Supplementary Abstract (following CONSORT 2010 guidelines for abstracts)

Background

Healthy development of the gut microbiome provides long-term health benefits. Children raised in countries with high infectious disease burdens, like Bangladesh, are frequently exposed to antibiotics and diarrheal pathogens, which perturb gut microbiome assembly. A recent double-blind cluster-randomized controlled in two low-income, densely populated communities in urban Bangladesh found automated water chlorination of shared taps to be an effective strategy for reducing child diarrhea and antibiotic use. Here, we performed exploratory analyses to evaluate the effect of this intervention on children's gut microbiota, including the bacterial pathogens and antibiotic resistance genes (ARGs) they harbored.

Methods

The trial was implemented from July 2015 – December 2016 in two low-income communities in urban Bangladesh: Tongi, a community outside Dhaka city, and Dhaka Uddan, a community within Dhaka city. In brief, 100 shared water taps that served as the primary source of drinking water for children younger than five years old were identified in both communities, then randomly assigned (1:1) to have their drinking water automatically chlorinated at the point of collection by a solid tablet chlorine doser (intervention group) or to be treated by a visually identical doser that supplied vitamin C (active control group). Approximately 500 children were enrolled in each group at baseline. Stool samples were collected one year after the start of the intervention. Following a child's stool production, caretakers were instructed to inoculate a small amount of stool in RNALater (a fecal preservative). Field staff then transported samples to the laboratory, where they were frozen at -80C upon arrival and remained frozen during subsequent shipment to the United States. Study staff stratified available RNALater-preserved stool samples by group, study site, and three pre-specified age strata (6-14 months, 15-30 months, 31 months and older) corresponding to distinct phases of gut microbiome development, then randomly selected samples for short-read, paired-end 150 bp sequencing of total stool DNA. The primary outcome was differentially abundant bacterial genera between treatment and control children across different phases of gut microbiome development. This analysis was not pre-specified in the original trial. Both study participants and researchers selecting samples, processing samples, and performing data analysis were unaware of which households were served by chlorinated taps (double-blinded). This trial is registered with ClinicalTrials.gov, number NCT02606981, and is completed.

Findings

We examined fecal metagenomes from 130 children from the control (n=64) and treatment groups (n=66). Water chlorination was associated with increased abundance of human enterobacteria, but shifts were small in magnitude. We observed no effects on the overall richness or diversity of taxa, and the prevalence of bacterial pathogens was similar across the two groups. However, several clinically relevant ARGs were relatively more abundant in the gut microbiomes of treatment children.

Interpretation

Water chlorination affected the developing gut microbiome of children in urban Bangladesh, including the resistance genes they harbored, though shifts in taxa abundance were generally small in magnitude. While further studies on the long-term health impacts of drinking chlorinated water would be valuable, we conclude that access to chlorinated water did not substantially impact child gut microbiome development in this setting, supporting the use of chlorination to increase global access to safe drinking water.

Funding

The Thrasher Research Fund and The World Bank Strategic Impact Evaluation Fund.

CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	Title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[i,ii]	See table 2	Supplementary Material
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	Pickering *et al.*, 2019: Methods, Participants Nadimpalli *et al.*, 2021: Not updated for this study
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	Introduction
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	Pickering *et al.*, 2019: Methods, Participants Nadimpalli *et al.*, 2021: Not updated for this study
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		Pickering *et al.*, 2019: N/A Nadimpalli *et al.*, 2021: Not updated for this study
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	Pickering *et al.*, 2019: Methods, Participants Nadimpalli *et al.*, 2021: Methods – Stool collection
	4b	Settings and locations where the data were collected		Pickering *et al.*, 2019: Methods, Study design Nadimpalli *et al.*, 2021: Methods – Stool Collection
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	Pickering *et al.*, 2019: Methods, Intervention delivery Nadimpalli *et al.*, 2021: Methods – Stool collection
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how	Whether outcome measures pertain to the cluster level, the individual participant level or both	Pickering *et al.*, 2019: Methods – outcomes

