

1 Linking community water and sanitation access to the global burden of antibiotic
2 resistance using human gut metagenomes from 26 countries

3
4

5 Erica R. Fuhrmeister^{1,2}, Abigail P. Harvey², Maya L. Nadimpalli^{1,3}, Karin Gallandat⁴, Argaw
6 Ambelu⁵, Benjamin F. Arnold⁶, Joe Brown⁷, Oliver Cumming⁴, Ashlee M. Earl⁸, Gagandeep
7 Kang⁹, Samuel Kariuki¹⁰, Karen Levy¹¹, Chris Pinto¹², Jenna M. Swarthout¹, Gabriel Trueba¹³,
8 Pablo Tsukayama¹⁴, Colin J. Worby⁸, Amy J. Pickering^{1,2,15}

9

10 ¹Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA

11 ²Department of Civil and Environmental Engineering, University of California, Berkeley, CA,
12 USA

13 ³Stuart B. Levy Center for Integrated Management of Antimicrobial Resistance, Tufts University,
14 Boston, MA, USA

15 ⁴Department of Disease Control, London School of Hygiene and Tropical Medicine, London,
16 United Kingdom

17 ⁵Faculty of Public Health, Jimma University, Jimma, Ethiopia

18 ⁶Francis I. Proctor Foundation, University of California, San Francisco, CA, USA

19 ⁷Department of Environmental Sciences and Engineering, Gillings School of Global Public
20 Health, University of North Carolina, Chapel Hill, NC, USA

21 ⁸Infectious Disease & Microbiome Program, Broad Institute, Cambridge, MA, USA

22 ⁹The Wellcome Trust Research Laboratory, Division of Gastrointestinal Sciences, Christian
23 Medical College, Vellore, India

24 ¹⁰Centre for Microbiology Research, Kenya Medical Research Institute, Nairobi, Kenya

25 ¹¹Department of Environmental and Occupational Health Sciences, School of Public Health,
26 University of Washington, Seattle, Washington, USA

27 ¹²Department of Global Health and Development, Faculty of Public Health and Policy, London
28 School of Hygiene and Tropical Medicine, London, United Kingdom

29 ¹³Instituto de Microbiología, Colegio de Ciencias Biológicas y Ambientales, Universidad San
30 Francisco de Quito, Quito, Ecuador

31 ¹⁴Laboratorio de Genómica Microbiana, Facultad de Ciencias y Filosofía, Universidad Peruana
32 Cayetano Heredia, Lima, Peru

33 ¹⁵Blum Center for Developing Economies, University of California, Berkeley, CA, USA

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49 **Abstract**

50 Background: Antibiotic resistance is a leading cause of death, with the highest burden in low-
51 resource settings. There is limited evidence on the potential for water, sanitation, and hygiene
52 (WASH) infrastructure to reduce the burden of antibiotic resistance in humans.

53 Methods: We used geospatially tagged human gut metagenomes and household survey datasets
54 to determine the association between antibiotic resistance gene (ARG) abundance and
55 community-level coverage of improved drinking water points and improved sanitation facilities.
56 Adjusted general linearized models with robust standard errors were used to estimate the
57 relationship between ARG abundance in the human gut and access to water and sanitation.

58 Findings: We identified 1589 publicly available metagenomes from 26 countries. The average
59 abundance of ARGs, in units of \log_{10} ARG reads per kilobase per million (RPKM) mapped reads
60 classified as bacteria, was highest in Africa compared to other World Health Organization (WHO)
61 regions (one-way ANOVA $p < 0.001$, post hoc Tukey HSD $p < 0.05$). Increased access to both
62 improved water and sanitation was associated with lower ARG abundance (effect estimate: -0.26,
63 95% CI [-0.44, -0.08]); the association was stronger in urban (-0.37 [-0.68, -0.07]) compared to
64 rural areas (-0.16 [-0.38, 0.07]). Improved sanitation alone was associated with reduced ARG
65 abundance (-0.16 [-0.32, 0.00]) while improved drinking water was not (-0.09 [-0.35, 0.16]).

66 Interpretation: While additional studies to investigate casual effects are needed, increasing
67 access to water and sanitation could be an effective strategy to curb the proliferation of antibiotic
68 resistance in low- and middle-income countries.

69 Funding: Bill & Melinda Gates Foundation

70

71 **Research in Context**

72 Evidence before this study

73 Antibiotic resistance is a growing global health threat that disproportionately affects low- and
74 middle-income countries (LMICs). In 2019, an estimated 5 million deaths were associated with

75 antibiotic resistance, with the highest death rate in western sub-Saharan Africa. Water, sanitation,
76 and hygiene (WASH) interventions (e.g., household drinking water treatment, flush toilet, hand
77 washing facilities with soap) can reduce diarrheal and respiratory infections, as reported in
78 previous meta-analyses. Estimates, based on probability modeling, suggest improvements in
79 water and sanitation could decrease antibiotic use for diarrheal disease treatment by 47-50% and
80 69-72%, respectively. Improving WASH infrastructure could theoretically contribute to the control
81 of antibiotic resistance by preventing the release of antibiotics, resistant organisms, or antibiotic
82 resistance genes (ARGs) into the environment, thus decreasing the burden of antibiotic-resistant
83 infections. One global analysis across 73 countries suggested that improved infrastructure,
84 including WASH services, was associated with reduced antibiotic resistance prevalence in
85 isolates, however the independent effect of WASH access was not assessed.

86
87 We searched PubMed for evidence on the impact of WASH interventions (excluding those related
88 to animals and agriculture) on antibiotic resistance using the following keyword chain: (*water OR*
89 *sanitation OR hygiene OR WASH*) *AND* (*antimicrobial OR antibiotic*) *AND* *resistance*) *NOT*
90 (*"OneHealth" OR "One Health" OR animal OR livestock*). We selected reviews and systematic
91 reviews (n=1420) to be screened for relevance to WASH and antibiotic resistance. The reference
92 lists of included reviews were then searched for individual studies. We also consulted international
93 agency guidelines and online resources from the Joint Programming Initiative on Antimicrobial
94 Resistance, the International Scientific Forum on Hygiene, ReAct, Resistomap, and the London
95 School of Hygiene and Tropical Medicine AMR Centre.

96
97 Studies focusing on centralized water or wastewater treatment technologies in high income
98 countries (HICs) reported variable removal of antibiotics (53 to >90%), antibiotic-resistant bacteria
99 (90-99.9%) and ARGs (90-99.9%) from waste streams. Other studies were conducted on hand
100 hygiene, which has proven effective at reducing human infections and antibiotic use. No studies

101 were identified on the effect of on-site sanitation systems (e.g., pour-flush toilets, pit latrines),
102 which serve an estimated 2.7 billion people globally, or fecal sludge management interventions
103 on antibiotic resistance.

104

105 Added value of this study

106 In this study, we used 1589 publicly available human gut metagenomes from around the world to
107 assess the abundance of ARGs as a function of access to improved drinking water and sanitation
108 infrastructure. This analysis provides new evidence of differences in the abundance of antibiotic
109 resistance in the human gut across the world and finds that decreased gut abundance of ARGs
110 is associated with increased access to improved drinking water and sanitation.

