemia i.commsancat i.datasetnumeric if toilet==0, or

***Adjusted regression***

xtlogit modsevanemia i.commsancat age age2 i.parity i.wealth i.educ hv025
i.datasetnumeric if toilet==1, or
xtlogit modsevanemia i.commsancat age age2 i.parity i.wealth i.educ hv025
i.datasetnumeric if toilet==0, or
```

R Script

```
library(MatchIt)
library(foreign)

###LOADING DATASETS, setting missing values to 9999###
d = read.dta("/Compiled dataset.dta")

### RECLASSIFYING MISSING VALUES ###
d[is.na(d)] = 9999

###EXACT MATCHING###
m.out = matchit(notoilet~dataset + communitywealthtertile +
communitywatertertile + primaryeduc + richpoor, data=d, method="exact")

###CREATING DATASET WITH MATCHED VARIABLE SUBCLASS###
m.data = match.data(m.out)

###EXPORTING DATASET TO DATAFRAME###
d.matched = data.frame(m.data)

###WRITING TEXT FILES
write.table(d.matched, file="/Compiled and matched.csv", sep=",",
col.names=TRUE, row.names=FALSE)

### PREGNANT WOMEN ###

###LOADING DATASETS, setting missing values to 9999###
d.prego = subset(d, preg==1)

###EXACT MATCHING###
m.out.prego = matchit(notoilet~dataset + communitywealthtertile +
communitywatertertile + primaryeduc + richpoor, data=d.prego, method="exact")

###CREATING DATASET WITH MATCHED VARIABLE SUBCLASS###
m.data.prego = match.data(m.out.prego)

###EXPORTING DATASET TO DATAFRAME###
d.matched.prego = data.frame(m.data.prego)

###WRITING TEXT FILES
write.table(d.matched.prego, file="/Compiled and matched anemia pregos.csv",
sep=",", col.names=TRUE, row.names=FALSE)
```