		and when they were assessed		Nadimpalli *et al.*, 2021: Methods – Data Analysis
	6b	Any changes to trial outcomes after the trial commenced, with reasons		Pickering *et al.*, 2019: Methods – outcomes Nadimpalli *et al.*, 2021: N/A for this study
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	Pickering *et al.*, 2019: Methods – statistical analysis Nadimpalli *et al.*, 2021: The original trial was powered to detect reductions in diarrhea. Sample size calculations were not performed for this sub study.
	7b	When applicable, explanation of any interim analyses and stopping guidelines		Pickering *et al.*, 2019: Procedures Nadimpalli *et al.*, 2021: Not updated for this study
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		Pickering *et al.*, 2019: Methods, Randomisation Nadimpalli *et al.*, 2021: Not updated for this study
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	Pickering *et al.*, 2019: Methods, Randomisation Nadimpalli *et al.*, 2021: Methods – Sample Selection for metagenomic sequencing
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	Pickering *et al.*, 2019: Methods, Randomisation Nadimpalli *et al.*, 2021: Not updated for this study
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	Pickering *et al.*, 2019: Methods, Randomisation Nadimpalli *et al.*, 2021: Not updated for this study
	10a		Who generated the random allocation sequence, who	Pickering *et al.*, 2019: Methods, Randomisation

			enrolled clusters, and who assigned clusters to interventions	Nadimpalli *et al.*, 2021: Not updated for this study
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	Pickering *et al.*, 2019: Methods, Randomisation Nadimpalli *et al.*, 2021: Not updated for this study
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	Pickering *et al.*, 2019: Methods, Study design Nadimpalli *et al.*, 2021: Not updated for this study
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		Pickering *et al.*, 2019: Methods, intervention delivery and masking Nadimpalli *et al.*, 2021: Methods – Stool collection
	11b	If relevant, description of the similarity of interventions		Pickering *et al.*, 2019: Methods, intervention delivery and masking Nadimpalli *et al.*, 2021: Not updated for this study
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	Pickering *et al.*, 2019: Methods – statistical analysis Nadimpalli *et al.*, 2021: Methods – Data analysis
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		Pickering *et al.*, 2019: Methods – statistical analysis Nadimpalli *et al.*, 2021: Methods – Data analysis, Results
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	Pickering *et al.*, 2019: Figure 1 Nadimpalli *et al.*, 2021: Table 1
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	Pickering *et al.*, 2019: Figure 1 Nadimpalli *et al.*, 2021: Not updated for this study

Recruitment	14a	Dates defining the periods of recruitment and follow-up		Pickering *et al.*, 2019: Results Nadimpalli *et al.*, 2021: Methods – Stool Collection
	14b	Why the trial ended or was stopped		Pickering *et al.*, 2019: Results and Discussion Nadimpalli *et al.*, 2021: Not updated for this study
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Baseline characteristics for the individual and cluster levels as applicable for each group	Pickering *et al.*, 2019: Table 1 Nadimpalli *et al.*, 2021: Table 1
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	Pickering *et al.*, 2019: Figure 1 Nadimpalli *et al.*, 2021: Table 1
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	Pickering *et al.*, 2019: Figure 1, Table 3 Nadimpalli *et al.*, 2021: Results, Figure 1, Figure 3, Table 2
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		Pickering *et al.*, 2019: Table 3 Nadimpalli *et al.*, 2021: Table 2
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		Pickering *et al.*, 2019: Results, Table 2, Table 3 Nadimpalli *et al.*, 2021: Results, Figure 2, Supplementary Material
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[iii])		Pickering *et al.*, 2019: Results Nadimpalli *et al.*, 2021: Not updated for this study, as focused on secondary sample analysis
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		Pickering *et al.*, 2019: Discussion Nadimpalli *et al.*, 2021: Discussion
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	Pickering *et al.*, 2019: Discussion

			Nadimpalli *et al.*, 2021: Results, Discussion
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Pickering *et al.*, 2019: Discussion Nadimpalli *et al.*, 2021: Discussion
Other information			
Registration	23	Registration number and name of trial registry	Pickering *et al.*, 2019: Methods Nadimpalli *et al.*, 2021: Not updated for this study
Protocol	24	Where the full trial protocol can be accessed, if available	Pickering *et al.*, 2019: Published on Open Science Framework, link in Methods Nadimpalli *et al.*, 2021: Not updated for this study
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	Pickering *et al.*, 2019: Abstract, Role of Funder section Nadimpalli *et al.*, 2021: Acknowledgements, Role of the Funding Source

** Note: page numbers optional depending on journal requirements*

Supplementary Tables

Table 1. Differentially abundant bacterial genera among 130 treatment and control children participating in an automated water chlorination intervention trial in Bangladesh, overall and by three age strata corresponding to distinct phases of gut microbiome development.