111

112 Implications of all the available evidence

113 Current approaches to controlling antibiotic resistance in humans predominantly focus on
114 antibiotic stewardship; however, this approach is challenging in LMICs where infectious illnesses
115 are generally more prevalent and unregulated antibiotic usage is common. Along with efforts to
116 provide other known social benefits, such as reducing infectious disease and improving gender
117 equality, improving access to safe drinking water and sanitation could contribute to reducing the
118 burden of antibiotic resistance. This work highlights improving access to adequate water and
119 sanitation as a potentially effective strategy, although additional studies designed to rigorously
120 investigate the casual relationship between WASH and antibiotic resistance are needed.

121

122 **Introduction**

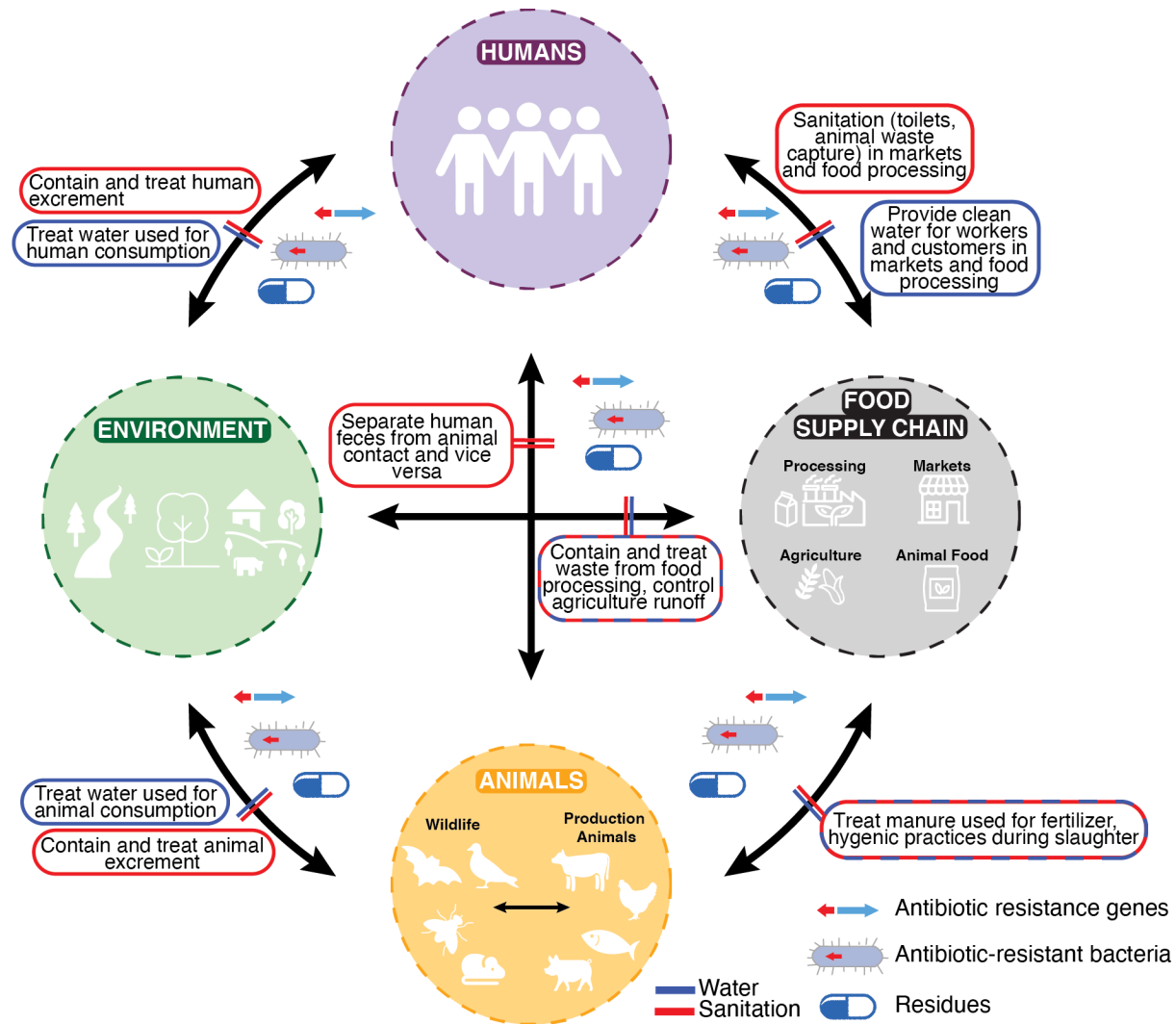
123 Increasing rates of antibiotic resistance among bacterial pathogens are a major public health
124 concern¹. Many of the most concerning resistance phenotypes currently observed in the clinical
125 setting are conferred through antibiotic resistance genes (ARGs), which can be acquired from
126 pathogenic or commensal bacteria.² In the past decade, our increased capacity to cost-effectively

127 sequence the DNA of microbial communities (*i.e.*, metagenomic sequencing) has revealed how
128 common ARGs are among bacteria that colonize humans, animals, and environmental niches,³
129 highlighting the scale of the challenge of preventing pathogens from acquiring ARGs.

130

131 Low-and middle-income countries (LMICs) have the highest burden of antibiotic-resistant
132 infections,¹ and some of the most concerning ARGs (e.g., *bla*_{NDM-1} conferring beta-lactam and
133 *mcr-1* conferring colistin resistance) are believed to have first mobilized into human pathogens in
134 LMICs.^{4,5} Inadequate water, sanitation and hygiene (WASH) infrastructure could exacerbate the
135 spread of resistance.⁶ Clean drinking water and flush toilets are nearly universal in most high-
136 income countries (HICs), but variable to non-existent in many LMIC settings. Access to improved
137 WASH could prevent the emergence, transfer and spread of antibiotic resistance in many ways.⁷
138 Here, we have depicted potential interventions to stop spill over between humans, animals, the
139 environment, and food supply systems (**Figure 1**). First, use of toilets (improved sanitation) and
140 wastewater treatment greatly limits the load of antibiotic-resistant bacteria and ARGs that could
141 otherwise be excreted into the environment. *E. coli*, which is enriched in excreta was implicated
142 in 23% of total global deaths attributable to antibiotic resistance in 2019.¹ Uninterrupted access
143 to clean drinking water provides an additional barrier to exposure. Second, by reducing the overall
144 infectious disease burden and accompanying need for antibiotic use, WASH services could
145 indirectly reduce the population-level antibiotic selection mechanisms that support the
146 propagation of resistant strains.⁸ Previous global analyses leveraging urban sewage samples
147 have found that country-level WASH access is associated with ARG burden, lending support to
148 this hypothesis.⁹ However, given that the majority of human fecal waste generated in LMICs does
149 not enter centralized sewer systems,¹⁰ investigations that rely on urban sewage samples are not
150 representative.

151



152
 153 **Figure 1.** Potential water (blue) and sanitation (red) interventions needed to block the transmission of
 154 antibiotic resistance genes and antibiotic-resistant bacteria as well as dissemination of antibiotic residues
 155 between humans, animals, the environment, and the food supply chain.
 156

157 Here, we leverage publicly available human gut metagenomes worldwide to describe regional
 158 trends and investigate the association between community levels of drinking water and sanitation
 159 access and ARG abundance in the human gut. Gut metagenome data from global populations
 160 have become increasingly available over the past decade, including from LMICs that lack
 161 centralized water and sanitation infrastructure. The overall goal of our analysis was to examine
 162 regional differences in ARG abundance and contribute new evidence on whether limited WASH
 163 infrastructure may be a critical driver of human antibiotic resistance carriage.

164

165 **Methods**

166 Metagenome Identification

167 We obtained publicly available human gut metagenomes, resulting from short-read (Illumina)
168 whole-genome sequencing, from the NCBI sequence read archives (SRA). We included all
169 available metagenomes from LMICs and a subset of metagenomes from HICs. Studies from HICs
170 were selected to span varying human ages and geographic locations. Longitude and latitude
171 coordinates were obtained from the SRA metagenome metadata and verified with the
172 corresponding research article. If coordinates were unavailable in the SRA, location was
173 determined based on the description in the research article. Additional human gut metagenomes
174 were contributed from Mozambique, Bangladesh, and Kenya by authors of these studies.
175 Metagenomes were excluded if location and combined access to both drinking water and
176 sanitation could not be determined.

177

178 Antibiotic Resistance Gene Identification and Normalization

179 ARGs were identified by mapping reads to the Comprehensive Antibiotic Resistance Database
180 (CARD)(v 3.0.9)¹¹ of ARG protein sequences using BLASTX in DIAMOND (v 0.9.30.131).¹²
181 Results were filtered using an alignment cutoff of 25 amino acids and an identity of 95%. Paired-
182 end reads were mapped separately, and duplicate mappings were removed. To normalize for
183 total bacteria in each sample, reads per kilobase per million mapped reads (RPKM) of ARGs were
184 calculated using the total number of reads classified as bacteria (using Bracken¹³) for total reads.

185

186 Taxonomy

187 Kraken2¹⁴ with the standard database was used for taxonomic classification of short reads.
188 Bracken¹³ corrected counts were used to calculate the total number of reads classified as bacteria

189 as well as the relative abundance of the five most abundant families among reads classified as
190 bacteria.

191

192 Survey Methods

193 Household variables, including access to improved drinking water and sanitation, were primarily
194 obtained from geospatially tagged, nationally representative household survey data sets.
195 Demographic and Health Surveys (DHS) (<https://www.dhsprogram.com>) were obtained for the
196 year closest to the study date (2010-2018). When DHS data were not available (El Salvador,
197 Mongolia, and Mexico), we obtained Multiple Indicator Cluster Surveys (MICS)
198 (<http://mics.unicef.org/surveys>) from the closest available year to the study date (appendix Table
199 A1). For the countries without DHS or MICS data (Ecuador, China), we obtained comparable
200 country-level or study-specific survey data (Table A1). If the study year was unknown (N=9), we
201 used the most recent survey available.

202

203 For countries with DHS data (which includes GPS coordinates for survey clusters), survey clusters
204 located within a 25 km radius of the metagenome coordinates were selected. Two additional radii
205 (50 km and 75 km) were used to conduct a sensitivity analysis of radius threshold. For countries
206 without DHS data (which contains information on the region or district), we determined the locally
207 defined administrative area containing the metagenome coordinate and filtered to households
208 within that administrative area. Within the defined geographic region (radius or administrative
209 area), we selected urban, rural, or both types of household data based on the authors' description
210 of the study site. From survey data joined to each metagenome dataset, we calculated the
211 proportion of households in the geographically defined radius with access to improved sanitation
212 (separates excreta from human contact), access to improved drinking water (design protects
213 source from contamination), and access to both improved sanitation and improved drinking water
214 (definitions in appendix Table A2). We were unable to investigate access to safely managed

215 sanitation and drinking water, as defined by the Sustainable Development Goals,¹⁵ due to the
216 absence of estimates for most LMICs in our analysis. Since access to improved drinking water
217 and sanitation could not be independently determined for HICs, we assigned a value of 99% for
218 combined access. We calculated the proportion of children using antibiotics based on reported
219 use in response to diarrhea or fever in the previous two weeks. In addition, we determined the
220 proportion of households owning livestock and other assets. We used consumption estimates
221 (Defined Daily Dose (DDDs) per 1000 people per day) to estimate country-level antibiotic usage.
222 Additional data obtained included antibiotic consumption estimates (DDDs per 1000 people per
223 day)¹⁶, gross domestic product (GDP) per capita¹⁷, income classification per country as
224 determined by World Bank designation in 2021, population density¹⁸, and animal antibiotic
225 consumption in 2010.¹⁹

226

227 Data Analysis

228 We used generalized linear models to model log₁₀-transformed, normalized abundance of ARGs
229 (in total, by drug class, and individually) as a function of local access to both improved sanitation
230 and improved drinking water in R (v 4.0.5). In separate analyses, total abundance was modeled
231 as a function of local access to improved drinking water and improved sanitation individually.
232 Robust standard errors were used to account for clustered data (multiple samples in one
233 georeferenced cluster). World Health Organization (WHO) region and population density were
234 included as covariates in all models. Two adjusted analyses were conducted. First, we included
235 additional covariates that were available for the full set of metagenomes (antibiotic usage in
236 humans, GDP per capita, library layout-paired or single-, and average read length). A second
237 adjusted analysis was conducted for the subset of samples with additional data available (shared
238 sanitation, walk time to water source, antibiotic consumption in children, animal antibiotic
239 consumption, livestock ownership-any animal-, age, sex, finished walls, floors, roof, electricity
240 access, watch/clock, radio, TV, mobile phone, refrigerator, bike, motorcycle, car/truck, and clean

241 fuel). The association between each covariate and outcome variable was assessed; those with a
242 p-value < 0.2 were screened for collinearity. A subset of prescreened variables was included in
243 the final model to reduce collinearity. Our prespecified analysis plan is available on Open Science
244 Framework (<https://osf.io/n4z7f/>).

245
246 Associations between abundance of ARGs and access to improved water and sanitation were
247 conducted in subsets of the data separated by age (0-18 vs.18+), World Bank income
248 classification (low, lower-middle vs. upper-middle, high), WHO region, animal antibiotic usage (<
249 50 vs. \geq 50 mg/population-corrected unit, PCU), urbanicity (rural vs. urban), relative abundance
250 of *Enterobacteriaceae* (< 2% vs. \geq 2%), and human antibiotic usage (< 12 vs. \geq 12 DDD per 1000
251 persons per day).

252
253 For models of individual gene abundance, genes present in less than 5% of samples were
254 excluded and p-values were corrected for multiple comparisons using the Bonferroni method.

255
256 We used one-way ANOVA followed by post hoc Tukey HSD to compare abundance of ARGs and
257 log₁₀-transformed relative abundance of *Enterobacteriaceae* by region. To compare total ARGs
258 by urbanicity in each WHO region, we used student t-tests and a Bonferroni correction.