Genera	Treatment Coefficient (95% CI)[a]	fdr-corrected *p-value*[b]
Overall		
Akkermansia	2.42 (1.87, 2.98)	4.55E-14
Escherichia	1.11 (0.67, 1.55)	2.20E-06
Flavonifractor	0.89 (0.5, 1.27)	1.83E-05
Phascolarctobacterium	2.11 (1.54, 2.69)	6.44E-11
Age 6-14 months		
Alysiella	0.99 (0.46, 1.52)	0.00148608
Aureimonas	0.26 (0.13, 0.4)	0.00097828
Bombilactobacillus	1.31 (0.74, 1.88)	0.00020148
Bremerella	1.14 (0.63, 1.65)	0.00025458
Candidatus Nitrotoga	1.22 (0.84, 1.6)	2.8404E-06
Candidatus Reidiella	-1.24 (-1.79, -0.68)	0.00025753
Candidatus Vampirococcus	-1.15 (-1.66, -0.63)	0.00025989
Chromohalobacter	-0.82 (-1.14, -0.51)	4.3267E-05
Ferriphaselus	1.41 (0.73, 2.09)	0.00055415
Furfurilactobacillus	-1.87 (-2.4, -1.33)	8.967E-07
Fusobacterium	2.55 (1.37, 3.74)	0.00037227
Inhella	0.67 (0.42, 0.93)	4.1354E-05
Jinshanibacter	1.11 (0.65, 1.58)	0.00013004
Lactobacillus	-4.59 (-5.8, -3.39)	2.3712E-07
Laribacter	1.23 (0.74, 1.73)	7.5377E-05
Latilactobacillus	-0.78 (-1.03, -0.53)	4.8752E-06
Leuconostoc	-3.22 (-4.52, -1.92)	8.3036E-05
Mariprofundus	0.48 (0.24, 0.71)	0.00064749
Mycetohabitans	2.12 (1.73, 2.51)	6.8067E-10
Natronoglycomyces	-2.06 (-3.13, -1)	0.00107337
Nitrosospira	-1.14 (-1.73, -0.55)	0.00104238
Paenarthrobacter	-1.16 (-1.77, -0.56)	0.00109374
Paraphotobacterium	1.79 (0.91, 2.67)	0.00068419
Phascolarctobacterium	2.68 (1.45, 3.92)	0.00036026
Plesiomonas	5.04 (3.31, 6.77)	1.1257E-05
Propionimicrobium	1.15 (0.76, 1.53)	8.2463E-06
Rothia	-1.06 (-1.7, -0.43)	0.00358964
Sphingosinithalassobacter	-0.97 (-1.29, -0.65)	6.1356E-06

Streptococcus	-1.66 (-2.55, -0.77)	0.00144412
Telmatocola	1.68 (1.11, 2.24)	9.4924E-06
Terricaulis	-2.13 (-3.06, -1.21)	0.00019193
Thermanaerovibrio	2.46 (1.3, 3.63)	0.0004472
Thermodesulfatator	-2.02 (-2.95, -1.09)	0.00035531
Usitatibacter	0.39 (0.21, 0.58)	0.0004252
Age 15-30 months		
Bergeyella	-1.11 (-1.62, -0.59)	1.22E-04
Casimicrobium	-0.95 (-1.5, -0.4)	1.40E-03
Ciceribacter	-1.46 (-2.29, -0.63)	1.26E-03
Comamonas	-1.49 (-2.2, -0.78)	1.54E-04
Erysipelatoclostridium	0.95 (0.49, 1.41)	2.19E-04
Flavonifractor	1.76 (1.27, 2.25)	9.23E-09
Fusobacterium	1.07 (0.46, 1.69)	1.36E-03
Lancefieldella	1.15 (0.61, 1.7)	1.46E-04
Leuconostoc	-2.95 (-4.02, -1.89)	2.24E-06
Methylacidimicrobium	-0.84 (-1.28, -0.39)	6.42E-04
Microcystis	-0.85 (-1.33, -0.37)	1.24E-03
Parabacteroides	1.19 (0.52, 1.85)	1.05E-03
Pasteurella	-0.65 (-0.96, -0.34)	1.55E-04
Phascolarctobacterium	2.3 (1.44, 3.16)	4.12E-06
Piscirickettsia	-0.95 (-1.45, -0.44)	6.16E-04
Psychrobacter	-0.61 (-0.94, -0.28)	7.69E-04
Tuwongella	-0.9 (-1.4, -0.4)	1.06E-03
Verrucomicrobium	-0.75 (-1.18, -0.32)	1.24E-03
Weissella	-1.76 (-2.77, -0.76)	1.28E-03
Age 31-61 months		
Akkermansia	3.03 (2.09, 3.98)	1.01E-07