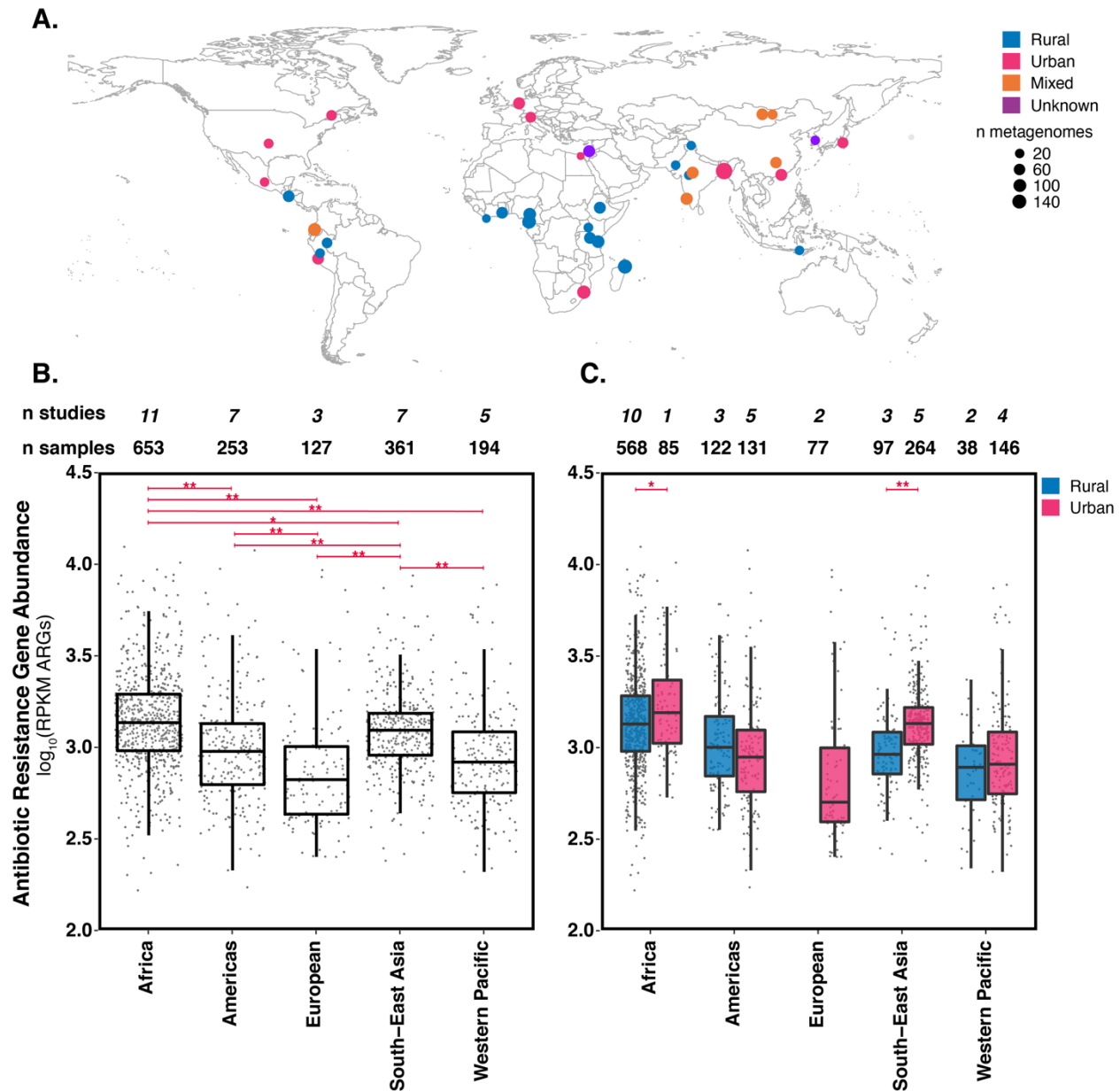
259
260

261 **Results**

262
263 Identified Study Characteristics

264 We identified 36 studies in total: 25 studies in the SRA containing metagenomes from LMICs,
265 three with metagenomes from LMICs from authors of this work, and eight from HICs. Three
266 studies (2 China, 1 Fiji) were excluded due to lack of data on both improved drinking water and

267 sanitation. One study (Brazil) was omitted due to a high relative abundance of non-bacterial taxa
 268 and no detectable ARGs in several metagenomes. In total, 1589 metagenomes from 26 countries
 269 (4 low, 8 lower-middle, 6 upper-middle, and 8 HIC) were included in the analysis (**Figure 2A**, and
 270 appendix Table A3).



271
 272 **Figure 2.** A) Location of studies included in this analysis. Urbanicity indicated by color and number of
 273 metagenomes indicated by size of dot. Additional study information in appendix Table A3. B) Abundance
 274 of antibiotic resistance genes in units of \log_{10} ARG reads per kilobase per million (RPKM) mapped reads
 275 classified as bacteria by WHO region. **Indicates $p < 0.001$ and * indicates $p < 0.05$ by ANOVA with post hoc
 276 Tukey HSD. C) Abundance of antibiotic resistance genes in rural and urban areas by WHO region.
 277 **Indicates $p < 0.001$ and * indicates $p < 0.05$ by student's t-test correcting for multiple comparisons. Boxplots

278 consist of 25th percentile, median, 75th percentile and whiskers extend to at most 1.5x the inner quartile
279 range. Eastern Mediterranean is not shown in B and C due to the presence of only 1 metagenome (Egypt).
280 Rural European is not shown in C, as no metagenomes were identified from rural Europe.
281

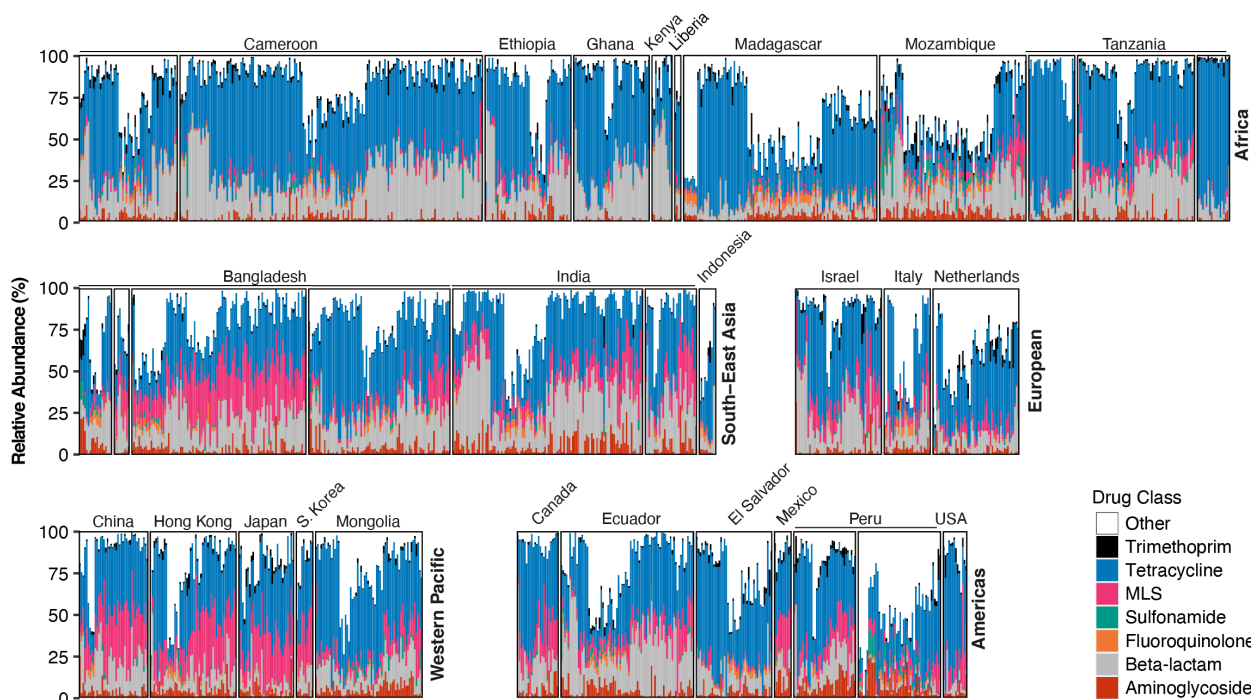
282 The most abundant bacterial families among all metagenomes were, *Prevotellaceae*,
283 *Bacteroidaceae*, *Oscillospiraceae* (commonly referred to as *Ruminococcaceae*),
284 *Lachnospiraceae*, and *Enterobacteriaceae* (Figure A1). On average, the most abundant family in
285 Africa and South-East Asia (SEA) was *Prevotellaceae* whereas Europe and the Western Pacific
286 were dominated by *Bacteroidaceae*. There was a high degree of variability in metagenomes in
287 the Americas with *Prevotellaceae* most abundant in El Salvador and Ecuador, *Bacteroidaceae*
288 most abundant in Canada, and *Oscillospiraceae* most abundant in Peru. Since
289 *Enterobacteriaceae* are known to harbor a disproportionately high number of ARGs and ARG
290 databases are dominated by *Enterobacteriaceae* alleles, we were interested in the variation in
291 relative abundance across WHO regions. The log₁₀-transformed average relative abundance of
292 *Enterobacteriaceae*, ranging from -0.03 ± 0.7 [mean \pm std] in Europe to 0.5 ± 0.6 in Africa, was
293 higher in Africa compared to all other WHO regions (one-way ANOVA $p < 0.001$; post hoc Tukey
294 HSD $p < 0.05$ for SEA and $p < 0.001$ for Americas, Europe, and Western Pacific). Notably, the
295 average relative abundance of *Enterobacteriaceae* was above 15% in Mozambique, Madagascar,
296 Italy, Brazil, Indonesia, and in one Bangladesh study.

297
298

299 Abundance of Antibiotic Resistance Genes

300 The average log₁₀-transformed, normalized abundance of ARGs was highest in Africa
301 (mean=3.14) compared to all other WHO regions (one-way ANOVA $p < 0.001$, post hoc Tukey
302 HSD $p < 0.001$ for all but SEA, $p = 0.03$ for SEA; **Figure 2B**), followed by SEA which was
303 significantly higher than the Americas, Europe, and Western Pacific ($p < 0.001$). Abundance trends
304 varied by drug class (appendix Figure A2). Beta-lactamase genes were highest in Africa and SEA
305 (appendix Figure A2). Tetracycline resistance was ubiquitous throughout all WHO regions (**Figure**

306 **3** and appendix Figure A3). Beyond genes conferring resistance to tetracyclines, the resistome
 307 (collection of all ARGs in a metagenome) was dominated by genes encoding resistance to beta-
 308 lactams in Africa, Europe, and some countries in the Americas. In SEA resistance to beta-lactams
 309 and macrolides, lincosamides, and streptogramins (MLS) were dominant, while MLS were
 310 abundant in the Western Pacific. Among metagenomes where urbanicity could be determined,
 311 there was a higher ARG abundance in urban areas in Africa and SEA (SEA $p < 0.001$, Africa
 312 $p = 0.01$; **Figure 1C**).



314 **Figure 3.** Relative abundance of antibiotic resistance genes by drug class. Metagenomes ordered by
 315 hierarchal clustering of relative abundance by drug class profiles and grouped by study as well as WHO
 316 region. Each bar represents the relative proportion of each drug class in one metagenome. Eastern
 317 Mediterranean is not shown due to the presence of only 1 metagenome (Egypt). Note: multiple studies
 318 conducted in Cameroon, Tanzania, Bangladesh, India, and Peru. MLS=Macrolide, lincosamide,
 319 streptogramin.

320

321 Association between ARG Abundance and Access to Improved Water and Sanitation

322 Overall, increased access to improved water and sanitation was associated with a decrease in
 323 ARG abundance (**Table 1**). A 100% increase in access to improved water and sanitation was
 324 associated with a 0.26 reduction in ARG abundance (-0.26 [-0.44, -0.08], $p < 0.01$) (estimate [95%

325 CI]) in the analysis adjusting for all covariates. A more realistic increase in access to water and
 326 sanitation of 25% was associated with a 0.07 reduction in ARG abundance. Relaxing the
 327 geographical linkage between survey data and reported sample location (using 50 and 75km radii)
 328 did not significantly impact results; nor did the incorporation of additional covariates where
 329 available in the adjusted analysis (appendix Tables A4 and A5). Across quartiles of water and
 330 sanitation access, significantly lower ARG abundance was found in regions with 75-100% access
 331 compared to 0-25% access (-0.17 [-0.34, 0.00], p=0.04).
 332

Table 1. Association between ARG abundance and combined access to both improved drinking water and sanitation. Adjusted analyses were conducted with WHO region and population density as covariates (Adjusted WHO Region and Population Density) and additional covariates of GDP per capita, antibiotic usage in humans, read length, and library layout (Adjusted All Covariates). Additional information on covariates provided in appendix Table A2. Effect estimates were generated across the full range and by quartiles of improved water and sanitation coverage. Robust standard errors were used to account for clustering (multiple metagenomes at one georeferenced cluster).