[a]Treatment coefficients generated by the R package *corncob*, representing the additive change in the logit-transformed relative abundance of bacterial genera between treatment and control children.

[b] *p*-value generated by *corncob*. The Benjamini-Hochberg method was used to correct for multiple comparisons.

Table 2. Effect of extended exposure to a water chlorination intervention (greater than or equal to 6 months) on differentially abundant genera among treatment and control children aged 15 months and older.

	15-61 months (n=103)		15-61 months and exposed to the intervention for at least 6 months (n=91)	
Genera	Treat. Coef. (95% CI)[a]	fdr-corrected *p-value*[b]	Treat. Coef. (95% CI)[a]	fdr-corrected *p-value*[b]
Akkermansia	2.44 (1.84, 3.03)	1.90E-12	0.64 (0.22, 1.07)	0.179423723
Flavonifractor	0.92 (0.5, 1.33)	0.00777148	0.83 (0.38, 1.28)	0.097962071
Phascolarctobacterium	1.97 (1.34, 2.59)	5.79E-07	1.95 (1.25, 2.64)	5.60E-05

[a]Treatment coefficients generated by the R package *corncob*, representing the additive change in the logit-transformed relative abundance of bacterial genera between treatment and control children.

[b] *p*-value generated by *corncob*. The Benjamini-Hochberg method was used to correct for multiple comparisons.

Table 3. Detection of 14 gastrointestinal pathogens in the stool of 527 children participating in a cluster-randomized automated water chlorination trial.

	Control n=278 (%)	Treatment n=249 (%)	RR (95% CI)	Adjusted *p*-value
Norovirus GI/GII	51 (18)	31 (12)	0.65 (0.41, 1.02)	0.56
Campylobacter	55 (20)	49 (20)	0.93 (0.63, 1.37)	0.78
Salmonella	94 (34)	79 (32)	0.93 (0.69, 1.26)	0.78
Enterotoxigenic *E. coli* (ETEC) LT/ST	77 (28)	66 (27)	0.93 (0.67, 1.30)	0.78
Shigella	71 (26)	73 (29)	1.15 (0.83, 1.60)	0.78
Giardia	105 (38)	102 (41)	1.05 (0.80, 1.39)	0.78
Pathogenic *E. coli*[a]	132 (48)	123 (49)	1.05 (0.82, 1.34)	0.78
Cryptosporidium	12 (4)	8 (3)	0.70 (0.27, 1.72)	0.78
C. difficile	12 (4)	9 (4)	0.94 (0.38, 2.23)	0.89
Shiga-like toxin-producing *E. coli* (STEC) stx1/stx2	21 (8)	9 (4)	0.52 (0.22, 1.10)	0.56
Adenovirus 40/41	8 (3)	5 (2)	0.69 (0.21, 2.11)	0.78
Rotavirus A	1 (0)	3 (1)	--	
Yersinia enterocolitica	--	--	--	
Vibrio cholerae	1 (0)	--	--	
Entamoeba histolytica	2 (1)	1 (0)	--	

Note: Relative risk ratios rates (RR) were calculated using Poisson regression models adjusted for child's age and study site. Resulting two-sided *p*-values were adjusted for multiple comparisons using the Benjamini–Hochberg method. RRs, associated 95% CIs, and adjusted p-values are only presented for pathogens that were detected among at least 1% of samples; models failed to converge below this threshold. [a]Defined as any of the following: ETEC, STEC, or *Shigella.*