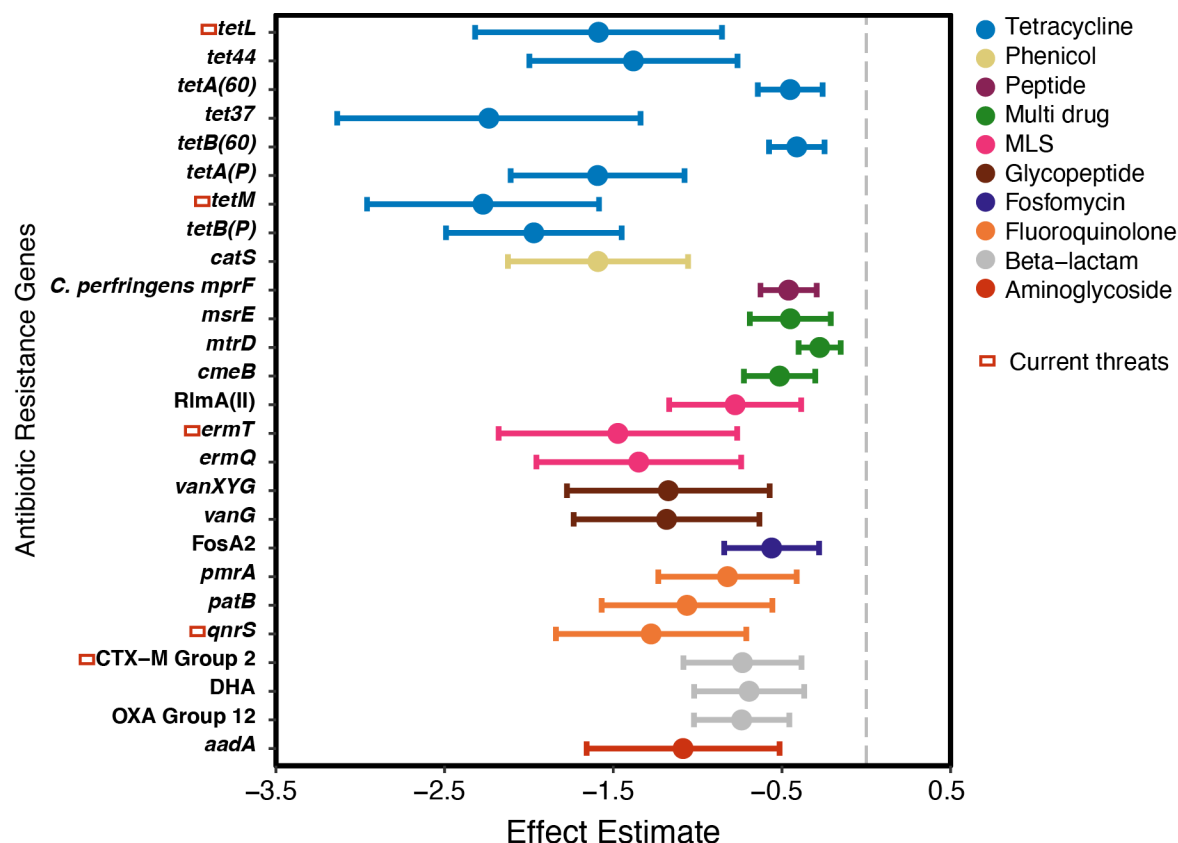
	Adjusted WHO Region and Population Density		Adjusted All Covariates		N
	Effect Estimate (95% CI)	p-value	Effect Estimate (95% CI)	p-value	
Water and Sanitation Access	-0.25 (-0.41, -0.10)	0.002	-0.26 (-0.44, -0.08)	0.005	1589
Quartiles					
0-25%	Reference		Reference		523
25-50%	-0.03 (-0.15, 0.09)	0.62	0.01 (-0.12, 0.14)	0.91	502
50-75%	-0.13 (-0.26, -0.01)	0.04	-0.09 (-0.24, 0.05)	0.20	171
75-100%	-0.19 (-0.36, -0.03)	0.02	-0.17 (-0.34, 0.00)	0.04	393

333
 334 The magnitude of association was higher for improved sanitation access alone compared to
 335 improved drinking water alone (sanitation: -0.16 [-0.32, 0.00] p=0.05; drinking water: -0.09 [-0.35,
 336 0.16], p=0.47; appendix Table A6).
 337
 338 Considering the most common drug classes (those detected in ≥80% of samples), we found
 339 increased access to improved water and sanitation was associated with a lower abundance of
 340 ARGs conferring resistance to tetracycline (-0.31 [-0.48, -0.13], p<0.001), fluoroquinolones (-0.74

341 [-1.40, -0.08], $p=0.03$) and trimethoprim (-1.00 [-1.70, -0.31], $p<0.01$) antibiotics (appendix Table
342 A7), while WASH was not associated with abundance of genes conferring resistance to beta-
343 lactams, aminoglycosides, or MLS ($p>0.05$ for all).

344
345 Improved water and sanitation was associated with a greater decrease in ARG abundance in
346 urban compared to rural areas (urban: -0.37 [-0.68, -0.07]; rural: -0.16 [-0.38, 0.07], $p_{\text{interaction}}=0.26$
347 **Table A8**). There was a greater, but non-significant, decrease, in higher-income compared to
348 lower-income countries, areas with higher antibiotic usage in animals, and areas with higher
349 antibiotic usage in humans. In the subset of metagenomes with higher relative abundance of
350 *Enterobacteriaceae*, there was a greater decrease in ARG abundance compared to
351 metagenomes with a lower relative abundance (RA) (RA $\geq 2\%$: -0.29 [-0.45, -0.12]; RA $< 2\%$: -
352 0.18 [-0.33, -0.03], $p_{\text{interaction}}=0.80$). Of the WHO regions investigated, improved water and
353 sanitation was associated with the greatest decrease in ARGs in the Americas (-0.56 [-0.91, -
354 0.21], $p<0.001$) and SEA (-0.28 [-0.52, -0.03], $p=0.03$).

355
356
357 Increased access to improved water and sanitation was associated with a decrease in the
358 abundance of 29 resistance genes conferring resistance to a range of antibiotic classes in the
359 analysis adjusted for all covariates (**Figure 4** and appendix Table A9). Of those 29 ARGs, *dfrA1*
360 (conferring resistance to trimethoprim), *tetL* and *tetM* (tetracycline), *ermT* (MLS), *qnrS*
361 (fluoroquinolone), and *bla*_{CTX-M} Group 2 (beta-lactamases) have been identified as current health
362 threats due to their presence in human-associated environments, carriage by mobile genetic
363 elements, and presence in the genomes of ESKAPE (*Enterococcus faecium*, *Staphylococcus*
364 *aureus*, *Klebsiella pneumoniae*, *Acinobacter baumannii*, *Pseudomonas aeruginosa*, and
365 *Enterobacter*) pathogens.²⁰



366
 367 **Figure 4.** Association between abundance of individual ARGs and combined access to both improved
 368 drinking water and sanitation. Effect estimate (change in \log_{10} ARG abundance for a 100% increase in
 369 access to improved water and sanitation) and 95% confidence intervals for antibiotic resistance genes
 370 present in at least 5% of metagenomes and significant at p -value < 0.05 in the fully adjusted model. Robust
 371 standard errors were used to account for clustering (multiple metagenomes at one georeferenced cluster).
 372 P -values were corrected for multiple comparisons using the Bonferroni method. Zero values were assigned
 373 a value of half the lowest gene abundance. ARGs that are enriched in human-associated environments,
 374 mobile, and present in ESKAPE pathogens are referred to as “Current threats”²⁰ and are indicated by the
 375 leftmost rectangle. Beta-lactam resistance genes included in each group are available in appendix Table
 376 A9. MLS=Macrolide, lincosamide, streptogramin.
 377

378 Discussion

379 We found that increased access to improved water and sanitation was associated with decreased
 380 ARG abundance, particularly in urban areas, and areas with higher antibiotic usage in humans
 381 and animals. We were unable to identify any metagenomes explicitly from rural sites in HICs,
 382 making it difficult to isolate the impact of urbanicity from economy type. Improved sanitation was
 383 associated with a greater decrease in ARG abundance compared to drinking water alone, but the
 384 greatest decrease was observed with access to both improved drinking water and sanitation. Our
 385 results suggest increasing access to improved water and sanitation at the community-level could

386 be an effective strategy to reduce the proliferation of antibiotic resistance. Additional studies are
387 needed to determine if there is a causal relationship between improving WASH and antibiotic
388 resistance burden. Also, we identified a relatively small number of human gut metagenomes from
389 around the world, especially in LMICs. There is a clear need for more metagenomic data from
390 LMICs where the burden of infectious illnesses is highest and microbiome-based research has
391 the potential to improve human health.

392
393 The abundance of ARGs was similar in magnitude in all WHO regions, but abundance was highest
394 in Africa. All studies in the WHO-defined African region were from sub-Saharan Africa, and thus
395 abundance estimates are not representative of North Africa which benefits from better WASH
396 conditions. Total ARGs were also highest in Africa using wastewater surveillance in urban areas,⁹
397 which is notable given most metagenomes included in our analysis from Africa were from rural
398 settings. We found the lowest abundance of ARGs in the Americas. Unlike Africa, the Americas
399 were relatively heterogeneous in terms of most abundant taxa (including *Enterobacteriaceae*) and
400 classes of ARGs. Previous urban wastewater surveillance efforts found the Americas had the
401 second highest ARG abundance,⁹ but significant variation in gut microbiomes by region and
402 urbanicity in the Americas could contribute to discrepancies in resistome characterization
403 between our work and others. Tetracycline resistance was abundant in all WHO regions,
404 consistent with a previous analysis of human gut metagenomes from 11 countries.²¹ The
405 widespread distribution of tetracycline resistance is not surprising given it has been used globally
406 for decades as a broad-spectrum antibiotic in humans and animals.²² Lastly, there was a relatively
407 high abundance of genes conferring resistance to beta-lactams in Africa and South-East Asia.
408 Traveler studies indicate that northern Africa and Asia are reservoirs of beta-lactam resistance,
409 including extended-spectrum β -lactamase-producing *Enterobacteriaceae*.²³

410

411 Our findings, in line with previous literature, suggest increasing access to improve water and
412 sanitation services within communities is a potential strategy to combat antibiotic resistance. A
413 metagenomic analysis of wastewater samples from 60 high- and low-income countries found that
414 lower open defecation rates, higher access to improved water sources in rural areas, and
415 increased water and sanitation investments were associated with decreased ARG abundance.⁹
416 We found similar associations between ARG abundance and water and sanitation access in our
417 analysis of human gut metagenomes from rural and urban areas in LMICs, which are likely poorly
418 represented in wastewater surveillance efforts given that 91% of rural and 37% of urban
419 populations are not connected to sewerage systems.¹⁵ Modeling studies have also reported
420 associations between antibiotic resistance and WASH. In a study of 73 countries around the world
421 with antibiotic consumption data, infrastructure improvements - including but not limited to
422 improved water and sanitation – were associated with reduced prevalence of antibiotic resistant
423 organisms (*Escherichia coli*, *Klebsiella* spp., and *S. aureus*) in clinical surveillance data.⁶

424
425 Similarly, combined access to both improved water and sanitation resulted in the greatest
426 reduction in ARG abundance in our analysis where we captured diversity of ARGs circulating in
427 community rather than clinical settings. A modelling-based investigation of the potential impact of
428 varying levels of sanitation improvements on abundance of ARGs in the environment in South-
429 East Asia found that moving from open defecation to improved sanitation, in general, provided
430 the greatest reduction in ARGs, with proportionally smaller reductions occurring as more
431 advanced wastewater treatment processes were added.²⁴ Lastly, in a country-specific,
432 experimental analysis (in contrast to the aforementioned global or WHO regional-scale,
433 observational studies) there was no association between WASH and ARG carriage, measured
434 using PCR one year after the intervention, among 120 children (<14 months old) in Maputo,
435 Mozambique.²⁵ Additional studies across diverse populations are needed to further investigate
436 the relationship between WASH and antibiotic resistance gene abundance in human guts. The

437 period of follow-up in future intervention studies should also consider amount of time needed to
438 clear pre-existing antibiotic resistance genes from the human gut.

439

440 An important limitation to our work is the observational, ecological design and the potential for
441 confounding. While our study adjusted for WHO region in all models to account for confounding
442 socioeconomic variables, it is possible our estimates are an overestimate of the association
443 between ARG abundance and water and sanitation. Other limitations include a lack of
444 standardized surveys collecting data on antibiotic consumption, both quantity and diversity, of
445 antibiotics across communities. Current sources are constrained to commercial sales data or
446 household reported consumption in sick children, both of which are limited in scope. Additionally,
447 *Enterobacteriaceae* growth in the absence of temperature control or preservation solutions during
448 sample collection and transport could impact our results, although most metagenomic studies
449 tightly control sample collection, transport, and testing. In the subgroup analysis there was a
450 stronger association between access to improved water and sanitation and ARG abundance in
451 metagenomes with higher relative abundance of *Enterobacteriaceae*. Also, identification of ARGs
452 through alignment-based methods is known to result in false positives. In a small subset of
453 metagenomes, we used ROCKER²⁶, a modeling-based strategy, to identify Class A beta-lactamase
454 genes and found 75% agreement between CARD identifications and ROCKER (data not shown).
455 While we chose to normalize our estimates of ARG abundance using total bacterial reads, this
456 does not account for bacterial genome size variation among samples and could provide higher
457 estimates when ARGs are carried on smaller genomes. Alternative approaches leveraging read
458 mapping to bacterial single copy core genes (SCGs)²⁷ attempt to correct for this, though may not
459 account for confounders such as plasmid copy number variability and incomplete carriage of the
460 chosen SCG set. To identify any potential discrepancies, we performed a secondary analysis
461 using a SCG normalization approach and did not find the results to differ considerably although
462 the magnitude of effect was somewhat diminished (SI pg 3). Finally, use of the CARD database

463 could overstate abundance of *Enterobacteriaceae* and presence of antibiotic resistance genes
464 does not imply phenotypic resistance.

465
466 Education and better regulated use of antimicrobials in humans and animals (i.e., stewardship)
467 are the main approaches to managing antibiotic resistance. Mitigating infectious diseases is also
468 a critical component of the WHO Global Action Plan on Antimicrobial Resistance, which
469 recommends strategies such as vaccinations and effective WASH.²⁸ However, the focus for
470 WASH has been on healthcare rather than community settings. In a review of the 77 action plans
471 on antimicrobial resistance in the WHO library, only 11 countries mention community-level
472 WASH.²⁹ While most surveillance on antimicrobial-resistant organisms has been limited to clinical
473 settings in LMICs, community acquired resistance is substantial.³⁰ Incorporating WASH
474 infrastructure in community settings into national strategies could be effective for curbing antibiotic
475 resistance in LMICs. We found the most significant reduction in ARGs was associated with access
476 to combined improvements in water and sanitation, suggesting comprehensive WASH access
477 may be more effective than single interventions.

478

479 **Contributors**

480 AJP, ERF, and MLN conceptualized the study. APH and ERF contributed to data curation. ERF
481 conducted the analysis which was replicated by APH and JMS. AJP, JB, and KL provided
482 additional data (metagenomes and/or metadata) beyond those publicly available at the time of
483 analysis. CP contributed to the conceptual diagram. KG conducted the literature review. ERF,
484 MLN, and KG contributed to the first draft of the manuscript. All co-authors provided input on
485 methods and edited the final manuscript. AJP and OC obtained funding for the study.

486

487 **Acknowledgments**

488 This work was supported by grant LSHTM INV-006556 from the Bill and Melinda Gates
489 Foundation. ERF was supported by the NSF Postdoctoral Research Fellowships in Biology
490 Program under Grant No. 1906957. MLN was supported by NIH award KL2TR002545 and the
491 Stuart B. Levy Center for Integrated Management of Antimicrobial Resistance at Tufts (Levy
492 CIMAR), a collaboration of Tufts Medical Center and the Tufts University Office of the Vice Provost
493 for Research (OVPR) Research and Scholarship Strategic Plan (RSSP). KL and GT contributions
494 were supported by NIAID grant numbers R01AI137679 and 1K01AI103544. AME and CJW's
495 contributions were supported by NIAID grant number U19AI110818. SK is supported by the
496 National Institute of Allergy and Infectious Diseases of the National Institutes of Health under
497 Award Number R01AI099525 and the Wellcome Trust Grant Number 215675-Z-19-Z. The
498 authors would also like to thank Konstantinos Konstantinidis for his consultation on bioinformatic
499 methods. Any opinions, findings, and conclusions or recommendations expressed in this material
500 are those of the author(s) and do not necessarily reflect the views of the funding organizations.

501

502 **Data Sharing**

503 SRA project numbers for studies used in this analysis are located in Table A3. Compiled data
504 used in our analysis are available on OSF.

505

506

507

508

509

510

511

512 References

- 513 1 Murray CJ, Ikuta KS, Sharara F, *et al.* Global burden of bacterial antimicrobial resistance in
514 2019: a systematic analysis. *The Lancet* 2022; **399**: 629–55.
- 515 2 U.S. Department of Health and Human Services, Centers for Disease Control and Prevention
516 (CDC). Antibiotic Resistance Threats in the United States, 2019. Atlanta, GA, 2019.
- 517 3 Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJ. The Structure and Diversity of
518 Human, Animal and Environmental Resistomes. *Microbiome* 2016; **4**: 54.
- 519 4 Kumarasamy KK, Toleman MA, Walsh TR, *et al.* Emergence of a new antibiotic resistance
520 mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological
521 study. *Lancet Infect Dis* 2010; **10**: 597–602.
- 522 5 Liu Y-Y, Wang Y, Walsh TR, *et al.* Emergence of plasmid-mediated colistin resistance
523 mechanism MCR-1 in animals and human beings in China: a microbiological and molecular
524 biological study. *Lancet Infect Dis* 2016; **16**: 161–8.
- 525 6 Collignon P, Beggs JJ, Walsh TR, Gandra S, Laxminarayan R. Anthropological and
526 socioeconomic factors contributing to global antimicrobial resistance: a univariate and
527 multivariable analysis. *Lancet Planet Health* 2018; **2**: e398–405.
- 528 7 Nadimpalli ML, Marks SJ, Montealegre MC, *et al.* Urban informal settlements as hotspots of
529 antimicrobial resistance and the need to curb environmental transmission. *Nat Microbiol*
530 2020; **5**: 787–95.
- 531 8 Lipsitch M, Samore MH. Antimicrobial Use and Antimicrobial Resistance: A Population
532 Perspective. *Emerg Infect Dis* 2002; **8**: 347–54.
- 533 9 Hendriksen RS, Munk P, Njage P, *et al.* Global monitoring of antimicrobial resistance based
534 on metagenomics analyses of urban sewage. *Nat Commun* 2019; **10**: 1124.
- 535 10 Jones ER, van Vliet MTH, Qadir M, Bierkens MFP. Country-level and gridded estimates of
536 wastewater production, collection, treatment and reuse. *Earth Syst Sci Data* 2021; **13**: 237–
537 54.
- 538 11 Alcock BP, Raphenya AR, Lau TTY, *et al.* CARD 2020: antibiotic resistance surveillance with
539 the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2020; **48**: D517–25.
- 540 12 Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat*
541 *Methods* 2015; **12**: 59–60.
- 542 13 Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in
543 metagenomics data. *PeerJ Comput Sci* 2017; **3**: e104.
- 544 14 Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*
545 2019; **20**: 257.

- 546 15WHO and UNICEF. Progress on drinking-water, sanitation and hygiene: 2017 update and
547 SDG baselines. 2017 <https://www.who.int/publications-detail-redirect/9789241512893>
548 (accessed June 23, 2021).
- 549 16Browne AJ, Chipeta MG, Haines-Woodhouse G, *et al.* Global antibiotic consumption and
550 usage in humans, 2000–18: a spatial modelling study. *Lancet Planet Health* 2021; **5**: e893–
551 904.
- 552 17The World Bank. GDP (current US\$) | Data.
553 <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD> (accessed June 23, 2021).
- 554 18Center For International Earth Science Information Network-CIESIN-Columbia University.
555 Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. 2017.
556 DOI:10.7927/H49C6VHW.
- 557 19Antibiotic use in livestock. Our World Data. [https://ourworldindata.org/grapher/antibiotic-use-](https://ourworldindata.org/grapher/antibiotic-use-in-livestock)
558 [in-livestock](https://ourworldindata.org/grapher/antibiotic-use-in-livestock) (accessed June 17, 2021).
- 559 20Zhang A-N, Gaston JM, Dai CL, *et al.* An omics-based framework for assessing the health
560 risk of antimicrobial resistance genes. *Nat Commun* 2021; **12**: 4765.
- 561 21Feng J, Li B, Jiang X, *et al.* Antibiotic resistome in a large-scale healthy human gut microbiota
562 deciphered by metagenomic and network analyses. *Environ Microbiol* 2018; **20**: 355–68.
- 563 22Roberts MC. Tetracycline Therapy: Update. *Clin Infect Dis* 2003; **36**: 462–7.
- 564 23Ruppé E, Andremont A, Armand-Lefèvre L. Digestive tract colonization by multidrug-resistant
565 Enterobacteriaceae in travellers: An update. *Travel Med Infect Dis* 2018; **21**: 28–35.
- 566 24Graham DW, Giesen MJ, Bunce JT. Strategic Approach for Prioritising Local and Regional
567 Sanitation Interventions for Reducing Global Antibiotic Resistance. *Water* 2019; **11**: 27.
- 568 25Berendes D, Knee J, Sumner T, *et al.* Gut carriage of antimicrobial resistance genes among
569 young children in urban Maputo, Mozambique: Associations with enteric pathogen carriage
570 and environmental risk factors. *PloS One* 2019; **14**: e0225464.
- 571 26Orellana LH, Rodriguez-R LM, Konstantinidis KT. ROcker: accurate detection and
572 quantification of target genes in short-read metagenomic data sets by modeling sliding-
573 window bitscores. *Nucleic Acids Res* 2017; **45**: e14.
- 574 27Lee K, Kim D-W, Lee D-H, *et al.* Mobile resistome of human gut and pathogen drives
575 anthropogenic bloom of antibiotic resistance. *Microbiome* 2020; **8**: 2.
- 576 28World Health Organization. Global action plan on antimicrobial resistance. 2015
577 <https://www.who.int/publications-detail-redirect/9789241509763> (accessed Aug 26, 2021).
- 578 29Essack S. Water, sanitation and hygiene in national action plans for antimicrobial resistance.
579 *Bull World Health Organ* 2021; **99**: 606–8.

580 30 Ingle DJ, Levine MM, Kotloff KL, Holt KE, Robins-Browne RM. Dynamics of antimicrobial
581 resistance in intestinal *Escherichia coli* from children in community settings in South Asia and
582 sub-Saharan Africa. *Nat Microbiol* 2018; **3**: 1063–73.

583